JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

VOLUME 12, NOMOR 1, JUNI 2020 ISSN 2086 – 4132
AKREDITASI NOMOR: 747/Akred/P2MI-LIPI/04/2016

PUSAT PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT
POLITEKNIK STATISTIKA STIS

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

Jurnal "Aplikasi Statistika dan Komputasi Statistik" memuat karya ilmiah hasil penelitian dan kajian teori statistik dan komputasi statistik yang diterapkan khususnya pada bidang ekonomi dan sosial kependudukan, serta teknologi informasi yang terbit dua kali dalam setahun setiap bulan Juni dan Desember.

Alamat Redaksi:
Politeknik Statistika STIS
Jl. Otto Iskandardinata 64C
Jakarta Timur 13330
Telp. 021-8191437

PENGANTAR REDAKSI

Puji syukur kehadirat Allah, Tuhan Yang Maha Esa, "Jurnal Aplikasi Statistika dan Komputasi Statistik" Volume 12, Nomor 1, Juni 2020 dapat diterbitkan. Jurnal ilmiah ini dapat terwujud atas partisipasi semua pihak, penulis internal dilingkungan Politeknik Statistika STIS maupun penulis eksternal, serta mitra bestari.

Semoga artikel dalam jurnal ini dapat menambah pengetahuan para pembaca tentang penggunaan metode statistika serta komputasi statistik pada berbagai jenis data. Redaksi terus menunggu artikel-artikel ilmiah selanjutnya dari Bapak/Ibu agar publikasi yang dihasilkan menjadi salah satu sarana untuk memberikan sosialisasi statistika bagi masyarakat.

Jakarta, Juni 2020

Ketua Dewan Redaksi,

Dr. Nasrudin

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

VOLUME 12, NOMOR 1, JUNI 2020
AKREDITASI NOMOR: 747/Akred/P2MI-LIPI/04/2016

DAFTAR ISI

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

(*Journal of Statistical Application & Statistical Computing*)

ISSN 2086 – 4132 **Volume 12, Nomor 1, Juni 2020**

DDC: 315.98

Iut Tri Utami, Fadjryani, dan Diah Daniaty

Perbandingan Klasifikasi Status Pendonor Darah dengan Menggunakan Regresi Logistik dan K-Nearest Neighbor

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 1 – 12

Abstrak

Donor Darah Sukarela (DDS) adalah orang yang dengan sukarela mentranfusikan darahnya kepada orang lain. Seseorang dapat menjadi pendonor darah jika memenuhi kriteria dari PMI dan lolos dalam pemeriksaan dokter. Syarat yang diberlakukan PMI menyebabkan calon pendonor darah dapat diklasifikasikan menjadi layak dan tidak layak dalam mendonorkan darahnya. Salah satu cara untuk menentukan pola prediksi status kelayakan calon pendonor darah di PMI adalah menggunakan regresi logistik biner dan k-Nearest Neighbor (kNN). Peubah yang signifikan mempengaruhi kelayakan calon pendonor darah adalah kadar Haemoglobin. Akurasi yang dihasilkan oleh metode regresi logistik biner dan kNN pada penelitian ini adalah 93% dan 79%.

Kata kunci: DDS, regresi logistik biner, k-Nearest Neighbor

DDC: 315.98

Nur'eni dan Lilies Handayani

Regresi Probit untuk Analisis Variabel-Variabel yang Mempengaruhi Perceraian di Sulawesi Tengah

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 13 – 22

Abstrak

Sulawesi Tengah adalah salah satu Provinsi di Indonesia yang memiliki permasalahan dalam perceraian. Tingkat perceraian di Sulawesi Tengah pada tahun 2016 sebesar 2,44%. Persentase tingkat perceraian di Sulawesi Tengah ini menjadi tingkat perceraian ketiga tertinggi di Indonesia. Pada penelitian ini diteliti faktor-faktor yang mempengaruhi kasus perceraian di Sulawesi Tengah. Metode yang digunakan adalah regresi probit biner dengan variabel respon adalah status perkawinan. Hasil penelitian menunjukkan bahwa variabel prediktor yang mempengaruhi perceraian secara signifikan di Provinsi Sulawesi Tengah adalah umur kawin pertama (X_2) kategori 1 (18-21 tahun) dan kategori 2 (>21 tahun), tingkat pendidikan (X_3) kategori 1 (SD) dan kategori 4 (di atas SMA), daerah tempat tinggal (X_4) kategori 1 (kota) dan jumlah pengeluaran rumah tangga (X_6) dengan tingkat ketepatan klasifikasi model sebesar 99,2%.

Kata kunci: regresi probit biner, perceraian, status perkawinan, ketepatan klasifikasi

DDC: 315.98

Yulinda Nurul Aeni

Proyeksi Penyerapan Tenaga Kerja Perikanan Berdasarkan Faktor Industrialisasi Menggunakan Metode Fungsi Transfer

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 23 – 36

Abstrak
FAO menempatkan Indonesia sebagai negara dengan potensi perikanan terbesar di dunia, namun potensi tersebut belum dimanfaatkan dengan optimal. Pemerintah telah membentuk program industrialisasi perikanan, namun pelaksanaannya yang belum terimplementasi dengan baik menyebabkan penyerapan tenaga kerja di sektor ini masih rendah. Penelitian ini akan membentuk proyeksi penyerapan tenaga kerja subsektor perikanan 2019-2024 menggunakan metode fungsi transfer dengan mempertimbangkan faktor industrialisasi perikanan sebagai prediktor terhadap indeks elastisitas penyerapan tenaga kerja perikanan. Industrialisasi perikanan diukur berdasarkan faktor perkembangan investasi dan pertumbuhan jumlah perusahaan perikanan. Hasil proyeksi menunjukkan bahwa industrialisasi perikanan di tahun mendatang belum mampu mendorong respon pertumbuhan penyerapan tenaga kerja subsektor perikanan.

Kata kunci: fungsi transfer, industrialisasi, perikanan, proyeksi, tenaga kerja.

DDC: 315.98

Vira Wahyuningrum

Penerapan Radial Basis Function Neural Network dalam Pengklasifikasian Daerah Tertinggal di Indonesia

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 37 – 54

Abstrak
Penetapan daerah tertinggal di Indonesia merupakan kasus pengklasifikasian dengan dua kategori pada variabel respon (biner). Pengklasifikasian dengan metode klasifikasi linier yang umum digunakan yaitu regresi logistik pada tahap eksplorasi data menghasilkan *misclassification* yang relatif besar, sehingga diperlukan suatu metode alternatif. Artificial Neural Network (ANN) merupakan alternatif yang menjanjikan untuk berbagai metode klasifikasi konvensional. Radial Basis Function Neural Network (RBFNN) merupakan salah satu arsitektur ANN yang populer digunakan dalam klasifikasi. Metode RBFNN menggunakan dua pendekatan yaitu *supervised* dan *unsupervised* serta dalam beberapa penelitian menghasilkan akurasi klasifikasi yang tinggi. Penelitian ini bertujuan menerapkan metode RBFNN untuk kasus klasifikasi daerah tertinggal di Indonesia untuk melihat arsitektur RBFNN yang terbentuk dan ketepatan klasifikasi yang dihasilkan. Hasil dari penelitian ini adalah penerapan RBFNN memberikan performa yang sangat baik yaitu nilai akurasi sebesar 93,48 persen, sensitivitas 81,10 persen dan spesifikasi 97,43 persen. Nilai *F-Measure* arsitektur RBFNN yang dihasilkan mencapai 85,36 persen.

Kata kunci: Neural Network, Radial Basis Function, klasifikasi, daerah tertinggal

DDC: 315.98

Muhammad Syafiudin dan Turro S. Wongkaren

Analisis Permintaan Pangan dan Nonpangan Rumah Tangga dengan Gangguan Kesehatan di Indonesia

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 55 – 70

Abstrak
Penelitian ini bertujuan menganalisis dampak tidak langsung gangguan kesehatan terhadap permintaan pangan dan non pangan rumah tangga. Dengan menggunakan data Susenas Panel tahun 2012 dan 2013 dan menerapkan *two step heckman selection model* untuk estimasi pendapatan dan *seemingly unrelated regression estimator* untuk estimasi konsumsi rumah tangga. Hasilnya menunjukkan bahwa gangguan kesehatan kepala rumah tangga akan menurunkan pendapatannya. Dampak ini akan lebih dirasakan oleh rumah tangga perempuan miskin dan bekerja di sektor pertanian. Penurunan pendapatan ini menyebabkan porsi pengeluaran konsumsi non pangan menurun, khususnya untuk pengeluaran pemeliharaan perumahan, namun pengeluaran untuk perawatan tubuh justru meningkat. Sedangkan untuk porsi konsumsi pangan tidak terpengaruh. Hal ini menunjukkan bahwa, gangguan kesehatan dapat menyebabkan penurunan tingkat kesejahteraan rumah tangga karena menyebabkan penurunan pendapatan dan peningkatan pengeluaran kesehatan. Oleh karena itu, diperlukan kebijakan yang dapat melindungi kesejahteraan rumah tangga ketika mengalami gangguan kesehatan, bisa berupa subsidi biaya kesehatan atau *cash transfer*.

Kata kunci: Gangguan Kesehatan, Konsumsi, Pangan, Non Pangan, Pendapatan Rumah Tangga

DDC: 315.98

Dara Puspita Anggraeni, Dedi Rosadi, Hermansah, dan Ahmad Ashril Rizal

Prediksi Harga Emas Dunia di Masa Pandemi Covid-19 Menggunakan Model ARIMA

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 71 – 84

Abstrak
Penelitian ini bertujuan memodelkan serta memprediksi harga emas dunia di masa pandemi COVID-19. Penelitian ini juga hanya memasukkan nilai masa lampau dari harga emas dunia tanpa adanya pengaruh faktor eksogen(independen) pada model. Model yang dipergunakan adalah model Autoregressive Integrated Moving Average (ARIMA). Adapun data yang dipergunakan pada permodelan sebanyak 240 data observasi dimana data merupakan data bulanan harga emas dunia bulan Agustus 2000 hingga Juli 2020. Model terbaik untuk harga emas dunia ini adalah ARIMA(0,1,1) dengan nilai Mean Absolute Percentage Error (MAPE) sebesar 3,70%. Hasil prediksi harga emas dunia untuk bulan Agustus 2020 hingga Januari 2021 berturut-turut adalah sebesar 1930,046; 1945,651; 1961,381; 1977,240; 1993,227; 2009,343 US$/Troy Ons emas. Prediksi ini menunjukkan tren naik dengan rata-rata peningkatan selama periode tersebut (Agustus 2020-Januari 2021) sebesar15,8594 US$/Troy ons per bulannya.

Kata kunci: Harga Emas, COVID-19, ARIMA, Prediksi

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

(*Journal of Statistical Application & Statistical Computing*)

ISSN 2086 – 4132 **Volume 12, Nomor 1, Juni 2020**

DDC: 315.98

Iut Tri Utami, Fadjryani, dan Diah Daniaty

Perbandingan Klasifikasi Status Pendonor Darah dengan Menggunakan Regresi Logistik dan K-Nearest Neighbor

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 1 – 12

Abstract

Voluntary Blood Donors are people who voluntarily transfer their blood to others. A person can become a blood donor if he meets the criteria of PMI and passes the doctor's examination The requirements imposed by PMI can cause prospective blood donors to be classified as feasible and improper in donating blood. One way to determine the pattern of predicting the eligibility status of prospective blood donors at PMI is to use k-Nearest Neighbor (kNN) and binary logistic regression. The significant variable influencing the eligibility of prospective blood donors is hemoglobin levels. The accuracy produced by the binary logistic regression and kNN method in this study was 93% and 79%.

Keywords: Voluntary blood donors, binary logistic regression, k-Nearest Neighbor

DDC: 315.98

Nur'eni dan Lilies Handayani

Determinan Partisipasi Sekolah Anak Penyandang Disabilitas di Indonesia Tahun 2015

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 13 – 22

Abstract

Central Sulawesi is one of province in Indonesia which has divorce trouble. The divorce rate in Central Sulawesi in 2016 is 2,44%, the third highest divorce rate in Indonesia. This research aimed the influence factors of divorce cases in Central Sulawesi by using binary probit regression with the respon variable is marriage status. The result shows that the predictor variable which influence the divorce in Central Sulawesi are age of first marriage (X_2) with age category 18-21 years and over 21 years, education level (X_3) with elementary school and over high school category, habitation (X_4) with urban categoryand household expanditure (X_6) with the accuracy rate is 99,2%.

Keywords: binary probit regression, divorce, marriage status, accuracy rate

DDC: 315.98

Yulinda Nurul Aeni

Proyeksi Penyerapan Tenaga Kerja Perikanan Berdasarkan Faktor Industrialisasi Menggunakan Metode Fungsi Transfer

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 23 – 36

Abstract
FAO places Indonesia as the country with the largest fishery potential in the world, but this potential has not been utilized optimally. The government had established a fisheries industrialization, but this program had not been implemented properly, resulting in low employment in this sector. This research would form projections of labor absorption in the fisheries subsector 2019-2024 using the transfer function method by considering the factor of fisheries industrialization as a predictor of the fisheries labor absorption elasticity index. The industrialization of fisheries is measured based on investment development factors and growth in the number of fishing companies. The projection results showed that the industrialization of fisheries in the coming year had not been able to encourage a response to the growth of labor absorption in the fisheries subsector.

Keywords: transfer function, industrialization, fisheries, projection, labor

DDC: 315.98

Vira Wahyuningrum

Penerapan Radial Basis Function Neural Network dalam Pengklasifikasian Daerah Tertinggal di Indonesia

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 37 – 54

Abstract
Determination of underdeveloped regency in Indonesia is the case with the classification of two categories on the response variable (binary). Classification with the linear classification method that is commonly used is logistic regression at the data exploration stage resulting in relatively large misclassification, so we need an alternative method. Artificial Neural Network (ANN) is a promising alternative to various conventional classification methods. Radial Basis Function Neural Network (RBFNN) is one of the popular ANN architectures used in classification. The RBFNN method uses two approaches namely supervised and unsupervised and in several studies produces high classification accuracy. This study aims to apply the RBFNN method for the classification case of underdeveloped regency in Indonesia to see the RBFNN architecture formed and the resulting classification accuracy. The results of this study are the application of RBFNN provides an excellent performance that is an accuracy value of 93.48 percent, sensitivity 81.10 percent and 97.43 percent specifications. The F-Measure value of RBFNN architecture reached 85.36 percent.

Keywords: Neural Network, Radial Basis Function, Classification, Underdeveloped regency

DDC: 315.98

Muhammad Syafiudin dan Turro S. Wongkaren

Analisis Permintaan Pangan dan Nonpangan Rumah Tangga dengan Gangguan Kesehatan di Indonesia

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 55 - 70

Abstract

This study aims to analyze the indirect effects of health problems on household food and non-food demand. This research uses Susenas Panel data for 2012 and 2013 and applies the ' two step heckman selection model' to estimate income, and 'seemingly unrelated regression estimator' to estimate household consumption. The results show that health issues or problems of the household will decrease their income. This impact will be worse experienced by the poor female household and work in the agricultural sector. The decrease in income has led to a decrease in non-food expenditure, especially on housing maintenance; but on the contrary, expenditure for body care has increased. However, it has not affected the expenditure on food consumption. This finding shows that health problems would be lowering household welfare as decreasing income and increasing health expenditure. Therefore, it is necessary to formulate policies to protect household welfare directly affected by health problems, for example by providing health subsidy or cash transfer.

Keywords: illhealth, illness, consumption, food and Non-food, household income

DDC: 315.98

Dara Puspita Anggraeni, Dedi Rosadi, Hermansah, dan Ahmad Ashril Rizal

Prediksi Harga Emas Dunia di Masa Pandemi Covid-19 Menggunakan Model ARIMA

Jurnal Aplikasi Statistika & Komputasi Statistik, Volume 12, Nomor 1, Juni 2020, hal 71 – 84

Abstract

This research aims to model and predict the world gold price during the COVID-19 pandemic. This research only includes the past values of world gold prices without the influence of exogenous (independent) factors on the model. The model used in this research is the Autoregressive Integrated Moving Average (ARIMA). The data used in the modeling are 240 observational data which were the monthly data on world gold prices from August 2000 to July 2020. The best model for this world gold price is ARIMA (0,1,1) with a Mean Absolute Percentage Error (MAPE) value of 3.70%. The prediction results of the world gold price from August 2020 to January 2021 are 1930,046 respectively; 1945,651; 1961,381; 1977,240; 1993,227; 2009,343 US $ / Troy Ounce of gold. This prediction shows an upward trend with the average increase 15,8594 US $ / Troy ounce per month during that period (August 2020-January 2021).

Keywords: Gold Price, COVID-19, ARIMA, Prediction

PERBANDINGAN KLASIFIKASI STATUS PENDONOR DARAH DENGAN MENGGUNAKAN REGRESI LOGISTIK DAN K-NEAREST NEIGHBOR

Iut Tri Utami[1], Fadjryani[2], Diah Daniaty[3]

Program Studi Statistika Jurusan Matematika FMIPA Universitas Tadulako[1,2], Badan Pusat Statistik[3]
e-mail: [1]triutami.iut@gmail.com, [2]fadjryani_mipauntad@yahoo.com, [3]diahdaniaty@gmail.com

Abstrak

Donor Darah Sukarela (DDS) adalah orang yang dengan sukarela mentranfusikan darahnya kepada orang lain. Seseorang dapat menjadi pendonor darah jika memenuhi kriteria dari PMI dan lolos dalam pemeriksaan dokter. Syarat yang diberlakukan PMI menyebabkan calon pendonor darah dapat diklasifikasikan menjadi layak dan tidak layak dalam mendonorkan darahnya. Salah satu cara untuk menentukan pola prediksi status kelayakan calon pendonor darah di PMI adalah menggunakan regresi logistik biner dan k-Nearest Neighbor (kNN). Peubah yang signifikan mempengaruhi kelayakan calon pendonor darah adalah kadar Haemoglobin. Akurasi yang dihasilkan oleh metode regresi logistik biner dan kNN pada penelitian ini adalah 93% dan 79%.

Kata kunci: DDS, regresi logistik biner, k-Nearest Neighbor

Abstract

Voluntary Blood Donors are people who voluntarily transfer their blood to others. A person can become a blood donor if he meets the criteria of PMI and passes the doctor's examination The requirements imposed by PMI can cause prospective blood donors to be classified as feasible and improper in donating blood. One way to determine the pattern of predicting the eligibility status of prospective blood donors at PMI is to use k-Nearest Neighbor (kNN) and binary logistic regression. The significant variable influencing the eligibility of prospective blood donors is hemoglobin levels. The accuracy produced by the binary logistic regression and kNN method in this study was 93% and 79%.
.

***Keywords**: Voluntary blood donors, binary logistic regression, k-Nearest Neighbor*

PENDAHULUAN

Latar Belakang

Donor darah mempunyai manfaat yang baik bagi kesehatan tubuh. Sayangnya, banyak orang takut donor darah dengan beragam alasan mulai dari takut jarum suntik, lemas dan kehabisan darah. Manfaat menyumbangkan darah tidak hanya dirasakan bagi penerima darah tetapi juga bagi pendonor darah. Ada berbagai hal positif yang bisa didapatkan ketika mendonorkan darah, mulai dari membantu membakar kalori, menjaga kesehatan jantung, meningkatkan produksi sel darah, hingga menurunkan risiko kanker. Sayangnya, jumlah pendonor darah masih belum banyak dan belum memenuhi target kebutuhan darah nasional.

Demi kesehatan tubuh pendonor dan penerima darah, tidak semua orang dapat mendonorkan darahnya. Beberapa hal yang disyaratkan PMI saat mau mendonorkan darahnya yaitu umur 17-60 tahun, berat minimal 45 kg, temperatur tubuh 36,6°C-37,5°C yang diukur secara oral, tekanan darah baik (Sistole = 110-160 mm Hg dan Diastole = 70-100 mm Hg), denyut nadi teratur 50-100 kali per menit, dan Hemoglobin wanita minimal = 12 gr % sedangkan pria minimal = 12,5 gr % (Depkes RI, 2009). Syarat yang diberlakukan PMI menyebabkan calon pendonor darah dapat diklasifikasikan menjadi layak dan tidak layak dalam mendonorkan darahnya (PMI,2009). Salah satu cara untuk menentukan pola prediksi calon pendonor darah di PMI adalah dengan menggunakan regresi logistik dan *k-Nearest Neighbor* (kNN).

Hosmer dan Lemeshow (2000) menjelaskan bahwa metode regresi logistik adalah suatu metode analisis statistika yang menganalisis hubungan antara peubah respon yang memiliki dua kategori atau lebih dengan satu atau lebih peubah penjelas. Salah satu model regresi logistik adalah regresi logistik biner. Metode regresi logistik biner juga dapat digunakan untuk menganalisis nilai ketepatan klasifikasi.

kNN merupakan suatu metode mengelompokkan suatu objek dengan mempertimbangkan kelas terdekat dari objek tersebut. kNN merupakan metode klasifikasi yang sangat sederhana, efisien dan efektif dalam bidang pengenalan pola, kategori teks, pengolahan objek dan mampu melakukan training data dalam jumlah yang besar (Bathia,2010). Meskipun sederhana, kNN dianggap menjadi salah satu dari sepuluh algoritma klasifikasi data mining yang terbaik (Wu et al, 2008). Salah satu masalah dari algoritma ini adalah efek yang sama terjadi dari semua atribut yang terdapat pada data baru dan data lama dalam data set pelatihan (Moradian dan Baraani, 2009).

Penelitian terdahulu tentang pengklasifikasian calon pendonor darah telah banyak dilakukan. Nugroho, dkk (2018) melakukan penelitian tentang klasifikasi pendonor darah menggunakan metode Support Vector Machine (SVM) pada dataset RFMTC yang menghasilkan akurasi sebesar 72.64%. Penelitian lain dilakukan Bayususetyo,dkk (2017) dengan mengklasifikasi calon pendonor darah dengan menggunakan metode Naive Bayes *Classifier* dengan studi kasus PMI di Kota Semarang. Sapriana (2017) menerapkan algoritma Naive Bayes *Classifier* pada klasifikasi status kelayakan pendonor darah di Unit Transfusi Darah Palang Merah Indonesia Kota Makassar dengan akurasi yang dihasilkan sebesar 90%.

Pada penelitian ini akan mengkaji ketepatan klasifikasi dan faktor-faktor yang mempengaruhi status kelayakan calon pendonor darah dengan menggunakan metode regresi logistik biner dan kNN. Perbandingan kedua metode dapat dilihat dari nilai akurasi nya. Semakin tinggi akurasi yang dihasilkan maka model semakin tepat dalam mengklasifikasikan. Adapun peubah penjelas yang akan digunakan adalah umur, jenis kelamin, berat badan, kadar HB, dan tekanan darah.

METODOLOGI

Tinjauan Referensi

Status Kelayakan Calon Pendonor Darah

Donor darah tidak hanya menguntungkan bagi penerima darah, tapi juga bermanfaat untuk pendonor. Tidak semua orang bisa mendonorkan darahnya, ada beberapa syarat donor darah yang perlu diketahui.

Syarat donor darah yang paling utama adalah kondisi fisik harus sehat. Usia juga menjadi syarat donor darah yaitu 17-60 tahun. Namun, untuk remaja usia 17 tahun diperbolehkan menjadi donor darah apabila mendapat izin tertulis dari orangtua. Calon pendonor baru dikatakan layak jika lolos pemeriksaan kesehatan sebelum mendonorkan darah.

Pendonor harus memiliki berat badan minimal 45 kilogram dan dalam kondisi sehat, baik jasmani maupun rohani. Selain itu, pendonor harus memiliki suhu tubuh 36, 6oC - 37,5oC. Tekanan darah pendonor harus berada pada angka 100-160 untuk *sistole* dan 70-100 untuk *diastole*. Denyut nadi saat pemeriksaan juga harus sekitar 50-100 kali per menit. Sementara itu, kadar haemoglobin pendonor harus minimal 12 gr/dL untuk wanita, dan minimal 12,5 gr/dL untuk pria (PMI, 2019).

Pendonor dapat mendonorkan darahnya paling banyak lima kali setahun dengan jangka waktu sekurang-kurangnya tiga bulan. Calon pendonor darah menjalani pemeriksaan pendahuluan, seperti kondisi berat badan, kadar HB, golongan darah, dan dilanjutkan dengan pemeriksaan dokter.

1. Sumber Data

Data yang akan digunakan pada penelitian ini adalah data sekunder yang diambil dari Unit Donor Darah (UDD) PMI Kabupaten Sigi, Provinsi Sulawesi Tengah. Jumlah data sebanyak 101 orang. Data akan dibagi menjadi dua yaitu data *training* dan data *testing*. Data *training* yang digunakan pada penelitian ini adalah 70% dari data keseluruhan, sisanya sebagai data *testing*. *Software* yang digunakan untuk mengolah data pada penelitian ini adalah R dengan paket *Class*, *Caret* dan *MASS*. Peubah penelitian yang akan digunakan pada penelitian ini adalah :

Tabel 1. Data Atribut

No	Atribut	Keterangan
1	Status Donor (Y)	1 = Layak Donor 2 = Tidak Layak
2	Jenis Kelamin (X_1)	1 = Laki-Laki 2 = Perempuan
3	Usia (X_2)	Numerikal
4	Berat Badan (X_3)	Numerikal
5	Tinggi Badan (X_4)	Numerikal
6	*Sistole* (X_5)	Tekanan darah atas
7	*Diastole* (X_6)	Tekanan darah bawah
8	Kadar HB(X_7)	Numerikal

2. Analisis Data

Tahapan analisis data yang dilakukan dalam penelitian ini yaitu :

a) Pengambilan data yang dilakukan di UDD PMI Kabupaten Sigi, Provinsi Sulawesi Tengah
b) Membagi data menjadi dua yaitu data *training* dan data *testing*. Data *training* digunakan untuk pembentukan model dan data *testing* digunakan untuk validasi model.
c) Pembentukan model kNN dan regresi logistik biner dengan menggunakan data *training*.
d) Menentukan prediksi klasifikasi calon pendonor darah dengan menggunakan data *testing*.
e) Membandingkan tingkat akurasi yang dihasilkan dari metode kNN dan regresi logistik biner.
f) Menentukan metode terbaik untuk mengklasifikasi calon pendonor darah.
g) Menarik kesimpulan

Langkah-langkah analisis data dengan menggunakan regresi logistik biner adalah :

a. Melakukan analisis deskriptif karakteristik status pendonor dan faktor-faktor yang mempengaruhi status kelayakan calon pendonor darah.
b. Menentukan model awal regresi logistik biner.

Menurut Agresti (2002) untuk menentukan estimasi parameter regresi logistik biner dapat digunakan metode

Maximum Likelihood (kemungkinan maksimum) yang membutuhkan turunan pertama dan turunan kedua dari fungsi Likelihood. Secara umum, model regresi logistik biner dengan E(Y=1|x) dapat dituliskan dengan:

$$\pi(X) = \frac{e^{\beta_0+\beta_1 x_1+\cdots+\sum_{k=1}^{n}\beta_{ik}D_{i+\beta_p x_p}}}{1+e^{\beta_0+\beta_1 x_1+\cdots+\sum_{k=1}^{n}\beta_{ik}D_{i+\beta_p x_p}}}$$

dimana $\pi(X)$ adalah peluang sukses suatu kejadian, x_i (untuk i =1,2,..., p) adalah faktor-faktor yang mempengaruhi peubah respon, p adalah banyaknya peubah penjelas yang digunakan, D adalah variabel *dummy* dan *k* adalah banyaknya variabel *dummy* yang digunakan. Dengan menggunakan transformasi logit, model tersebut dapat dituliskan dengan:

$$g(x) = ln\left(\frac{\pi(X)}{1-\pi(X)}\right)$$

$$= \beta_0 + \beta_1 x_1 + \cdots + \sum_{k=1}^{n}\beta_{ik}\,D_i + \beta_p x_p + \varepsilon$$

dimana $\beta_i,\ i = 1,2,\ldots p$

c. Melakukan pengujian koefisien parameter model secara simultan.

Hipotesis yang digunakan adalah:

$H_0 :\ \beta_1 = \beta_2 = \cdots = \beta_p = 0$

$H_1 =$ minimal ada satu $\beta_i \neq 0$

Statistik uji yang digunakan yaitu :

$$G = -2\ln\left[\frac{\left(\frac{n_1}{n}\right)^{n_1}\left(\frac{n_0}{n}\right)^{n_0}}{\prod_{i=1}^{n}\pi_i^{y_i}(1-\pi_i)^{1-y_i}}\right]$$

dengan y_i adalah variabel respon, n_1 adalah $\sum y_i$, n_0 adalah $\sum(1-y_i)$ dan n adalah $n_0 + n_1$. Statistik uji G mengikuti sebaran χ^2 dengan derajat bebas $p-1$, dimana p adalah jumlah parameter yang digunakan.

d. Melakukan pengujian koefisien parameter model secara parsial.

Hipotesis yang digunakan :

$H_0 :\ \beta_i = 0$ (peubah penjelas ke-i tidak berpengaruh terhadap peubah respon)

$H_1 :\ \beta_i \neq 0$ (peubah penjelas ke-i berpengaruh terhadap peubah respon)

Statistik uji yang digunakan :

$$W = \frac{\widehat{\beta_\iota}}{SE(\widehat{\beta_\iota})}$$

dengan $\widehat{\beta_\iota}$: penduga parameter β_i dan $SE(\widehat{\beta_\iota})$: standard error dari $\widehat{\beta_\iota}$. Statistik uji Wald mengikuti distribusi chi-kuadrat dengan derajat bebas satu dimana H_0 ditolak saat $|W| > \chi^2_{(\alpha,1)}$.

d. Melakukan pengujian koefisien parameter model secara parsial.

Hipotesis yang digunakan :

H_0 : β_i=0 (peubah penjelas ke-i tidak berpengaruh terhadap peubah respon)

H_1 : β_i≠0 (peubah penjelas ke-i berpengaruh terhadap peubah respon)

Statistik uji yang digunakan :

W=(β_i)ˆ/SE((β_i)ˆ)

dengan $\widehat{\beta_\iota}$: penduga parameter β_i dan $SE(\widehat{\beta_\iota})$: standard error dari $\widehat{\beta_\iota}$. Statistik uji Wald mengikuti distribusi chi-kuadrat dengan derajat bebas satu dimana H_0 ditolak saat $|W| > \chi^2_{(\alpha,1)}$.

e. Menguji kelayakan model.

Pengujian kelayakan (*goodness of fit*) pada model regresi logistik menggunakan uji Hosmer-Lemeshow. Uji Hosmer-Lemeshow didasarkan pada pengelompokan pada nilai dugaan peluangnya yang menyebar Chi- Kuadrat (Hosmer & Lemeshow, 2000).

Hipotesis yang digunakan :

H_0 : model yang dibangun layak

H_1 : model yang dibangun tidak layak

Statistik uji yang digunakan :

$$\hat{C} = \sum_{k=1}^{g}\frac{(O_k - n'_k\overline{\pi_k})^2}{n'_k\overline{\pi_k}(1-\overline{\pi_k})}$$

dengan $\hat{C}$ adalah statistik Hosmer-Lemeshow, $O_k \sum_{j=1}^{c_k} y_i$ adalah jumlah nilai peubah respon pada kelompok ke-k, g adalah banyaknya amatan dalam kelompok ke-k, n'_k adalah jumlah sampel pada kelompok ke- k, c_k adalah banyaknya kombinasi peubah bebas pada kelompok ke-k dan $\overline{\pi_k}$ adalah rata-rata dari $\hat{\pi}$ untuk kelompok ke-k. Statistik $\hat{C}$ menyebar

mengikuti sebaran Chi-Kuadrat dengan derajat bebas $g-2$ (Hosmer dan Lemeshow, 2000). Kesimpulan menolak hipotesis nol jika nilai $C_{\text{Hitung}} > \chi^2_{\alpha(g-2)}$.

e. Menguji kelayakan model.

Pengukuran kinerja klasifikasi dilakukan dengan matriks konfusi (*confusion matrix*). Matriks konfusi merupakan tabel yang mencatat hasil kinerja klasifikasi.

Tabel 2. Matriks Konfusi

Hasil observasi	Taksiran	
	y_1	y_2
y_1	n_{11}	y_1
y_2	n_{21}	y_2

Keterangan :

n_{11} : jumlah subjek dari y_1 tepat diklasifikasikan sebagai y_1

n_{12} : jumlah subjek dari y_1 salah diklasifikasikan sebagai y_2

n_{12} : jumlah subjek dari y_2 salah diklasifikasikan sebagai y_1

n_{11} : jumlah subjek dari y_2 tepat diklasifikasikan sebagai y_2.

Perhitungan nilai akurasi merupakan proporsi observasi yang diprediksi benar oleh fungsi klasifikasi, digunakan rumus :

$$\text{Akurasi} = \frac{n_{11} + n_{22}}{n_{11} + n_{12} + n_{21} + n_{22}}$$

Adapun langkah- langkah dari algoritma kNN adalah:

a) Tentukan parameter *k*

Penentuan *k* terbaik menggunakan *k-fold cross validation.*

b) Hitung jarak antara data testing dan data training.

Jika data berbentuk numerik maka menggunakan jarak Euclid seperti pada persamaan berikut:

$$D(x_1, y_1) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Keterangan :

x_i : data training

y_i : data testing

n : dimensi data

Perhitungan jarak antar observasi berupa data campuran kuantitatif dan kualitatif dapat dilakukan dengan menggunakan koefisien kemiripan umum Gower. Koefisien kemiripan Gower dapat digunakan untuk melihat kemiripan antar observasi dengan melakukan perhitungan jarak pada setiap peubah acak yang ada sesuai dengan skala pengukuran perubah acak tersebut (Gower, 1971). Secara umum, persamaan kemiripan Gower ditunjukkan pada persamaan berikut :

$$s(x_i, x_j) = \frac{\sum_{k=1}^{p} s_k(x_{ik}, x_{jk}) \delta(x_{ik}, x_{jk}) w_k}{\sum_{k=1}^{p} \delta(x_{ik}, x_{jk}) w_k}$$

$$\delta(x_{ik}, x_{jk}) = \begin{cases} 1; x_{ik}, x_{jk} \in \mathbb{R} \\ 0; lainnya \end{cases}$$

Keterangan :

$s(x_i, x_j)$: koefisien kemiripan Gower antara observasi ke – i dan j

$s_k(x_{ik}, x_{jk})$: koefisien kemiripan antara observasi ke-i dan j pada peubah k

$\delta(x_{ik}, x_{jk})$: kemungkinan perbandingan peubah k obsevasi ke-i dan j

w_k : bobot pilihan yang menyatakan kepentingan variabel, $w_k = 1$

x_{ik}, x_{jk} : nilai observasi ke-i dan j pada peubah ke-k

Koefisien kemiripan s_k pada setiap variabel dihitung berdasarkan skala pengukuran peubah *k*. Perhitungan nilai s_k pada skala nominal, ordinal, interval dan rasio secara berurutan ditunjukkan pada persamaan berikut :

Nominal : $s_k(x_{ik}, x_{jk}) = \begin{cases} 1; x_{ik} = x_{jk} \\ 0; x_{ik} \neq x_{jk} \end{cases}$

Ordinal : $s_k(x_{ik}, x_{jk}) = 1 - \frac{|r_k(x_{ik}) - r_k(x_{jk})|}{max_m\{r_k(x_{mk})\} - min_m\{r_k(x_{mk})\}}$

Interval ; Rasio : $s_k(x_{ik}, x_{jk}) = 1 - \frac{|x_{ik} - x_{jk}|}{max_m\{x_{mk}\} - min_m\{x_{mk}\}}$

Keterangan :

x_{mk} : nilai observasi ke-*m* peubah *k*

$r_k(x_{mk})$: rank dari nilai observasi ke-*m* peubah ordinal *k*

$max_m\{x_{mk}\}$:nilai maksimum dari seluruh nilai peubah *k*

$min_m\{x_{mk}\}$: nilai minimum dari seluruh nilai peubah *k*

$max_m\{r_k(x_{mk})\}$: rank maksimum dari seluruh nilai peubah ordinal *k*

$min_m\{r_k(x_{mk})\}$: rank minimum dari seluruh nilai peubah ordinal *k*

Penentuan tetangga terdekat kNN diperoleh berdasarkan nilai jarak ketakmiripan antar observasi sehingga nilai koefisien kemiripan Gower perlu ditransformasi menjadi nilai koefisien ketakmiripan dengan menggunakan persamaan :

$$d_k(x_i, x_j) = 1 - s_k(x_i, x_j)$$

Keterangan :

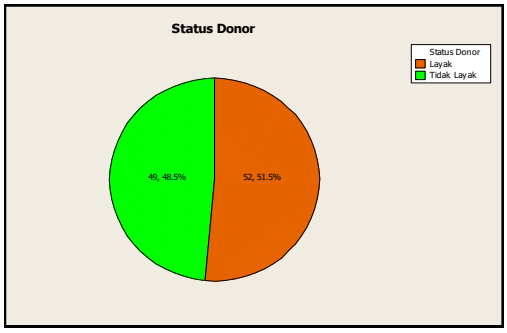
$d_k(x_i, x_j)$: koefisien ketakmiripan Gower observasi ke–*i* dan *j* pada peubah k.

c) Memilih jarak terdekat sampai pada parameter k
d) Memilih jumlah kelas terbanyak sebanyak k lalu diklasifikasikan
e) Menentukan nilai akurasi klasifikasi.

HASIL DAN PEMBAHASAN

Gambaran Karakteristik Status Pendonor Darah

Informasi masing-masing peubah dapat diperoleh dengan melakukan eksplorasi data. Status pendonor sebagai peubah respon dengan kategori (1) layak mendonorkan darahnya dan (0) tidak layak mendonorkan darahnya. Berdasarkan Gambar 1, status pendonor yang layak mendonorkan darahnya sebesar 51.5% yaitu sebanyak 52 orang dan yang tidak layak sebesar 48.5% yaitu sebanyak 49 orang.

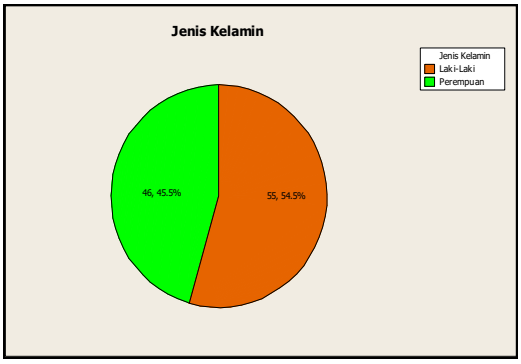


Gambar 1. Karakteristik Status Pendonor Darah

Gambaran Karakteristik Jenis Kelamin

Jenis Kelamin mempunyai skala pengukuran kategori dengan label (1) laki-laki dan (2) perempuan. Jumlah pendonor laki-laki sebanyak 55 orang dan perempuan sebanyak 46 orang. Informasi untuk peubah jenis kelamin ditampilkan seperti gambar berikut:

Berdasarkan Gambar 2 di atas, dari 101 responden diperoleh hasil persentase calon pendonor laki-laki sebanyak 54.5% dan pendonor perempuan sebanyak 45.5%.



Gambar 2. Karakateristik Jenis Kelamin

Gambaran Karakteristik Usia, Berat Badan, Tinggi Badan, Sistole, Diastole dan Kadar Haemoglobin

Peubah usia, berat badan, tinggi badan, sistole, diastole dan kadar HB bertipe numerikal sehingga informasi tentang karakteristik datanya dapat dilihat dari statistik deskriptif. Berikut adalah gambaran karakteristik peubah usia, berat badan, tinggi badan, sistole, diastole dan kadar HB :

Tabel 3. Statistik Deskriptif

	Min	Max	Mean	Std Dev
Usia	19	54	31.89	8.233
Berat badan	50	89	65.01	7.349
Tinggi badan	152	180	164.65	5.317
Sistole	100	160	118.42	20.530
Diastole	50	96	69.81	12.844
Kadar HB	46	83	61.14	9.881

Berdasarkan Tabel 3 dapat dilihat bahwa untuk peubah usia memiliki rata-rata usia calon pendonor sekitar 31 tahun dengan standar deviasi sebesar 8.233. Rata-rata berat badan dan tinggi badan pendonor adalah 65 kg dan 164 cm dengan standar deviasi 7.349 dan 5.317. Tekanan darah

terdiri dari dua yaitu tekanan darah atas (*sistole*) dan tekanan darah bawah (*diastole*). Rata-rata *sistole* dan *diastole* pendonor adalah 118.42 mm Hg dan 69.81 mm Hg dengan standar deviasi 20.53 dan 12.844. Peubah kadar_HB memiliki nilai rata-rata 61.14 g/dL dan standar deviasi 9.881 g/dL dengan kadar Hb tertinggi yaitu 83 g/dL dan terendah yaitu 46 g/dL.

1. Regresi Logistik

Regresi logistik biner merupakan suatu metode statistik yang digunakan untuk menganalisis hubungan antara peubah penjelas (X) dengan peubah respon (Y) yang berupa data kategori dikotomi. Nilai variabel Y=1 menyatakan adanya suatu karakteristik dan Y=0 menyatakan tidak adanya suatu karakteristik. Tahapan analisis data pada regresi logistik adalah :

Penentuan Model Awal Regresi Logistik Biner

Langkah pertama dalam analisis regresi logistik biner adalah menentukan model awal regresi logistik biner (Model 1). Pada penelitian ini regresi logistik biner dilakukan dengan meregresikan peubah status kelayakan calon pendonor darah dengan peubah jenis kelamin, usia, berat badan, tinggi badan, s*istole*, *diastole*, dan kadar Haemoglobin. Secara umum, model awal regresi logistik biner yang diperoleh yaitu :

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

dengan nilai

$$g(x) = 1.7532 - 0.4779X_{12} + 0.0670X_2 - 0.0313X_3 + 0.0928X_4 - 0.0016X_5 - 0.0024X_6 - 0.2817X_7$$

Pengujian Koefisien Parameter secara Simultan

Hipotesis yang digunakan adalah

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$

$H_1 =$ minimal ada satu $\beta_i \neq 0$

Nilai G yang diperoleh pada pengujian secara simultan adalah 52.943. Karena nilai G = 52.943 > $\chi^2_{(0.05,7)}$=14,0671 maka H_0 ditolak. Berdasarkan keputusan bahwa H_0 ditolak maka disimpulkan bahwa peubah penjelas yang terdapat pada model berpengaruh nyata secara simultan.

Pengujian Koefisien Parameter secara Parsial

Hipotesis yang digunakan :

$H_0 : \beta_i = 0$ (peubah penjelas ke-*i* tidak berpengaruh terhadap peubah respon)

$H_1 : \beta_i \neq 0$ (peubah penjelas ke-*i* berpengaruh terhadap peubah respon)

Berdasarkan Tabel 4 diperoleh bahwa

Tabel 4. Nilai Wald Model 1

Peubah	Wald	*Sig*	Keputusan
Jenis Kelamin Perempuan	0.282	0.595	Tidak signifikan
Usia	1.168	0.280	Tidak signifikan
Berat Badan	0.284	0.594	Tidak signifikan
Tinggi Badan	1.788	0.181	Tidak signifikan
Sistole	0.010	0.919	Tidak signifikan
Diastole	0.003	0.954	Tidak signifikan
Kadar HB	11.577	0.001	Signifikan

peubah yang berpengaruh nyata terhadap status kelayakan calon pendonor darah adalah kadar Haemoglobin, hal ini ditunjukkan dengan nilai *sig* = 0.001 < α = 0.05. Model 1 masih didapatkan peubah yang tidak berpengaruh signifikan terhadap model, karena memiliki nilai *Sig* > α = 5% sehingga peubah yang tidak berpengaruh harus dihilangkan. Langkah selanjutnya adalah menghilangkan peubah jenis kelamin perempuan, usia, berat badan, tinggi badan, *sistole*, dan *diastole* dari model regresi logistik biner.

Metode AIC adalah metode yang dapat digunakan untuk memilih model regresi terbaik yang ditemukan oleh Akaike dan Schwarz (Grasa, 1989). Menurut metode AIC, model regresi terbaik adalah model regresi yang mempunyai nilai AIC terkecil. Adapun nilai AIC untuk setiap model 1 dan 2 adalah :

Tabel 5. Nilai AIC Kedua Model

	Model	AIC
1	Model 1	62.81449
2	Model 2	55.60601

Dari hasil output diatas dapat dilihat bahwa Model 2 merupakan model yang terbaik karena memiliki nilai AIC terkecil yaitu 55.60601.

Model kedua diperoleh dari menghilangkan peubah-peubah yang tidak signifikan pada model pertama. Hasil yang diperoleh sebagai berikut :

$$g(x) = -15.176 + 0.257X_1$$

Pengujian koefisien regresi secara simultan didapatkan nilai G sebesar 48.152. Nilai tersebut lebih besar daripada $\chi^2_{(0.05,1)}$=3.84 sehingga peubah penjelas yang terdapat pada model berpengaruh nyata secara serentak.

Nilai Wald yang diperoleh pada model kedua sebesar 22.058 dengan sig = 0.000. Nilai sig model kedua < α = 5% sehingga peubah kadar haemoglobin berpengaruh nyata terhadap status kelayakan calon pendonor darah.

Pengujian Kelayakan Model

Hipotesis yang digunakan :
H_0 : model yang dibangun layak
H_1: model yang dibangun tidak layak

Berdasarkan uji kesesuaian model diperoleh nilai $\hat{C}$ = 11.270 < $\chi^2_{(0.05,8)}$ = 15.5073 sehingga H_0 diterima. Jadi, model regresi logistik biner yang terbentuk layak atau tidak ada perbedaan antara observasi dengan kemungkinan hasil prediksi.

Setelah dilakukan uji signifikansi terhadap model, baik secara keseluruhan maupun individual serta dilakukan uji kesesuaian model maka diperoleh model akhir sebagai berikut:

$$\pi(x) = \frac{\exp(-15.176 + 0.257)}{1 + \exp(-15.176 + 0.257)}$$

Odds Ratio untuk kadar Haemoglobin (exp(0.257)) = 1.293 artinya semakin tinggi kadar Haemoglobin calon pendonor darah maka kecenderungannya semakin layak untuk mendonorkan darahnya.

Ketepatan Klasifikasi

Hasil prediksi pada data *testing* selanjutnya dibandingkan dengan data sebenarnya pada respon data testing. Nilai akurasi diperoleh dari fungsi confusion Matrix () dengan menggunakan paket caret pada software R. Berikut adalah hasil ketepatan klasifikasi status kelayakan calon pendonor darah :

Tabel 6. Matriks Konfusi

Aktual	Prediksi	
	Layak	Tidak
Layak	15	0
Tidak	2	12

Berdasarkan Tabel 6, dapat dihitung nilai akurasi sebesar 93% yang didapat dari membandingkan jumlah dari calon pendonor darah yang tepat diklasifikasikan dengan jumlah seluruh observasi.

Output yang dihasilkan R untuk analisis regresi logistik biner adalah sebagai berikut:

```
Call:
glm(formula = Status.Donor ~ ., family = binomial(link = "logit"),
    data = training)

Deviance Residuals:
     Min       1Q   Median       3Q      Max
-2.32715  -0.34092  -0.05505   0.50735   2.33770

Coefficients:
                Estimate  Std. Error  z value  Pr(>|z|)
(Intercept)     1.753221  12.313344   0.142    0.886777
JKPerempuan    -0.477929   0.899929  -0.531    0.595368
Usia            0.067030   0.062017   1.081    0.279769
Bb             -0.031323   0.058774  -0.533    0.594079
Tb              0.092768   0.069381   1.337    0.181196
Diastole       -0.001653   0.016309  -0.101    0.919273
Sistole        -0.002425   0.042318  -0.057    0.954300
Kadar          -0.281751   0.082807  -3.402    0.000668

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 99.758  on 71  degrees of freedom
Residual deviance: 46.814  on 64  degrees of freedom
```

AIC: 62.814

Number of Fisher Scoring iterations: 6

Output Confusion Matrix yang dihasilkan software R yaitu :

Confusion Matrix and Statistics

Reference
Prediction Layak tidak
Layak 15 2
tidak 0 12

Accuracy : 0.931
95% CI : (0.7723, 0.9915)
No Information Rate : 0.5172
P-Value [Acc > NIR] : 1.901e-06
Kappa : 0.8612
Mcnemar's Test P-Value : 0.4795

Sensitivity : 1.0000
Specificity : 0.8571
Pos Pred Value : 0.8824
Neg Pred Value : 1.0000
Prevalence : 0.5172
Detection Rate : 0.5172
Detection Prevalence : 0.5862
Balanced Accuracy : 0.9286

'Positive' Class: Layak

2. k-Nearest Neighbor (KNN)

Menurut Prasetyo (2012) k-Nearest Neighbor (k-NN) adalah metode yang melakukan klasifikasi berdasarkan kedekatan lokasi (jarak) suatu data dengan data lain. Nilai k pada kNN berarti k-data terdekat dari data testing.

Metode k-NN cukup sederhana, tidak ada asumsi mengenai distribusi data dan mudah diaplikasikan. Pemilihan nilai k (jumlah data/tetangga terdekat) ditentukan dengan menggunakan cross validation. Cross validation adalah sebuah teknik validasi model untuk menilai bagaimana hasil statistik analisis akan menggeneralisasi kumpulan data independen. Teknik ini utamanya digunakan untuk melakukan prediksi model dan memperkirakan seberapa akurat sebuah model prediktif ketika dijalankan dalam praktiknya. Salah satu teknik dari validasi silang adalah k-fold cross validation, yang mana memecah data menjadi k bagian set data dengan ukuran yang sama. Penggunaan k-fold cross validation untuk menghilangkan bias pada data. Pemilihan nilai k ini bisa mempengaruhi tingkat akurasi prediksi yang dikerjakan (Santosa, 2007).

Langkah-langkah dari perhitungan kNN adalah sebagai berikut:

a) Normalisasi data calon pendonor darah.

Normalisasi data linier adalah proses penskalaan nilai atribut data sehingga bisa jatuh pada range tertentu. Tujuan dari normalisasi data adalah untuk mempersempit atau mengecilkan nilai range pada data tersebut. Keuntungan dari metode ini adalah keseimbangan nilai perbandingan antara data saat sebelum dan sesudah nilai normalisasi. Kekurangannya adalah jika ada data baru metode ini akan memungkinkan terjebak pada *out of bound error*. Normalisasi dihitung menggunakan persamaan berikut:

$$Nor\ (X^*) = \frac{X - \min X}{(maks\ X - \min X)}$$

b) Menghitung jarak data training ke data acuan (data testing) calon pendonor darah menggunakan *euclidian distance*.

c) Selanjutnya mengurutkan hasil jarak *Euclidean* dari yang terkecil ke terbesar.

d) Proses menentukan nilai k.

e) Proses mencari label atau kelas mayoritas sebanyak nilai K sesudah normalisasi.

Output yang dihasilkan *software* R untuk mendapatkan *k* terbaik adalah sebagai berikut :

k-Nearest Neighbors

72 samples

7 predictor

2 classes: 'Layak', 'tidak'

Pre-processing: centered (7), scaled (7)

Resampling: Cross-Validated (10 fold, repeated 3 times)

Summary of sample sizes: 66, 65, 64, 64, 66, 64, ...

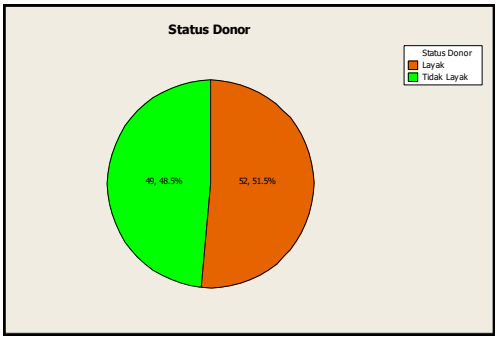
Resampling results across tuning parameters:

k	Accuracy	Kappa
5	0.8261905	0.6512836
7	0.8103175	0.6173531
9	0.8067460	0.6107975
11	0.8158730	0.6301212
13	0.8075397	0.6134545
15	0.8035714	0.6066169
17	0.8049603	0.6086768
19	0.7946429	0.5903862
21	0.7912698	0.5830190
23	0.8160714	0.6308053
25	0.8222222	0.6445058
27	0.8263889	0.6528392
29	0.8305556	0.6611725
31	0.8472222	0.6936942
33	**0.8486111**	**0.6964720**
35	0.8438492	0.6863869
37	0.8299603	0.6600450
39	0.8450397	0.6899165
41	0.8347222	0.6687541
43	0.8305556	0.6604207

Accuracy was used to select the optimal model using the largest value.

The final value used for the model was k = 33.

Akurasi digunakan untuk memilih model yang optimal dengan menggunakan nilai yang tertinggi. Berdasarkan output di atas, nilai k terbaik berdasarkan akurasi yang tertinggi adalah 33. Pada k = 33 nilai akurasi yang dihasilkan adalah 84.86%, artinya kemampuan model dalam menebak status kelayakan calon pendonor darah adalah sebesar 84.86%. Pemilihan k terbaik juga dapat dilihat dari plot yang dihasilkan dari *Cross Validation* yang berulang-ulang. Berikut adalah tampilan plot yang menununjukkan nilai k terbaik :



Gambar 3. Plot Akurasi berdasarkan *k-fold Cross validation*

Berdasarkan Gambar 3 banyaknya *k* terbaik yang dihasilkan dari plot akurasi adalah 33. Hasil *k* yang sama didapatkan dengan menggunakan k-*fold Cross Validation*. Iterasi *k* yang akan digunakan pada model dengan berbagai nilai tingkat akurasi serta nilai parameter yang lain dapat dilihat pada tabel *confusion Matrix*. Berikut adalah output yang dihasilkan *software* R :

fusion Matrix and Statistics

Prediction	Reference: Layak	Reference: tidak
Layak	14	5
tidak	1	9

Accuracy : 0.7931

95% CI : (0.6028, 0.9201)

No Information Rate : 0.5172

P-Value [Acc > NIR] : 0.002088

Kappa : 0.5817

Mcnemar's Test P-Value : 0.220671

Sensitivity : 0.9333

Specificity : 0.6429

Pos Pred Value : 0.7368

Neg Pred Value : 0.9000

Prevalence : 0.5172

Detection Rate : 0.4828

Detection Prevalence : 0.6552

Balanced Accuracy : 0.7881

'Positive' Class : Layak

Berdasarkan output di atas didapatkan nilai akurasi sebesar 79.31%, artinya kemampuan model kNN menebak status pendonor darah sebesar 79.31%. Sedangkan berdasarkan data aktual orang yang memiliki status layak mendonorkan darahnya, model dapat menebak dengan benar sebesar 93.3%. Berdasarkan data aktual orang yang memiliki status pendonor tidak layak mendonorkan darahnya, model dapat menebak dengan benar sebesar 64.29%.

3. Perbandingan Akurasi Klasifikasi kNN dan Regresi Logistik Biner

Berdasarkan hasil analisis yang telah dilakukan yaitu klasifikasi menggunakan kNN dan regresi logistik biner didapatkan hasil akurasi masing-masing metode, hasil akurasi tersebut digunakan untuk membandingkan kedua metode tersebut. Semakin tinggi nilai akurasinya maka semakin tinggi ketepatan klasifikasi suatu model. Berikut adalah perbandingan akurasi dari kedua metode:

Tabel 7. Perbandingan Akurasi

Model	Akurasi
kNN	79.31%
Regresi logistik biner	93.1%

Jika dilihat dari perbandingan nilai akurasi metode regresi logistik biner dan kNN, maka metode terbaik untuk mengklasifikasikan status kelayakan pendonor darah adalah metode regresi logistik biner. Metode regresi logistik biner memiliki kemampuan dalam memprediksi benar dari data aktual pendonor yang layak mendonorkan darahnya lebih baik daripada metode kNN karena memiliki nilai akurasi lebih tinggi dari pada metode kNN.

KESIMPULAN

proksi dari pendapatan per kapita Berdasarkan hasil dan pembahasan yang telah dipaparkan sebelumnya, maka dapat diambil kesimpulan bahwa akurasi yang dihasilkan metode kNN sebesar 79.31% sedangkan metode regresi logistik biner adalah 93.1%. Hasil perbandingan ketepatan klasifikasi antara kNN dan regresi logistik biner dapat dilihat dari nilai akurasinya. Semakin tinggi tingkat akurasi maka model akan semakin baik dalam mengklasifikasikan. Nilai akurasi klasifikasi yang diperoleh pada metode regresi logistik biner lebih tinggi daripada metode kNN. Sehingga metode terbaik untuk mengklasifikasikan status calon pendonor darah adalah regresi logistik biner karena memiliki keakuratan klasifikasi yang lebih baik daripada kNN.

Hasil lain yang diperoleh dari penelitian ini adalah peubah yang berpengaruh nyata pada status kelayakan calon pendonor darah adalah peubah kadar Haemoglobin. Hasil ini didapatkan dari pengujian pada regresi logistik biner.

DAFTAR PUSTAKA

Agresti, A. 2002. *Categorical Data Analysis*. New York : John Wiley&Sons.

Bayususetyo, D., Santoso, R., dan Tarno. 2017. Klasifikasi Calon Pendonor Darah Menggunakan Metode Naive Bayes Classifier (Studi Kasus : Calon Pendonor Darah di Kota Semarang). *Jurnal Gaussian* 6(2) : 193-200.

Bhatia, M., Vandana., 2010. Survey of Nearest Neighbor Techniques. *International Journal of Computer Science and Information Security* 8, 1947-5500.

Departemen Kesehatan RI. 2009. Donor Darah. Tersedia pada http://kemenkes.go.id/

Gower, JC. 1971. A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27(4): 857-871.

Grasa, A. 1989 Econometric Model Selection: A New. Approach. Kluwer. Springer Science and Bussiness Media.

Hosmer, D.W. and Lemeshow, S. 2000. *Applied Logistic Regression*. New York: John Wiley & Son, Inc.

Moradian, M., and Baraani, A. 2009. K-Nearest Neighbor Based Association Algorithm. *Journal of Theoretical and Applied Information Technology* 6, 123 – 129.

Nugroho, EB., Furqon, MT. dan Hidayat, N. 2018. Klasifikasi Pendonor Darah Menggunakan Metode Support Vector Machine (SVM) pada Dataset RFMTC. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 2(10) : 3860-3864.

Palang Merah Indonesia. (2019, September 11). Pelayanan Donor Darah. Tersedia pada http://www.pmi.or.id/

Prasetyo, E. 2012. Data Mining Konsep dan Aplikasi Menggunakan MATLAB. Yogyakarta: Penerbit ANDI Yogyakarta.

Santosa, B. 2007. *Data Mining Terapan*. Yogyakarta : Graha Ilmu.

Sapriana, B.M. 2017. Penerapan Algoritma Naive Bayes Classifier Pada Klasifikasi Status Kelayakan Pendonor Darah di Unit Transfusi Darah Palang Merah Indonesia (UTD PMI) Kota Makassar. Skripsi. Makassar: Universitas Negeri Makassar.

Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Yu PS *et al*. 2008. Top 10 Algorithms in Data Mining. *Journal of Knowledge and Information Systems* 14(1): 1-37.

REGRESI PROBIT UNTUK ANALISIS VARIABEL-VARIABEL YANG MEMPENGARUHI PERCERAIAN DI SULAWESI TENGAH

Nur'eni[1], Lilies Handayani[2]

[1,2] Program Studi Statistika, Universitas Tadulako
e-mail: [1]eniocy@yahoo.com, [2]lilies.stath@gmail.com

Abstrak

Sulawesi Tengah adalah salah satu Provinsi di Indonesia yang memiliki permasalahan dalam perceraian. Tingkat perceraian di Sulawesi Tengah pada tahun 2016 sebesar 2,44%. Persentase tingkat perceraian di Sulawesi Tengah ini menjadi tingkat perceraian ketiga tertinggi di Indonesia. Pada penelitian ini diteliti faktor-faktor yang mempengaruhi kasus perceraian di Sulawesi Tengah. Metode yang digunakan adalah regresi probit biner dengan variabel respon adalah status perkawinan. Hasil penelitian menunjukkan bahwa variabel prediktor yang mempengaruhi perceraian secara signifikan di Provinsi Sulawesi Tengah adalah umur kawin pertama (X2) kategori 1 (18-21 tahun) dan kategori 2 (>21 tahun), tingkat pendidikan (X3) kategori 1 (SD) dan kategori 4 (di atas SMA), daerah tempat tinggal (X4) kategori 1 (kota) dan jumlah pengeluaran rumah tangga (X6) dengan tingkat ketepatan klasifikasi model sebesar 99,2%.

Kata kunci: regresi probit biner, perceraian, status perkawinan, ketepatan klasifikasi

Abstract

Central Sulawesi is one of province in Indonesia which has divorce trouble. The divorce rate in Central Sulawesi in 2016 is 2,44%, the third highest divorce rate in Indonesia. This research aimed the influence factors of divorce cases in Central Sulawesi by using binary probit regression with the respon variable is marriage status. The result shows that the predictor variable which influence the divorce in Central Sulawesi are age of first marriage (X2) with age category 18-21 years and over 21 years, education level (X3) with elementary school and over high school category, habitation (X4) with urban categoryand household expanditure (X6) with the accuracy rate is 99,2%.

Keywords*: binary probit regression, divorce, marriage status, accuracy rate*

PENDAHULUAN

Demografi pertumbuhan penduduk Indonesia sangat dipengaruhi oleh adanya fertilitas. Perkawinan merupakan salah satu variabel yang mempengaruhi tinggi rendahnya tingkat fertilitas, sehingga secara tidak langsung mempengaruhi pertumbuhan penduduk. Perkawinan adalah ikatan lahir batin antara seorang pria dan seorang wanita sebagai suami istri dengan tujuan membentuk keluarga (rumah tangga) yang bahagia dan kekal berdasarkan ketuhanan Yang Maha Esa menurut UU Perkawinan No. 1 Tahun 1974. Perkawinan jika dilakukan pada umur yang "tepat" akan membawa kebahagian bagi keluarga dan pasangan (suami dan istri) yang menjalankan perkawinan tersebut. Perkawinan yang dilakukan pada usia yang terlalu muda (dini) akan membawa banyak konsekuensi pada pasangan, antara lain adalah kesehatan, pendidikan, dan ekonomi. Pada kesehatan khususnya dalam hal kejiwaan, dimana perkawinan yang dilakukan pada usia dini akan lebih mudah berakhir dengan kegagalan karena ketiadaan kesiapan mental menghadapi dinamika kehidupan berumah tangga (Dariyo, 2004).

Perceraian merupakan suatu peristiwa perpisahan secara resmi antara pasangan suami istri dan mereka berketetapan untuk tidak menjalankan tugas dan kewajiban sebagai suami istri. Mereka tidak lagi hidup dan tinggal serumah bersama, karena tidak ada ikatan yang resmi. Mereka yang telah bercerai tetapi belum memiliki anak, maka perpisahan tidak menimbulkan dampak traumatis terhadap psikologis anak-anak. Namun mereka yang telah memiliki keturunan, tentu saja perceraian menimbulkan masalah psiko-emosional bagi anak-anak (Dariyo, 2004). Menurut para ahli, seperti Nakamura (1990), Turner & Helms (1995), Sudarto & Wirawan (2001), terdapat beberapa faktor penyebab perceraian yaitu a) kekerasan verbal, b) masalah ekonomi, c) keterlibatan dalam perjudian, d)keterlibatan dalam penyalahgunaan minuman keras, e) perselingkuhan.

Perceraian merupakan salah satu masalah yang dihadapi oleh negara. Di Indonesia perceraian merupakan masalah yang penting untuk diatasi agar rumah tangga yang dijalani rukun dan damai. Menurut Badan Pusat Statistik (BPS, 2015; BPS, 2016) tingkat perceraian di Indonesia pada tahun 2015 sebesar 1,91% dan pada tahun 2016 meningkat sebesar 1,93%.

Sulawesi Tengah adalah salah satu Provinsi di Indonesia yang juga memiliki permasalahan dengan perceraian. Menurut Badan Pusat Statistik (BPS) tingkat perceraian di Sulawesi Tengah tahun 2016 sebesar 2,44%, mengalami kenaikan dibanding tingkat perceraian pada tahun 2015 sebesar 1,97%. Persentase perceraian di Sulawesi Tengah ini menjadi tingkat perceraian ketiga tertinggi setelah Nusa Tenggara Barat dan Kalimantan Selatan. Sehingga perlu dilakukan penelitian mengenai variabel-variabel yang mempengaruhi perceraian di Provinsi Sulawesi Tengah.

Analisis regresi dapat digunakan untuk menjelaskan variabel yang mempengaruhi perceraian di Provinsi Sulawesi Tengah. Analisis regresi merupakan metode analisis data yang menggambarkan hubungan sebab akibat antara variabel respon dan variabel prediktor. Jika variabel responnya berskala interval atau *ratio*, maka digunakan regresi linear. Analisis statistik yang dapat menjelaskan hubungan antara variabel respon dan variabel prediktor dimana variabel respon berupa data kualitatif atau kategori yaitu model logit dan model probit (Gujarati, 2004). Namun terdapat perbedaan dari kedua metode tersebut yaitu *link function* dan interpretasi model. Metode regresi probit merupakan metode yang menggunakan *link function* distribusi normal dengan interpretasi model menggunakan nilai efek marginal yang merupakan kelebihan dari regresi probit, sedangkan metode regresi logistik menggunakan *link function* distribusi logistik dan interpretasi model menggunakan nilai *odds ratio* (Masitoh & Ratnasari, 2016).

Penelitian sebelumnya yang berkaitan dengan perceraian dilakukan oleh Tresia pada tahun 2006 di wilayah Sumbar. Penelitian tersebut menggunakan metode regresi logistik. Diperoleh kesimpulan bahwa variabel yang berpengaruh signifikan terhadap perceraian adalah tingkat pendidikan, jenis pekerjaan, pendapatan, dan jumlah anak. Penelitian yang berkaitan dengan perceraian juga dilakukan oleh Riduan pada tahun 1998 dengan menggunakan model gabungan (logistik dan *piecewice proportional hazard*) diperoleh hasil bahwa variabel yang berpengaruh secara signifikan terhadap perceraian adalah kehadiran anak, umur kawin pertama, selisih umur dan perbedaan pendidikan. Berdasarkan teori di atas maka penelitian ini bertujuan untuk menerapkan metode regresi probit biner untuk mengetahui variabel apa saja yang berpengaruh terhadap kasus perceraian di Provinsi Sulawesi Tengah.

METODOLOGI

1. Landasan Teori

Regresi probit biner adalah metode regresi yang digunakan untuk menganalisis variabel dependen yang bersifat kualitatif dan beberapa variabel independen yang bersifat kualitatif, kuantitatif, atau gabungan dari kualitatif dan kuantitatif dengan pendekatan CDF *(cumulative distribution function)* distribusi normal (Gujarati, 2004). Model regesi probit biner adalah sebagai berikut :

$$\boldsymbol{Y}^{*} = \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{x}_i + \boldsymbol{\varepsilon}$$

dimana $\boldsymbol{Y}^{*}$merupakan vektor variabel respon diskrit, $\boldsymbol{\beta}$ merupakan vektor parameter koefisien dengan $\boldsymbol{\beta} = \left[\beta_0, \beta_1, \beta_2, \dots, \beta_p\right]^T$, $\boldsymbol{x}$ merupakan vektor variabel prediktor dengan $\boldsymbol{x} = \left[1,\ x_1,\ x_2, \dots, x_p\right]^T$, $\boldsymbol{\varepsilon}$merupakan *error* yang diasumsikan berdistribusi $N(0,1)$ (Greene, 2008).

Model probit untuk $Y = 0$ adalah probabilitas gagal $= q(x_i)$:

$$P(Y = 0|\boldsymbol{x}) = \Phi(\gamma - \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{x}_i) = q(\boldsymbol{x}_i)$$

Sedangkan model probit $Y = 1$ adalah probabilitas sukses $= p(x_i)$:

$$P(Y = 1|\boldsymbol{x}) = 1 - q(\boldsymbol{x}_i) = p(\boldsymbol{x}_i)$$

dimana $\Phi(\gamma - \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{x}_i)$merupakan fungsi distribusi kumulatif distribusi normal dengan rumus sebagai berikut :

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right) dx$$

Tabel 1. Variabel Respon dan Variabel Prediktor

Simbol	Nama Variabel	Kategori	Skala Data
Y	Status perkawinan	0 : cerai hidup	Nominal
		1 : kawin	
X_1	Status bekerja	0 : tidak bekerja	Nominal
		1 : bekerja	
X_2	Umur kawin pertama	0 : < 18 tahun	Ordinal
		1 : 18-21 tahun	
		2 : > 21 tahun	
X_3	Tingkat pendidikan	0 : tidak memiliki ijazah	Ordinal
		1 : SD	
		2 : SMP	
		3 : SMA	
		4 : di atas SMA	
X_4	Daerah tempat tinggal	0 : desa	Nominal
		1 : kota	
X_5	Status kepemilikan anak	0 : belum memiliki anak	Nominal
		1 : memiliki anak	
X_6	Jumlah pengeluaran rumah tangga	-	Rasio

Interpretasi model regresi probit biner tidak berdasarkan nilai koefisien model akan tetapi menggunakan efek marginal. Efek marginal menyatakan besarnya pengaruh tiap variabel prediktor yang signifikan terhadap probabilitas tiap kategori pada variabel respon dengan rumus sebagai berikut :

$$\frac{\partial P = (Y = 0|\boldsymbol{x}_i)}{\partial \boldsymbol{x}_i} = -\phi(\gamma - \boldsymbol{\beta^T x}_i)\boldsymbol{\beta}$$

$$\frac{\partial P = (Y = 1|\boldsymbol{x}_i)}{\partial \boldsymbol{x}_i} = \phi(\gamma - \boldsymbol{\beta^T x}_i)\boldsymbol{\beta}$$

(Greene, 2008).

2. Metode Analisis

Penelitian ini menggunakan data sekunder yang berasal dari Survei Sosial Ekonomi Nasional (SUSENAS) Provinsi Sulawesi Tengah tahun 2016. Pada Tabel 1 merupakan variabel-variabel yang digunakan sebagai variabel respon (Y) dan variabel-variabel prediktor (X) yang digunakan dalam penelitian.

Analisis pada penelitian ini dilakukan dengan menggunakan aplikasi R dan SPSS. Adapun tahapan pemodelan variabel-variabel yang mempengaruhi kasus perceraian menggunakan metode regresi probit biner dengan langkah-langkah sebagai berikut :

1. Membuat model regresi probit biner dengan meregresikan variabel status perkawinan (Y) dengan variabel X_1 hingga X_6 dimana parameter model diestimasi dengan menggunakan metode MLE (*Maximum Likelihood Estimation*)
2. Menguji signifikansi parameter secara serentak dengan statistik uji G^2 pada persamaan berikut :

$$G^2 = -2\ln\left[\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right] = 2\ln L(\hat{\Omega}) - 2\ln L(\hat{\omega})$$

Daerah penolakan dari statistik uji G^2 adalah tolak H_0 jika nilai $G^2 > \chi^2_{(db;\alpha)}$ dimana db adalah derajat bebas atau *p-value* < α. (Hosmer & Lemeshow, 2000).

3. Jika didapatkan kesimpulan bahwa minimal terdapat satu variabel prediktor yang signifikan maka dilakukan uji parsial dengan statistik uji *Wald* pada persamaan :

$$W_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

Uji *Wald* memiliki daerah penolakan yaitu nilai W dibandingkan dengan Z_{tabel} pada taraf signifikan α yang digunakan. Jika $|W| > Z_{\alpha/2}$atau *p-value* < α. maka diputuskan untuk tolak H_0 (Hosmer & Lemeshow, 2000).

4. Menguji kesesuaian model regresi probit biner dengan statistik uji *Deviance* yang ditunjukkan oleh persamaan :

$$D = -2\sum_{i=1}^{n}\left[y_i\, ln\left(\frac{\hat{P}_i}{y_i}\right) + (1-y_i)\, ln\left(\frac{1-\hat{P}_i}{1-y_i}\right)\right]$$

Keputusan H_0ditolak yaitu jika nilai $D > \chi^2_{(db;\alpha)}$atau *p-value* < α (Hosmer & Lemeshow, 2000).

5. Melakukan evaluasi ketepatan klasifikasi model yang dilakukan dengan melihat peluang kesalahan klasifikasi model atau APER (*Apparent Error Rate*), formula untuk menghitung ketepatan klasifikasi model regresi probit biner dengan persamaan :

$$1 - APER = 1 - \left(\frac{n_{12} + n_{21}}{N}\right) \times 100\%$$

6. Menarik kesimpulan dari hasil penelitian

HASIL DAN PEMBAHASAN

Model regresi probit biner pada penelitian ini dibentuk melalui variabel respon Y (status perkawinan) yang bersifat kualitatif dengan dua kategori yaitu cerai hidup dan kawin, sedangkan variabel prediktor X yang digunakan untuk pemodelan regresi probit biner adalah status bekerja (X_1), umur kawin pertama (X_2), tingkat pendidikan (X_3), daerah tempat tinggal (X_4), status kepemilikan anak (X_5) dan jumlah pengeluaran rumah tangga (X_6).

Langkah pertama untuk melakukan pemodelan dengan metode probit biner adalah melalukan uji signifikansi parameter secara serentak. Pengujian ini bertujuan untuk mengetahui setidaknya terdapat satu variabel prediktor yang signifikan terhadap model dengan hipotesis $H_0 : \beta_1 = \beta_2 =$

$\cdots = \beta_6 = 0$ dan H_1 : minimal ada satu $\beta_j \neq 0$. Uji yang digunakan untuk menguji signifikansi model secara serentak menggunakan uji rasio *likelihood*, hasil yang diperoleh adalah sebagai berikut :

1. Metode Analisis

Metode analisis yang digunakan dalam penelitian ini adalah analisis deskriptif dan analisis inferensia. Analisis deskriptif untuk mengetahui gambaran karakteristik anak penyandang disabilitas berdasarkan status partisipasi sekolahnya. Sedangkan analisis inferensia untuk menganalisis variabel-variabel yang memengaruhi dan kecenderungan berpartisipasi sekolah anak penyandang disabilitas. Analisis inferensia yang digunakan dalam penelitian ini adalah analisis multilevel regresi logistik biner dua level (bilevel regresi logistik biner). Level satu untuk tingkat individu dan level dua untuk tingkat provinsi. Menurut Hox (2010) analisis multilevel merupakan analisis yang digunakan untuk mengatasi masalah data dengan struktur hirarki. Model multilevel yang digunakan adalah model multilevel dengan *random intercept* karena diasumsikan bahwa pengaruh variabel bebas setiap kelompok adalah sama. Analisis multilevel regresi logistik biner digunakan karena variabel respons memiliki dua kategori yaitu anak penyandang disabilitas yang tidak bersekolah (y = 0) dan anak penyandang disabilitas yang bersekolah (y = 1).

Tahapan analisis diawali dengan uji kebebasan *Chi-square* dan uji *U Mann-Whitney*. Uji ini bertujuan untuk mengetahui variabel-variabel penjelas yang signifikan berhubungan dengan variabel partisipasi sekolah anak penyandang disabilitas untuk analisis bilevel regresi logistik biner. Kemudian melakukan pengujian signifikansi *random effect* dengan *Likelihood Ratio Test* untuk mengetahui apakah model multilevel regresi logistik biner lebih cocok digunakan daripada model regresi logistik biner satu level. Selanjutnya melakukan penghitungan variasi antar unit di level-2 menggunakan nilai *Intraclass Corelation Coefficient* (ICC).

Lalu dilakukan pengujian simultan yang bertujuan untuk mengetahui pengaruh seluruh variabel penjelas secara bersama-sama terhadap partisipasi sekolah anak penyandang disabilitas. Jika pengujian parameter secara simultan memberikan kesimpulan bahwa terdapat paling sedikit satu variabel penjelas yang memengaruhi partisipasi sekolah anak penyandang disabilitas, maka tahapan selanjutnya menguji variabel penjelas secara parsial. Pengujian parsial digunakan untuk mengetahui variabel penjelas yang signifikan memengaruhi partisipasi sekolah anak penyandang disabilitas secara parsial. Setelah mengetahui variabel-variabel yang signifikan memengaruhi partisipasi sekolah anak penyandang disabilitas, tahap selanjutnya adalah menginterpretasikan nilai *Odds Ratio* (OR).

HASIL DAN PEMBAHASAN

Model regresi probit biner pada penelitian ini dibentuk melalui variabel respon Y (status perkawinan) yang bersifat kualitatif dengan dua kategori yaitu cerai hidup dan kawin, sedangkan variabel prediktor X yang digunakan untuk pemodelan regresi probit biner adalah status bekerja (X_1), umur kawin pertama (X_2), tingkat pendidikan (X_3), daerah tempat tinggal (X_4), status kepemilikan anak (X_5) dan jumlah pengeluaran rumah tangga (X_6).

Langkah pertama untuk melakukan pemodelan dengan metode probit biner adalah melalukan uji signifikansi parameter secara serentak. Pengujian ini bertujuan untuk mengetahui setidaknya terdapat satu variabel prediktor yang signifikan terhadap model dengan hipotesis $H_0 : \beta_1 = \beta_2 = \cdots = \beta_6 = 0$ dan H_1 : minimal ada satu $\beta_j \neq 0$. Uji yang digunakan untuk menguji signifikansi model secara serentak menggunakan uji rasio *likelihood*, hasil yang diperoleh adalah seperti pada Tabel 2.

Berdasarkan tabel 2, hasil pengujian signifikansi parameter secara serentak menunjukkan bahwa nilai G^2 yang dihasilkan sebesar $65{,}529$ dengan *p-value*

Tabel 2. Uji Serentak Model Regresi Probit Linier

Statistik Uji	*Chi-Square*	*p-value*	Kesimpulan
G^2	65,529	0,000	Tolak H_0

Tabel 3. Uji Parsial Model Regresi Probit Linier

Prediktor	Koefisien Regresi	SE	*W*	*P-value*
Konstanta	0,325339	0,146011	2,23	0,026
X_1 (status bekerja)				
1 (bekerja)	0,114305	0,0747391	1,53	0,126
X_2 (umur kawin pertama)				
1 (18-21 tahun)	0,194059	0,0972509	2,00	0,046*
2 (> 21 tahun)	0,264884	0,0989808	2,68	0,007*
X_3 (tingkat pendidikan)				
1 (SD)	0,360575	0,0989029	3,65	0,000*
2 (SMP)	0,107861	0,140483	0,77	0,443
3 (SMA)	0,204419	0,109184	1,87	0,061
4 (di atas SMA)	0,361750	0,155663	2,32	0,020*
X_4 (daerah tempat tinggal)				
1 (kota)	0,199568	0,0782641	2,55	0,011*
X_5 (status kepemilikan anak)				
1 (memiliki anak)	-0,0532595	0.0734891	-0,72	0,469
X_6 (jumlah pengeluaran rumah tangga)	-0,0000003	0,0000001	-4,87	0,016*

sebesar $0{,}000$ dengan α yang digunakan sebesar $0{,}05$. Karena *p-value* $(0{,}000)$ lebih kecil dari $\alpha\,(0{,}05)$ sehingga keputusan yang diambil adalah tolak H0. Dengan demikian dapat disimpulkan bahwa pada regresi probit biner dengan tingkat kepercayaan 95% minimal ada satu parameter yang signifikan pada model. Selanjutnya dilakukan pengujian parameter secara parsial yang hasilnya dapat dilihat pada tabel 3.

Pengujian signifikansi parameter parsial dilakukan untuk mengetahui variabel-variabel prediktor mana saja yang signifikan terhadap model dengan hipotesis $H_0 : \beta_j = 0$ dan $H_1 : \beta_j \neq 0$ untuk $j = 1, 2, \ldots, 6$. Uji parsial dilakukan dengan statistik uji *Wald*. Dari tabel di atas, diketahui bahwa nilai mutlak statistik uji *W* pada variabel prediktor X_2 (umur kawin pertama) kategori 1 (18-21 tahun) dan kategori 2 (> 21 tahun), X_3 (tingkat pendidikan) kategori 1 (SD) dan kategori 4 (di atas SMA), X_4 (daerah tempat tinggal) kategori 1 (kota) dan X_6 (jumlah pengeluaran rumah tangga) lebih besar dari nilai tabel $Z_{0,05/2} = 1{,}96$ atau dapat dilihat dari nilai *p-value* pada masing-masing prediktor yang nilainya kurang dari $\alpha(0{,}05)$, sehingga keputusan yang diambil adalah tolak H_0. Dengan demikian dapat disimpulkan bahwa variabel yang signifikan terhadap perceraian adalah X_2 (umur kawin pertama) kategori 1 (18-21 tahun) dan kategori 2 (>21 tahun), X_3 (tingkat pendidikan) kategori 1 (SD) dan kategori 4 (di atas SMA), X_4 (daerah tempat tinggal) kategori 1 (kota) dan X_6 (jumlah pengeluaran rumah tangga).

Berdasarkan pengujian signifikansi parameter secara parsial, maka dibentuk model regresi probit biner dengan menggunakan nilai koefisien regresi yang signifikan tersebut sebagai berikut :

$$y^* = 0{,}325339 + 0{,}194059X_{2(1)} + 0{,}264884X_{2(2)} + 0{,}360575X_{3(1)} + 0{,}361750X_{3(4)} + 0{,}199568X_{4(1)} - 0{,}0000003X_6$$

Dari model di atas dapat dihitung nilai prediksi probabilitas setiap variabel.

Sehingga didapatkan persamaan probabilitas respoden yang masuk ke dalam kategori bercerai adalah sebagai berikut :

$$P(y = 0 \mid x) = \Phi(0{,}325339 + 0{,}194059X_{2(1)} + 0{,}264884X_{2(2)} + 0{,}360575X_{3(1)} + 0{,}361750X_{3(4)} + 0{,}199568X_{4(1)} - 0{,}0000003X_{6}$$

Besar pengaruh variabel umur kawin pertama (X_2), tingkat pendidikan (X_3), daerah tempat tinggal (X_4) dan jumlah pengeluaran rumah tangga (X_6) dapat dilihat melalui nilai *marginal effect.* Pada responden pertama untuk data perceraian dihitung *marginal effect* untuk mengetahui besar pengaruh keempat variabel dalam menggolongkan responden pertama ke kategori $y = 0$ atau kategori bercerai. Sebagai ilustrasi, karakteristik untuk responden pertama pada variabel X_2 (umur kawin pertama) adalah kategori 1 (18-21 tahun), variabel X_3 (tingkat pendidikan) yaitu kategori 1 (SD), variabel X_4 (daerah tempat tinggal) yaitu kategori 1 (desa) dan variabel X_6 (jumlah pengeluaran rumah tangga) sebesar Rp 754.595. Hasil perhitungan *marginal effect* adalah sebagai berikut :

1. Nilai *marginal effect* untuk variabel X_2

   $$\frac{\partial P=(Y=0|\boldsymbol{x})}{\partial X_2} = -\phi(\gamma - \boldsymbol{\beta^T x})\boldsymbol{\beta_2} = 0{,}4724$$

   Nilai *marginal effect* variabel X_2 (umur kawin pertama) adalah 0,4724. Hal ini menunjukkan bahwa jika variabel X_2 (umur kawin pertama) berkategori 1 atau umur perkawinan pertamanya antara 18 sampai 21 tahun maka akan menaikkan kecenderungan responden pertama masuk kategori $y = 0$ atau kategori bercerai sebesar 0,4724 atau 47,24%.
2. Nilai *marginal effect* untuk variabel X_3

   $$\frac{\partial P=(Y=0|\boldsymbol{x})}{\partial X_3} = -\phi(\gamma - \boldsymbol{\beta^T x})\boldsymbol{\beta_3} = 0{,}8777$$

   Jika variabel X_3 (tingkat pendidikan) berkategori 1 atau tingkat pendidikannya SD, maka variabel tersebut akan meningkatkan kecenderungan responden pertama terklasifikasikan ke dalam kategori bercerai sebesar 0,8777 atau 87,77%.
3. Nilai *marginal effect* untuk variabel X_4

   $$\frac{\partial P=(Y=0|\boldsymbol{x})}{\partial X_4} = -\phi(\gamma - \boldsymbol{\beta^T x})\boldsymbol{\beta_4} = 0{,}4858$$

   Variabel X_4 (daerah tempat tinggal) yang berkategori 1 atau daerah tempat tinggal berada di kota akan meningkatkan kecenderungan responden masuk pada ketegori bercerai adalah sebesar 0,4858 atau 48,58%.
4. Nilai *marginal effect* untuk variabel X_6

   $$\frac{\partial P=(Y=0|\boldsymbol{x})}{\partial X_6} = -\phi(\gamma - \boldsymbol{\beta^T x})\boldsymbol{\beta_6} = -0{,}0000007$$

   Dari perhitungan *marginal effect* X_6 (jumlah pengeluaran rumah tangga), diketahui bahwa dengan kenaikan jumlah pengeluaran rumah tangga satu satuan akan menurunkan responden pertama masuk ke kategori bercerai sebesar 0,0000007 atau 0,00007%.

Selanjutnya dilakukan pengujian kesesuaian model guna mengetahui apakah terdapat perbedaan yang signifikan antara hasil pengamatan dengan hasil prediksi model, dengan hipotesis H_0 : tidak terdapat perbedaan antara hasil observasi dengan hasil prediksi (model sesuai) dan H_1 : terdapat perbedaan antara hasil observasi dengan hasil prediksi (model tidak sesuai). Uji yang digunakan dalam penelitian ini adalah uji *Deviance.* Hasil yang diperoleh adalah pada tabel 4.

Berdasarkan tabel 4, hasil pengujian signifikansi kesesuian model menunjukkan bahwa nilai *Deviance* yang dihasilkan sebesar 2091,82 dan *p-value* sebesar 0,004 dengan α yang digunakan yaitu 0,05.

Tabel 4. Uji Kesesuaian Model

Statistik Uji	***Chi-Square***	***p-value***	**Kesimpulan**
Deviance	2091,82	0,004	Tolak H_0

Tabel 5. Klasifikasi Model Regresi Probit Biner

Kelompok Aktual	Kelompok Prediksi		Total
	$y = 0$	$y = 1$	
$y = 0$	459	12	471
$y = 1$	4	1525	1529
Total	463	1537	2000

Karena *p-value* $(0{,}004) < \alpha\ (0{,}05)$ sehingga keputusan yang diambil adalah tolak H_0. Dengan demikian dapat disimpulkan bahwa terdapat perbedaan antara hasil observasi dengan hasil prediksi (model tidak sesuai).

Selanjutnya akan dibentuk tabel ketepatan klasifikasi atau *confusion matrix yang* digunakan untuk menggambarkan ukuran ketepatan antara data aktual dan data prediksi. Berikut adalah hasil pengelompokkan data aktual dengan data prediksi.

Berdasarkan tabel 5 dapat dihitung nilai persentase ketepatan klasifikasi dengan nilai APER yaitu :

$$APER = \left(\frac{12+4}{2000}\right) = 0{,}008$$

Sehingga diperoleh ketepatan klasifikasi :

$$(1 - APER) \times 100\% = 99{,}2\%$$

Berdasarkan perhitungan yang diperoleh menunjukkan bahwa model regresi probit biner memiliki kemampuan mengklasifikasikan pengamatan dengan benar sebesar 99,2%. Hal ini menunjukkan bahwa model regresi probit yang dibentuk sangat baik dalam mengklasifikasikan kasus perceraian di Provinsi Sulawesi Tengah berdasarkan variable-variabel yang berpengaruh.

KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dilakukan sebelumnya, maka dapat diperoleh kesimpulan bahwa berdasarkan model regresi probit biner yang dibentuk menghasilkan empat variabel yang signifikan terhadap kasus perceraian di Provinsi Sulawesi Tengah yaitu : X2 (umur kawin pertama) kategori 1 (18-21 tahun) dan kategori 2 (> 21 tahun), X3 (tingkat pendidikan) kategori 1 (SD) dan kategori 4 (di atas SMA), X4 (daerah tempat tinggal) kategori 1 (kota) dan X6 (jumlah pengeluaran rumah tangga). Hal ini dapat menjadi bahan pertimbangan bagi pemerintah dan masyarakat agar lebih memperhatikan faktor-faktor yang berpengaruh tersebut sehingga kasus percerain di Provinsi Sulawesi Tengah dapat diminimalisir.

DAFTAR PUSTAKA

Badan Pusat Statistik [BPS]. 2015. *Statistik Kesejahteraan Rakyat Indonesia*. Jakarta: BPS Indonesia.

Badan Pusat Statistik [BPS]. 2016. *Statistik Kesejahteraan Rakyat Indonesia*. Jakarta: BPS Indonesia.

Dariyo, A. 2004. Memahami Psikologi Perceraian dalam Kehidupan Keluarga. *Jurnal Psikologi*. Jakarta: Universitas Indonusa Esa Unggul.

Greene, W. H. 2008. *Econometrics Analysis (6th Edition)*. New Jersey: Pretince Hall, Inc.

Gujarati, D. 2004. *Basic Econometrics (4th Edition)*. New York: The McGraw-Hill.

Hosmer, D and Lemeshow, S. 2000. *Applied Logistic Regression (2nd Edition)*. New Jersey: John Wiley & Sons.

Masitoh, F dan Ratnasari, V. 2016. Pemodelan Status Ketahanan Pangan di Provinsi Jawa Timur dengan Pendekatan Metode Regresi Probit Biner. *Jurnal Institut Teknologi Sepuluh Nopember*. Surabaya: ITS.

Nakamura, H. 1990. *Perceraian Orang Jawa*. Yogyakarta: UGM Press.

Riduan. 1998. Penerapan Model Gabungan (Logistik dan Piecewise Proportional Hazard). *Tesis*.: Bogor: IPB.

Sudarto, L dan Wirawan, H. E. 2001. Penghayatan Makna Hidup Perempuan Bercerai. *Jurnal Ilmiah Psikologi*. Jakarta: Universitas Tarumanegara.

Turner, J. S and Helms, D. B. 1995. *Life-Span Development (5th Edition)*. New York: Holt, Rinehart & Winston.

Tresia, D. 2006. Faktor-faktor yang Mempengaruhi Perceraian di Sumatera Barat. *Jurnal Universitas Andalas*. Padang: Universitas Andalas.

PROYEKSI PENYERAPAN TENAGA KERJA PERIKANAN BERDASARKAN FAKTOR INDUSTRIALISASI MENGGUNAKAN METODE FUNGSI TRANSFER

Yulinda Nurul Aeni[1]

Pusat Penelitian Kependudukan – Lembaga Ilmu Pengetahuan Indonesia (P2K-LIPI)
e-mail: [1]yulindaaini@gmail.com

Abstrak

FAO menempatkan Indonesia sebagai negara dengan potensi perikanan terbesar di dunia, namun potensi tersebut belum dimanfaatkan dengan optimal. Pemerintah telah membentuk program industrialisasi perikanan, namun pelaksanaannya yang belum terimplementasi dengan baik menyebabkan penyerapan tenaga kerja di sektor ini masih rendah. Penelitian ini akan membentuk proyeksi penyerapan tenaga kerja subsektor perikanan 2019-2024 menggunakan metode fungsi transfer dengan mempertimbangkan faktor industrialisasi perikanan sebagai prediktor terhadap indeks elastisitas penyerapan tenaga kerja perikanan. Industrialisasi perikanan diukur berdasarkan faktor perkembangan investasi dan pertumbuhan jumlah perusahaan perikanan. Hasil proyeksi menunjukkan bahwa industrialisasi perikanan di tahun mendatang belum mampu mendorong respon pertumbuhan penyerapan tenaga kerja subsektor perikanan.

Kata kunci: fungsi transfer, industrialisasi, perikanan, proyeksi, tenaga kerja

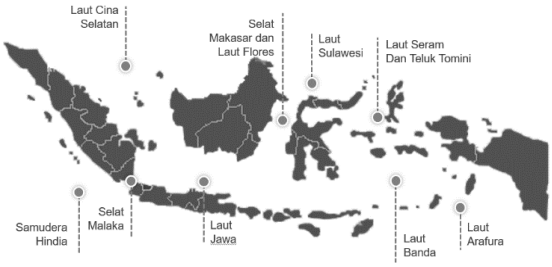
Abstract

FAO places Indonesia as the country with the largest fishery potential in the world, but this potential has not been utilized optimally. The government had established a fisheries industrialization, but this program had not been implemented properly, resulting in low employment in this sector. This research would form projections of labor absorption in the fisheries subsector 2019-2024 using the transfer function method by considering the factor of fisheries industrialization as a predictor of the fisheries labor absorption elasticity index. The industrialization of fisheries is measured based on investment development factors and growth in the number of fishing companies. The projection results showed that the industrialization of fisheries in the coming year had not been able to encourage a response to the growth of labor absorption in the fisheries subsector.

Keywords*: transfer function, industrialization, fisheries, projection, labor*

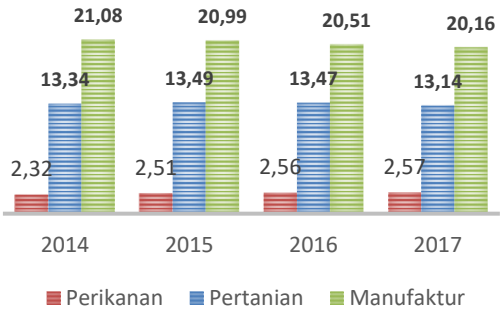
PENDAHULUAN

Sumber daya perikanan laut Indonesia tersebar di perairan wilayah Indonesia dan perairan Zona Ekonomi Eksklusif Indonesia (ZEEI) yang terbagi dalam sembilan Wilayah Pengelolaan Perikanan (KKP, 2015) dengan rincian pada Gambar 1 berikut.



Gambar 1. Wilayah Pengelolaan Perikanan (WPP) di Indonesia

Posisinya yang dikelilingi lautan, menjadikan potensi subsektor perikanan Indonesia melimpah sehingga subsektor ini menjadi salah satu subsektor yang menjanjikan bagi perekonomian Indonesia. Bahkan pada tahun 2012, *Food Agriculture Organization* (FAO) menempatkan Indonesia sebagai negara dengan potensi perikanan terbesar di dunia (FAO, 2012).



Gambar 2. Distribusi PDB Perikanan, Pertanian, dan Manufaktur 2014-2017 (Persen)

Namun, permasalahan utama dari negeri yang 75% wilayahnya berupa lautan ini adalah kesenjangan antara potensi besar tersebut dengan kemampuan memanfaatkan secara optimal. Dengan potensi perikanan yang mencapai 68 juta ton pertahun, hanya 35% yang bisa dimanfaatkan (Pusat Data Statistik dan Informasi, Kelautan dan Perikanan Dalam Angka 2018, 2018). Hal ini menjadikan subsektor perikanan memiliki kontribusi relatif rendah terhadap PDB, yaitu 2,57% di tahun 2017 (Gambar 2). Angka ini jauh lebih rendah dibanding kontribusi subsektor lain seperti pertanian yang menyumbang PDB sebesar 13,14% dan manufaktur 20,16% (BPS, Produk Domestik Bruto (PDB) Atas Dasar Harga Konstan Menurut Lapangan Usaha, 2017).

Salah satu faktor penyebab pemanfaatan potensi perikanan di Indonesia yang belum optimal adalah karena rendahnya kualitas dan kuantitas produksi yang dihasilkan oleh nelayan yang disebabkan oleh aksesibilitas dan ketersediaan infrastruktur yang belum memadai. Di sisi lain, prosedur pelayanan perizinan usaha yang dianggap sulit membuat minimnya investasi di subsektor ini sehingga membuat industri sulit bangkit (Mariza, Wicaksono, dn Octavia, 2016)

Pemanfaatan potensi perikanan yang belum optimal menjadi penghambat pemerintah yang di tahun 2014 telah membentuk program untuk menjadikan Indonesia sebagai poros maritim dunia (Kominfo, 2018). Selanjutnya, pemerintah menurunkan berbagai kebijakan dalam program-program aksi, salah satunya adalah pelaksanaan industrialisasi kelautan dan perikanan yang bertujuan untuk mempercepat pembangunan subsektor kelautan dan perikanan di Indonesia (KKP, 2012). Tidak lanjut dari kebijakan tersebut, pemerintah mengeluarkan rencana aksi percepatan pembangunan industri perikanan nasional melalui Peraturan Presiden No.3 Tahun 2017. Dalam pelaksanaanya, rencana aksi tersebut belum terimplementasi dengan efektif karena sedikitnya program nyata pemerintah untuk merealisasikan isi yang tertuang dalam perpres tersebut. Padahal apabila pemerintah memberikan perhatian yang serius, subsektor ini akan memberikan kontribusi yang lebih besar terhadap pembangunan ekonomi, terutama untuk masyarakat nelayan dan petani ikan, serta dapat menyerap tenaga kerja lebih banyak (Subri, 2007).

Dalam aspek ketenagakerjaan, dengan pertumbuhan ekonomi yang kurang dari 5%, hanya sekitar 15 juta tenaga kerja yang terserap. Subsektor ini baru berkontribusi sebesar 13,1% terhadap total jumlah tenaga kerja di semua subsektor yang pada akhir 2015 tercatat sebanyak 114,8 juta orang. Jumlah tersebut relatif kecil dibanding kontribusi subsektor lain, seperti pertanian yang menyerap sebanyak 37,75 juta tenaga kerja (32,88%) (BPS, 2015). Selain itu, jumlah nelayan dan rumah tangga perikanan (RTP) juga mengalami pertumbuhan yang kurang signifikan. Dari tahun 2010-2014, pertumbuhan tenaga kerja di subsubsektor perikanan tangkap hanya sebesar 1,045% dengan kenaikan jumlah tenaga kerja sebanyak 119 ribu nelayan. Angka ini tidak jauh berbeda dengan subsubsektor perikanan budidaya yang mengalami pertumbuhan tenaga kerja sebesar 1,20% dari tahun 2010-2014 dengan kenaikan jumlah tenaga kerja sebanyak 648,5 ribu pembudidaya (KKP, 2015).

Terdapat beberapa penelitian yang membahas mengenai industrialisasi perikanan, diantaranya adalah penelitian yang dilakukan oleh Poernomo dan Heruwati (2011) mengenai dampak positif dan negatif pelaksanaan program industrialisasi perikanan dan penelitian oleh Huda, dkk (2015) mengenai industrialisasi perikanan dalam pengembangan wilayah Jawa Timur. Penelitian-penelitian tersebut lebih banyak membahas mengenai pelaksanaan industrialisasi perikanan, namun masih sedikit yang membahas mengenai dampaknya terhadap penyerapan tenaga kerja di sektor tersebut. Karena itu, penelitian ini akan membahas mengenai penyerapan tenaga kerja berdasarkan faktor industrialisasi.

Berdasarkan latar belakang di atas, penelitian ini bertujuan untuk membuat proyeksi penyerapan tenaga kerja subsektor perikanan dalam lingkup nasional tahun 2019-2024 berdasarkan faktor industrialisasi perikanan. Proyeksi digunakan untuk melihat pengaruh perkembangan jumlah investasi dan jumlah perusahaan perikanan dalam meramalkan penyerapan tenaga kerja subsektor perikanan nasional. Proyeksi tenaga kerja perikanan yang tepat diperlukan oleh pengambil kebijakan untuk memperkirakan penyerapan tenaga kerja di masa depan sehingga pemanfaatan potensi perikanan di Indonesia akan optimal dan mendukung kebijakan program-program pembangunan sektor perikanan.

METODOLOGI PENELITIAN

1. Konsep dan Definisi

Rao

Berikut merupakan beberapa konsep dan definisi variabel yang digunakan dalam penelitian.

a. Menurut KKP (2015), penyerapan tenaga kerja pada subsektor perikanan dibagi pada kegiatan perikanan tangkap, perikanan budidaya, pengolahan, dan pemasaran, serta jasa penunjang lainnya yang meliputi tenaga kerja yang terlibat pada program-program pemberdayaan di subsektor perikanan.
b. Elastisitas penyerapan tenaga kerja merupakan rasio dari jumlah tenaga kerja dan PDB subsektor perikanan nasional (Dumairy, 2004). Data PDB yang digunakan adalah PDB atas dasar harga konstan tahun 2000 karena telah menghilangkan pengaruh inflasi sehingga angka yang dihasilkan mencerminkan pertumbuhan riil.
c. Industrialisasi subsektor perikanan sebagai prediktor diproksi menggunakan variabel pertumbuhan jumlah perusahaan dan perkembangan investasi perikanan nansional. Adapun perkembangan jumlah investasi dihitung dari akumulasi Penanaman Modal Asing (PMA) dan Penanaman Modal Dalam Negeri (PMDN).

2. Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder berupa data *time series* tahun 2000-2016. Keterangan mengenai nama variabel, sumber, tipe dan skala data ditampilkan pada Tabel 1 berikut.

Tabel 1. Jenis dan Sumber Data Penelitian

1. Nama Variabel	*2.* Sumber Data	Tipe, Skala Data
Tenaga Kerja Subsektor Perikanan Nasional Tahun 2000-2016 (Juta Orang)	Dirjen Perikanan Tangkap dan Perikanan Budidaya, Kementerian Kelautan dna Perikanan (KKP)	Numerik, Rasio
Produk Domestik Bruto (PDB) Subsektor Perikanan Nasional 2000-2016 (%)	Badan Pusat Statistik (BPS)	Numerik, Rasio
Perkembangan Investasi Subsektor Perikanan Nasional Tahun 2000-2016 (Juta Rupiah)	*National Single Windows of Investment (NSWI*), Badan Koordinasi Penanaman Modal (BKPM	Numerik, Rasio
Pertumbuhan Jumlah Perusahaan Perikanan 2000-2016 (Instansi)	Badan Pusat Statistik (BPS)	Numerik, Rasio

3. Metode Analisis

Penelitian mengenai proyeksi penyerapan tenaga kerja pernah dilakukan oleh Kementerian Pertanian untuk tahun 2013-2019 menggunakan metode *Moving average* (MA) dan geometri (Pusat Data dan Informasi Pertanian, 2013). Selain itu, proyeksi ketenagakerjaan juga pernah dilakukan menggunakan metode pertumbuhan eksponensial dan fungsi derivasi *Cobb-Douglas* 2 *input* (Junaidi dan Zulfanetti, 2016). Penelitian lain mengenai proyeksi tenaga kerja menggunakan metode geometri dan ekstrapolasi juga pernah dilakukan. Penelitian ini mengasumsikan bahwa pertumbuhan dan elastisitas penyerapan tenaga kerja tidak berubah dari tahun ke tahun (Kumalasari, 2012). Dari beberapa penelitian tersebut, metode yang paling banyak digunakan adalah metode deterministik. Hanya sedikit penelitian yang menggunakan metode probabilistik berdasarkan pola dan karakteristik data serta melihat pengaruh dari variabel lain yang mungkin mempengaruhi, seperti metode fungsi transfer. Untuk itu, dalam penelitian ini, akan digunakan metode fungsi transfer dengan melihat pengaruh industrialisasi perikanan terhadap penyerapan tenaga kerja.

Fungsi transfer merupakan gabungan dari model *Autoregressive Integrated Moving average* (ARIMA) *univariate* dan analisis regresi berganda sehingga menjadi satu model yang mencampurkan pendekatan deret berkala dengan pendekatan kausal. Seluruh sistem tersebut adalah sistem yang dinamis, dengan kata lain deret *input* memberikan pengaruhnya kepada deret *output* melalui fungsi transfer (Makridakis, 1999). Pemodelan dilakukan secara serentak menggunakan deret *input* perkembangan jumlah perusahaan perikanan dan jumlah investasi terhadap elastisitas penyerapan tenaga kerja, sehingga model fungsi transfer yang digunakan adalah multi *input*.

Berikut merupakan langkah analisis dalam membentuk pemodelan fungsi transfer multi *input* (Wei, 2006).

1. Mengidentifikasi deret *input* dan *output*.
2. Uji stasioneritas *mean* menggunakan uji *Augmented Dickey Fuller* dan uji stasioneritas varians menggunakan *Box-Cox Transformation.* Jika deret *input* atau deret *output* belum stasioner dalam varians maka dilakukan transformasi, dan jika belum stasioner dalam *mean* akan dilakukan *differencing.* Deret data yang telah stasioner disebut x_t dan y_t.
3. Identifikasi model ARIMA untuk seluruh deret *input* dengan melihat plot *Auto-Correlation Function* (ACF) dan *Partial Auto-Correlation Function* (PACF) untuk mengetahui lag yang signifikan dan ada/tidaknya periode musiman, sehingga orde model ARIMA dapat diketahui (Tabel 2).
4. Uji signifikansi parameter model ARIMA.
5. Uji asumsi residual yang meliputi uji *white noise* dan uji normalitas.
6. Pemilihan model ARIMA terbaik berdasarkan *nilai Root Mean Square Error* (RMSE) dan *Mean Absolute Persentage Error* (MAPE) terkecil, serta nilai *R-squared* tertinggi.

Tabel 2. Karakteristik Plot ACF dan PACF

Proses	ACF	PACF
AR (*p*)	Garis cenderung turun cepat mengikuti pola gelombang sinus	Garis terputus ketika menuju angka nol setelah lag-p
MA (*q*)	Garis terputus ketika menuju angka nol setelah lag-p	Garis cenderung turun cepat mengikuti pola gelombang sinus
ARIMA (*p*,d,*q*)	Garis secara terus menerus menuju nol	Garis secara terus menerus menuju nol

7. Melakukan *prewhitening* deret *input* dan *output* untuk menghilangkan seluruh pola yang diketahui agar yang tertinggal hanya *white noise.*

   Saat deret *input* dalam kondisi stasioner, maka dilanjutkan dengan penentuan model ARIMA pada deret *input*. Setelah model ARIMA yang sesuai untuk deret *input* terbentuk, maka dilakukan proses *prewhitening.* Model untuk deret *input* yang telah di-*prewhitening* adalah sebagai berikut.

$$\frac{\phi_x(B)}{\theta_x(B)} x_t = \alpha_t \qquad (1)$$

   Keterangan:
   ϕ_x : Parameter *autoregressive* deret *input*
   θ_x : Parameter *moving average* deret *input*
   x_t: Deret *input* yang stasioner
   α_t : Deret *noise*

   Deret *output* dimodelkan menggunakan parameter ARIMA yang terbentuk pada deret *input*. *Prewhitening* pada deret *output* dilakukan dengan cara yang sama pada deret *input*, yaitu sebagai berikut.

$$\frac{\phi_y(B)}{\theta_y(B)} y_t = \beta_t \qquad (2)$$

   Keterangan:
   ϕ_y : Parameter *autoregressive* deret *output*
   θ_y : Parameter *moving average* deret *output*
   y_t: Deret *output* yang stasioner
   β_t: Deret *noise*

8. Melakukan deteksi hubungan antara variabel *input* dan *output* dengan menggunakan *Cross-Correlation Function* (CCF).

   CCF digunakan untuk mengukur kekuatan dan arah hubungan diantara dua variabel random x_t (variabel *input*) dan y_t (variabel *output*) yang masing-masing merupakan proses *univariate* yang stationer. Fungsi *cross correlation* antara x_t dan y_t dapat ditulis sebagai berikut.

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y} \qquad (3)$$

   Keterangan:
   σ_x : Standar deviasi x_t
   σ_y : Standar deviasi y_t
   ρ_{xy} : Fungsi *Cross Correlation*
   γ_{xy} : Covarian x_t dan y_t

9. Identifikasi awal model fungsi transfer dengan menentukan orde model fungsi transfer (*b, r, s*) berdasarkan plot korelasi silang.
   a. Nilai *b* menjelaskan bahwa y_t tidak dipengaruhi oleh nilai x_t sampai periode *t+b*, besarnya *b* sama dengan jumlah bobot *respon impuls* yang tidak berbeda dari nol secara signifikan.
   b. Nilai *s* menyatakan berapa lama deret *output* y_t secara terus menerus dipengaruhi oleh nilai-nilai baru dari deret *input* (x_t). y_t dipengaruhi oleh x_{t+b}, x_{t+b+1},..., x_{t+b+s},. Nilai *s* adalah jumlah bobot *respon impuls* sebelum terjadinya pola menurun.
   c. Nilai *r* menunjukkan bahwa y_t berkaitan dengan nilai-nilai masa lalu dari y_t, yaitu y_{t-1}, y_{t-2},..., y_{t-r}. Terdapat tiga kondisi pada nilai *r* yang mempunyai indikasi pemodelan berbeda, yaitu *r=0*, bila jumlah bobot *respon impuls* hanya terdiri dari beberapa lag yang kemudian

langsung terpotong, r=1, bila bobot *respon impuls* menunjukkan suatu pola eksponensial yang menurun, dan r=2, bila bobot *respon impuls* menunjukkan suatu pola eksponensial menurun dan mengikuti pola sinusoidal.

10. Identifikasi *noise function* dengan melihat plot ACF dan PACF dari identifikasi awal model fungsi transfer.

$$n_t = y_t - \sum_{j=1}^{m}\left[\delta_j(\beta)\right]^{-1}\omega_j(B)^{x}_{j,t-b_j} \quad (4)$$

Keterangan:
δ_j dan ω_j : konstansa fungsi transfer
n_t : Deret *noise*

11. Identifikasi akhir model fungsi transfer multi *input* dengan mengkombinasikan model awal dan *noise function.*

$$y_t = \frac{\omega_s(B)}{\delta_r(B)}x_{t-b} + \frac{\theta(B)}{\phi(B)}a_t. \quad (5)$$

12. Melakukan proyeksi menggunakan model fungsi transfer multi *input*.

HASIL ANALISIS DAN PEMBAHASAN

Industrialisasi Kelautan dan Perikanan

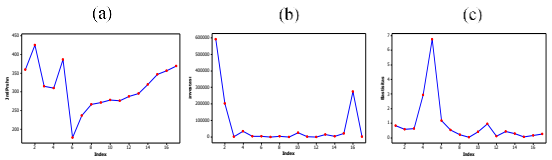
Dalam teori Colin Clark & Simon Kuznets, industrialisasi dianggap sebagai proses pertumbuhan ekonomi dalam wujud akselerasi investasi dan tabungan. Jika tabungan cukup tinggi, maka kemampuan sebuah negara untuk mengadakan investasi juga meningkat sehingga target pertumbuhan ekonomi dan penciptaan lapangan kerja lebih mungkin dicapai secara cepat (Hakim, 2009).

Di bidang kelautan dan perikanan, KKP telah membentuk peraturan terkait indus-trialisasi. Dalam peraturan tersebut dijelaskan bahwa industrialisasi kelautan dan perikanan merupakan integrasi sistem produksi hulu dan hilir untuk meningkatkan skala dan kualitas produksi, produktivitas, daya saing, dan nilai tambah sumber daya kelautan dan perikanan secara berkelanjutan (KKP, 2012). Industrialisasi perikanan ini bertujuan untuk meningkatkan kesejahteraan nelayan pem-budidaya, pengolah, dan pemasar hasil perikanan, menyerap tenaga kerja, serta meningkatkan devisa negara. Industrialisasi perikanan ini berprinsip untuk mendorong penguatan struktur industri melalui pe-ningkatan jumlah dan kualitas industri perikanan dan pembinaan hubungan antar entitas sesama industri pada semua tahapan rantai nilai *(value chain*) (Republik Indonesia, 2016). Selain itu, kebijakan industrialisasi perikanan dan investasi akan diarahkan untuk mendorong kemitraan usaha yang saling menguntungkan melalui pengembangan komoditas nasional dan produk-produk inovatif dan kompetitif di pasar global (Bappenas, 2016).

Pengembangan usaha dan investasi yang menjadi salah satu strategi pelaksanaan indus-trialisasi perikanan bertujuan untuk men-dorong kemitraan usaha yang saling meng-untungkan antara usaha skala mikro, kecil, dan menengah dengan usaha skala besar melalui pengembangan komoditas nasional dan produk-produk yang kompetitif di pasar global. Dengan pengembangan usaha dan investasi, industri/perusahaan perikanan skala kecil dan menengah diharapkan akan dapat berkembang sehingga dapat memperkuat basis industri perikanan secara nasional. Hal ini menunjukkan bahwa pelaksanaan industri-alisasi perikanan dapat diukur berdasarkan pertumbuhan jumlah perusahan/industri per-ikanan serta perkembangan investasi di sektor ini.

Identifikasi Model Deret Input dan Deret Output

1. Identifikasi Pola Variabel



Gambar 3. Plot *Time Series* (a) Variabel Jumlah Perusahaan, (b) Jumlah Investasi, (c) Elastisitas Penyerapan Tenaga Kerja

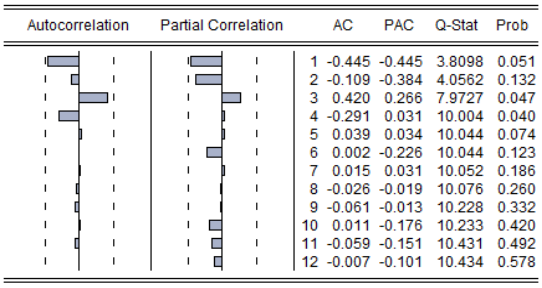
Identifikasi pola variabel digunakan untuk melihat ada tidaknya pola tertentu, seperti pola musiman, tren, dan lain-lain yang nantinya akan berpengaruh terhadap model fungsi transfer yang digunakan.

Identifikasi pola variabel menggunakan *time series plot* dengan hasil pada Gambar 3.

Hasil identifikasi pada Gambar 3 menunjukkan bahwa pola data untuk variabel jumlah perusahaan, jumlah investasi, dan elastisitas penyerapan tenaga kerja subsektor perikanan berfluktuasi dan tidak membentuk pola tertentu, sehingga pemilihan model fungsi transfer dalam melakukan proyeksi telah sesuai. Adapun model fungsi transfer yang digunakan adalah fungsi transfer multi *input non-seasonal*, karena pola data tidak mengindikasikan pola musiman.

2. Identifikasi Model Deret *Input* Jumlah Perusahaan

Dalam identifikasi model untuk masing-masing deret *input*, terdapat beberapa tahapan yang dilakukan, seperti uji stasioner *means* dan *variances*, serta penentuan dan pemilihan model ARIMA terbaik.



Gambar 4. Plot ACF dan PACF Deret *Input* Jumlah Perusahaan

Tahapan pertama adalah penentuan ordo model ARIMA berdasarkan plot ACF dan PACF pada Gambar 4. Dari identifikasi plot ACF dan PACF, jumlah lag yang signifikan pada grafik PACF sebanyak 2 (p=2) dan pada grafik ACF sebanyak 1 (p=1). Dengan jumlah *differencing* sebanyak 1, maka terdapat 5 kombinasi model tentatif ARIMA yang terbentu, yaitu (2,1,1), (2,1,0), (1,1,1), (1,1,0), dan (0,1,1).

Selanjutnya dilakukan pengujian pada residual yang meliputi uji *white noise* menggunakan uji *Ljung-Box* dan uji normalitas menggunakan uji *Kolmogorov-Smirnov* dengan hasil pada Tabel 3 berikut.

Tabel 3 menunjukkan bahwa semua model tentatif ARIMA memiliki nilai *p-value* lebih dari taraf sinifikansi 0.05 (5%), sehingga residual untuk masing-masing model ARIMA telah memenuhi asumsi *white noise*. Namun berdasarkan nilai *p-value* uji normalitas, diketahui bahwa hanya model ARIMA (2,1,1) dan ARIMA (2,1,0) yang memenuhi asumsi distribusi Normal. Selanjutnya, dari beberapa model tentatif dilakukan pemilihan model terbaik berdasarkan nilai *Akaike's Information Criterion* (AIC), *Schwartz Bayesian Criterion* (SBC), dan *Mean Square Error* (MSE) pada Tabel 4 berikut.

Tabel 4 menunjukkan bahwa model ARIMA (2,1,1) merupakan model terbaik dan layak digunakan untuk peramalan karena memiliki nilai AIC dan SIC terkecil dibandingkan model ARIMA lainnya.

Tabel 3. Pengujian Residual Model

Model	Uji *White noise*		Uji Normalitas	
	P-value	Keputusan	*P-value*	Keputusan
ARIMA (2,1,1)	0,438	*White noise*	>0,15	Normal
ARIMA (2,1,0)	0,567	*White noise*	0,038	Normal
ARIMA (1,1,1)	0,697	*White noise*	<0,01	Tidak Normal
ARIMA (1,1,0)	0,764	*White noise*	<0,01	Tidak Normal
ARIMA (0,1,1)	0,739	*White noise*	<0,01	Tidak Normal

Tabel 4. Nilai AIC, SBC, dan MSE Model ARIMA Deret Input Jumlah Perusahaan

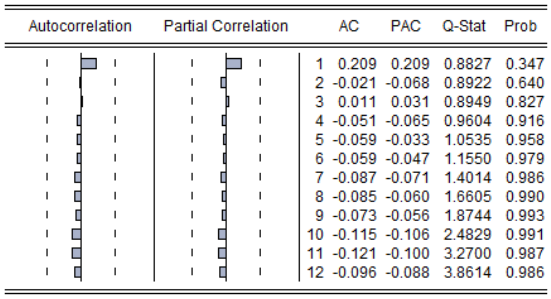
Model	**AIC**	**SIC**	**MSE**
ARIMA (2,1,1)	9,9060	10,089	3741,3
ARIMA (2,1,0)	11,453	11,500	3640,4
ARIMA (1,1,1)	11,211	11,307	4071,9
ARIMA (1,1,0)	11,362	11,410	3977,1
ARIMA (0,1,1)	13,270	13,319	3873,5

Berikut merupakan mode ARIMA (2,1,1) untuk deret *input* jumlah perusahaan.

$$X_t = 2{,}098X_{t-1} + 1{,}545X_{t-2} - 0{,}447X_{t-3} + a_t - 0{,}968a_{t-1}$$

3. Identifikasi Model Deret *Input* Jumlah Investasi

Berikut merupakan tahapan-tahapan identifikasi model ARIMA untuk deret *input* jumlah investasi. Tahapan pertama adalah penentuan ordo model ARIMA berdasarkan plot ACF dan PACF di Gambar 5.



Gambar 5. Plot ACF dan PACF Deret Input Jumlah Investasi

Berdasarkan Gambar 5, plot ACF signifikan pada lag ke-1 dan plot PACF signifikan pada lag ke-1, sehingga model tentatif ARIMA yang terbentuk adalah ARIMA (1,0,1), AR(1), dan MA(1). Berikut merupakan uji residual dari masing-masing model tentatif.

Tabel 5 menunjukkan bahwa semua model tentatif ARIMA memiliki nilai *p-value* lebih dari taraf sinifikansi 0.05 (5%), sehingga residual untuk masing-masing model ARIMA telah memenuhi asumsi *white noise* dan berdistribusi Normal. Untuk itu dilakukan pemilihan model terbaik berdasarkan nilai AIC, SIC, dan MSE terkecil.

Berdasarkan Tabel 6, terlihat bahwa model dengan nilai AIC, SIC, dan MSE terkecil adalah ARIMA (1,0,0), sehingga model ini merupakan model tentatif terbaik untuk meramalkan deret *input* jumlah investasi. Berikut merupakan persamaan model ARIMA (1,0,0) untuk deret *input* jumlah investasi.

$$y_t = 0{,}270y_{t-1} + \alpha_t$$

4. Pre-whitening Deret Input dan Output

Prewhitening dilakukan berdasarkan identifikasi model ARIMA pada masing-masing deret input dan deret output. Model prewhitening deret input jumlah perusahaan adalah sebagai berikut.

$$\alpha_t = x_t - 0{,}2098x_{t-1} - 1{,}545x_{t-2} + 0{,}447x_{t-3} + 0{,}968\alpha_{t-1}$$

Dengan cara yang sama, model prewhitening untuk deret input jumlah investasi (y_t) adalah sebagai berikut.

$$\alpha_t = y_t - 0{,}270y_{t-1}$$

Prewhitening deret output elastisitas penyerapan tenaga kerja adalah sebagai berikut.

$$\beta_t = z_t - 2{,}098z_{t-1} - 1{,}545z_{t-2} + 0{,}447z_{t-3} + 0{,}968\beta_{t-1}$$

$$\beta_t = z_t - 0{,}270z_{t-1}$$

5. Menghitung Korelasi Silang Antar Variabel

Deret input dan deret output yang telah melalui proses prewhitening, selanjutnya dilakukan perhitungan korelasi silangnya. Nilai dari korelasi silang digunakan untuk mengidentifikasi orde model awal fungsi transfer (b,r,s).

6. Identifikasi Awal Model Fungsi Transfer

Identifikasi model awal fungsi transfer dilakukan dengan melihat plot korelasi silang antara deret *input* dan deret

Tabel 5. Pengujian Residual Model

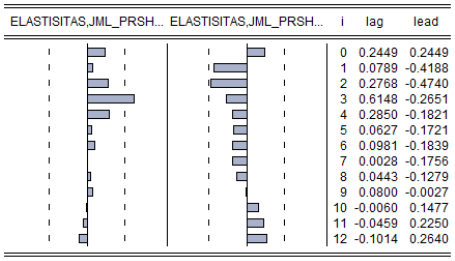
Model	Uji *White noise*		Uji Normalitas	
	P-value	Keputusan	*P-value*	Keputusan
ARIMA (1,0,1)	0,542	*White noise*	>0,150	Normal
ARIMA (1,0,0)	0,552	*White noise*	>0,150	Normal
ARIMA (0,0,1)	0,514	*White noise*	>0,150	Normal

Tabel 6. Nilai AIC, SBC, dan MSE Model ARIMA Deret Input Jumlah Investasi

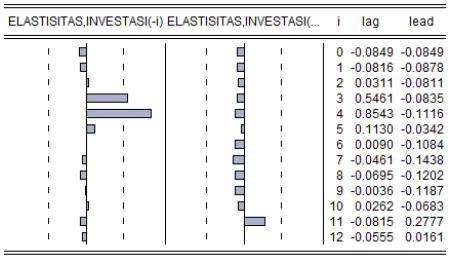
Model	AIC	SBC	MSE
ARIMA (1,0,1)	25,420	25,517	7,334
ARIMA (1,0,0)	25,349	25,397	7,043
ARIMA (0,0,1)	26,271	26,320	7,071

output pada Gambar 6. Untuk orde fungsi transfer deret *input* jumlah perusahan adalah b=3, r=0, dan s=0. Selanjutnya dilakukan *overfitting* model untuk memperoleh model fungsi transfer terbaik. Hasil dari kandidat model beserta nilai *R-squared*, RMSE, dan MAPE model dapat dilihat pada Tabel 7 berikut.

(a)



(b)



Gambar 6. Korelasi Silang Elastisitas Penyerapan Tenaga Kerja dan (a) Jumlah Perusahaan, (b) Jumlah Investasi

Tabel 7. Identifikasi Model Fungsi Transfer untuk Deret Input Jumlah Perusahaan

Model	R^2	RMSE	MAPE
b=1, s=0, r=0	0,539	1,340	272,197
b=2, s=0, r=0	0,635	1,316	370,322
b=3, s=0, r=0	0,703	1,344	275,680
b=3, s=1, r=0	0,701	1,456	188,708
b=3, s=2, r=0	0,055	2,837	480,732
b=3, s=3, r=0	0,113	3,073	420,327

Tabel 7 menunjukkan bahwa model fungsi transfer (3,1,0) merupakan model terbaik karena memiliki nilai RMSE dan MAPE terkecil serta memiliki nilai R^2 di atas 0,7. Sehingga model awal fungsi transfer untuk deret *input* jumlah perusahaan adalah sebagai berikut.

$$z_t = \frac{\omega(B)}{\delta(B)} x_{t-3} + n_t$$
$$z_t = (\omega_0 - \omega_0 B) x_{t-3} + n_t$$
$$z_t = (0{,}019 - 0{,}014B) x_{t-3} + n_t$$
$$z_t = 0{,}019 x_{t-3} - 0{,}014 x_{t-4} + n_t$$

Adapun orde fungsi transfer deret *input* jumlah investasi adalah b=3, r=2, s=1. Berikut merupakan identifikasi model fungsi transfer terbaik berdasarkan *overfitting model*.

Tabel 8. Identifikasi Model Fungsi Transfer untuk Deret Input Jumlah Investasi

Model	R^2	RMSE	MAPE
b=1, s=1, r=2	0,746	1,040	274,348
b=2, s=1, r=2	0,890	0,730	180,418
b=3, s=1, r=0	0,951	0,512	166,262
b=3, s=2, r=0	0,974	0,395	144,327
b=3, s=3, r=0	0,975	0,420	129,379
b=3, s=0, r=1	0,981	0,305	77,3470
b=3, s=1, r=2	0,423	0,363	141,989

Tabel 8 menunjukkan bahwa model fungsi transfer (3,0,1) memiliki nilai MAPE dan RMSE terkecil, serta memiliki nilai R^2 tertinggi. Sehingga model awal fungsi transfer untuk deret *input* jumlah investasi adalah sebagai berikut.

$$z_t = 0{,}000007763 y_{t-3} + n_t$$

Setelah diperoleh model awal fungsi transfer untuk masing-masing deret *input*, dilakukan pendugaan model awal fungsi transfer bersama antara x_t, y_t, dan z_t. Sehingga model fungsi transfer *input* ganda awal adalah sebagai berikut.

$$z_t = 0{,}019 x_{t-3} - 0{,}014 x_{t-4} + 0{,}000007763 y_{t-3} + n_t$$

7. Identifikasi *Noise Function*

Model fungsi transfer *input* ganda awal selanjutnya digunakan untuk menghitung *noise function* (n_t) dari model. Namun sebelumnya, dilakukan pengujian asumsi residual *white noise* dan

Tabel 9. Uji Asumsi Deret Noise

Deret *Input*	Uji *White Noise*		Uji Normalitas	
	P-value	Keterangan	P-value	Keterangan
Jumlah Perusahaan	0,732	*White noise*	0,233	Normal
Jumlah Investasi	0,283	*White noise*	0,068	Normal

berdistribusi Normal dengan hasil sebagai berikut.

Tabel 9 menunjukkan bahwa deret *noise* telah memenuhi asumsi *white noise* dan berdistribusi Normal untuk masing-masing deret *input*. Selanjutnya, dilakukan penghitungan nilai n_t dengan melakukan transformasi terhadap model awal, sehingga diperoleh persamaan berikut.

$$n_t = z_t - 0{,}019x_{t-3} + 0{,}014x_{t-4} - 0{,}000007763y_{t-3}$$

Dari *noise function* tersebut selanjutnya dilakukan identifikasi model ARIMA dari deret sisaan dengan beberapa alternatif model sebagai berikut.

Tabel 10 menunjukkan bahwa model ARIMA (1,0,1) merupakan model terbaik yang memiliki nilai MAPE dan RMSE kecil serta memiliki nilai R^2 yang tinggi, sehingga model ARIMA deret sisaannya adalah sebagai berikut.

$$\phi_p(B)n_t = \theta_q(B)\alpha_t$$

$$n_t = \frac{\theta_q(B)}{\phi_p(B)}\alpha_t$$

$$n_t = \frac{(1-0{,}966B)}{(1-0{,}999B}\alpha_t$$

8. Pendugaan Akhir Model Fungsi Transfer

Identifikasi model akhir fungsi transfer dilakukan dengan mengkombinasikan model awal fungsi transfer *input* ganda dengan model ARIMA deret sisaan. Berikut merupakan model fungsi transfer yang diperoleh.

$$z_t = 0.019x_{t-3} - 0.014x_{t-4} + 0.000007763y_{t-3} + \frac{(1-0.966B)}{(1-0.999B)}\alpha_t$$

Berdasarkan model tersebut, dapat diketahui bahwa elastisitas penyerapan tenaga kerja subsektor perikanan pada tahun ke-*t* dipengaruhi oleh selisih perkembangan jumlah perusahaan pada tiga tahun (dengan koefisien 0,019) dan empat tahun (dengan koefisien -0,014) sebelum tahun t, dengan asumsi nilai dari variabel lain tetap. Dengan persamaan tersebut, elastisitas penyerapan tenaga kerja subsektor perikanan di tahun ke-*t* akan tergolong elastis jika jumlah perusahaan subsektor perikanan di atas 200 perusahaan dengan asumsi tidak ada selisih jumlah perusahaan di tahun ke *t-3* dan *t-4*. Hasil penelitian mengenai pengaruh perkembangan jumlah perusahaan perikanan terhadap penyerapan tenaga kerja juga dikemukakan oleh Napitupulu (2016) yang menjelaskan bahwa dari 8 variabel prediktor yang mem-pengaruhi penyerapan tenaga kerja baik perikanan tangkap maupun budidaya, 3 diantaranya signifikan dan memberikan pengaruh positif, yaitu nilai produksi, jumlah kapal, dan jumlah perusahaan perikanan. Untuk itu, strategi pengembangan subsektor perikanan agar dapat berkontribusi terhadap penyerapan tenaga kerja dapat dilakukan dengan menambah nilai produksi perikanan, jumlah kapal penangkap ikan, dan jumlah perusahaan perikanan tangkap dan budidaya.

Selain jumlah perusahaan perikanan, elastisitas penyerapan tenaga kerja pada tahun ke-*t* juga dipengaruhi oleh perkembangan jumlah investasi tiga tahun sebelum tahun ke-*t* dengan koefisien 0,000007763, artinya elastisitas penyerapan tenaga kerja di tahun ke-*t* akan memiliki kategori elastis jika perkembangan jumlah investasi di tahun t-3 sebesar 129 milyar atau lebih, dengan asumsi variabel lain bernilai tetap. Beberapa hasil penelitian juga menyatakan pengaruh positif jumlah investasi terhadap penyerapan tenaga kerja, diantaranya adalah penelitian oleh Luhur, dkk (2014) yang menyatakan bahwa kebijakan meningkatkan investasi pada sektor industri perikanan melalui pembangunan dan per-baikan infrastruktur, institusi, dan sumber daya manusia bisa menjadi salah satu fokus kebijakan untuk mendorong kinerja yang lebih baik pada

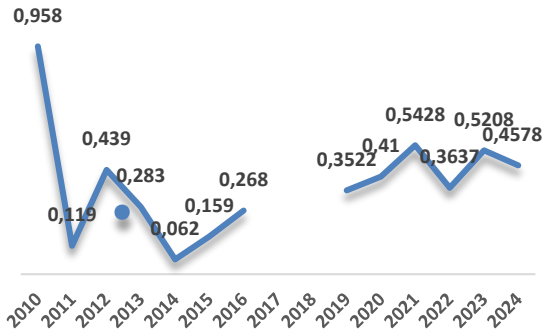
Tabel 10. Identifikasi Model ARIMA Deret Sisaan

Deret *Input*	Deret *inputt*	Model tentatif	R^2	RMSE	MAPE
b=3	b=3	ARIMA (1,0,0)	0,860	1,164	329,814
s=1	s=0	ARIMA (0,0,1)	0,856	1,178	324,437
r=0	r=1	ARIMA (1,0,1)	0,874	1,274	191,679

usaha perikanan. Penelitian lain oleh Putra (2011) menjelaskan bahwa penambahan investasi sebesar 100 milyar rupiah pada sektor perikanan berdampak pada peningkatan total *output* perekonomian yang juga menyebabkan adanya tambahan terhadap kebutuhan tenaga kerja total sekitar 78,6%.

Proyeksi Penyerapan Tenaga Kerja Subsektor Perikanan Nasional

Setelah memperoleh model akhir fungsi trasfer, dilakukan penghitungan proyeksi deret *output* elastisitas penyerapan tenaga kerja tahun 2019-2024 dengan hasil sebagai berikut.



Gambar 7. Perkembangan Indeks Elastisitas Penyerapan Tenaga Kerja

Dari tahun 2010 hingga 2016, pertumbuhan ekonomi subsektor perikanan memang cukup baik, berkisar antara 5-7,8%. Namun, pertumbuhan ekonomi yang baik tersebut tidak diimbangi dengan penyerapan tenaga kerja yang baik juga, karena nilai median elastisitas penyerapan tenaga kerja di tahun 2010-2016 hanya sebesar 0,268. Hal ini menjadikan elastisitas penyerapan tenaga kerja subsektor perikanan tergolong inelastis. Dengan indeks elastisitas tersebut, maksimum tenaga kerja yang dapat terserap pada subsektor perikanan sejumlah 100 ribu pekerja setiap tahunnya.

Meskipun demikian, Gambar 7 menunjukkan bahwa indeks elastisitas penyerapan tenaga kerja di tahun 2019-2024 cenderung mengalami pertumbuhan dibanding tahun-tahun sebelumnya. *Forecast* elastisitas penyerapan tenaga kerja menunjukkan angka yang berfluktuasi antara 0,35-0,55. Koefisien elastisitas untuk masing-masing tahun adalah positif yang menunjukkan bahwa laju pertumbuhan PDB dan penyerapan tenaga kerja subsektor perikanan akan mengalami peningkatan setiap tahunnya. Namun, karena kategori elastisitas tergolong inelastis, maka pertumbuhan ekonomi di tahun 2019-2024 tidak terlalu mendorong respon pertumbuhan penyerapan tenaga kerja subsektor perikanan. Subsektor ini akan mengalami peningkatan rata-rata penyerapan tenaga kerja setiap tahunnya, yaitu sekitar 120-180 ribu tenaga kerja dari setiap 1% pertumbuhan ekonomi di sektor ini.

Angka elastisitas ini sebenarnya jauh di bawah target yang telah ditetapkan pemerintah. Tahun 2016, pemerintah menargetkan angka elastisitas sebanyak 350 ribu serapan tenaga kerja dari setiap 1% pertumbuhan ekonomi, karena pemerintah berkeyakinan terhadap prospek perekonomian yang membaik dan implementasi program pemerintah yang bertumpu pada sektor padat karya. Namun, meskipun realisasi indeks elastisitas meleset, elastisitas penyerapan tenaga kerja tetap menjadi salah satu tolok ukur pertumbuhan berkualitas dalam masa pemerintahan 2014-2019 (Baihaqi, 2016).

Dalam pelaksanaan industrialisasi perikanan yang bertujuan untuk meng-optimalkan pemanfaatan potensi SDA perikanan dan penyerapan tenaga kerja subsektor perikanan dan kelautan, ada tiga aktor yang menjadi *triple bottom line*, yaitu pemerintah sebagai pembuat kebijakan, pelaku industri/perusahaan, serta pekerja. SDA perikanan & kelautan dimiliki dan dikelola atas nama rakyat oleh pemerintah, sehingga tujuan ekonomi dalam kerangka kerja perikanan mengharuskan pemerintah untuk memastikan manfaat ekonomi sumber daya perikanan dimaksimalkan. Untuk itu, pemerintah perlu memprioritaskan langkah-langkah manajemen yang selaras dengan tujuan ekonomi dengan melibatkan kelompok kepentingan khusus, yaitu industri perikanan dan investor (Laxe, Bermudez, dan Palmero, 2018).

Investasi menjadi salah satu langkah strategis untuk mengembangkan industrialisasi subsektor perikanan dari hulu sampai hilir. Penciptaan industri di hulu sampai hilir tentunya harus mempertimbangkan keteram-pilan yang tersedia secara lokal. Untuk itu, perlu dibuat kebijakan bertingkat dari lokal, regional, hingga nasional untuk menyesuaikan perbedaan geografis dalam struktur dan komposisi lapangan kerja subsektor perikanan dan kelautan (Putten, Cvitanovic, dan Fulton, 2016). Lebih lanjut, kebijakan di bidang investasi dan industri perikanan dan kelautan harus didasarkan pada partisipasi dan transparansi untuk mendukung penerapan strategi tata kelola sumber daya perikanan dan kelautan. Kebijakan yang memadai dan pelibatan aktor ini selanjutnya akan mengarahkan pada pencapaian pertumbuhan ekonomi (Emery dkk, 2017). Dengan demikian, PDB yang menjadi salah satu indikator pertumbuhan ekonomi perikanan bisa meningkat dan secara tidak langsung akan meningkatkan penyerapan tenaga kerja di subsektor ini.

KESIMPULAN

Hasil proyeksi elastisitas penyerapan tenaga kerja subsektor perikanan menggunakan metode fungsi transfer multi *input* dengan pertumbuhan jumlah perusahaan perikanan dan perkembangan jumlah investasi sebagai deret *input* menunjukkan bahwa di tahun 2019-2024 elastisitas penyerapan tenaga kerja lebih banyak masuk dalam kategori inelastis dengan penyerapan sebesar 120-180 ribu tenaga kerja dari setiap 1% pertumbuan ekonomi. Untuk itu, diperlukan strategi pengembangan subsektor perikanan agar dapat berkontribusi terhadap penyerapan tenaga kerja, salah satunya adalah dengan meningkatkan jumlah dan skala industri perikanan tangkap dan budidaya agar memiliki nilai tambah yang lebih tinggi. Pembinaan hubungan antar entitas sesama industri pada semua tingkatan juga diperlukan untuk mendorong kemitraan usaha yang saling menguntungkan melalui pengembangan komoditas nasional dan produk-produk inovtasi dan kompetitif di pasar global. Selain itu, kebijakan meningkatkan investasi pada sektor industri perikanan melalui pembangunan dan perbaikan infrastruktur, institusi, dan sumber daya manusia bisa menjadi salah satu fokus kebijakan untuk mendorong kinerja yang lebih baik pada usaha perikanan.

DAFTAR PUSTAKA

Baihaqi, M. (2016, 01 27). *Target Elastisitas Penyerapan Tenaga Kerja Dipertahankan*. (Harian Ekonomi Neraca) Dipetik 09 09, 2019, dari http://www.neraca.co.id/article/64765/target-elastisitas-penyerapan-tenaga-kerja-dipertahankan

Bappenas. (2016). Kajian Strategis Industrialisasi Perikanan Untuk Mendukung Pembangunan Ekonomi Wilayah. Jakarta: Bappenas.

BPS. (2015). *Keadaan Angkatan Kerja di Indonesia 2015*. Jakarta: Badan Pusat Statistik. Diambil kembali dari https://www.bps.go.id/publication/2015/11/30/311dc33e7624d47529ec4800/keadaan-angkat an-kerja-di-indonesia-agustus-2015.html

BPS. (2017). *Produk Domestik Bruto (PDB) Atas Dasar Harga Konstan Menurut Lapangan Usaha*. Dipetik Mei 2018, 07, dari Badan Pusat Statistik (BPS): https://www.bps.go.id/statictable/2009/07/02/1200/-seri-2000-pdb-atas-dasar-harga-konstan-2000-menu rut-lapangan-usaha-miliar-rupiah-2000-2014.html

Dumairy. (2004). Matematika Terapan untuk Bisnis dan Ekonomi. Yogyakarta: BPFE.

Emery, T. J., Gardner, C., Hartmann, K., & Cartwright, I. (2017). Beyond sustainability: is government obliged to increase economic benefit from fisheries in the face industry resistance? *Marine Policy, 76*, 48-54. doi:https://doi.org/10.1016/j.marpol.2016.11.018

FAO. (2012). *The State of World Fisheries and Aquaculture*. Rome: Food and

Agriculture Organization of The United States.

Hakim, M. (2009). Industrialisasi Di Indonesia : Menuju Kemitraan yang Islami. *Jurnal Hukum Islam, 7*(1), 106-121.

Huda, H. M., Purnamadewi, Y. L., & Firdaus, M. (2015). Industrialisasi Perikanan Dalam Pengembangan Wilayah Jawa Timur. *Jurnal Tata Loka, 17*(2), 99-112.

Junaidi dan Zulfanetti. (2016). Analisis Kondisi dan Proyeksi Ketenagakerjaan Di Provinsi Jambi. *Jurnal Perspektif Pembiayaan dan Pembangunan Daerah, 3*(3), 141-150. doi:https://doi.org/10.22437/ppd.v3i3.3516

Kementerian Pertanian. (2013). *Analisis dan Proyeksi Tenaga Kerja Sektor Pertanian 2013-2019.* Jakarta: Sekretariat Jenderal Kementerian Pertanian Republik Indonesia.

KKP. (2012). *Pedoman Umum Industrialisasi Kelautan dan Perikanan.* Jakarta: Kementerian Kelautan dan Perikanan Republik Indonesia. Diambil kembali dari http://jdih.kkp.go.id/peraturan/per-27-men-2012.pdf

KKP. (2012). Peraturan Menteri Kelautan dan Perikanan No.27 Tahun 2012 Tentang Pedoman Umum Industrialisasi Kelautan dan Perikanan Indonesia.

Kominfo. (2018). *Menuju Poros Maritim Dunia*. Dipetik Juli 15, 2018, dari Kementerian Komunikasi dan Informatika Republik Indonesia: https://www.kominfo.go.id/content/detail/8231/menuju-poros-maritim-dun ia/0/kerja_nyata

Kumalasari, R. (2012). Analisis Proyeksi Tenaga Kerja Daerah di Kabupaten Magetan Tahun 2011-2014. Surakarta: Penerbit Universitas Sebelas Maret.

Laxe, F. G., Bermudez, F. M., & Palmero, M. F. (2018). Governance of the fishery industry: A new global context. *Ocean and Coastal Management, 153*, 33-45. doi:https://doi.org/10.1016/j.ocecoaman.2017.12.009

Makridakis, H. (1999). *Forcasting Methods and Its Application.* Jakarya: Binarupa Aksara.

Mariza, N., Wicaksono, B., & Octavia, J. (2016). *Kebijakan Percepatan Pembangunan Industri Perikanan Nasional.* Jakarta: Center for Public Policy Transformation.

Mongabay. (2018). *Kapan Industri Perikanan Nasional Kuat Lagi ?* Dipetik 15 Juli, 2018, dari http://www.mongabay.co.id/2018/06/14/kapan-industri-perikanan-nasional-kuat-lagi/

Musyaffa, A. (2018). *Potensi Pemanfaatan Laut Indonesia Belum Optimal*. Dipetik Mei 2018, 07, dari https://www.aa.com.tr/id/ekonomi/kadin-pemanfaatan-potensi-laut-in donesia-belum-maksimal/1027120

Poernomo, A., & Heruwati, E. S. (2011). Industrialisasi Perikanan : Suatu Tantangan Untuk Perubahan. *Squalen, 6*(3), 87-94.

Pusat Data dan Informasi Pertanian. (2013). *Analisis dan Proyeksi Tenaga Kerja Sektor Pertanian 2013-2019.* Jakarta: Sekretariat Jenderal Kementerian Pertanian Republik Indonesia.

Pusat Data Statistik dan Informasi. (2015). *Kelautan dan Perikanan Dalam Angka.* Jakarta: Kementerian Kelautan dan Perikanan Republik Indonesia.

Pusat Data Statistik dan Informasi. (2018). *Kelautan dan Perikanan Dalam Angka 2018.* Diambil kembali dari Kementerian Kelautan dan Perikanan Republik Indonesia: https://kkp.go.id/setjen/satudata/page/1453-kelautan-dan-perikanan-dalam-angka

Putten, I., Cvitanovic, C., & Fulton, E. (2016). A changing marine sector in Australian coastal communities: An analysis of inter and intra sectoral industry connections and employment. *Ocean & Coastal*

Management, 131, 1-12. doi:https://doi.org/ 10.1016/j.ocecoaman.2016.07.010

Republik Indonesia. (2016). Instruksi Presiden Nomor 7 Tahun 2016 Mengenai Percepatan Pembangunan Industri Perikanan Nasional. Jakarta: Sekretariat Negara.

Republik Indonesia. (2017). Peraturan Presiden Nomor 3 Tahun 2017 Mengenai Rencana Aksi Percepatan Pembangunan Industri Perikanan Nasional. Jakarta: Sekretariat Negara.

Subri, M. (2007). *Ekonomi Kelautan.* Jakarta: PT. Raja Grafindo Persada.

Wei, W. (2006). *Time Series Analysis.* United States of America: Pearson Addison Wisley.

PENERAPAN RADIAL BASIS FUNCTION NEURAL NETWORK DALAM PENGKLASIFIKASIAN DAERAH TERTINGGAL DI INDONESIA

Vira Wahyuningrum[1]

[1]Badan Pusat Statistik Provinsi Jawa Barat
e-mail: [1]vira.2w@bps.go.id

Abstrak

Penetapan daerah tertinggal di Indonesia merupakan kasus pengklasifikasian dengan dua kategori pada variabel respon (biner). Pengklasifikasian dengan metode klasifikasi linier yang umum digunakan yaitu regresi logistik pada tahap eksplorasi data menghasilkan *misclassification* yang relatif besar, sehingga diperlukan suatu metode alternatif. Artificial Neural Network (ANN) merupakan alternatif yang menjanjikan untuk berbagai metode klasifikasi konvensional. Radial Basis Function Neural Network (RBFNN) merupakan salah satu arsitektur ANN yang populer digunakan dalam klasifikasi. Metode RBFNN menggunakan dua pendekatan yaitu *supervised* dan *unsupervised* serta dalam beberapa penelitian menghasilkan akurasi klasifikasi yang tinggi. Penelitian ini bertujuan menerapkan metode RBFNN untuk kasus klasifikasi daerah tertinggal di Indonesia untuk melihat arsitektur RBFNN yang terbentuk dan ketepatan klasifikasi yang dihasilkan. Hasil dari penelitian ini adalah penerapan RBFNN memberikan performa yang sangat baik yaitu nilai akurasi sebesar 93,48 persen, sensitivitas 81,10 persen dan spesifikasi 97,43 persen. Nilai *F-Measure* arsitektur RBFNN yang dihasilkan mencapai 85,36 persen.

Kata kunci: Neural Network, Radial Basis Function, klasifikasi, daerah tertinggal

Abstract

Determination of underdeveloped regency in Indonesia is the case with the classification of two categories on the response variable (binary). Classification with the linear classification method that is commonly used is logistic regression at the data exploration stage resulting in relatively large misclassification, so we need an alternative method. Artificial Neural Network (ANN) is a promising alternative to various conventional classification methods. Radial Basis Function Neural Network (RBFNN) is one of the popular ANN architectures used in classification. The RBFNN method uses two approaches namely supervised and unsupervised and in several studies produces high classification accuracy. This study aims to apply the RBFNN method for the classification case of underdeveloped regency in Indonesia to see the RBFNN architecture formed and the resulting classification accuracy. The results of this study are the application of RBFNN provides an excellent performance that is an accuracy value of 93.48 percent, sensitivity 81.10 percent and 97.43 percent specifications. The F-Measure value of RBFNN architecture reached 85.36 percent.

Keywords*: Neural Network, Radial Basis Function, Classification, Underdeveloped regency*

PENDAHULUAN

Pembangunan Daerah Tertinggal (PDT) merupakan suatu proses, upaya, dan tindakan secara terencana untuk meningkatkan kualitas masyarakat dan wilayah yang merupakan bagian integral dari pembangunan nasional. Percepatan PDT dimulai dengan pengidentifikasian daerah tertinggal oleh Kementerian Negara Pembangunan Daerah Tertinggal (KNPDT) dengan menggunakan kriteria yang telah ditetapkan. Status daerah sebagai daerah tertinggal ditetapkan oleh Presiden melalui Peraturan Pemerintah. Berdasarkan penetapan tersebut, KNPDT kemudian bekerja sama dengan kementerian-kementerian lain dan pemerintah daerah untuk menyusun strategi percepatan PDT.

Pengklasifikasian daerah tertinggal di Indonesia menggunakan variabel yang telah digunakan sebelumnya oleh KNPDT dengan suatu metode statistik merupakan kajian yang menarik. Hasil pengklasifikasian dengan metode statistik dapat dibandingkan ketepatannya dengan hasil penetapan daerah tertinggal oleh KNPDT. Klasifikasi merupakan pengelompokan objek ke beberapa kelompok berdasarkan variabel yang diamati. Klasifikasi daerah tertinggal merupakan kasus klasifikasi dengan dua kategori pada variabel respon (biner/dikotomi). Data yang digunakan dalam pengklasifikasian daerah tertinggal melibatkan 27 variabel dalam enam kriteria yang sangat kompleks.

Metode klasifikasi yang sering digunakan untuk data dengan respon kategorik atau biner adalah metode regresi logistik. Namun, pada regresi logistik dan metode klasifikasi klasik lainnya diperlukan asumsi awal yang harus dipenuhi agar diperoleh hasil klasifikasi yang optimal. Metode regresi logistik memiliki syarat pemenuhan asumsi diantaranya tidak terjadinya multikolinieritas pada variabel prediktornya. Dalam penelitian di bidang sosial, masalah multikolinieritas seringkali tidak bisa dihindari sehingga asumsi independensi tidak terpenuhi. Berdasarkan hasil eksplorasi data untuk mendeteksi multikolinieritas dengan nilai Variance Inflation Factor (VIF), terdapat 5 variabel prediktor yang mengandung multikolinieritas pada model regresi daerah tertinggal. Berkaitan dengan hasil eksplorasi data tersebut, metode regresi logistik memiliki keterbatasan untuk digunakan dalam pengklasifikasian daerah tertinggal dengan keseluruhan variabel yang ada. Penggunaan metode regresi logistik dengan pelanggaran asumsi akan mempengaruhi hasil klasifikasi yaitu terjadinya *misclassification* atau *error rate* yang relatif besar. Oleh karena itu, diperlukan metode alternatif yang dapat menjadi solusi dalam kasus klasifikasi ini.

Beberapa metode modern dikembangkan sebagai metode alternatif untuk membantu menyelesaikan masalah klasifikasi, salah satunya metode berbasis *machine learning* yaitu Artificial Neural Network (ANN) atau dikenal juga sebagai Jaringan Syaraf Tiruan (JST). Dalam dekade terakhir, ANN telah muncul sebagai alat menarik untuk pemodelan proses nonlinier, terutama dalam situasi di mana pengembangan fenomenologis atau model regresi konvensional menjadi tidak praktis atau rumit (Živković, 2008). ANN merupakan metode yang dapat menjadi solusi karena mampu melakukan ekstraksi hubungan antara *input* dan *output* tanpa asumsi awal dan bekerja berdasarkan informasi dari data yang tersedia. ANN memiliki kemampuan memodelkan permasalahan nonlinier kompleks yang sulit dipecahkan dengan persamaan matematis biasa (Haykin,2008).

ANN dapat diaplikasikan untuk pengklasifikasian daerah tertinggal dengan pertimbangan banyaknya variabel yang digunakan, pola hubungan antar variabel prediktor dan variabel respon yang tidak diketahui dengan jelas dan jumlah observasi yang cukup banyak (491 kabupaten). Selain itu, penelitian-penelitian sebelumnya yang menunjukkan kelebihan metode ANN dalam menyelesaikan masalah klasifikasi sangat mendukung untuk menggunakan metode tersebut dalam penelitian ini.

Penelitian mengenai klasifikasi daerah tertinggal dilakukan oleh Naibaho (2016) menggunakan metode Back Propagation Neural Network (BPNN) dan Learning Vector Quatization (LVQ) yang menghasilkan kesimpulan bahwa metode LVQ memberikan tingkat akurasi prediksi yang lebih baik daripada BPNN. Selanjutnya, penulis juga telah menerapkan metode Probalilistic Neural Network (PNN) dalam kasus klasifikasi yang sama (Wahyuningrum, 2017). Sitamahalakshmi, dkk (Sitamahalakshmi, dkk, 2011) membandingkan metode Radial Basis Function Neural Network (RBFNN) dan PNN dalam pengenalan karakter Telugu (pola tulisan bahasa Telugu di India) dan menghasilkan kesimpulan bahwa akurasi metode RBFNN secara keseluruhan lebih baik dibandingkan PNN. Metode RBFNN dipilih dalam penelitian ini karena merupakan arsitektur yang populer digunakan dalam klasifikasi. RBFNN bersifat *feed forward*, tipe jaringan (*multilayer*) serta keunggulan fungsi aktivasi yang digunakan. Berdasarkan uraian tersebut, penelitian ini bertujuan menerapkan metode RBFNN dalam pengklasifikasian daerah tertinggal di Indonesia untuk mendapatkan model/arsitektur dan melihat performa klasifikasi yang dihasilkan oleh metode tersebut.

METODOLOGI PENELITIAN

1. Tinjauan Referensi

Klasifikasi Daerah Tertinggal

Daerah tertinggal adalah daerah kabupaten yang wilayah serta masyarakatnya kurang berkembang dibandingkan dengan daerah lain dalam skala nasional. Pemerintah menetapkan 122 kabupaten di Indonesia sebagai daerah tertinggal tahun 2015-2019 dalam Peraturan Pemerintah Nomor 131 Tahun 2015. Penetapan daerah tertinggal dilakukan lima tahun sekali dan dilakukan evaluasi pada setiap tahun. Dalam Keputusan Menteri Negara PDT RI No.001/KEP/M-PDT/I/2005, suatu daerah dikategorikan sebagai daerah tertinggal dikarenakan beberapa faktor penyebab, antara lain faktor geografis, sumber daya alam, sumber daya manusia, sarana dan prasarana, daerah rawan bencana dan konflik sosial, serta kebijakan pembangunan.

Penetapan kriteria daerah tertinggal dilakukan dengan menggunakan pendekatan berdasarkan pada 6 kriteria dasar, yaitu : (i) perekonomian masyarakat, (ii) sumber daya manusia, (iii) prasarana (infrastruktur), (iv) kemampuan keuangan daerah, (v) aksesibilitas, dan (vi) karakteristik daerah (KNPDT dalam Naibaho, 2016). Keenam kriteria tersebut dibagi menjadi beberapa variabel sebagai dasar penetapan daerah tertinggal dengan total sebanyak 27 variabel (Lampiran). Klasifikasi daerah termasuk tertinggal atau tidak tertinggal ditentukan berdasarkan hasil perhitungan indeks dari nilai 27 *standardized indicators* masing-masing daerah.

Berdasarkan Panduan Penjelasan Penetapan Daerah Tertinggal yang disusun oleh KNPDT, secara ringkas tahapan penentuan daerah tertinggal adalah sebagai berikut (Naibaho, 2016):

1. Data dasar yang digunakan adalah data yang bersumber dari Podes, Susenas dan Kemampuan Keuangan Daerah (KKD) sebanyak 27 variabel.
2. Dilakukan standarisasi variabel karena satuan masing-masing variabel tidak sama. Tujuan standarisasi agar dapat diperbandingkan antar variabel.
3. Dari hasil standardisasi veriabel, selanjutnya dilakukan perkalian dengan bobot untuk masing-masing variabel.
4. Indikator dari hasil perkalian standarisasi variabel dengan masing-masing bobot variabel ditentukan arahnya berupa tanda positif atau negatif. Variabel seperti jumlah prasarana atau kemampuan keuangan daerah mempunyai arah yang negatif, dan sebaliknya. Selanjutnya dilakukan penjumlahan indeks.
5. Hasil total indeks ini yang dijadikan patokan penetapan daerah tertinggal, dimana daerah-daerah yang memiliki

total indeks di atas 0 merupakan daerah tertinggal.

Artificial Neural Network (ANN)

ANN atau dikenal juga dengan Neural Network (NN) atau JST merupakan suatu jaringan syaraf yang dibangun untuk meniru cara kerja otak manusia. ANN adalah sistem pemroses informasi yang memiliki karakteristik mirip dengan jaringan syaraf biologi (Fausett, 1994). ANN tercipta dari suatu generalisasi model matematis dari pemahaman manusia (*human cognition*) yang berdasarkan pada asumsi berikut:

1. Pemrosesan informasi terjadi pada elemen sederhana (*neuron*).
2. Sinyal mengalir di antara *neuron* melalui suatu penghubung.
3. Setiap penghubung antara *neuron* satu dengan lainnya mempunyai bobot/*weight*.
4. Setiap *neuron* akan menerapkan suatu fungsi aktivasi (biasanya nonlinier) terhadap inputnya (hasil penjumlahan sinyal yang dibobot) untuk menentukan sinyal outputnya.

Karakteristik dari ANN antara lain meliputi bentuk/pola hubungan antar *neuron* yang disebut sebagai arsitektur, metode untuk menentukan bobot hubungan yang disebut sebagai pelatihan (*training*)/algoritma dan fungsi aktivasi. Arsitektur NN merupakan pengaturan *neuron* ke dalam lapisan, pola hubungan dalam lapisan, dan di antara lapisan (Fausset, 1994). Arsitektur NN antara lain jaringan dengan lapisan tunggal (*Single Layer Net*), jaringan dengan banyak lapisan (*Multilayer Net*) dan jaringan Kompetitif (*Competitive Layer*).

Secara umum proses pembelajaran pada NN digolongkan menjadi dua, yaitu:

1. *Supervised learning* (pembelajaran dengan pengawasan), yaitu jaringan diberikan target yang akan dicapai sebagai dasar untuk mengubah bobot pada jaringan.
2. *Unsupervised learning* (pembelajaran tanpa pengawasan), yaitu jaringan mengorganisasikan sendiri bobot dalam jaringan berdaarkan parameter tanpa adanya target.

Beberapa metode/arsitektur yang dapat digunakan dalam pengklasifikasian menggunakan ANN diantaranya BPNN, PNN, RBFNN dan LVQ.

2. Metode Analisis

Radial Basis Function Neural Network (RBFNN)

RBFNN merupakan arsitektur dari ANN yang bersifat *feed forward* dan dapat melakukan proses klasifikasi dengan waktu singkat. RBFNN melakukan proses pembelajaran yang sangat cepat karena *neuron* disetel secara lokal (Sitamahalakshmi, dkk, 2011). RBFNN merupakan metode jaringan syaraf tiruan yang menggunakan fungsi aktivasi radial basis dan umum dipakai dalam kasus klasifikasi dan prediksi/peramalan. Dalam beberapa penelitian, metode RBFNN dimodifikasi dengan pendekatan *K-means cluster* dan fungsi aktivasi Gaussian sehingga meningkatkan keakurasian hasil klasifikasi. Waktu pelatihan (*training*) pada jaringan RBFNN sangat cepat dan memiliki kemampuan generalisasi yang baik (Sitamahalaksmi,dkk, 2011). RBFNN dapat mengatasi beberapa keterbatasan BPNN karena menggunakan *hidden layer* tunggal untuk pemodelan fungsi nonlinier, sehingga mampu melatih data lebih cepat dari BPNN (Ghaderzadeh, dkk, 2013). RBFNN adalah teknik alternatif yang menarik untuk masalah klasifikasi (Kotsiantis, dkk, 2007).

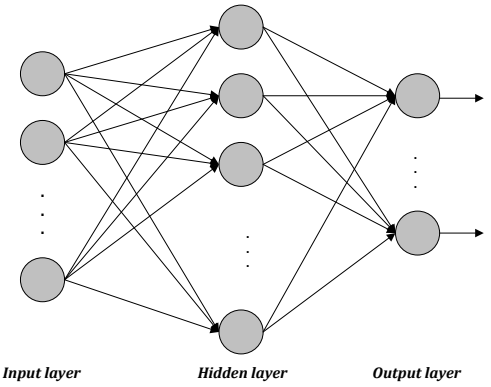
RBFNN menggunakan radial basis function yaitu Gaussian. Fungsi aktivasi Gaussian merupakan fungsi yang memperhitungkan jarak atau kedekatan antara data dengan pusat data. Fungsi Gaussian mempunyai sifat lokal, yaitu bila *input* dekat dengan rata-rata (pusat), maka fungsi akan menghasilkan nilai 1, sedangkan bila *input* jauh dari rata-rata maka fungsi akan memberikan nilai nol. Selain itu fungsi Gaussian merupakan salah satu radial basis function yang memberikan hasil terbaik dalam pengenalan pola. Jaringan RBFNN termasuk dalam tipe jaringan ANN dengan banyak lapisan atau biasa disebut dengan *multilayer*. RBFNN

dalam proses klasifikasi menghitung bobot jarak Gaussian dengan pusat *cluster*.

RBFNN pertama kali diformulasikan oleh D.S. Broomhead dan David Lowe pada sebuah penelitian tahun 1988. RBFNN mempunyai satu lapisan tersembunyi/*hidden layer* dengan fungsi aktivasi radial basis dan lapisan *output*. Jaringan RBFNN merupakan jaringan yang unik karena menggunakan dua pendekatan dalam metodenya yaitu *supervised* dan *unsupervised*. Proses pembelajaran RBFNN hanya dilakukan satu arah dan sekali saja. Jaringan syaraf tiruan fungsi *radial basis* biasanya membutuhkan jumlah *neuron* yang lebih banyak daripada jaringan *feedforward* (Kusumadewi, 2003). Pada RBFNN, *input* akan diolah oleh suatu fungsi aktivasi dan bukan merupakan hasil penjumlahan terbobot dari data *input*, namun berupa vektor jarak antara vektor bobot dan vektor *input* yang dikalikan dengan bobot bias.

RBFNN merupakan arsitektur ANN yang dikenal sangat cepat dalam melakukan proses pembelajaran (*training*). Kelebihan jaringan RBFNN antara lain waktu pelatihan yang lebih cepat dibandingkan Multilayer Perceptron (MLP) dan interpretasi *hidden layer* pada RBFNN lebih mudah dibandingkan *hidden layer* pada MLP. Sedangkan kelemahannya, pada RBFNN generalisasi tidak begitu baik dan terlalu fleksibel, sehingga menyebabkan terjadinya *noise* pada saat pelatihan dan menyebabkan kesalahan prediksi (Devega, 2013).

RBFNN terdiri dari tiga lapisan/*layer* yaitu *input layer*, *hidden layer* dan *output layer*. Setiap *neuron*/unit pada *input layer* sesuai dengan komponen dari vektor masukan. *Hidden layer* adalah satu-satunya lapisan tersembunyi dalam RBFNN yang berlaku transformasi nonlinier dari *input layer* ke dalam *hidden layer* dengan menggunakan fungsi aktivasi nonlinier. Dengan menggunakan algoritma *K-Mean Cluster*, pola pelatihan pada *hidden layer* akan terkelompok ke dalam jumlah yang wajar. *Output layer* terdiri dari *neuron* linier terhubung ke semua *neuron* tersembunyi. Jumlah *neuron* pada *output layer* sama dengan jumlah kelas data target. Arsitektur RBFNN digambarkan sebagai berikut (Sitamahalakshmi, dkk, 2011):



Gambar 1. Arsitektur RBFNN

Output dari setiap *neuron* pada *hidden layer* dihitung dengan menggunakan fungsi radial basis Gaussian sebagai berikut (Sitamahalakshmi, dkk, 2011):

$$G\left(\|x-\mu_i\|\right)=exp\left(-\frac{\|x-\mu_i\|^2}{2\sigma^2}\right). \qquad (1)$$

di mana,

x : sampel pelatihan

μ_i: pusat dari *neuron* ke i pada *hidden layer*

σ : lebar *neuron* (*spread*)

Neuron output didefinisikan oleh fungsi penjumlahan berikut (Sitamahalakshmi, dkk, 2011):

$$Y\left(x\right)=\sum_i wG\left(\|x-\mu_i\|\right)+b. \qquad (2)$$

Di mana b adalah bobot bias dan w adalah vektor bobot yang dihitung dengan mengalikan *pseudoinverse* dari matriks G dengan vektor target (d) dari data *training* dengan rumus (Sitamahalakshmi, dkk, 2011):

$$w=\left(G^TG\right)^{-1}G^Td. \qquad (3)$$

Hal yang khusus pada RBFNN adalah pemrosesan sinyal dari *input layer* ke *hidden layer* bersifat nonlinier, sedangkan dari *hidden layer* ke *ouput layer* bersifat linier. Pada *hidden layer* digunakan sebuah fungsi aktivasi berbasis radial untuk membawa *input* menuju lapisan (*layer*) berikutnya. Dari beberapa fungsi aktivasi, fungsi Gaussian merupakan fungsi aktivasi yang paling sering digunakan dalam metode

RBFNN karena mempunyai sifat lokal, berdasarkan jarak/ kedekatan *input* dengan rata-rata (*pusat cluster*). *Hidden layer* pada RBFNN dapat dilihat sebagai fungsi yang memetakan pola *input* dari ruang nonlinier ke ruang linier. Pada *hidden layer,* kekuatan diskriminatif jaringan ditentukan oleh pusat *Radial Basis Function* (RBF). Ada beberapa metode yang berbeda untuk memilih pusat/*centroid,* namun yang paling sering digunakan adalah *K-means cluster*.

Tahap pembelajaran *unsupervised* pada RBFNN adalah menentukan *mean* dan standar deviasi dari variabel *input* pada setiap *node* pada unit *hidden layer*. Metode *K-Mean Cluster* adalah salah satu dari beberapa metode *unsupervised* pada pemodelan RBFNN. Metode *K-means cluster* digunakan untuk menemukan satu set pusat yang lebih akurat mencerminkan distribusi titik data. Jumlah pusat diputuskan di awal dan setiap pusat seharusnya mewakili sekelompok titik data. Pada metode *K-means cluster,* data dipartisi kedalam K *cluster*. Penentuan nilai rata-rata dari setiap *cluster* dilakukan dengan iterasi.

Dalam metode RBFNN terdapat sejumlah parameter (*weight*) yang harus ditaksir. Untuk mendapatkan model RBFNN yang sesuai, perlu menentukan kombinasi yang tepat antara jumlah variabel *input*, jumlah *node* (*cluster*) pada unit *hidden layers*, *mean* dan standar deviasi (skala atau *width*) dari variabel *input* pada setiap *node*, yang berimplikasi pada jumlah parameter yang optimal.

Performa Klasifikasi

Salah satu cara untuk mengevaluasi performa suatu metode klasifikasi/ klasifikator adalah dengan menggunakan confusion matrix. Dari confusion matrix ini dapat dihitung beberapa parameter untuk mengukur performa klasifikasi. Menurut Han (Han dkk., 2011), confusion matrix merupakan alat yang sangat berguna untuk menganalisis seberapa baik klasifikator mengenali objek dari kelas yang berbeda. Elemen dari confusion matrix didefinisikan sebagai TP:True Positive (benar positif); FP: False Positive (salah positif); FN: False Negative (salah negatif) ;dan TN: True Negative (benar negatif).

Parameter penting dalam pengukuran performa klasifikasi yang dapat diukur berdasarkan *confusion matrix* adalah akurasi, sensitivitas dan spesifikasi. Akurasi didefinisikan sebagai perbandingan antara jumlah diagonal utama *confusion matrix* dengan jumlah keseluruhan objek. Nilai akurasi yang semakin mendekati 100 % menunjukkan performa klasifikator semakin tinggi/ baik. Sensitivitas atau disebut juga *recall* merupakan ukuran nilai ketepatan dari suatu kejadian yang diinginkan (benar positif). Sedangkan spesifikasi adalah persentase dari suatu kejadian yang tidak diinginkan (benar negatif). Formulasi penghitungan akurasi, sensitivitas dan spesifikasi berdasarkan *confusion matrix* dituliskan dalam persamaan berikut (Ghaderzadeh,dkk, 2013) :

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} x100\% \quad (4)$$

$$Sensitivitas = \frac{TP}{TP + FN} x100\% \quad (5)$$

$$Spesifikasi = \frac{TN}{FP + TN} x100\% \quad (6)$$

Nilai *error rate* juga dipakai untuk menentukan performa hasil klasifikasi di mana nilai ini menunjukkan tingkat kesalahan klasifikasi (*misclassification*)

Tabel 1. *Confusion Matrix* dengan Respon Biner untuk Klasifikasi Daerah Tertinggal

Prediksi Aktual	Daerah Tertinggal (Y=1)	Daerah Tidak Tertinggal (Y=0)	Total
Daerah Tertinggal (Y=1)	TP	FN *(type II error)*	$TP + FN$
Daerah Tidak Tertinggal (Y=0)	FP *(type I error)*	TN	$FP + TN$
Total	$TP + FP$	$FN + TN$	$TP+TN+FP+FN$

atau kelas prediksi yang tidak sesuai dengan kelas aktualnya. Semakin semakin rendah nilai *misclassification* maka semakin baik klasifikator, dan sebaliknya.

$$missclassification\ (\%) = 100 - akurasi \quad (7)$$

F-Measure dapat digunakan sebagai ukuran tunggal dari uji performa klasifikasi yang merupakan ukuran rata-rata harmonik dari sensitivitas (*recall*) dan *precision*. *Recall* dan *precision* adalah dua ukuran penting untuk mengevaluasi kebenaran algoritma klasifikasi. *F-Measure* diformulasikan sebagai berikut (Sitamahalakshmi, dkk, 2011):

$$F - Measure = \frac{2 * precision * recall}{(precision + recall)} x100\% \quad (8)$$

Penghitungan *precision* menggunakan rumus berikut:

$$precision = \frac{TP}{TP + FP} x100\% \quad (9)$$

F-Measure menggabungkan ukuran *recall* dan *precision* dengan nilai maksimum adalah 1. Bila nilai tersebut semakin mendekati 1 atau 100 persen, maka kinerja *classifier* (algoritma klasifikasi) dianggap semakin baik.

Sumber Data dan Variabel

Penelitian ini menggunakan data sekunder meliputi 27 variabel yang digunakan oleh Kementerian Negara Pembangunan Daerah Tertinggal dan Transmigrasi (KNPDT) dalam penetapan daerah tertinggal dan tidak tertinggal di Indonesia. Variabel *input* yang digunakan diantaranya berasal dari data Potensi Desa dan data Survei Sosial Ekonomi Nasional (Susenas) dari BPS serta data Kemampuan Keuangan Daerah (KKD) dari Kemenkeu. Penelitian ini menggunakan data sebanyak 491 kabupaten/kota di Indonesia. Variabel *output* (Y) merupakan *output* biner yaitu daerah tertinggal (Y=1) dan daerah tidak tertinggal (Y=0).

Sebelum dilakukan pengolahan data, perlu dilakukan *pre-processing* data untuk menyamakan satuan pada 27 variabel *input* melalui standarisasi. Standarisasi dilakukan dengan mengurangkan setiap nilai variabel x dengan rata-rata pada variabel tersebut dan membagi dengan standar deviasinya (s). Standarisasi nilai *input* akan menghasilkan nilai pada range[-1,1].

Tahapan Penelitian

Langkah-langkah yang akan dilakukan dalam pengklasfikasian daerah tertinggal dengan metode RBFNN sebagai berikut:

1. Menyiapkan *pre-processing* data yang akan digunakan (standarisasi variabel *input*)

Tabel 2. Arsitektur RBFNN pada Klasifikasi Daerah Tertinggal

Parameter	Detail
Data Sampel	Data *Training* (90%), Data *Testing* (10%)
Cross Validation	k-*Fold* (k=10)
Standarisasi Nilai *Input*	*Range* [-1,1]
Pembelajaran	Kombinasi *Supervised* (terawasi) dan *Unsupervised*
Fungsi Aktivasi	Gaussian $G\left(\left\|x-\mu_i\right\|\right) = exp\left(-\frac{\left\|x-\mu_i\right\|^2}{2\sigma^2}\right).$
Jumlah *Layer*	1 *input layer*, 1 *hidden layer*, 1 *output layer*
Jumlah *Neuron*	*Input layer* = 27 (jumlah variabel *input*) *Hidden Layer* = jumlah *cluster* (*k-means cluster*) *Output Layer* = 2 (jumlah kelas data *output*)

2. Membagi data menjadi 10 bagian (k=10) untuk *k-Fold cross validation*.
3. Menentukan jumlah *cluster* pada *hidden layer* dan melakukan klasifikasi menggunakan metode RBFNN pada data *training* dan data *testing* untuk masing-masing *fold*/bagian.
4. Menghitung performa klasifikasi meliputi akurasi, sensitivitas, dan spesifikasi dan *F-Measure*
5. Kesimpulan penelitian

HASIL DAN PEMBAHASAN

Penelitian ini meggunakan metode *k-fold cross validation* yaitu membagi data ke dalam k bagian (*fold*) secara acak dengan ukuran *fold* yang mendekati sama, di mana k adalah jumlah bagian yang ditentukan. Nilai k yang sering digunakan dalam penelitian adalah 10 (Priddy, 2005). Pembagian data *training* dan data *testing* berdasarkan 10-*fold cross validation* untuk 491 sampel yang digunakan dan selanjutnya dilakukan klasifikasi menggunakan RBFNN pada data *training* dan data *testing*.

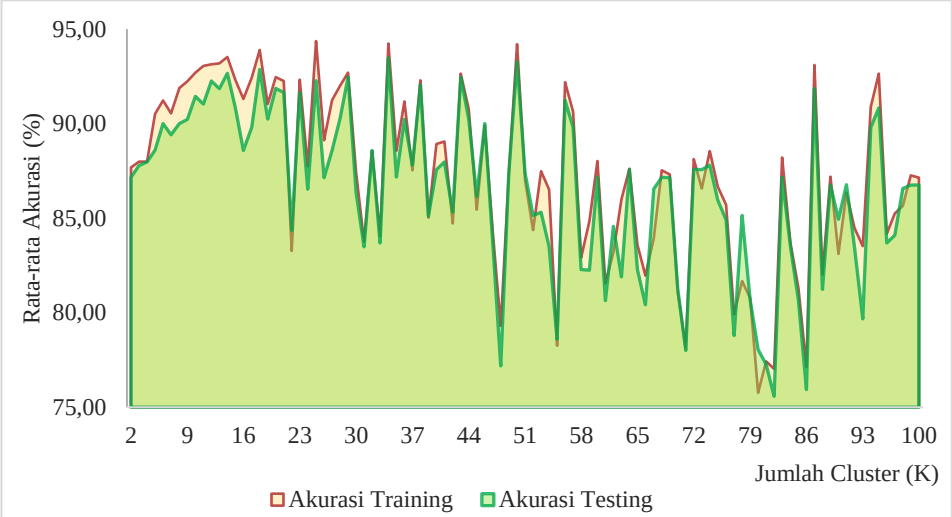
RBFNN merupakan jaringan dengan arsitektur yang terdiri atas 3 *layer* meliputi *input layer, hidden layer* dan *output layer*. Pada *input layer* terdapat sejumlah *neuron* yang merupakan jumlah variabel yang digunakan dalam pengklasifikasian yaitu sebanyak 27 variabel. Banyaknya *neuron* pada *output layer* RBFNN sama dengan jumlah kelas data yaitu 2 *neuron*. Banyak *neuron* pada *hidden layer* sebanyak *cluster* yang digunakan dalam model pengklasifikasian. *Center cluster* pada penelitian ini ditentukan dengan algoritma *K-means cluster*, dan jumlah *cluster* (K) yang diuji ditentukan pada nilai K=2 sampai dengan K=100. Rata-rata akurasi paling tinggi dihasilkan pada jumlah *cluster* k=34 yaitu sebesar 93,48 persen, sedangkan akurasi paling rendah pada jumlah *cluster* sebanyak 82 yaitu 75,57 persen. Oleh karena itu, jumlah *cluster* yang dipilih dalam arsitektur RBFNN ini adalah 34.

Jumlah *cluster* yang digunakan menggambarkan banyaknya *neuron* pada *hidden layer* RBFNN. Dengan demikian, arsitektur RBFNN yang dipilih dalam pengklasifikasian daerah tertinggal adalah RBFNN (27-34-2) tersusun oleh 27 *neuron* pada *input layer*, 34 *neuron* pada *hidden layer* dan 2 *neuron* pada *output layer*.

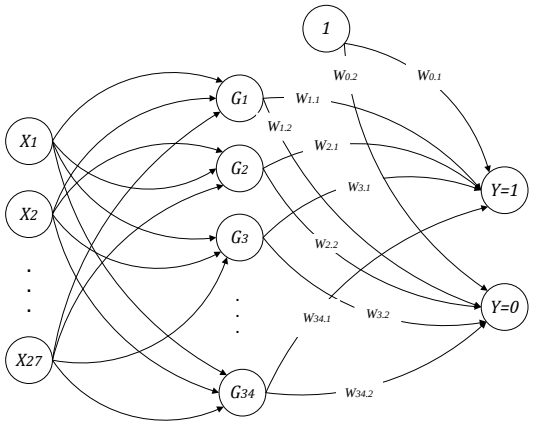
Akurasi yang dihasilkan dengan arsitektur RBFNN (27-34-2) dari data *training* melalui 10 percobaan *10-fold cross validation* berkisar antara 89,80 hingga 97,96 persen. Hasil *testing* dari 10 *fold* menunjukkan akurasi paling tinggi pada

Tabel 3. Pembagian Data Menggunakan 10-*Fold cross validation*

Data Sampel	1	2	3	4	5	6	7	8	9	10
Data *Training*	441	442	442	442	442	442	442	442	442	442
Data *Testing*	50	49	49	49	49	49	49	49	49	49



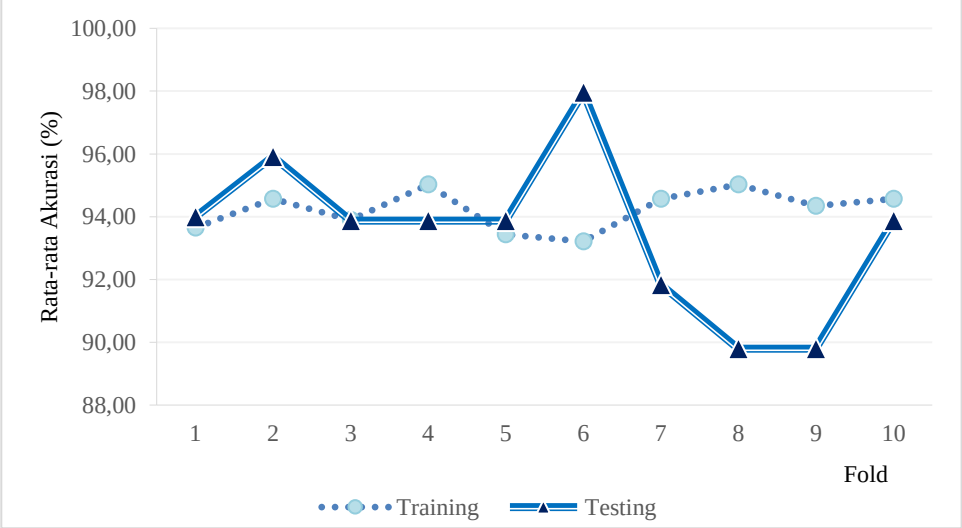
Gambar 2. Akurasi Klasifikasi RBFNN Menurut Jumlah *Cluster* (persen)



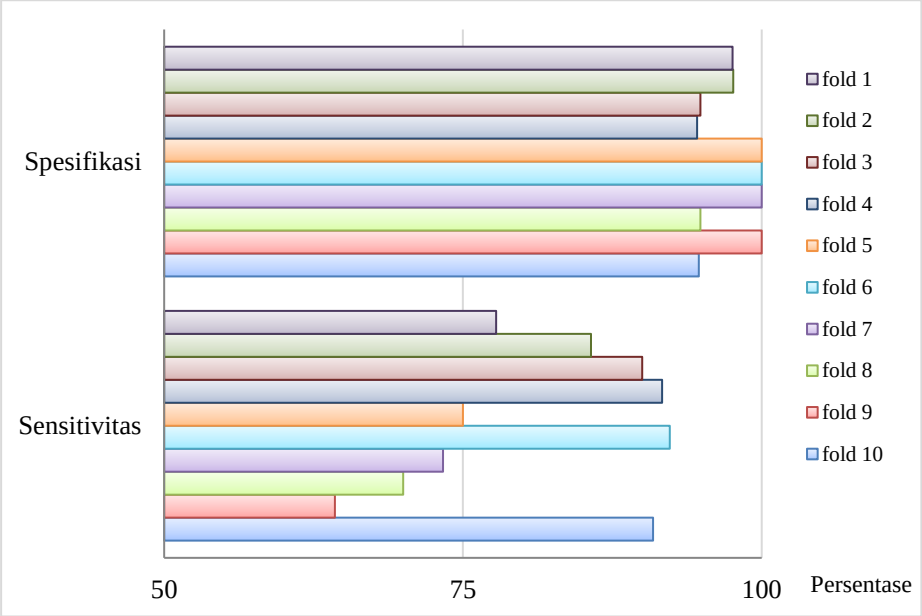
Gambar 3. Arsitektur RBFNN (27-34-2)

fold 6, artinya dengan menggunakan arsitektur RBFNN, dari 49 data *testing* dalam *fold* tersebut, 48 daerah berhasil diklasifikasikan sesuai data aktualnya dan hanya 1 daerah yang diklasifikasikan berbeda dengan data aktualnya. *Overfitting* (akurasi hasil klasifikasi data *training* lebih tinggi daripada data *testing*) arsitektur RBFNN (27-34-2) ini terjadi pada 6 dari 10 percobaan.

Sensitivitas arsitektur RBFNN (27-34-2) dari data *testing fold* 6 menunjukkan persentase paling tinggi yaitu sebesar 92,31



Gambar 4. Ketepatan Hasil Klasifikasi dengan RBFNN (27-34-2)



Gambar 5. Sensitivitas dan Spesifikasi Hasil Klasifikasi RBFNN Menurut Fold

Tabel 4. Rata-rata *Misclassification* RBFNN (27-34-2) Hasil Klasifikasi

Metode RBFNN	Rata-rata Akurasi	Rata-rata *Misclassification* (100-Akurasi)%
Data *Training*	94,23	5,77
Data *Testing*	93,48	6,52

Tabel 5. *F-Measure* Arsitektur RBFNN (27-34-2)

Metode	*True Positive*	*False Positive*	*Precison*	Sensitivitas *(Recall)*	*F-Measure*
RBFNN	91	10	90,10	81,10	85,36

persen yang berarti bahwa arsitektur RBFNN memberikan ketepatan klasifikasi sebesar 92,31 persen dalam mengklasifikasikan daerah tertinggal. Spesifikasi atau ketepatan klasifikasi daerah tertinggal dengan model RBFNN pada *fold* 5, *fold* 6, *fold* 7 dan *fold* 9 mencapai 100 persen, yang berarti keseluruhan daerah tidak tertinggal dalam keempat *fold* ini diklasifikasikan secara tepat sesuai dengan aktualnya. *Testing fold* 10 memberikan spesifikasi terendah, yaitu 94,74 persen dari daerah tidak tertinggal yang diklasifikasikan dengan tepat sesuai status sebenarnya sebagai daerah tidak tertinggal.

Rata-rata akurasi yang dihasilkan berdasarkan 10-*fold cross validation* pada arsitektur RBFNN (27-34-2) mencapai 93,48 persen. Rata-rata sensitivitas yang menggambarkan kemampuan jaringan alam mengklasifikasikan daerah tertinggal sesuai data aktualnya memberikan hasil sebesar 81,10 persen. Untuk spesifikasi, RBFNN mampu menghasilkan ketepatan klasifikasi daerah tidak tertinggal sebesar 97,43 persen.

Secara umum, berdasarkan evaluasi performa klasifikasi dengan menggunakan arsitektur RBFNN (27-34-2) diperoleh hasil performa klasifikasi yang baik. RBFNN memiliki kelemahan pada sensitivitas namun unggul dalam spesifikasi. Rata-rata *misclassification* dari model RBFNN sebesar 6,52 persen dengan hasil *misclassification* sebanyak 32 daerah. *Misclassification* pada RBFNN meliputi 22 daerah tertinggal yang diklasifikasikan sebagai daerah tidak tertinggal dan 10 daerah tidak tertinggal yang diklasifikasikan sebagai daerah tertinggal. Daftar daerah yang mengalami *misclassification* dari total 10 *fold* data *testing* dapat dilihat pada Lampiran.

Misclassification merupakan besarnya kesalahan klasifikasi pada hasil prediksi dibandingkan data aktual. Rata-rata *Misclassification* yang dihasilkan dari data *training* sebesar 5,77 persen, sedangkan pada data *testing* sedikit meningkat menjadi sebesar 6,52 persen.

F-Measure merupakan ukuran tunggal dari uji performa klasifikasi berdasarkan sensitivitas (*recall*) dan *precision*. *Nilai F-Measure* dari RBFNN sebesar 85,36 persen, artinya kinerja algoritma RBFNN dianggap baik (mendekati 100 persen semakin baik). Dalam sudut pandang metode RBFNN sebagai metode yang relatif akurat untuk pengklasifikasian daerah tertinggal di Indonesia, data *misclassification* yang dihasilkan dapat menjadi perhatian untuk dikaji lebih lanjut ataupun menjadi pertimbangan dalam penentuan kebijakan bagi *stakeholder* dan pihak yang terkait. Data hasil klasifikasi dengan metode RBFNN ini dapat menjadi pembanding untuk penentuan status daerah tertinggal pada periode mendatang.

KESIMPULAN DAN SARAN

Pada pengklasifikasian dengan RBFNN, dengan algoritma *k-means cluster* didapatkan jumlah *cluster* sebanyak 34 *cluster*, sehingga arsitektur yang terbentuk adalah RBFNN (27-34-2). Arsitektur ini tersusun oleh 27 *neuron* pada *input layer*,

34 *neuron* pada *hidden*n *layer*, dan 2 *neuron* pada *output layer*. Dari segi ukuran arsitektur model, RBFNN cukup efisien dengan lebih sedikit *neuron* pada *hidden layer* (sesuai jumlah *cluster*). Rata-rata performa klasifikasi berdasarkan 10-*fold cross validation* yang diperoleh dengan RBFNN (27-34-2) yaitu akurasi 93,48 persen, sensitivitas 81,10 persen, spesifikasi 97,43 persen. Berdasarkan evaluasi performa klasifikasi dari arsitektur RBFNN dapat disimpulkan bahwa metode RBFNN memiliki keunggulan pada spesifikasi. Untuk performa klasifikasi secara keseluruhan, RBFNN menghasilkan nilai *F-Measure* sebesar 85,36 persen.

Untuk pengembangan penelitian, dapat dilakukan pengklasifikasian dengan menggunakan penggabungan (*hybrid*) model, misalnya model Radial Basis Probabilistic Neural Network (RBPNN) yang merupakan kombinasi dari arsitektur jaringan RBFNN dengan PNN. RBPNN memadukan keunggulan dari kedua model yang digabungkan sehingga diharapkan dapat meningkatkan performa klasifikasi yang dihasilkan. Metose RBFNN ini juga dapat diterapkan pada pemodelan klasifikasi dengan lebih banyak variabel dan cakupan data.

DAFTAR PUSTAKA

Devega, Mariza. 2013. Diagnosa Kerusakan Bantalan Gelinding Menggunakan Metode RBFNN. Tesis. Semarang: Universitas Diponegoro.

Fausett, L. 1994. *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*. Prentice-Hall, New Jersey, USA.

Ghaderzadeh, M., Fein, R., Standring, A. 2013. *Early Detection of Cancer from Benign Hyperplasia of Prostate. Applied Medical Informatics* Vol. 33, No. 3, pp: 45-54

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining Concepts and Techniques Third Edition*. Waltham: Elsevier Inc.

Haykin, S. 2008. *Neural Networks and Learning Machines*. New Jersey: Prentice Hall.

Kotsiantis, S., Zaharakis, I., Pintelas, P. 2007. *Supervised machine learning: A review of classification techniques. Frontiers in Artificial Intelligence and Applications*.

Kusumadewi, S. 2003. *Artificial Intelligence* (Teknik dan Aplikasinya). Yogyakarta (ID): Graha Ilmu.

Naibaho, E. 2016. Perbandingan Back Propagation Neural Network dan Learning Vector Quatization. Tesis. Bandung: Universitas Padjajaran

Priddy, K.L., Keller, P.E. 2005. *Artificial Neural Network : An Introduction*. Washington: Spie Press

Sitamahalakshmi, T., et al. 2011. *Performance Comparison of Radial Basis Function Networks and Probabilistic Neural Networks for Telugu Character Recognition. Global Journal of Computer Science and Technology Volume 11 Issue 4*

Wahyuningrum, V.2017. *Classification of Underdeveloped Regency using Probabilistic Neural Networks. Proceeding International Conference (ICAS), Vol.2 No 1 (2017)*

Živković, Ž., Mihajlović,I., Nikolić,D. 2008. *Artificial Neural Network Method Applied on The Nonlinear Multivariate Problems.*University of Belgrade, Serbia

LAMPIRAN

Daftar Variabel *Output* (Y) dan Variabel *Input* (X)

Variabel	Kode	Nama Variabel
Output (Y)	Y	Status Daerah Y=1 Daerah Tertinggal Y=0 Daerah Tidak Tertinggal
Input (X)		**Kriteria Perekonomian Masyarakat**
	X_1	Persentase penduduk miskin (BPS,Susenas)
	X_2	Pengeluaran konsumsi per kapita (BPS,Susenas)
		Kriteria Sumber Daya Manusia
	X_3	Angka harapan hidup (BPS,Susenas)
	X_4	Rata-rata lama sekolah (BPS,Susenas)
	X_5	Angka melek huruf (BPS,Susenas)
		Kriteria Infrastruktur
	X_6	Jumlah desa dengan jenis permukaan jalan terluas aspal/beton (BPS,Podes)
	X_7	Jumlah desa dengan jenis permukaan jalan terluas diperkeras (BPS,Podes)
	X_8	Jumlah desa dengan jenis permukaan jalan terluas tanah (BPS,Podes)
	X_9	Jumlah desa dengan jenis permukaan jalan terluas lainnya (BPS,Podes)
	X_{10}	Persentase rumah tangga pengguna listrik (BPS,Podes)
	X_{11}	Persentase rumah tangga pengguna telepon (BPS,Podes)
	X_{12}	Persentase rumah tangga pengguna air bersih (BPS,Podes)
	X_{13}	Jumlah desa yang memiliki pasar tanpa bangunan permanen (BPS,Podes)
	X_{14}	Jumlah prasarana kesehatan per 1000 penduduk (BPS,Susenas)
	X_{15}	Jumlah dokter per 1000 penduduk (BPS,Susenas)
	X_{16}	Jumlah SD dan SMP per 1000 penduduk (BPS,Susenas)
		Kriteria Kemampuan Keuangan Daerah
	X_{17}	Kemampuan keuangan daerah (Kemenkeu)
		Kriteria Aksesibilitas
	X_{18}	Rata-rata jarak dari kantor desa/kelurahan ke kantor kabupaten yang membawahi (BPS,Podes)
	X_{19}	Jumlah desa dengan akses ke pelayanan kesehatan >5 km (BPS,Podes)
	X_{20}	Jarak desa ke pelayanan pendidikan dasar (BPS,Podes)
		Kriteria Karakteristik Daerah
	X_{21}	Persentase desa gempa bumi (BPS,Podes)
	X_{22}	Persentase desa tanah longsor (BPS,Podes)
	X_{23}	Persentase desa banjir (BPS,Podes)
	X_{24}	Persentase desa bencana lainnya (BPS,Podes)
	X_{25}	Persentase desa di kawasan hutan lindung (BPS,Podes)
	X_{26}	Persentase desa berlahan kritis (BPS,Podes)
	X_{27}	Persentase desa konflik satu tahun terakhir (BPS,Podes)

Hasil Eksplorasi Data: Uji Multikolinieritas dan *Output* Regresi Logistik

UJI MULTIKOLINIERITAS
VIF >5 artinya variabel mengandung masalah multikolinieritas

```
> data<-read.csv("dataNN.csv")
>                                                                    reg.log<-
glm(Y_2015~X1+X2+X3+X4+X5+X6+X7+X8+X9+X10+X11+X12+X13+X14+X15+X16+X17+
X18+X19+X20+X21+X22+X23+X24+X25+X26+X27,data=data)
> vif(reg.log)
X1       X2       X3       X4       X5       X6       X7       X8
2.831022 2.253291 1.562333 5.540024 5.655349 3.148160 1.517814 3.439646
 X9       X10      X11      X12      X13      X14      X15      X16
2.274809 4.922362 2.752009 2.334973 1.463256 7.959401 3.319661 6.485056
X17      X18      X19      X20      X21      X22      X23      X24
1.577984 2.019335 5.429571 4.280363 1.194668 1.327243 1.241577 1.219430
X25      X26      X27
1.858747 1.489184 1.243659
```

REGRESI LOGISTIK

```
data<-read.csv("dataNN.csv")
kab <- data[,1]
data[,1] <- NULL
train2<-sample(nrow(data),round(0.9*nrow(data)))
data_train2<-data$Y_2015[train2]
data_test<-data$Y_2015[-train2]
fit<-glm(Y_2015~.,data=data[train2,],family="binomial")
logistics.train2<-fit$fitted.value ##prediction
mat.logis.train2<-table(round(data$Y_2015[train2]),round(logistics.train2,0)) #confusion matrix
acc.logis.train2<-sum(diag(mat.logis.train2))/sum(mat.logis.train2)  #accuracy
logistics.test<-predict(fit,newdata=data[-train2,],type="response") ##prediction
mat.logis.test<-table(round(data$Y_2015[-train2]),round(logistics.test,0)) #confusion matrix
acc.logis.test<-sum(diag(mat.logis.test))/sum(mat.logis.test) #accuracy
result.acc.training2<-c(acc.logis.train2) ##akurasi training
result.acc.testing2<-c(acc.logis.test)
result<-rbind(result.acc.training2,result.acc.testing2)
colnames(result)<-c("Logistics Regression")
rownames(result)<-c("Training", "Testing")
result
```

```
> result
         Logistics Regression
Training            1.0000000
Testing             0.8571429
```

```
> result
         Logistics Regression
Training            0.9841629
Testing             0.8571429
```

Ringkasan Data Penelitian

No. Urut	Nama Kabupaten	Y	X_1	X_2	...	X_{27}
1	SIMEULUE	0	20,57	628,09	...	0,00
2	ACEH SINGKIL	1	18,73	620,40	...	3,33
3	ACEH SELATAN	0	13,44	616,71	...	0,38
4	ACEH TENGGARA	0	14,39	609,76	...	4,16
5	ACEH TIMUR	0	16,59	599,27	...	0,58
. .	. .	. .	. .	. .	. .	. .

.	.	.	.	.	.	.
490	DEIYAI	1	47,52	593,06	...	0,00
491	JAYAPURA	0	16,19	650,99	...	5,13

Sumber: BPS dan Kemenkeu, diolah

Ringkasan Data Hasil Standarisasi Variabel *Input* (X)

No. Urut	Nama Kabupaten	X_1	X_2	...	X_{27}
1	SIMEULUE	0,41119	0,81892	...	0,00000
2	ACEH SINGKIL	0,37099	0,78362	...	0,08772
3	ACEH SELATAN	0,25541	0,76669	...	0,01012
4	ACEH TENGGARA	0,27616	0,73478	...	0,10936
5	ACEH TIMUR	0,32423	0,68663	...	0,01536
. . .	. . .	. . .	. . .	. . .	. . .
490	DEIYAI	1,00000	0,65813	...	0,00000
491	JAYAPURA	0,31549	0,92403	...	0,13495

Sumber: Hasil Pengolahan

Pembagian Data Berdasarkan *10-Fold cross validation*

Sampel / *Fold*	1	2	3	4	5	6	7	8	9	10
Fold 1	50	50	50	50	50	50	50	50	50	50
Fold 2	49	49	49	49	49	49	49	49	49	49
Fold 3	49	49	49	49	49	49	49	49	49	49
Fold 4	49	49	49	49	49	49	49	49	49	49
Fold 5	49	49	49	49	49	49	49	49	49	49
Fold 6	49	49	49	49	49	49	49	49	49	49
Fold 7	49	49	49	49	49	49	49	49	49	49
Fold 8	49	49	49	49	49	49	49	49	49	49
Fold 9	49	49	49	49	49	49	. 49	. 49	. 49	. 49
Fold 10	. 49	. 49	. 49	. 49	. 49	. 49	. 49	. 49	. 49	. 49

Data Sampel	Data *Training*	Data *Testing*
1	441	50
2	442	49
3	442	49
4	442	49
5	442	49
6	442	49
7	442	49
8	442	49
9	442	49
10	442	49

116.

4.Data *Training*	5.Data *Testing*

Rata-rata Akurasi *Training* dan *Testing* RBFNN Menurut Jumlah *Cluster* (*k-means cluster*)

Cluster (k)	Akurasi *Training*	Akurasi *Testing*
2	0,8769	0,8716
3	0,8798	0,8778
4	0,8798	0,8798
5	0,9052	0,8859
6	0,9122	0,9001
7	0,9054	0,8940
8	0,9188	0,9001
9	0,9224	0,9022
10	0,9269	0,9144

Cluster (k)	Akurasi *Training*	Akurasi *Testing*
35	0,8857	0,8718
36	0,9118	0,9022
37	0,8753	0,8781
38	0,9228	0,9206
39	0,8502	0,8512
40	0,8891	0,8756
41	0,8905	0,8798
42	0,8473	0,8531
43	0,9265	0,9246

Cluster (k)	Akurasi *Training*	Akurasi *Testing*
68	0,8753	0,8716
69	0,8731	0,8715
70	0,8106	0,8126
71	0,7812	0,7800
72	0,8812	0,8759
73	0,8656	0,8756
74	0,8855	0,8780
75	0,8667	0,8596
76	0,8570	0,8491

11	0,9305	0,9103	44	0,9076	0,9024	77	0,7991	0,7879
12	0,9314	0,9226	45	0,8545	0,8614	78	0,8167	0,8514
13	0,9319	0,9185	46	0,8989	0,9001	79	0,8078	0,8070
14	0,9353	0,9266	47	0,8430	0,8369	80	0,7576	0,7804
15	0,9231	0,9083	48	0,7930	0,7718	81	0,7743	0,7727
16	0,9131	0,8859	49	0,8724	0,8736	82	0,7703	0,7557
17	0,9242	0,8981	50	0,9421	0,9328	83	0,8821	0,8716
18	0,9389	0,9287	51	0,8701	0,8737	84	0,8368	0,8351
19	0,9104	0,9025	53	0,8436	0,8514	85	0,8131	0,8065
20	0,9246	0,9187	53	0,8749	0,8532	86	0,7713	0,7593
21	0,9226	0,9165	54	0,8651	0,8350	87	0,9310	0,9185
22	0,8327	0,8435	55	0,7826	0,7859	88	0,8201	0,8123
23	0,9233	0,9164	56	0,9219	0,9123	89	0,8719	0,8675
24	0,8774	0,8654	57	0,9063	0,8981	90	0,8312	0,8493
25	0,9437	0,9227	58	0,8291	0,8228	91	0,8635	0,8678
26	0,8912	0,8715	59	0,8484	0,8226	92	0,8448	0,8348
27	0,9124	0,8858	60	0,8803	0,8718	93	0,8352	0,7968
28	0,9201	0,9028	61	0,8156	0,8065	94	0,9088	0,8981
29	0,9269	0,9246	62	0,8321	0,8456	95	0,9265	0,9083
30	0,8744	0,8636	63	0,8599	0,8189	96	0,8418	0,8369
31	0,8375	0,8348	64	0,8762	0,8756	97	0,8525	0,8409
32	0,8853	0,8858	65	0,8359	0,8228	98	0,8565	0,8657
33	0,8405	0,8369	66	0,8197	0,8042	99	0,8726	0,8675
34	**0,9423**	**0,9348**	67	0,8394	0,8654	100	0,8715	0,8676

Sumber: Hasil Pengolahan

Confision matrix Hasil Prediksi RBFNN (27-34-2) Menurut *10-fold cross validation*

Aktual		Prediksi			
		RBFNN (27-34-2)			
		Training		*Testing*	
		1	0	1	0
Fold 1	1	85	19	7	2
	0	9	328	1	40
Fold 2	1	90	16	6	1
	0	8	328	1	41
Fold 3	1	83	20	9	1
	0	7	332	2	37
Fold 4	1	85	16	11	1
	0	6	335	2	35
Fold 5	1	79	22	9	3
	0	7	334	0	37
Fold 6	1	79	21	12	1
	0	9	333	0	36
Fold 7	1	82	16	11	4
	0	8	336	0	34
Fold 8	1	87	16	7	3
	0	6	333	2	37
Fold 9	1	82	17	9	5
	0	8	335	0	35
Fold 10	1	84	18	10	1
	0	6	334	2	36
Total	1			91	22
	0			10	368

Sumber: Hasil Pengolahan

Performa Klasifikasi Data *Training* dan *Testing* dengan Arsitektur RBFNN (27-34-2)

Percobaan	*Training*					*Testing*			
	M	AK	SV	SP	Lama	M	AK	SV	SP
1	6,35	93,65	81,73	97,33	4,989	6,00	94,00	77,78	97,56
2	5,43	94,57	84,91	97,62	4,924	4,08	95,92	85,71	97,62
3	6,11	93,89	80,58	97,94	4,973	6,12	93,88	90,00	94,87
4	4,98	95,02	84,16	98,24	4,950	6,12	93,88	91,67	94,59
5	6,56	93,44	78,22	97,95	5,320	6,12	93,88	75,00	100,00
6	6,79	93,21	79,00	97,37	5,414	2,04	97,96	92,31	100,00
7	5,43	94,57	83,67	97,67	23,150	8,16	91,84	73,33	100,00
8	4,98	95,02	84,47	98,23	5,008	10,20	89,80	70,00	94,87
9	5,66	94,34	82,83	97,67	4,930	10,20	89,80	64,29	100,00
10	5,43	94,57	82,35	98,24	5,008	6,12	93,88	90,91	94,74
Rata-rata	5,77	94,23	82,19	97,82	6,867	6,52	93,48	81,10	97,43

Sumber: Hasil Pengolahan

Keterangan: M (%*Misclassification*); AK (%Akurasi); SV (%Sensitivitas); SP (%Spesifikasi); Lama (Waktu Komputasi dalam detik)

Daftar Daerah dengan Prediksi *Misclassification* per *Testing Fold* Menggunakan RBFNN (27-34-2)

No.	*Testing Fold*	Nama Daerah	Aktual	Prediksi
1	1	KEPULAUAN MERANTI	Tidak Tertinggal	Tertinggal
2	1	MUSI RAWAS	Tertinggal	Tidak Tertinggal
3	1	BOALEMO	Tertinggal	Tidak Tertinggal
4	2	SIMEULUE	Tidak Tertinggal	Tertinggal
5	2	LEBAK	Tertinggal	Tidak Tertinggal
6	3	HALMAHERA UTARA	Tidak Tertinggal	Tertinggal
7	3	BIAK NUMFOR	Tertinggal	Tidak Tertinggal
8	3	KATINGAN	Tidak Tertinggal	Tertinggal
9	4	MANOKWARI	Tidak Tertinggal	Tertinggal
10	4	KAPUAS HULU	Tertinggal	Tidak Tertinggal
11	4	MALUKU TENGGARA	Tidak Tertinggal	Tertinggal
12	5	TOLI-TOLI	Tertinggal	Tidak Tertinggal
13	5	BONDOWOSO	Tertinggal	Tidak Tertinggal
14	5	KEEROM	Tertinggal	Tidak Tertinggal
15	6	NIAS	Tertinggal	Tidak Tertinggal
16	7	SITUBONDO	Tertinggal	Tidak Tertinggal
17	7	HULU SUNGAI UTARA	Tertinggal	Tidak Tertinggal
18	7	BENGKAYANG	Tertinggal	Tidak Tertinggal
19	7	POHUWATO	Tertinggal	Tidak Tertinggal
20	8	NIAS BARAT	Tertinggal	Tidak Tertinggal
21	8	POSO	Tidak Tertinggal	Tertinggal
22	8	NUNUKAN	Tertinggal	Tidak Tertinggal
23	8	SELUMA	Tertinggal	Tidak Tertinggal
24	8	HALMAHERA TENGAH	Tidak Tertinggal	Tertinggal
25	9	PANDEGLANG	Tertinggal	Tidak Tertinggal
26	9	KONAWE	Tertinggal	Tidak Tertinggal
27	9	LAMPUNG BARAT	Tertinggal	Tidak Tertinggal
28	9	BOMBANA	Tertinggal	Tidak Tertinggal
29	9	PASAMAN BARAT	Tertinggal	Tidak Tertinggal

30	10	KOLAKA UTARA	Tidak Tertinggal	Tertinggal
31	10	PROBOLINGGO	Tidak Tertinggal	Tertinggal
32	10	NABIRE	Tertinggal	Tidak Tertinggal

Sumber: Hasil Pengolahan

ANALISIS PERMINTAAN PANGAN DAN NONPANGAN RUMAH TANGGA DENGAN GANGGUAN KESEHATAN DI INDONESIA

Muhammad Syafiudin[1], Turro S. Wongkaren[2]

[1]Badan Pusat Statistik, [2]Universitas Indonesia
e-mail:, [1] apikudin@bps.go.id

Abstrak

Penelitian ini bertujuan menganalisis dampak tidak langsung gangguan kesehatan terhadap permintaan pangan dan non pangan rumah tangga. Dengan menggunakan data Susenas Panel tahun 2012 dan 2013 dan menerapkan *two step heckman selection model* untuk estimasi pendapatan dan *seemingly unrelated regression estimator* untuk estimasi konsumsi rumah tangga. Hasilnya menunjukkan bahwa gangguan kesehatan kepala rumah tangga akan menurunkan pendapatannya. Dampak ini akan lebih dirasakan oleh rumah tangga perempuan miskin dan bekerja di sektor pertanian. Penurunan pendapatan ini menyebabkan porsi pengeluaran konsumsi non pangan menurun, khususnya untuk pengeluaran pemeliharaan perumahan, namun pengeluaran untuk perawatan tubuh justru meningkat. Sedangkan untuk porsi konsumsi pangan tidak terpengaruh. Hal ini menunjukkan bahwa, gangguan kesehatan dapat menyebabkan penurunan tingkat kesejahteraan rumah tangga karena menyebabkan penurunan pendapatan dan peningkatan pengeluaran kesehatan. Oleh karena itu, diperlukan kebijakan yang dapat melindungi kesejahteraan rumah tangga ketika mengalami gangguan kesehatan, bisa berupa subsidi biaya kesehatan atau *cash transfer*.

Kata kunci*:* Gangguan Kesehatan, Konsumsi, Pangan, Non Pangan, Pendapatan Rumah Tangga

Abstract

This study aims to analyze the indirect effects of health problems on household food and non-food demand. This research uses Susenas Panel data for 2012 and 2013 and applies the ' two step heckman selection model' to estimate income, and 'seemingly unrelated regression estimator' to estimate household consumption. The results show that health issues or problems of the household will decrease their income. This impact will be worse experienced by the poor female household and work in the agricultural sector. The decrease in income has led to a decrease in non-food expenditure, especially on housing maintenance; but on the contrary, expenditure for body care has increased. However, it has not affected the expenditure on food consumption. This finding shows that health problems would be lowering household welfare as decreasing income and increasing health expenditure. Therefore, it is necessary to formulate policies to protect household welfare directly affected by health problems, for example by providing health subsidy or cash transfer.

Keywords: *illhealth, illness, consumption, food and Non-food, household income*

PENDAHULUAN

Sebagai salah satu agen ekonomi, tingkat kesejahteraan rumah tangga sangat rentan terhadap fluktuasi tingkat pendapatan. Salah satu penyebabnya adalah gangguan kesehatan (*ill health*) yang dialami oleh anggota rumah tangga (Gertler & Gruber, 2002). Menurut Genoni (2012), terdapat biaya ekonomi yang harus ditanggung individu atau rumah tangga ketika mengalami gangguan kesehatan yaitu pertama, gangguan kesehatan membatasi kemampuan individu untuk bekerja dan kedua dapat menimbulkan tambahan pengeluaran untuk perawatan kesehatan.

Di negara berkembang, dimana pada umumnya cakupan program jaminan kesehatan masih relatif rendah, dampak gangguan kesehatan terhadap kesejahteraan bisa lebih besar (Russell, 2004; Wagstaff, 2007). Hal ini terjadi karena rumah tangga harus menanggung sebagian besar biaya pelayanan kesehatan yang harus dikeluarkan (*Out Of Pocket*). Dimana jumlahnya bisa sangat besar dan membebani perekonomian rumah tangga. Sehingga hal ini bisa menjadi salah satu pemicu terjadinya kemiskinan (Hoogeveen, Tesliuc, Vakis, & Dercon, 2004).

Sebagai salah satu negara berkembang, hal ini juga terjadi di Indonesia. Hak setiap warga negara untuk mendapatkan pelayanan kesehatan yang baik merupakan tanggung jawab negara. Sebagaimana telah di atur di dalam Undang-undang Dasar 1945 amandemen keempat pasal 28H (1). Namun pada implementasinya, pelaksanaan amanat undang-undang tersebut masih belum berjalan dengan baik. Meskipun sejak awal tahun 2014 pemerintah telah menjalankan Program Jaminan Kesehatan Nasional (JKN), yang merupakan salah satu implementasi dari Undang-Undang Dasar 1945 tersebut. Namun dalam pelaksanaannya masih banyak ditemui kekurangan baik pada kualitas pelayanan maupun kuantitasnya yang belum menjangkau seluruh penduduk di Indonesia.

Berdasarkan data Kementerian Kesehatan Republik Indonesia hingga akhir tahun 2016 cakupan kepesertaan Jaminan Kesehatan Nasional (JKN) baru mencapai 66,46 persen. Dan berdasarkan hasil riset kesehatan dasar tahun 2013 (Riskesdas 2013), dari seluruh rumah tangga pada kuintil terbawah berdasarkan indeks kepemilikan , baru sekitar 37,1 persen yang mendapatkan pelayanan kesehatan gratis. Dari sisi pembiayaannya, 53,5 persen dari seluruh rawat inap dan 67,9 persen dari seluruh rawat jalan masih menggunakan biaya sendiri (Riskesdas 2013). Itu artinya sumber biaya yang digunakan untuk semua fasilitas kesehatan di Indonesia masih didominasi oleh biaya sendiri (*out of pocket*).

Padahal berdasarkan laporan *The Global Medical Trend Survey Report* yang dirilis oleh *Tower Watson & Co*, biaya kesehatan di Indonesia terus meningkat dan peningkatannya selalu berada diatas tingkat inflasi umum. Pada tahun 2012 biaya kesehatan di Indonesia meningkat sebanyak 11,5%, kemudian melonjak menjadi 12,5% pada tahun 2013. Kenaikan ini melebihi tingkat inflasi secara umum yang hanya sebesar 4,3% pada tahun 2012 dan 8,4% pada tahun 2013. Sehingga tidak membuat heran jika masalah gangguan kesehatan dapat menjadi sebuah ancaman serius bagi kesejahteraan rumah tangga di Indonesia.

Berdasarkan W*orld Statistic Health* 2018 yang diterbitkan oleh WHO, pada tahun 2015 terdapat sebanyak 3,3 persen dari total penduduk di Indonesia memiliki pengeluaran untuk kesehatan lebih dari 10 persen dari total pendapatan yang dimiliki, dengan rata-rata pengeluaran kesehatan perkapita sebesar US$ 112. Sebuah nilai yang cukup besar khususnya bagi rumah tangga ekonomi menengah kebawah. Dimana sebagian besar pendapatan mereka dialokasikan untuk kebutuhan konsumsi sehari-hari, khususnya konsumsi pangan.

Padahal disisi lain, gangguan kesehatan yang dialami akan menyebabkan penurunan pendapatan. Gertler & Gruber (2002) dalam penelitiannya di Indonesia menemukan bahwa, gangguan kesehatan yang dialami oleh kepala rumah tangga

akan menyebabkan penurunan penawaran kerja sehingga berdampak terhadap penurunan pendapatan rumah tangga. Kedua hal ini tentu akan menjadi beban ekonomi bagi rumah tangga, terjadi peningkatan pengeluaran untuk kesehatan dan disaat yang sama terjadi penurunan pendapatan, tentu hal ini akan mempengaruhi tingkat konsumsi rumah tangga.

Namun demikian, dari beberapa hasil penelitian terdahulu, pengaruh gangguan kesehatan terhadap tingkat konsumsi rumah tangga menunjukkan hasil yang beragam bahkan saling bertolak belakang. Misalnya penelitian yang dilakukan oleh Cochrane (1991) dan Townsend (1994) menunjukkan bahwa, tingkat konsumsi rumah tangga tidak terpengaruh oleh gangguan kesehatan yang dialami oleh rumah tangga. Sedangkan Gertler & Gruber (2002) dan Sparrow et al. (2014) justru menunjukkan bahwa gangguan kesehatan yang dialami rumah tangga akan mempengaruhi tingkat konsumsi rumah tangga karena adanya peningkatan pengeluaran untuk kesehatan. Sementara itu Asfaw & Braun (2004), justru menunjukkan bahwa gangguan kesehatan tidak sepenuhnya mempengaruhi konsumsi rumah tangga. Menurutnya secara total, tingkat konsumsi rumah tangga tidak terpengaruh oleh gangguan kesehatan. Gangguan kesehatan hanya mempengaruhi konsumsi rumah tangga secara parsial yaitu hanya pada konsumsi pangan yang dibeli oleh rumah tangga.

Genoni (2012) berpendapat bahwa perbedaan hasil penelitian ini terjadi karena rumitnya mengidentifikasi dan mengukur dampak dari gangguan kesehatan. Menurut Strauss & Thomas (1998) kesehatan adalah sesuatu yang multidimensi, sehingga sebuah indikator kesehatan tidak mungkin dapat menangkap semua dampak yang ditimbulkan, karena bisa saja setiap indikator kesehatan memiliki dampak yang berbeda terhadap individu atau rumah tangga. Oleh karena itu, untuk mengukur dampak gangguan kesehatan terhadap tingkat konsumsi rumah tangga harus dimulai dengan mendefinisikan gangguan kesehatan tersebut. Oleh karena itu, dengan mengadaptasi penelitian Cochrane, (1991) & Sparrow et al., (2014) dalam mendefinisikan gangguan kesehatan, maka penelitian ini mendefinisikan gangguan kesehatan dengan menggunakan lama hari aktifitas sehari-hari terganggu karena keluhan kesehatan yang dialami . Definisi gangguan kesehatan ini bersifat umum dan tidak terbatas pada suatu gejala gangguan kesehatan tertentu. Namun tetap dapat menggambarkan biaya ekonomi yang akan ditanggung akibat gangguan kesehatan yang dialami sebagaimana yang dimaksud oleh Genoni (2012), yaitu berupa penurunan pendapatan akibat penurunan penawaran kerja karena gangguan kesehatan.

Selain itu, untuk mengukur dampak dari gangguan kesehatan terhadap tingkat konsumsi rumah tangga, maka bagaimana mekanisme gangguan kesehatan mempengaruhi konsumsi rumah tangga juga penting untuk diketahui. Sehingga analisa dampak gangguan kesehatan terhadap tingkat konsumsi yang dihasilkan akan lebih baik karena dapat melihat dampak dari setiap proses yang terjadi dari mekanisme tersebut. Menurut Nguyet & Mangyo (2010), penyebab utama terjadinya penurunan tingkat konsumsi adalah akibat penurunan pendapatan. Sehingga dalam hal ini penurunan pendapatan merupakan *intermediate relationship* antara gangguan kesehatan dan tingkat konsumsi.

Oleh karena itu, untuk mengukur dampak gangguan kesehatan terhadap tingkat konsumsi, maka penelitian ini memulainya dengan mengukur dampak gangguan kesehatan terhadap tingkat pendapatan, kemudian melihat bagaimana dampak perubahan tingkat pendapatan tersebut terhadap tingkat konsumsi rumah tangga. Sehingga dalam hal ini gangguan kesehatan memiliki dampak tidak langsung terhadap tingkat konsumsi rumah tangga. Dimana pada penelitian-penelitian sebelumnya, gangguan kesehatan dianggap berpengaruh langsung terhadap perubahan tingkat konsumsi rumah tangga, sehingga bagaimana mekanisme gangguan kesehatan mempengaruhi gangguan kesehatan belum dapat dijelaskan dengan baik (lihat Asfaw

& Braun (2004), Wang, Zhang, & Hsiao (2006); Wagstaff (2007); Nguyet & Mangyo (2010) dan Sparrow et al. (2014)).

Selain itu, menurut data BPS, dalam sepuluh tahun terakhir persentase penduduk miskin di Indonesia telah mengalami penurunan cukup signifikan yaitu dari 15,42 persen pada tahun 2008 menjadi sebesar 9,66 persen tahun 2018. Meskipun secara total mengalami penurunan yang cukup besar, namun jika dilihat berdasarkan lapangan usaha yang dilakoni, maka tingkat kemiskinan di Indonesia masih didominasi oleh rumah tangga pertanian. Sedangkan berdasarkan jenis kelamin kepala rumah tangga, penurunan tingkat kemiskinan pada rumah tangga perempuan (RTP) lebih lambat daripada rumah tangga laki-laki (RTL). Dari kedua fakta miris ini, menarik untuk dikaji bagaimana pola konsumsi rumah tangga, ketika mereka sudah dalam kondisi sulit secara ekonomi, namun harus mengalami fluktuasi pendapatan, dimana salah satunya karena gangguan kesehatan (Gertler & Gruber, 2002). Sehingga dengan demikian, penelitian ini akan menguji secara empiris hipotesis hubungan antara gangguan kesehatan dan tingkat konsumsi dengan memperhatikan hubungan tidak langsung diantara keduanya dan kemudian mengkaitkannya dengan isu kemiskinan yang terjadi di Indonesia ini.

TINJAUAN REFERENSI

1. Teori Penawaran Kerja dan Pendapatan Rumah Tangga

Teori penawaran kerja individu di dalam rumah tangga merupakan pengembangan dari teori permintaan neoklasik yaitu dengan menambahkan waktu bersantai (*leisure*) sebagai salah satu barang yang dikonsumsi oleh rumah tangga. Dengan asumsi bahwa *leisure* dan barang yang dikonsumsi baik makanan (X_1) maupun non-makanan dan jasa (X_2) adalah barang normal, dan rumah tangga akan mendapatkan kepuasan (*utility*) ketika mengkonsumsi *leisure* (L) kedua jenis barang tersebut. Maka Fungsi utilitas dari mengkonsumsi kedua barang tersebut dapat dituliskan sebagai berikut:

$$U = U(L, X_1, X_2; \tau) \quad (1)$$

Semakin banyak *leisure* atau barang yang dikonsumsi, maka tingkat kepuasan yang diperoleh semakin tinggi. Namun tingkat konsumsi seseorang terhadap barang/jasa dan *leisure* dibatasi oleh waktu dan tingkat pendapatan yang dimiliki. Tingkat pendapatan seseorang terdiri dari pendapatan upah (*labor income*) dan pendapatan non upah (*Non Labor Income*). Dan semua pendapatan yang dimiliki rumah tangga ini akan digunakan oleh rumah tangga untuk membeli X_1 dan X_2 pada tingkat harga P_1 dan P_2. Sehingga fungsi pendapatan individu menjadi sebagai berikut

$$F = W \cdot h + V = P_1 \cdot X_1 + P_2 \cdot X_2 \quad (2)$$

Jika tingkat upah adalah tetap dan waktu setiap orang habis untuk bekerja (h) dan bersantai (L), maka total waktu yang dimiliki menjadi ($T = L + h$). Maka persamaan (2) dapat ditulis menjadi

$$F = W \cdot T + V = W \cdot L + P_1 \cdot X_1 + P_2 \cdot X_2 \quad (3)$$

maka persamaan (3) adalah persamaan *Beckerian full income constraint*.

Dengan menggunakan fungsi *lagrange* maka kita dapat memaksimumkan fungsi kepuasan pada persamaan (1) dengan kendala anggaran pada persamaan (3), sehingga menjadi sebagai berikut

$$max\ \mathcal{L} = U(L, X_1, X_2) + \lambda(W \cdot T + V - W \cdot L - P_1 \cdot X_1 - P_2 \cdot X_2) \quad (4)$$

Dengan menggunakan *first order condition* dari persamaan (4) akan diperoleh *reduce form* dari persamaan penawaran jam kerja (S_h) dan persamaan permintaan barang dan jasa rumah tangga (D_{X_i}) berikut.

$$h^* = T - L^* = S_h(W, P_1, P_2, V, \tau) = S_h(W, P_1, P_2, F, \tau) \quad (5)$$

$$X_i^* = D_{X_i}(W, P_i, V, \tau) = D_{X_i}(W, P_i, F, \tau), \quad i=1,2 \quad (6)$$

Persamaan (5) dan (6) merupakan fungsi dari tingkat upah (W), tingkat harga (P_1 dan P_2) dan tingkat pendapatan (V atau F). Dengan asumsi bahwa *leisure* dan barang atau jasa yang dikonsumsi tersebut adalah barang normal, maka tingkat

konsumsi dan tingkat pendapatan berhubungan positif ($\partial X_i/\partial F > 0$), dimana tingkat pendapatan tersebut ditentukan oleh jumlah jam kerja yang ditawarkan, semakin tinggi jam kerja yang ditawarkan, maka semakin tinggi tingkat pendapatan yang bisa dihasilkan ($\partial F/\partial h) > 0$). Dalam konteks penelitian ini, ketika individu mengalami gangguan kesehatan akan menyebabkan penurunan jam kerja sehingga akan berdampak negatif terhadap tingkat pendapatan. Penurunan tingkat pendapatan ini akan berpengaruh negatif terhadap tingkat konsumsi rumah tangga. Sehingga secara tidak langsung, gangguan kesehatan akan berdampak negatif terhadap tingkat konsumsi rumah tangga melalui penurunan pendapatan.

2. Teori *Added Worker Effect*

Sebagai salah satu agen ekonomi, rumah tangga memiliki tujuan untuk memaksimumkan kepuasan seluruh anggota rumah tangga dengan asumsi bahwa kepuasan tersebut diperoleh melalui konsumsi dan waktu bersantai (*leisure*) dari setiap anggota rumah tangga dan tingkat pendapatan dan tingkat pengeluaran rumah tangga merupakan akumulasi dari semua anggota rumah tangga, maka rumah tangga akan berusaha memaksimumkan kepuasannya dengan kendala pendapatan dari seluruh anggota rumah tangga (*Family Utility Budget Constrait Model*). Ashenfelter & Heckman (1974) memodelkan tujuan dan kendala yang dimiliki rumah tangga ini menjadi sebagai berikut:

$$Max\ U = U(L_m, L_f, X) \qquad (7)$$

Jika L_m dan L_f adalah waktu yang dihabiskan diluar pasar tenaga kerja (*Non Work Time*) yang dilakukan oleh anggota rumah tangga laki-laki dan perempuan. Sedangkan X adalah komposit dari seluruh barang yang dikonsumsi oleh rumah tangga dengan asumsi bahwa tidak ada perubahan harga pada seluruh barang tersebut. Maka untuk memaksimumkan kepuasannya maka rumah tangga terkendala pada anggaran pendapatan yang dimiliki

$$W_mT + W_fT + V = W_mL_m + W_fL_f + PX \qquad (8)$$

W_m dan W_f adalah tingkat upah yang diterima oleh anggota rumah tangga, V adalah pendapatan non-upah. Sedangkan P adalah tingkat harga dari barang-barang yang dikonsumsi, T adalah total waktu yang dimiliki oleh setiap anggota rumah tangga yang diasumsikan bahwa alokasi waktu hanya digunakan untuk bekerja dan bersantai (*leisure*). Sehingga waktu untuk bekerja untuk setiap anggota rumah tangga adalah $h_i = T - L_i, i = m, f$. Maka sisi kiri persamaan (8) adalah total pendapatan rumah tangga yang merupakan akumulasi dari seluruh pendapatan anggota rumah tangga yang bekerja, sedangkan sisi kanannya adalah total pengeluaran rumah tangga. Sehingga persamaan (8) diatas akan menjadi:

$$W_m(T - L_m) + W_f(T - L_f) + V = PX \qquad (9)$$

Persamaan (9) menunjukkan bahwa total pendapatan sama dengan total pengeluaran. Maka dengan menggunakan *lagrange* multiplier dan *first order condition* untuk memaksimumkan fungsi utilitas (7) dengan kendala fungsi anggaran pada persamaan (8), akan diperoleh bentuk *reduce form* dari persamaan penawaran kerja dari setiap anggota rumah tangga (h_i) yang merupakan fungsi dari tingkat harga, tingkat upah dan pendapatan non upah yang diterima oleh rumah tangga.

$$h_i = R_i(W_m, W_f, P, V) \qquad (10)$$

Dari persamaan ini bisa dilihat bahwa fungsi penawaran kerja salah satu anggota rumah tangga akan dipengaruhi oleh tingkat upah anggota rumah tangga yang lain.

Menururt Ashenfelter & Heckman (1974) perubahan penawaranan kerja salah satu anggota rumah tangga akibat perubahan tingkat upah anggota rumah tangga yang lain terjadi karena adanya efek substitusi dan efek pendapatan. Efek substitusi terdiri atas dua macam yaitu pertama perubahan penawaran kerja karena perubahan tingkat upah pada diri sendiri (*own substitution effect*) yang bernilai

positif (S_{ii}) dan efek susbstitusi yang kedua terjadi karena terjadinya perubahan tingkat upah anggota rumah tangga yang lain (*cross Substitution Effect*) yang bisa bernilai positif atau negatif (S_{ij}), tergantung pada hubungan jam kerja antar anggota rumah tangga, apakah bersifat saling menggantikan (substitusi) atau melengkapi (komplementer).

$$S_{ii} > 0, \ (i = m, f) \quad (11)$$

$$\frac{dh_i}{dW_j} = S_{ij} + R_j \frac{dR_i}{dV} \quad (12)$$

Sedangkan perubahan penawaran kerja karena efek pendapatan $\left(R_j \frac{dR_i}{dV}\right)$ selalu bernilai negatif. Hal ini karena diasumsikan bahwa peningkatan pendapatan non-upah (*Non Labor Income*) dirumah tangga akan menurunkan penawaran kerja anggota rumah tangga. Oleh karena itu, jika tidak ada *cross Substitution Effect*, maka perubahan penawaran kerja anggota rumah tangga ke-*i* karena perubahan tingkat upah anggota rumah tangg ke-*j* hanya akan dipengaruhi oleh efek pendapatan yang bernilai negatif. Sehingga dapat disimpulkan bahwa terdapat hubungan negatif antara peningkatan penawaran kerja salah satu anggota rumah tangga dengan perubahan tingkat upah anggota rumah tangga yang lain.

Karena itu dalam konteks adanya gangguan kesehatan yang menimpa salah satu anggota rumah tangga yang menyebabkan penurunan penawaran kerjanya sehingga menyebabkan penurunan tingkat pendapatan, maka hal ini akan direspon oleh anggota rumah tangga yang lain dengan cara meningkatkan penawaran kerjanya sebagai kompensasi penurunan penawaran kerja anggota rumah tangga yang mengalami gangguan kesehatan tersebut, sehingga secara total pendapatan rumah tangga tetap bisa dipertahankan. Konsep ini dikembangkan oleh Mincer (1962) di dalam Maloney (1987) dan dikenal dengan *Added Worker Effect (*AWE).

3. Hukum Engel dan Pangsa Pengeluaran Pangan

Dalam ilmu ekonomi, terdapat hukum yang menjelaskan hubungan antara tingkat pendapatan dan tingkat konsumsi rumah tangga, yaitu hukum Engel yang dipopulerkan oleh seorang ahli statistik dari Jerman (1821-1896). Hukum Engel menyatakan bahwa ketika pendapatan meningkat maka porsi pengeluaran untuk konsumsi pangan akan semakin menurun sedangkan porsi untuk konsumsi non pangan akan meningkat. Hukum Engel tidak menyatakan bahwa ketika pendapatan meningkat maka akan terjadi penurunan tingkat konsumsi pangan. Namun yang terjadi adalah peningkatan pengeluaran untuk konsumsi pangan lebih lambat dibandingkan dengan peningkatan pendapatan. Hal ini menyebabkan elastisitas pendapatan terhadap makanan adalah inelastis karena nilai elastisitasnya berada diantara 0 dan 1. Sehingga dalam hal ini terdapat pergeseran porsi pengeluaran antara konsumsi pangan dan non pangan ketika terjadi perubahan pendapatan.

Salah satu penerapan hukum Engel adalah untuk melihat standar hidup suatu negara. Negara yang lebih miskin, akan memiliki koefisien Engel yang lebih besar. Artinya secara umum pengeluaran didalam negara tersebut masih didominasi oleh konsumsi pangan. Dan ini biasa terjadi pada negara-negara miskin. Namun sebaliknya, jika koefisien Engel lebih kecil, maka negara tersebut cenderung memiliki standar hidup yang lebih tinggi. Dalam penelitian ini, negara diidentikkan dengan rumah tangga. Hukum Engel merupakan penemuan empiris yang sangat konsisten. Menurut (Muellbauer, 1980), porsi (*share*) pengeluaran konsumsi untuk pangan terhadap total pengeluaran dapat dijadikan indikator tidak langsung terhadap tingkat kesejahteraan.

METODE ANALISIS

1. Sumber Data

Penelitian ini menggunakan data yang bersumber dari Survei Sosial Ekonomi Nasional (SUSENAS) triwulan pertama tahun 2012 sampai dengan 2013 yang dilakukan oleh Badan Pusat Statistik. Survei ini dilakukan pada sampel rumah

tangga terpilih di seluruh wilayah Indonesia hingga level kabupaten/kota. Pada periode survei ini, sampel rumah tangga dipilih secara acak dan didata secara *longitudinal*, artinya rumah tangga yang sama akan di data kembali setiap tahun selama periode survei. Observasi penelitian ini yaitu seluruh rumah tangga sampel terpilih yang dapat ditemui pada saat pendataan pada dua periode survei, dengan jumlah observasi sebanyak 6.893 rumah tangga. Data tahun 2012 digunakan untuk mendapatkan informasi tentang kejadian gangguan kesehatan yang dialami oleh kepala rumah tangga pada periode sebelumnya yang diduga mempengaruhi tingkat pendapatan tahun 2013.

Untuk mengatasi adanya data pendapatan yang kosong, maka kekosongan data tersebut diimputasi dengan mengikuti cara yang dilakukan oleh (Gertler & Gruber, 2002) yaitu dengan menggunakan rata-rata pendapatan kepala rumah tangga pada setiap propinsi berdasarkan jenis kelamin, kelompok umur (<25, 25-49, 50+) dan tingkat pendidikan (lama sekolah=0 (tidak berpendidikan), 1 thn <= lama sekolah <= 5 tahun, lama sekolah = 6 tahun dan lama sekolah >= 7 thn). Bagi kepala rumah tangga yang tidak memiliki data pendapatan maka akan diimputasi dengan data rata-rata pendapatan yang sesuai karakteristik tersebut pada setiap propinsi.

2. Model Empiris

Model empiris yang digunakan dalam penelitian ini terdiri dari persamaan pendapatan kepala rumah tangga dan persamaan konsumsi rumah tangga. Metode analisis yang digunakan dalam penelitian ini dilakukan dengan dua tahap yaitu tahap pertama estimasi pendapatan kepala rumah tangga dan tahap kedua estimasi konsumsi rumah tangga.

3. Persamaan Pendapatan Kepala Rumah Tangga

Unit analisis pada persamaan ini adalah kepala rumah tangga yang mengalami gangguan kesehatan pada tahun 2012 dan 2013. Sehingga dalam hal ini, pemilihan sampel kepala rumah tangga tersebut menjadi tidak acak sehingga dapat menyebabkan hasil estimasi menjadi bias (*sample selection bias*). Untuk mengatasi ini, maka digunakan metode estimasi *two step heckman selection model*, yaitu dengan menambahkan variabel bebas IMR (*Invers Mills Ratio*) pada model utama. Tahap pertama metode estimasi ini yaitu dengan menghitung peluang kepala rumah tangga mengalami gangguan kesehatan dengan menggunakan model probit, dimana variabel terikat yang digunakan yaitu variabel *dummy* apakah kepala rumah tangga mengalami gangguan kesehatan atau tidak selama dua periode survei. Sedangkan variabel bebas yang digunakan berupa karakteristik kepala rumah tangga dan kondisi perumahan tempat tinggal yang dapat mempengaruhi peluang kepala rumah tangga mengalami gangguan kesehatan. Sehingga model probit yang digunakan adalah sebagai berikut:

$P_i = \frac{1}{1+e^{-Z_i}} = \frac{e^{Z_i}}{1+e^{Z_i}}$ dimana $Z_i = \alpha_i + \sum_k \boldsymbol{\beta_k X_{ik}} + \boldsymbol{\varepsilon_i}$, *i* adalah kepala rumah tangga ke-*i* , sedangkan $\boldsymbol{X_{ik}}$ adalah variabel bebas ke-*k* (yaitu berupa umur, jenis kelamin, lama sekolah, jam kerja, luas lantai tempat tinggal, kepemilikan toilet dan kepemilikan akses listrik) dari kepala rumah tangga ke-*i*. Dari hasil estimasi ini maka diperoleh estimasi peluang kepala rumah tangga mengalami gangguan kesehatan ($\widehat{P_i}$), sehingga nilai IMR diperoleh dengan membagi nilai *probability density function* (PDF) dengan *cumulative distribution function* (CDF) dalam distribusi standar normal, dengan rumus:

$$IMR_i = \left(\frac{1}{\sqrt{2\pi}} e^{\frac{-p^2}{2}}\right) \Big/ \left(\frac{1}{\sqrt{2\pi}} \int_{-\sim}^{Z_i} e^{\frac{-p^2}{2}}\, dp\right) \qquad (13)$$

Nilai IMR inilah yang akan menjadi salah satu variabel bebas pada persamaan pendapatan kepala rumah tangga pada persamaan 14. Sehingga dengan menggunakan metode ini maka bias yang terjadi karena pemilihan sampel dapat dikoreksi dan estimasi yang dihasilkan menjadi lebih baik.

Selanjutnya yaitu estimasi persamaan pendapatan kepala rumah tangga pada persamaan 14. Variabel terikat yang digunakan adalah nilai pendapatan kepala

rumah tangga selama sebulan terakhir dari pekerjaan utama pada tahun 2013 dan dinyatakan dalam bentuk logaritma dari pendapatan ril perbulan . Sehingga secara umum persamaan pendapatan kepala rumah tangga adalah sebagai berikut:

$$\boldsymbol{income_i = \alpha + \beta HS_i + \sum_k \gamma_k X_{ik} + \varepsilon_i} \quad (14)$$

Variabel bebas utama yang digunakan pada persamaan 14 yaitu berupa variabel gangguan kesehatan yang dialami oleh kepala rumah tangga (HS_i) dan diduga secara langsung mempengaruhi pendapatan kepala rumah tangga ($income_i$). Variabel gangguan kesehatan didefinisikan sebagai akumulasi dari lama hari terganggunya aktifitas sehari-hari karena gangguan kesehatan yang dialami kepala rumah tangga pada tahun 2012 dan tahun 2013. Dengan asumsi bahwa, kejadian gangguan kesehatan yang terjadi berulang yang dialami oleh kepala rumah tangga mengindikasikan bahwa gangguan kesehatan yang dialami merupakan gangguan kesehatan serius yang dapat mempengaruhi pendapatan yang bisa dihasilkan oleh kepala rumah tangga.

Selain variabel gangguan kesehatan, karakteristik yang melekat pada kepala rumah tangga juga dapat mempengaruhi tingkat pendapatannya, sehingga dapat digunakan sebagai variabel kontrol pada persamaan pendapatan kepala rumah tangga (X_{ik}). Diantara adalah umur, jenis kelamin, tingkat pendidikan (lama sekolah), lapangan usaha dimana kepala rumah tangga bekerja (pertanian, industri dan jasa), apakah kepala rumah tangga pernah mengalami pindah pekerjaan dalam satu tahun terakhir, kepemilikan jaminan kesehatan dan jumlah anak yang dimiliki.

Selain itu, persamaan ini juga menggunakan variabel interaksi antara gangguan kesehatan dan jenis kelamin dan variabel interaksi antara gangguan kesehatan dan kepemilikan jaminan kesehatan. Kedua variabel ini untuk menangkap perbedaan dampak dari gangguan kesehatan berdasarkan jenis kelamin dan kepemilikan jaminan kesehatan terhadap tingkat pendapatan kepala rumah tangga.

Selain karakteristik kepala rumah tangga, persamaan ini juga menggunakan variabel lokasi tempat tinggal (desa/kota) dan diwilayah mana kepala rumah tangga tersebut tinggal (berdasarkan pulau besar yang ada diIndonesia). Variabel ini selain untuk mengontrol perubahan yang terjadi pada level agregat/tingkat wilayah, yang juga dapat mempengaruhi pendapatan rumah tangga seperti pertumbuhan ekonomi, perubahan infrastruktur atau perubahan faktor lingkungan (Genoni, 2012) namun juga dapat menangkap perbedaan kondisi pasar tenaga kerja pada setiap wilayah.

4. Persamaan Konsumsi Rumah Tangga

Berdasarkan hasil estimasi pada persamaan (14), maka tahap selanjutnya adalah estimasi dampak perubahan pendapatan tersebut terhadap konsumsi rumah tangga. Pada tahap estimasi tingkat konsumsi rumah tangga ini, variabel bebas utama yang digunakan adalah variabel pendapatan hasil estimasi (*predicted value*) secara parsial dari persamaan (14), yaitu dengan hanya menggunakan variabel gangguan kesehatan, tidak termasuk variabel kontrol individu dan wilayah maupun variabel interaksi lainnya. Sehingga variabel estimasi pendapatan tersebut menggambarkan tingkat pendapatan yang dipengaruhi oleh gangguan kesehatan yang dialami oleh kepala rumah tangga (Arends-Kuenning, et al. 2019). Dengan metode ini, maka perubahan tingkat konsumsi yang terjadi didalam rumah tangga terjadi karena perubahan tingkat pendapatan yang disebabkan oleh gangguan kesehatan yang dialami oleh kepala rumah tangga. Sehingga persamaan konsumsi rumah tangga adalah sebagai berikut:

$$\boldsymbol{Cons_{il} = \alpha_l + \beta_l \widehat{income}_i + \sum_k \delta_k X_{ik} + \varepsilon_i} \quad (15)$$

Dimana $\boldsymbol{Cons_{il}}$ adalah porsi (*share*) pengeluaran rumah tangga untuk kelompok pengeluaran ke-*l* yang terdiri dari pengeluaran konsumsi pangan, non pangan (selain kesehatan) dan pengeluaran

kesehatan, terhadap total pengeluaran rumah tangga pada rumah tangga ke-*i*. Penggunaaan porsi pengeluaran rumah tangga khususnya porsi pengeluaran untuk konsumsi pangan, dapat digunakan sebagai pendekatan tidak langsung untuk tingkat kesejahteraan rumah tangga (Muellbauer, 1980). Perubahan porsi pengeluaran untuk konsumsi pangan dan konsumsi non pangan dapat menggambarkan bagaimana pengaruh fluktuasi pendapatan yang dialami oleh rumah tangga terhadap tingkat kesejahteraan mereka.

Dari persamaan (15) diketahui bahwa $\sum_l \boldsymbol{Cons_l} = 1$ untuk setiap rumah tangga ke-*i*. Sehingga perubahan porsi pengeluaran pada suatu kelompok pengeluaran maka akan mempengaruhi perubahan porsi pada kelompok pengeluaran yang lain. Karena itu sebagai suatu sistem persamaan, maka hubungan antar persamaan pada persamaan konsumsi rumah tangga akan terjadi pada adanya korelasi antar galat (*error*). Oleh karena itu berdasarkan pada hal ini maka, estimasi persamaan (15) dilakukan dengan menggunakan metode *Seemingly Unrelated Regression Estimation* (SURE). Model SURE merupakan suatu sistem persamaan linear yang terdiri dari beberapa persamaan dimana galat (*error*) antar persamaan tersebut memiliki korelasi. Menurut (Faharudin, Mulyana, etc. 2015), untuk mengestimasi persamaan (15) dip erlukan retriksi parameter agar konsisten dengan teori utilitas, retriksi tersebut yaitu:

$$\sum_l \boldsymbol{\alpha_l} = \mathbf{1}\,;\ \sum_l \boldsymbol{\beta_l} = \mathbf{0}\,;\ \sum_l \boldsymbol{\delta_{kl}} = \mathbf{0} \qquad (16)$$

Sehingga dengan restriksi ini maka jika terdapat *n* persamaan dalam suatu sistem persamaan, maka persamaan yang diestimasi hanya sebanyak *n*-1 persamaan (Poi, 2002).

Selain dipengaruhi oleh variabel tingkat pendapatan, tingkat konsumsi rumah tangga juga dipengaruhi oleh faktor-faktor lain (X_{ik}), yaitu jumlah anggota rumah tangga, yang dapat menggambarkan beban yang harus ditanggung oleh rumah tangga (Faharuddin, et.al. 2015), apakah terdapat anggota rumah tangga lansia diatas 60 tahun dan anak usia kurang dari lima belas tahun (Nguyet & Mangyo, 2010). Dimana keduanya selain merupakan tanggungan dari kepala rumah tangga namun kuantitas konsumsi keduanya yang relatif berbeda didalam rumah tangga dapat mempengaruhi tingkat konsumsi rumah tangga. Selain itu, pendapatan dari anggota rumah tangga yang lain juga dapat mempengaruhi tingkat konsumsi rumah tangga. Ketika pendapatan kepala rumah tangga menurun karena gangguan kesehatan yang dialami, pendapatan dari anggota rumah tangga yang lain dapat digunakan untuk memenuhi konsumsi rumah tangga. Kemudian variabel tingkat harga kesehatan. Dalam hal ini, menggunakan indeks harga sektor kesehatan tahun 2013 untuk semua propinsi di Indonesia dari BPS. Variabel ini dapat mempengaruhi tingkat konsumsi rumah tangga karena terkait dengan besarnya biaya kesehatan yang harus dikeluarkan ketika gangguan kesehatan terjadi.

HASIL DAN ANALISA

1. Estimasi Pendapatan Kepala Rumah Tangga

Tabel 1 dan tabel 2 dibawah ini menyajikan hasil estimasi dampak gangguan kesehatan yang dialami oleh kepala rumah tangga terhadap tingkat pendapatan ril perbulan kepala rumah tangga. Hasilnya diketahui bahwa gangguan kesehatan berdampak negatif dan signifikan mempengaruhi tingkat pendapatan ril perbulan kepala rumah tangga. Satu hari kepala rumah tangga mengalami gangguan kesehatan dapat menurunkan pendapatan ril perbulan sebesar 3,5 persen.

Dari hasil estimasi pada tabel 1, juga dapat dilihat bahwa kepala rumah tangga yang bekerja di sektor pertanian merasakan dampak negatif yang lebih besar terhadap tingkat pendapatannya ketika mengalami gangguan kesehatan dibandingkan dengan kepala rumah tangga yang bekerja pada sektor industri dan sektor jasa. Hal ini dapat dilihat dari nilai koefisien dari variabel *dummy* sektor lapangan usaha untuk sektor industri dan jasa yang bertanda positif dan signifikan. Hal ini menunjukkan bahwa

Tabel 1. Dampak Gangguan Kesehatan Terhadap Tingkat Pendapatan Ril Kepala Rumah Tangga

Variabel Bebas		Variabel Terikat Pendapatan Ril Kepala Rumah Tangga
Observasi Seluruh Rumah Tangga		
Constanta		-2.909***
		(1.060)
Lama Hari Gangguan Kesehatan (*longterm13*)		-0.035**
		(0.013)
Dummy Bekerja Di Sektor Industri (*industri13)*		0.177*
		(0.102)
Dummy Bekerja Di Sektor Jasa (*jasa13*)		0.276***
		(0.087)
Interaksi antara variabel gangguan kesehatan dan variabel kelamin (*longJK*)	*dummy* jenis	0.040***
		(0.014)
Interaksi antara variabel gangguan kesehatan dan variabel kepemilikan Jaminan Kesehatan (*Longjamkes*)	dummy	0.001
		(0.007)

- *Standard error didalam kurung*
- **** p<0.01, ** p<0.05, * p<0.1*

Tabel 2. Dampak Gangguan Kesehatan Terhadap Tingkat Pendapatan Ril Kepala Rumah Tangga Menurut Jenis Kelamin dan Kelompok Pendapatan

Variabel Bebas (Lama Hari Gangguan Kesehatan)	Variabel Terikat Pendapatan Ril Kepala Rumah Tangga
Sub Group 1 (Jenis Kelamin KRT)	
KRT Laki-Laki	0.005
	(0.004)
KRT Perempuan	-0.040***
	(0.012)
Sub Group 2 (Kelompok Pendapatan)	
Kwartil 1	-0.024*
	(0.013)
Kwartil 2	0.020
	(0.045)
Kwartil 3	-0.014
	(0.031)
Kwartil 4	-0.083**
	(0.041)

- Standard error didalam kurung

*** p<0.01, ** p<0.05, * p<0.1

secara rata-rata tingkat pendapatan kepala rumah tangga yang bekerja pada sektor industri dan sektor jasa lebih tinggi dibandingkan kepala rumah tangga yang bekerja pada sektor pertanian, ketika mereka mengalami gangguan kesehatan.

Gangguan kesehatan tentu akan menyebabkan penurunan produktifitas dalam bekerja. Fakta miris sektor pertanian di Indonesia, mulai dari produktifitas yang rendah karena kepemilikan lahan yang relatif sempit (petani gurem), petani berusia lanjut, sistem pertanian yang masih tradisional dan tingkat pendapatan yang rendah semakin memperburuk kondisi ini. Sehingga sangat wajar ketika rumah tangga mengalami guncangan pendapatan akibat gangguan kesehatan, maka dampak yang dirasakan pada sektor pertanian akan lebih besar dibandingkan dengan sektor lainnya

seperti sektor industri dan sektor jasa. Nguyet & Mangyo (2010) dalam penelitiannya juga menemukan bahwa dampak negatif gangguan kesehatan terhadap pendapatan rumah tangga pertanian lebih besar dibandingkan dengan rumah tangga non pertanian.

Dari tabel 1 juga bisa diketahui bahwa kepemilikan jaminan kesehatan tidak berpengaruh signifikan terhadap tingkat pendapatan kepala rumah tangga ketika mengalami gangguan kesehatan. Hal ini karena masih sedikitnya rumah tangga yang memiliki jaminan kesehatan pada tahun 2013. Pada saat itu, sistem jaminan kesehatan masih belum terintegrasi dibawah koordinasi BPJS. Sehingga partisipasi masyarakat dalam keanggotaan jaminan kesehatan pada masih sangat rendah. Karena itu manfaat dari adanya jaminan kesehatan tersebut belum dapat dirasakan secara luas dimasyarakat.

Selain itu, dari tabel 1 juga bisa dilihat bahwa, kepala rumah tangga laki-laki yang mengalami gangguan kesehatan terkena dampak penurunan pendapatan yang lebih rendah dari pada kepala rumah tangga perempuan. Hasil ini juga semakin diperkuat oleh hasil estimasi yang dihasilkan ketika observasi kepala rumah tangga dibedakan berdasarkan jenis kelamin sebagaimana yang ditampilkan pada tabel 2 berikut ini.

Berdasarkan tabel 2 diatas terlihat bahwa dampak gangguan kesehatan hanya berpengaruh negatif dan signifikan terhadap pendapatan ketika kepala rumah tangga adalah perempuan. Sedangkan pada rumah tangga laki-laki dampak tersebut tidak berpengaruh signifikan. Menurut laporan World Bank (2013), umumnya rumah tangga dengan kepala rumah berjenis kelamin perempuan (RTP) merupakan rumah tangga miskin dan hanya memiliki satu orang dewasa pencari nafkah (*single earner*).

Fakta ini sejalan dengan data BPS pada tabel 3 dibawah ini yang menunjukkan bahwa rumah tangga dengan kepala rumah tangga perempuan sebagian besar berstatus tidak memiliki pasangan, baik karena cerai hidup atau cerai mati. Sehingga sebagai *single parent* dan *single earner,* kepala rumah tangga perempuan tersebut harus menjalankan fungsi ganda yaitu sebagai pencari nafkah juga sebagai pengurus rumah tangga.

Kondisi ini diperparah dengan rata-rata usia mereka yang relatif tua, sehingga secara fisik memiliki kondisi tubuh yang lebih rentan untuk mengalami gangguan kesehatan. Sehingga, ketika gangguan kesehatan terjadi menimpa mereka, maka tidak ada anggota rumah tangga lain yang dapat menggantikan posisinya untuk bekerja sebagai kompensasi penurunan penawaran kerja tersebut. Karena itu kejadian gangguan kesehatan yang dialami menyebabkan penurunan pendapatan yang signifikan pada rumah tangga perempuan dibandingkan dengan rumah tangga laki-laki.

Hasil estimasi pada tabel 2 juga menunjukkan bahwa, jika dilihat berdasarkan kelompok pendapatan, maka pendapatan rumah tangga yang berada pada kwartil pertama akan mengalami dampak negatif yang signifikan akibat gangguan

Tabel 3. Persentase Rumah Tangga menurut Kelompok Umur, Jenis Kelamin Kepala Rumah Tangga, dan Status Perkawinan Tahun 2013

Kelompok Umur RT	Perempuan				Laki-Laki			
	Belum kawin	Kawin	Cerai hidup	Cerai mati	Belum kawin	Kawin	Cerai hidup	Cerai mati
10-24	83.17	10.84	5.53	0.46	40.94	58.33	0.69	0.05
25-44	10.83	31.92	27.22	30.03	2.39	96.06	1.08	0.47
45-59	2.58	8.36	16.05	73.02	0.54	95.40	1.19	2.86
60+	1.11	1.91	5.52	91.46	0.35	87.33	0.97	11.35
Total	7.12	10.32	13.40	69.16	2.25	93.73	1.09	2.92

Sumber: BPS

Tabel 4. Persentase Rumah Tangga menurut Status/Kedudukan Dalam Pekerjaan Utama dan Kelompok Pendapatan Tahun 2013

Status/Kedudukan Dlm Pekerjaan Utama	Kwartil Pendapatan			
	1	2	3	4
Berusaha sendiri	28.25	26.08	26.78	18.27
Berusaha dibantu buruh tdk tetap/tdk bayar	32.09	32.91	22.79	18.14
Berusaha dibantu buruh tetap/dibayar	2.14	3.28	4.32	11.38
Buruh/karyawan/pegawai	15.76	22.73	34.91	49.32
Pekerja bebas	20.55	13.46	10.15	2.28
Pekerja keluarga/tdk dibayar	1.20	1.53	1.05	0.62
Total	100	100	100	100

Sumber: BPS

kesehatan yang terjadi pada kepala rumah tangga. Satu hari kepala rumah tangga mengalami gangguan kesehatan akan menurunkan pendapatan ril perbulan sebesar 2,4 persen. Sebagaimana diketahui, rumah tangga pada kwartil pertama ini merupakan rumah tangga miskin, dan umumnya merupakan rumah tangga perempuan miskin dan hanya memiliki satu orang dewasa pencari nafkah (*single earner*). Maka dari itu dampak negatif dari gangguan kesehatan yang dialami oleh kepala rumah tangga akan berpengaruh signifikan terhadap tingkat pendapatannya. Sedangkan bagi rumah tangga pada kwartil kedua dan ketiga, tidak ada pengaruh yang signifikan terhadap tingkat pendapatan kepala rumah tangga ketika mengalami gangguan kesehatan.

Sedangkan bagi rumah tangga pada kwartil keempat yaitu rumah tangga dengan tingkat pendapatan yang lebih tinggi, maka satu hari gangguan kesehatan yang dialami oleh kepala rumah tangga akan menyebabkan penurunan pendapatan sebesar 8,3 persen. Sebagian besar kepala rumah tangga pada kwartil keempat ini merupakan buruh/pegawai. Sehingga ketika mereka tidak dapat bekerja karena gangguan kesehatan yang dialami maka mereka akan menerima konsekuensi berupa pemotongan pendapatan dari tempat mereka bekerja yang menyebabkan pendapatan mereka menjadi menurun.

2. Estimasi Tingkat Konsumsi Rumah Tangga

Berdasarkan hasil estimasi dampak gangguan kesehatan terhadap tingkat pendapatan kepala rumah tangga sebelumnya, diketahui bahwa secara umum gangguan kesehatan berdampak negatif terhadap tingkat pendapatan kepala rumah tangga. Selanjutnya dengan menggunakan hasil estimasi tersebut, maka akan dilihat bagaimana dampak penurunan pendapatan tersebut terhadap konsumsi rumah tangga.

Dengan menggunakan metode *seemingly unrelated regression*, tabel 5 dibawah ini menunjukkan arah hubungan dan besarnya dampak perubahan tingkat pendapatan kepala rumah tangga ketika mengalami gangguan kesehatan terhadap porsi pengeluaran untuk konsumsi pangan dan non pangan rumah tangga.

Berdasarkan tabel 5 berikut, terlihat bahwa penurunan pendapatan akibat gangguan kesehatan yang dialami oleh kepala rumah tangga akan menyebabkan penurunan porsi pengeluaran untuk konsumsi non pangan rumah tangga. Dalam hal ini, penurunan satu persen pendapatan rumah tangga karena gangguan kesehatan akan menurunkan porsi pengeluaran untuk konsumsi non makanan sebesar 5,3 persen.

Sedangkan untuk konsumsi pangan, hasil estimasi menunjukkan bahwa tidak ada pengaruh yang signifikan terhadap porsi pengeluaran untuk konsumsi pangan rumah tangga akibat penurunan pendapatan tersebut. Pola yang sama juga terjadi jika estimasi dilakukan pada rumah tangga perempuan dan rumah tangga pada kwartil pertama dan kwartil keempat. Yaitu penurunan pendapatan yang disebabkan oleh gangguan kesehatan yang dialami oleh kepala rumah tangga yang berjenis kelamin perempuan dan rumah tangga yang berada pada kwartil pertama dan keempat, rasio pengeluaran untuk konsumsi non pangan

Tabel 5. Dampak Perubahan Tingkat Pendapatan Akibat Gangguan Kesehatan Terhadap Rasio Pengeluaran Konsumsi Pangan dan Non Pangan Rumah Tangga (%)

Variabel Bebas (Estimasi Tingkat Pendapatan)	**Variabel Terikat**	
	Rasio Pengeluaran Pangan	**Rasio Pengeluaran Non Pangan**
Seluruh Rumah Tangga	-0.001	0.053***
	(0.008)	(0.009)
Sub Group (Berdasarkan Jenis Kelamin KRT)		
KRT Laki-Laki	0.006	-0.373***
	(0.059)	(0.063)
KRT Perempuan	(0.001)	0.047***
	(0.007)	(0.008)
Sub Group (Kelompok Pendapatan)		
Kwartil 1	0.014	0.062***
	(0.021)	(0.023)
Kwartil 2	(0.030)	-0.050*
	(0.026)	(0.028)
Kwartil 3	(0.004)	0.120***
	(0.035)	(0.037)
Kwartil 4	0.002	0.038***
	(0.012)	(0.012)

- Standard error didalam kurung
- *** p<0.01, ** p<0.05, * p<0.1
- Hasil estimasi pada rumah tangga dg KRT laki-laki dan rumah tangga pada kwartil 2 dan 3, meskipun menunjukkan hasil yang signifikan, namun karena hasil estimasi pada persamaan pendapatan kepala rumah tangga menunjukkan hasil yang tidak signifikan, maka pada dasarnya hasil estimasi ini juga tidak signifikan sehingga tidak bisa dianalisis lebih lanjut.

akan menurun, sedangkan rasio pengeluaran untuk konsumsi pangan tidak terkena dampak oleh penurunan pendapatan tersebut.

Hasil ini menunjukkan bahwa komoditas pangan sifatnya tidak elastis terhadap perubahan pendapatan. Sebagaimana hukum Engel juga menyatakan bahwa nilai elastisitas komoditas pangan akan terletak diantara 0 dan 1. Karena merupakan kebutuhan dasar bagi manusia, maka perubahan yang terjadi pada tingkat pendapatan tidak memberikan dampak berarti terhadap perubahan tingkat konsumsi pangan. Artinya bahwa, ketika terjadi fluktuasi pendapatan yang dialami oleh rumah tangga, misalnya karena gangguan kesehatan maka rumah tangga akan lebih memilih untuk mempertahankan tingkat konsumsi pangannya dengan cara mengurangi porsi pengeluaran untuk konsumsi non pangan.

Kemudian jika ditelisik lebih jauh, tabel 6 dibawah ini menunjukkan komoditas non pangan yang mengalami perubahan karena penurunan pendapatan akibat gangguan kesehatan yang dialami oleh kepala rumah tangga . Terlihat bahwa penurunan porsi pengeluaran konsumsi non pangan terjadi pada pengeluaran untuk perawatan dan pemeliharan rumah (termasuk biaya sewa/kontrak rumah), sedangkan pengeluaran untuk peralatan kebersihan, perawatan tubuh dan kebutuhan sehari lainnya (misal: sabun cuci/mandi, surat kabar, alat tulis, kosmetik dll) justru mengalami peningkatan. Hal ini diduga terjadi karena adanya peningkatan pengeluaran untuk barang-barang yang terkait dengan perawatan kesehatan bagi kepala rumah tangga yang sedang mengalami gangguan kesehatan.

Sedangkan pengeluaran untuk sumber energi dan telekomunikasi serta pengeluaran untuk pendidikan dan transportasi tidak terpengaruh oleh penurunan pendapatan akibat gangguan kesehatan . Hasil estimasi yang sama juga diperoleh pada sub sampel rumah tangga perempuan dan rumah tangga yang berada

Tabel 6. Dampak Perubahan Tingkat Pendapatan Akibat Gangguan Kesehatan Terhadap Rasio Pengeluaran Konsumsi Non Pangan Rumah Tangga (%)

Variabel Bebas (Estimasi Tingkat Pendapatan)	**Pengeluaran Konsumsi Non Pangan (Selain Kesehatan)**			
	Perawatan dan Pemeliharaan Rumah	**Sumber Energi, Pos dan Telekomunikasi**	**Perawatan Tubuh**	**Pendidikan dan Transportasi**
Sub Sampel				
Seluruh Rumah Tangga	0.018*	0.000	-0.008***	0.007
	(0.010)	(0.006)	(0.003)	(0.009)
Rumah Tangga Perempuan	0.016*	0.000	-0.007***	0.006
	(0.008)	(0.005)	(0.003)	(0.008)
RumahTangga Kwartil Pertama	0.054**	-0.002	-0.015*	-0.014
	(0.026)	(0.018)	(0.009)	(0.025)
Rumah Tangga Kwartil 4	0.005	0.014*	0.004	0.007
	(0.013)	(0.008)	(0.004)	(0.013)

Standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

pada kwartil pertama. Sedangkan untuk rumah tangga di kwartil keempat, penurunan pengeluaran non pangan terjadi pada pengeluaran energi dan telekomunikasi (misal: bahan bakar, air, telekomunikasi dll). Sedangkan untuk pengeluaran non pangan yang lain tidak terpengaruh oleh adanya penurunan pendapatan akibat gangguan kesehatan yang terjadi.

KESIMPULAN DAN SARAN

Gangguan kesehatan yang dialami oleh kepala rumah tangga akan menyebabkan penurunan tingkat pendapatan ril perbulan. Penurunan pendapatan ini akan lebih dirasakan oleh rumah tangga yang berada pada kelompok pendapatan terendah, dengan kepala rumah tangga berjenis kelamin perempuan. Kepala rumah tangga yang bekerja pada sektor pertanian juga akan merasakan dampak yang lebih besar dibandingkan dengan kepala rumah tangga yang bekerja pada sektor industri dan jasa.

Akibat penurunan pendapatan kepala rumah tangga tersebut, maka rumah tangga akan menurunkan porsi pengeluaran konsumsi untuk non pangan. Sedangkan porsi pengeluaran untuk pangan tidak terdampak oleh adanya penurunan pendapatan akibat gangguan kesehatan tersebut. Artinya ketika pendapatan menurun, rumah tangga akan berusaha untuk menjaga tingkat konsumsi pangannya dengan cara menurunkan pengeluaran untuk konsumsi non pangan.

Perubahan pengeluaran konsumsi untuk non pangan tersebut terjadi untuk pengeluaran perawatan dan pemeliharaan rumah yang semakin menurun. Sedangkan pengeluaran untuk perawatan tubuh cenderung meningkat, yang dilakukan untuk memenuhi kebutuhan untuk barang-barang yang diperlukan untuk perawatan kesehatan bagi kepala rumah tangga yang sakit. Sementara itu, pengeluaran untuk pendidikan dan transportasi serta pengeluaran untuk kebutuhan sumber energi dan telekomunikasi tidak terpengaruh oleh adanya penurunan pendapatan akibat gangguan kesehatan yang dialami oleh kepala rumah tangga.

Beberapa keterbatasan penelitian ini diantaranya yaitu pertama, periode penelitian ini menggunakan data tahun 2013, yang dirasa sudah terlalu lama (*out of date*). Sehingga kurang bisa

menggambarkan kondisi terkini yang bisa sangat berbeda. Meskipun hal ini terjadi karena ketersediaan data panel tahunan yang terbatas. Kedua yaitu lemahnya ukuran gangguan kesehatan yang digunakan sehingga tidak dapat menangkap dengan baik tingkat keparahan dari gangguan kesehatan yang dialami oleh kepala rumah tangga, yang mungkin akan memberikan dampak yang berbeda terhadap pendapatan dan konsumsi rumah tangga. Yang terakhir yaitu, penelitian ini juga belum mempertimbangkan gangguan kesehatan yang terjadi pada anggota rumah tangga yang lain baik yang memiliki kontribusi terhadap pendapatan rumah tangga maupun yang tidak. Dimana hal ini bisa saja turut mempengaruhi pendapatan dan konsumsi rumah tangga.

Karena itu untuk penelitian berikutnya disarankan untuk menggunakan data yang terbaru yang lebih dapat menggambarkan kondisi terkini. Kemudian dengan menggunakan definisi gangguan kesehatan yang dapat menangkap derajat keparahan dari gangguan kesehatan yang terjadi tidak hanya pada kepala rumah tangga namun juga pada anggota rumah tangga lain, baik yang memiliki kontribusi terhadap pendapatan rumah tangga atau tidak. Sehingga diharapkan bisa melengkapi studi-studi empiris yang sudah ada sebelumnya.

Sedangkan bagi perumus kebijakan, kebijakan yang dapat melindungi kesejahteraan rumah tangga ketika mereka mengalami gangguan kesehatan harus lebih ditingkatkan, khususnya pada rumah tangga miskin disektor pertanian dan dengan kepala rumah tangga perempuan. Mengingat dampak negatif akibat gangguan kesehatan terhadap kesejahteraan yang dirasakan oleh rumah tangga tersebut lebih besar. Misalnya dalam bentuk subsidi langsung biaya pengobatan atau dalam bentuk *cash transfer*. Sehingga melalui kebijakan tersebut dapat mencegah penurunan kesejahteraan rumah tangga akibat penurunan pendapatan atau peningkatan pengeluaran akibat gangguan kesehatan yang dialami.

DAFTAR PUSTAKA

Arends-Kuenning, M., Baylis, K., & Garduño-Rivera, R. (2019). The effect of NAFTA on internal migration in Mexico: a regional economic analysis. *Applied Economics*, *51*(10), 1052–1068. https://doi.org/10.1080/00036846.2018.1524976

Asfaw, A., & Braun, J. von. (2004). Is Consumption Insured against Illness? Evidence on Vulnerability of Households to Health Shocks in Rural Ethiopia. *Economic Development and Cultural Change*, *53*(1), 115–129. https://doi.org/10.1086/423255

Ashenfelter, B. Y. O., & Heckman, J. (1974). The Estimation of Income and Substitution Effects in a Model of Family Labor Supply. *Econometrica, Vol. 42, No. 1 (Jan., 1974), Pp. 73-85*, *42*(1), 73–85.

Cochrane, J. H. (1991). A Simple Test of Consumption Insurance, *99*(5), 957–976.

Faharuddin, F., Mulyana, A., Yamin, M., & Yunita Yunita. (2015). Nutrient elasticities of food consumption: the case of Indonesia. *Journal of Agribusiness in Developing and Emerging Economies*, *5*(2), 57–75. https://doi.org/10.1108/JADEE-04-2017-0048

Faharudin, Mulyana, A., Yamin, M., & Yunita. (2015). ANALISIS POLA KONSUMSI PANGAN DI SUMATERA SELATAN 2013 : PENDEKATAN QUADRATIC ALMOST IDEAL DEMAND SYSTEM Analysis of Food Consumption Patterns in South Sumatra in 2013 : A Quadratic Almost Ideal Demand System Approach.

Genoni, M. E. (2012). Health Shocks and Consumption Smoothing: Evidence from Indonesia. *Economic Development and Cultural Change*, *60*(3), 475–506. https://doi.org/10.1086/664019

Gertler, P., & Gruber, J. (2002). Insuring consumption against illness.

American, The Review, Economic Database, Social Science.

Hoogeveen, J., Tesliuc, E., Vakis, R., & Dercon, S. (2004). A Guide to the Analysis of Risk Vulnerability and Vulnerable Groups.

Kementerian Kesehatan Republik Indonesia. (2017). *Profil Kesehatan Indonesia Tahun 2016*. https://doi.org/10.1111/evo.12990

Maloney, T. (1987). Employment Constraints and the Labor Supply of Married Women A Reexamination of the Added Worker Effect. *The Journal of Human Resources*, *22*(1), 51–61. https://doi.org/10.2307/145866

Muellbauer, A. D. and J. (1980). An Almost Ideal Demand System. *American Economic Association An*, *70*(2–3), 312–326. https://doi.org/S0277-9536(13)00479-6 [pii]\r10.1016/j.socscimed.2013.08.027

Nguyet, N. T. N., & Mangyo, E. (2010). Vulnerability of households to health shocks: An Indonesian study. *Bulletin of Indonesian Economic Studies*, *46*(2), 213–235. https://doi.org/10.1080/00074918.2010.486108

Poi, B. P. (2002). From the Help Desk: Demand System Estimation. *The Stata Journal: Promoting Communications on Statistics and Stata*, *2*(4), 403–410. https://doi.org/10.1177/1536867x0200200406

Russell, S. (2004). The Economic Burden Of Illness For Households In Developing Countries: A Review Of Studies Focusing On Malaria, Tuberculosis, And Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome. *The American Journal of Tropical Medicine and Hygiene*, *71*(2 suppl), 147–155. Retrieved from http://www.ajtmh.org/content/71/2_suppl/147

Sparrow, R., Poel, E. Van De, Hadiwidjaja, G., Yumna, A., Warda, N., & Suryahadi, A. (2014). Coping With The Economic Consequences Of Ill Health In Indonesia. *Health Econ*, *23*(July 2013), 719–728. https://doi.org/10.1002/hec

Strauss, J., & Thomas, D. (1998). Health , Nutrition , and Economic Development, *36*(2), 766–817.

Townsend, R. M. (1994). Risk and Insurance in Village India Author. *The Econometric Society*, *62*(3), 539–591.

Wagstaff, A. (2007). The economic consequences of health shocks: Evidence from Vietnam. *Journal of Health Economics*, *26*(1), 82–100. https://doi.org/10.1016/j.jhealeco.2006.07.001

Wang, H., Zhang, L., & Hsiao, W. (2006). Ill health and its potential influence on household consumptions in rural China. *Health Policy*, *78*(2–3), 167–177. https://doi.org/10.1016/j.healthpol.2005.09.008

Widyanti, W., Suryahadi, A., Sumarto, S., & Yumna, A. (2010). The Relationship Between Chronic Poverty and Household Dynamics: Evidence from Indonesia. *Ssrn*, (January). https://doi.org/10.2139/ssrn.1537080

World Bank. (2013). *Indonesia - Gender equality (English). Indonesia Gender policy brief ; no. 1.* Washington DC: World Bank.

PREDIKSI HARGA EMAS DUNIA DI MASA PANDEMI COVID-19 MENGGUNAKAN MODEL ARIMA

Dara Puspita Anggraeni [1], Dedi Rosadi[2], Hermansah[3], Ahmad Ashril Rizal[4]

[1]Universitas Nahdlatul Wathan Mataram, [2]Universitas Gadjah Mada Yogyakarta, [3]Universitas Riau Kepulauan,
[4]STMIK Syaikh Zainuddin NW Anjani Lombok timur
e-mail: [1]darapuspitaanggraeni40@gmail.com, [2]dedirosadi@gadjahmada.edu, [3]hermansah@mail.ugm.ac.id,
[4]ashril.rizal@gmail.com

Abstrak

Penelitian ini bertujuan memodelkan serta memprediksi harga emas dunia di masa pandemi COVID-19. Penelitian ini juga hanya memasukkan nilai masa lampau dari harga emas dunia tanpa adanya pengaruh faktor eksogen(independen) pada model. Model yang dipergunakan adalah model *Autoregressive Integrated Moving Average* (ARIMA). Adapun data yang dipergunakan pada permodelan sebanyak 240 data observasi dimana data merupakan data bulanan harga emas dunia bulan Agustus 2000 hingga Juli 2020. Model terbaik untuk harga emas dunia ini adalah ARIMA(0,1,1) dengan nilai *Mean Absolute Percentage Error* (MAPE) sebesar 3,70%. Hasil prediksi harga emas dunia untuk bulan Agustus 2020 hingga Januari 2021 berturut-turut adalah sebesar 1930,046; 1945,651; 1961,381; 1977,240; 1993,227; 2009,343 US$/Troy Ons emas. Prediksi ini menunjukkan tren naik dengan rata-rata peningkatan selama periode tersebut (Agustus 2020-Januari 2021) sebesar15,8594 US$/Troy ons per bulannya.

Kata kunci: Harga Emas, COVID-19, ARIMA, Prediksi

Abstract

This research aims to model and predict the world gold price during the COVID-19 pandemic. This research only includes the past values of world gold prices without the influence of exogenous (independent) factors on the model. The model used in this research is the Autoregressive Integrated Moving Average (ARIMA). The data used in the modeling are 240 observational data which were the monthly data on world gold prices from August 2000 to July 2020. The best model for this world gold price is ARIMA (0,1,1) with a Mean Absolute Percentage Error (MAPE) value of 3.70%. The prediction results of the world gold price from August 2020 to January 2021 are 1930,046 respectively; 1945,651; 1961,381; 1977,240; 1993,227; 2009,343 US $ / Troy Ounce of gold. This prediction shows an upward trend with the average increase 15,8594 US $ / Troy ounce per month during that period (August 2020-January 2021).

***Keywords**: Gold Price, COVID-19, ARIMA, Prediction*

PENDAHULUAN

Investasi diminati oleh sebagian besar orang di dunia. Investasi atau penanaman modal sering sekali dilakukan dengan tujuan mendapatkan keuntungan di masa depan. Jenis investasi yang paling diminati di saat ini adalah emas, saham, obligasi dan properti. Tiap jenis investasi ini memiliki keuntungan dan risiko yang berbeda-beda. Menurut Warsono (2010) pada umumnya investasi pada aset riil mempunyai nilai satuan yang relatif besar dan mempunyai likuiditas relatif rendah, sedangkan aset keuangan mempunyai nilai satuan yang relatif kecil dan pada umumnya mempunyai likuiditas yang tinggi. Investasi yang relatif mudah untuk dilakukan saat ini adalah pada aset keuangan. Salah satu prinsip dalam berinvestasi adalah *higher return higher risk*. Suatu investasi dengan pengembalian diharapkan sangat tinggi, maka risiko yang dihadapi oleh investor juga sangat tinggi. Sebaliknya, jika angin berinvestasi pada aset keuangan dengan risiko rendah, maka pengembalian yang diharapkan juga rendah(Yulianti & Silvy, 2013).

Emas atau logam mulia menjadi jenis investasi yang sering kali disebut sebagai investasi aman dibandingkan jenis instrumen investasi lannya. Investasi emas dikatakan mudah karena emas ini, tidak harus dimiliki seseorang yang mempunyai penghasilan besar ataupun seseorang yang mempunyai jabatan khusus. Emas ini juga dapat dimanfaatkan oleh siapapun dari berbagai macam kalangan masyarakat. Selain mudah, investasi emas juga merupakan salah satu investasi yang sangat menguntungkan karena emas merupakan satu-satunya logam mulia yang harga jualnya tidak terpengaruh oleh inflasi yang terjadi. Dapat dibuktikan sebagai contoh, satu koin dinar yang memiliki berat ± 4,25 gram misalnya setara dengan harga 1 ekor kambing pada masa Rasulullah SAW, dan sampai sekarang pun masih berlaku seperti itu karena emas tidak terpengaruh oleh inflasi. Yang berubah hanyalah daya beli emas dengan uang kertas seperti Rupiah yang semakin lama semakin menurun. Selain itu seperti yang kita ketahui bahwasanya harga emas cenderung terus menerus mengalami kenaikan setiap tahunnya(Fauziah & Surya, 2016) terlebih lagi disaat pandemi COVID-19 yang menelan korban sebanyak 15.785.641 kasus dan 640.016 kematian per 26 Juli 2020 (WHO, 2020b) sehingga sangat mengganggu sektor kesehatan juga sektor perekonomian, termasuk dunia investasi saham (Baker et al., 2020) ditunjukkan dengan nilai indeks saham gabungan yang menurun di beberapa negara yang mengakibatkan beralihnya investasi yang diminati oleh calon investor dari bentuk saham ke dalam bentuk investasi emas(Ji et al., 2020). Hal ini diperkuat oleh hasil penelitian yang telah dilakukan oleh Ji, Zhang dan Zhao ini dengan melakukan analisa pada beberapa jenis investasi diluar saham sebagai investasi teraman selama pandemi Covid-19, diperoleh hasil yang menunjukkan emas sebagai investasi teraman(Ji et al., 2020). Oleh karenanya sangat menarik untuk memodelkan harga emas dunia yang dapat bermanfaat baik bagi para praktisi maupun peneliti.

Pada penelitian ini peneliti akan menggunakan model ARIMA dikarenakan model ini telah terbukti cocok untuk memodelkan serta memprediksi harga emas dunia. Penelitian sebelumnya diantara lain dilakukan oleh Abdullah (2012), Khan (2013), Bandyopadhyay (2016) dan Yang (2019). George Box dan Gwilym Jenskin adalah penemu model ARIMA pada tahun 1976. Box dan Jenkins menggunakan model-model ARIMA untuk deret waktu satu variable (*univariate*). Model ARIMA (*p,d,q*), dimana *p* menyatakan orde dari proses *autoregressive* (AR), *d* menyatakan pembeda (*differencing*) dan *q* menyatakan orde dari proses *moving average* (MA). Dasar dari model ARIMA dilakukan dengan empat tahap strategi pemodelan yaitu identifikasi model, penaksiran parameter, pemeriksaan diagnostik dan prediksi (Rosadi, 2011).

Adapun tujuan dari penelitian ini adalah untuk memodelkan serta memprediksi harga emas dunia di masa mendatang dengan akurat.

METODE

1. Tinjauan Referensi

Emas merupakan *safe haven* dikarenakan ketersediaannya yang langka, banyak diminati dan sangat berharga secara intrinsik terlebih selama masa pandemi COVID-19 (Ji et al., 2020). Penelitian tentang model dan prediksi harga emas dunia di masa pandemi COVID-19 belum banyak dilakukan sehingga peneliti tertarik untuk melakukan penelitian ini dengan salah satu tujuannya adalah membantu praktisi(misalkan calon investor) dalam pengambilan keputusan investasi emas.

Penelitian tentang prediksi harga emas dunia menggunakan model ARIMA merujuk pada penelitian terdahulu dilakukan oleh Guha Bandyopadhyay yang memperoleh kesimpulan bahwa prediksi menggunakan model ARIMA sangat akurat digunakan untuk prediksi jangka pendek (Bandyopadhyay, 2016). ARIMA juga sangat baik menggambarkan data yang fluktuatif (Abdullah, 2012), dan cocok digunakan untuk menggambarkan serta meramalkan pergerakan harga emas dunia (Khan, 2013; Yang, 2019)

2. Metode Analisis

Time Series Data

Time Series (Runtun waktu) *data* yakni jenis data yang dikumpulkan menurut urutan waktu dalam suatu rentang waktu tertentu. Jika waktu dipandang bersifat diskrit (waktu dapat dimodelkan bersifat kontinu), maka frekuensi pengumpulan selalu sama (equidistant). Dalam kasus diskrit, frekuensi dapat berupa misalnya detik, menit, jam, hari, minggu, bulan atau tahun. Model yang digunakan adalah model-model time series, yang menjadi fokus dari perkuliahan ini (Rosadi, 2006).

Kestasioneran Data

Stasioneritas merupakan suatu keadaan jika proses pembangkitan yang mendasari suatu deret berkala didasarkan pada nilai tengah konstan dan nilai varians konstan. Dalam suatu data kemungkinan data tersebut tidak stationer hal ini dikarenakan *mean*(rata-rata) tidak konstan atau variannya tidak konstan sehingga untuk menghilangkan ketidakstasioneran terhadap *mean*, maka data tersebut dapat dibuat lebih mendekati stasioner dengan cara melakukan penggunaan metode pembedaan atau *differencing*. Perilaku data yang stasioner antara lain tidak mempunyai variasi yang terlalu besar dan mempunyai kecenderungan untuk mendekati nilai rata-ratanya, dan sebaliknya untuk data yang tidak stasioner(Gujarati, 2004).

1. Stasioner dalam variasi

Pada data yang tidak stasioner dalam variasi dapat dilakukan transformasi untuk membuat data tersebut stasioner.Box dan Cox pada tahun 1964 memperkenalkan transformasi pangkat (*power transformation*) sebagai berikut:

$$Z_t' = \begin{cases} \dfrac{Z_t^{\lambda} - 1}{\lambda}, \lambda \neq 0 \\ \ln(Z_t), \lambda = 0 \end{cases} \quad (1)$$

Dengan Z_t adalah deret waktu periode ke-t dan λ adalah parameter transformasi.

2. Stasioner dalam rata-rata

Data yang tidak stasioner dalam rata-rata dapat distasionerkan melalui proses *differensing*. Pengujian hipotesis yang yang sering digunakan untuk melakukan pengecekan kestasioneran data runtun waktu dalam rata-rata adalah uji *Augmented Dickey-Fuller* (ADF). Uji ini merupakan salah satu uji yang paling sering digunakan dalam pengujian stasioneritas dari data, yakni dengan melihat apakah di dalam model terdapat unit *root* atau tidak(Rahmawati et al., 2019).

Pengujian dilakukan dengan menguji hipotesis $H_0: \rho = 0$ (terdapat akar unit) dalam persamaan regresi

$$\Delta Y_t = \alpha + \delta t + \rho Y_{t-1} + \sum_{j=1}^{k} \emptyset_j Y_{t-j} + e_t \quad (2)$$

Hipotesis nol ditolak jika nilai statistik uji ADF memiliki nilai kurang (lebih negatif) dibandingkan dengan nilai daerah kritik. Jika hipotesis nol ditolak, data bersifat stasioner (Rosadi, 2011).

Fungsi Autokorelasi dan Fungsi Autokorelasi Parsial

Tahap identifikasi model pada model data runtun waktu dibutuhkan hasil perhitungan fungsi autokorelasi dan fungsi autokorelasi parsial. Adapun perhitungannya adalah sebagai berikut:

1. Fungsi Autokorelasi (*Autocorrelation Function*/ACF)

Konsepsi autokorelasi setara (identik) dengan korelasi Pearson untuk data bivariat. Menurut Mulyana (2004) persamaan koefisien autokorelasi $\rho_k$adalah:

$$\rho_k = \frac{\sum_{t=1}^{n-k}(Z_t-\bar{Z})(Z_{t+k}-\bar{Z})}{\sum_{t=1}^{n}(Z_t-\bar{Z})^2} \qquad (3)$$

Keterangan:

n: jumlah observasi

k: selisih waktu (*lag*)

Z_t: adalah data pada waktu-*t*

Z_{t+k}: adalah data pada waktu ke- t+k

$\bar{Z}$: rata-rata dari Z_t

Plot (grafik) ACF yang menurun menjadi salah satu indicator data belum stasioner.

2. Fungsi Autokorelasi Parsial (*Partial Autocorrelation function*/PACF)

Nilai PACF dapat dihitung dengan cara sebagai berikut:

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}} \qquad (4)$$

(Mulyana, 2004).

Keterangan: formula diatas merupakan perhitungan untuk PACF untuk *lag* k dan j=1,2,3,…,k

Model ARIMA

Model AR(p) dan MA(q) merupakan model data runtun waktu stasioner dan saling berkebalikan, sehingga keduanya dapat digabungkan dengan cara dijumlahkan, dan model yang diperoleh dinamakan model autoregresi rata-rata bergerak, disingkat ARMA(p,q). Karena AR(p) dan MA(q) adalah model data runtun waktu stasioner, maka ARMA(p,q) juga model data runtun waktu stasioner. Jika data tidak stasioner, maka dapat distasionerkan melalui proses stasioneritas, yang berupa proses diferensi jika trendnya linier, dan proses linieritas dengan proses diferensi pada data hasil proses linieritas, jika trend data tidak linier (Mulyana, 2004). Model ARMA(p,q) untuk data hasil proses diferensi dinamakan model autoregresi integrated rata-rata bergerak disingkat ARIMA(p,d,q) dengan persamaan:

Bentuk umum dari persamaan model ARIMA:

$$\left(1-\phi_1 B-\cdots-\phi_p B^p\right)(1-B)^d Z_t = \mu+\varepsilon_t+\theta_1\varepsilon_{t-1}+\cdots+\theta_q\varepsilon_{t-q}, \varepsilon_t \sim IID(0,\sigma^2) \qquad (5)$$

dengan B yang merupakan operator balik (*backward*), yakni $(B^j Z)_t = Z_{t-j}$ (Rosadi, 2011).

Persamaan diatas dapat pula dituliskan dalam bentuk:

$$\left(1-\phi_1 B-\cdots-\phi_p B^p\right)(1-B)^d Z_t = \mu+\left(1+\theta_1 B+\cdots+\theta_q B^q\right)\varepsilon_t, \varepsilon_t \sim IID(0,\sigma^2) \qquad (6)$$

Prapemrosesan Data

Dalam tahap awal dilakukan identifikasi model runtun waktu yang mungkin digunakan untuk memodelkan sifat-sifat data. Identifikasi secara sederhana dilakukan secara visual dengan melihat plot data, untuk melihat adanya tren, komponen musiman, nonstasioneritas dalam variasi dan lain-lain. Beberapa teknik prapemrosesan data yang umum dilakukan adalah membuang pencilan dari dalam data, penyaringan data dengan model/teknik statistika tertentu, transformasi data (seperti transformasi logaritma atau yang lebih umum transformasi Box-Cox), melakukan operasi deferens, detren (membuang tren), *deseasonalize* (membuang komponen musiman) dan lain-lain(Rosadi, 2011).

Identifikasi Model Stasioner

Bentuk model ARMA yang tepat dalam menggambarkan sifat-sifat data dapat ditentukan dengan plot sampel ACF/PACF dengan sifat-sifat fungsi

Tabel 1. Bentuk plot sampel ACF/PACF dari model ARMA

Proses	Sampel ACF	Sampe PACF
White noise (galat acak)	Tidak ada yang melewati batas interval pada $lag > 0$	Tidak ada yang melewati batas interval pada $lag > 0$
AR(p)	Meluruh menuju nol secara eksponensial	Di atas batas interval maksimum sampai lag ke p dan di bawah batas pada $lag > p$
MA(q)	Di atas batas interval maksimum sampai lag ke q dan di bawah batas pada $lag > q$	Meluruh menuju nol secara eksponensial
ARMA(p,q)	Meluruh menuju nol secara eksponensial	Meluruh menuju nol secara eksponensial

ACF/PACF teoritis dari model ARMA. Rangkuman bentuk plot sampel ACF/PACF dari model ARMA diberikan pada tabel 1 (Rosadi, 2011).

Penaksiran Parameter/Estimasi Model

Estimasi dari model ARMA dapat dilakukan dengan metode *Maksimum Likelihood Estimator* (MLE), *Least Square*, Hannan Rissanen, metode Whittle dan lain-lain(Rosadi, 2011).

Pengujian Signifikansi Parameter

Pengujian apakah koefisien hasil estimasi signifikan atau tidak dengan uji hipotesis:

- Uji konstanta pada model:

  $H_0 : \mu = 0$

  $H_1 : \mu \neq 0$

  H_0 diterima jika $-t_{tabel} \leq t_{hitung} \leq t_{tabel}$, sebaliknya jika $t_{hitung} \leq -t_{tabel}$ atau $t_{hitung} \geq t_{tabel}$ maka H_0 ditolak dan H_1 diterima (parameter signifikan).

  Adapun perhitungan statistik uji

  $$t_{hitung} = t = (\hat{\mu} - 0)/SE(\hat{\mu}) \quad (7)$$

  dan statistik tabel $t_{tabel} = t(df = n - 1; \alpha = 2{,}5\%)$.

- Uji Parameter untuk model AR(p)

  $H_0 : \phi_i = 0, i = 1{,}2 \ldots, p$

  $H_1 : \phi_i \neq 0, i = 1{,}2 \ldots, p$

  H_0 diterima jika $-t_{tabel} \leq t_{hitung} \leq t_{tabel}$, sebaliknya jika $t_{hitung} \leq -t_{tabel}$ atau $t_{hitung} \geq t_{tabel}$ maka H_0 ditolak dan H_1 diterima (parameter signifikan).

  Adapun perhitungan statistik uji

  $$t_{hitung} = t = (\widehat{\phi}_i - 0)/SE(\widehat{\phi}_i) \quad (8)$$

  dan statistik tabel $t_{tabel} = t(df = n - 1; \alpha = 2{,}5\%)$.

- Uji Parameter untuk model MA(q)

  $H_0 : \theta_i = 0, i = 1{,}2 \ldots, q$

  $H_1 : \theta_i \neq 0, i = 1{,}2 \ldots, q$

  H_0 diterima jika $-t_{tabel} \leq t_{hitung} \leq t_{tabel}$, sebaliknya jika $t_{hitung} \leq -t_{tabel}$ atau $t_{hitung} \geq t_{tabel}$ maka H_0 ditolak dan H_1 diterima (parameter signifikan).

  Adapun perhitungan statistik uji

  $$t_{hitung} = t = (\widehat{\theta}_i - 0)/SE(\widehat{\theta}_i) \quad (9)$$

  dan statistik tabel $t_{tabel} = t(df = n - 1; \alpha = 2{,}5\%)$.

Pada formula diatas n adalah banyak data yang dipergunakan dalam membangun model. Jika terdapat koefisien yang tidak signifikan, koefisien/orde *lag* tersebut dapat dibuang dari model dan model diestimasi kembali tanpa mengikutkan orde yang tidak signifikan

Pemeriksaan Diagnosis

Jika model merupakan model yang tepat, data yang dihitung dengan model (*fitted value*) akan memiliki sifat-sifat yang mirip dengan data asli . Dengan demikian, residual yang dihitung berdasarkan model

yang telah diestimasi mengikuti asumsi dari galat model teoritis, seperti sifat *white noise*, normalitas dari residual (walaupun asumsi ini dapat diabaikan, tidak sepenting asumsi white noise dari galat0 dan lain-lain. Untuk melihat apakah residual bersifat *white noise* , du acara bisa digunakan yakni:

Melihat apakah plot sampel ACF/PACF yang terstandarisasi (residual dibagi estimasi deviasi standar residual) telah memenuhi sifat-sifat proses *white noise* dengan mean 0 dan variansi 1

Melakukan uji korelasi parsial, yakni dengan menguji hipotesis:

H_0: $\rho_1 = \rho_2 = \cdots = \rho_k, k < n$(tidak terdapat korelasi serial dalam residual sampai $lag - k,\ k < n$)

Uji ini dapat dilakukan dengan statistik uji Box-Pierce:

$$Q = n\sum_{j=1}^{k} \hat{\rho}(j)^2 \quad (10)$$

Atau Ljung-Box:

$$Q = n(n+)\sum_{j=1}^{k} \hat{\rho}(j)^2/(n-j) \quad (11)$$

yang akan berdistribusi $\chi^2\big(k-(p+q)\big), k > (p+q)$. Di sini $\hat{\rho}(j)$ menunjukkan nilai sampel ACF residual pada *lag-j*, sedangkan p dan q menunjukkan orde dari model ARMA (p,q). Apabila hipotesis cek diagnostic ditolak, model yang telah diidentifikasi si atas tidak dapat digunakan dan selanjutnya model yang mungkin sesuai untuk data dapat diidentifikasi kembali(Rosadi, 2011).

Pemilihan Model Terbaik

Selanjutnya, dalam praktikan ada banyak model yang memenuhi pengujian diagnostic di atas. Untuk model terbaik, pilih model yang meminimalkan ukuran kriteria informasi, seperti *Aike Information Criteria (AIC)*

$$AIC = n\ln\big(\hat{\sigma_\varepsilon}^2\big) + 2(p+q+1) \quad (12)$$

$$\hat{\sigma_\varepsilon}^2 = SSE/n \quad (13)$$

Dengan *sum of squared error* (SSE) yang akan diestimasi dari jumlahan kuadrat semua nilai residual. Akan tetapi diketahui untuk model autoregresif, kriteria AIC tidak memberikan orde p yang konsisten sehingga untuk pembanding kita bisa menggunakan kriteria informasi lain, seperti *Schwarzt Bayesian Information Criteria* (*SBC*)

$$SBC = n\ln\big(\hat{\sigma_\varepsilon}^2\big) + (p+q+1)\ln n \quad (14)$$

Atau bentuk-bentuk kriteria informasi lain yang diusulkan di dalam literatur(Rosadi, 2011).

Pengukuran Ketepatan Model Prediksi

Dalam analisis runtun waktu, sering kali data dibagi menjadi dua bagian yang disebut data *in sample*, yakni data-data yang digunakan untuk memilih model terbaik dengan langkah-langkah pemodelan di atas dan data *out sample*, yakni bagian data yang digunakan untuk memvalidasi keakuratan prediksi dari model terbaik yang diperoleh berdasarkan data *in sample*. Model yang baik tentunya diharapkan model terbaik untuk penyesuaian (*fitting*) data *in sample* dan sekaligus model yang baik untuk prediksi dalam data *out sample*.

Beberapa ukuran kebaikan penyuaian atau prediksi dapat dikenalkan, seperti ukuran *Mean Square Error* (MSE), *root of MSE* (RMSE), *Median atau Mean Absolute Deviation* (MAD) dan lain-lain.

Jika $Z_1, \ldots, Z_n$ menyatakan keseluruhan data, data *in sample* dapat dinyatakan sebagai $Z_1, \ldots, Z_m, m < n$. Jika nilai hasil penyuaian disebut $\hat{Z}_1, \ldots, \hat{Z}_m, m < n$, MSE, RMSE dan MAD untuk data *in sample* didefinisikan sebagai (Rosadi, 2011):

$$MSE = \frac{\sum_{i=1}^{m}(Z_i - \hat{Z}_i)}{n}, m < n \quad (15)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m}(Z_i - \hat{Z}_i)}{n}}, m < n \quad (16)$$

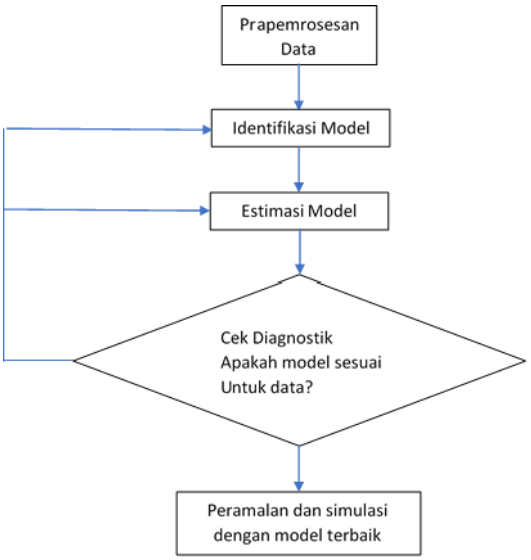
$$MAD = \frac{\sum_{i=1}^{m}|Z_i - \hat{Z}_i|}{n}, m < n \quad (17)$$

Lalu menurut Aswi dan Sukarna(2006), pengukuran ketepatan model dapat menggunakan perhitungan

Mean Absolute Percentage Error (*MAPE*) dengan formula sebagai berikut:

$$MAPE = \frac{\sum_{i=1}^{m}|Z_i - \hat{Z}_i|}{Z_i}, m < n \quad (18)$$

Gambar 1 merupakan diagram alir metodelogi pemodelan Box-Jenkins.



Gambar 1. Diagram alir metodelogi pemodelan Box-Jenkins(Rosadi, 2011).

Pandemi COVID-19

Pandemi COVID-19 yaitu masa dimana secara global terdapat penyakit menular yang disebabkan oleh virus korona yang baru ditemukan. Kebanyakan orang yang terinfeksi virus COVID-19 akan mengalami penyakit pernapasan ringan hingga sedang atau pulih tanpa memerlukan perawatan khusus. Manula dan mereka yang memiliki masalah medis mendasar seperti penyakit kardiovaskular, diabetes, penyakit pernapasan kronis dan kanker lebih berpeluang untuk terserang virus(WHO, 2020a). Pada saat ini, tidak ada vaksin atau perawatan khusus untuk COVID-19. Namun, ada banyak uji klinis yang sedang dilakukan untuk menemukan pengobatan untuk penyakit Coronavirus(WHO, 2020a). Secara global per 26 Juli 2020 terdapat 15.785.641 kasus dan 640.016 kematian(WHO, 2020b).

Data Penelitian

Data yang digunakan dalam penelitian ini adalah data harga emas dunia per Agustus 2000 hingga Juli 2020 yang diperoleh dari website Yahoo Finance. Data yang diambil merupakan data bulanan yakti harga emas dunia tiap awal bulan sehingga total banyak data yang dipergunakan adalah 241 data. Satuan data harga emas dunia adalah US$/Troy ons, dimana 1 Troy ons setara dengan 31,1035 gram.

HASIL DAN PEMBAHASAN

1. Statistika Deskriptif

Pembahasan akan diawali dengan membuat statistika deskriptif sebagai berikut:

Tabel 2. Statistika Deskriptif

Rata-rata	978,62
Standar Deviasi	468,03
Nilai Minimum	257,9
Nilai Maksimum	1900,3
Skewness	-0,14
Kurtosis	-1,29

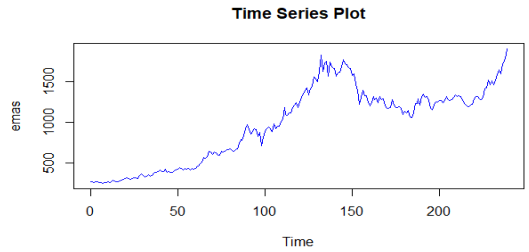
Informasi yang diperoleh dari Tabel 2 adalah rata-rata harga emas dunia adalah US$978,62/Troy ons, dimana data menyebar sebesar US$468,03/Troy ons dari rata-rata. Harga emas dunia paling rendah sebesar US$257,9/Troy ons dan paling tinggi sebesar US$1916,8/Troy ons, dengan nilai *kurtosis* sebesar -1,29 maka dapat dikatakan kurva *Platikurtik*, merupakan distribusi yang memiliki puncak hampir mendatar (nilai keruncingan < 3), serta *skewness* menunjukkan nilai negatif sebesar -0,14 atau nilai-nilai terkonsentrasi pada sisi sebelah kiri (terletak di sebelah kiri Mo), sehingga kurva memiliki ekor memanjang ke kiri, kurva menceng ke kiri atau menceng negatif untuk periode bulan Agustus 2000 hingga Juli 2020.

2. Time Series Plot

Time series plot dapat digunakan untuk melakukan perkiraan kasar dari bentuk model yang mungkin sesuai untuk

data dengan melihat plot data harga emas dunia.

Berdasarkan Gambar 2, diketahui bahwa harga emas dunia mengalami fluktuasi meskipun demikian secara garis besar terlihat bahwa harga emas dunia mengikuti pola tren naik. Hal ini menunjukkan emas sebagai investasi aman yang perlu diramalkan dengan model ARIMA.



Gambar 2. *Time Series Plot*

Metode Prediksi ARIMA

Adapun tahapan untuk mendapatkan model ARIMA terbaik yaitu: prapemrosesan data, identifikasi model, estimasi model serta pengecekan diagnostik kesesuaian model dengan data serta pemilihan model terbaik.

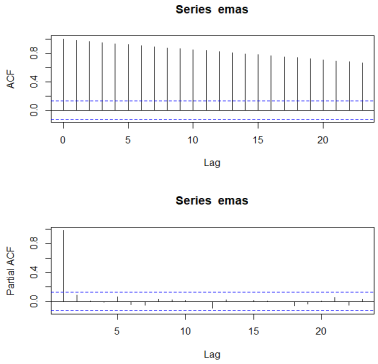
Pada Gambar 2, terlihat data mengandung tren linier, yang selanjutnya dapat dikonfirmasi dengan uji akar unit dengan uji *Augmented Dickey-Fuller/ADF* (yang menyatakan adanya akar unit) atau dengan plot ACF/PACF, seperti berikut:

Hipotesis Uji akar unit dengan uji *Augmented Dickey-Fuller/ADF*

H_0: $\rho = 0$ (terdapat akar unit)

H_0 ditolak jika τ memiliki nilai kurang dari (lebih negatif) dari $\tau_{\alpha;db}$, namun karena τ memiliki nilai lebih dari (kurang negatif) dari $\tau_{\alpha;db}$ maka H_0 diterima. Terdapat akar unit yang menandakan bahwa data belum stasioner.

Uji ADF menunjukkan bahwa

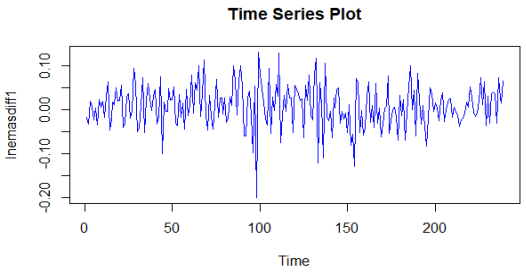


Gambar 3. *Plot ACF dan PACF data harga emas dunia (Agustus 2000-Juli 2020)*

hipotesis nol adanya akar unit dalam data (data tidak stasioner) diterima yang selanjutnya terkonfirmasi dari plot ACF yang meluruh secara lambat menuju nol.

Dengan demikian perlu dilakukan proses *differencing* dengan sebelumnya melakukan transformasi logaritma natural pada 240 data observasi harga emas dunia periode Agustus 2000 sampai dengan Juli 2020.

Gambar 4 berikut *plot* data hasil transformasi logaritma natural pada 240 data observasi harga emas dunia periode



Gambar 4. Data hasil proses *differencing* logaritma natural harga emas dunia

Tabel 3. Hasil pengujian akar unit dengan uji *Augmented Dickey-Fuller/ADF*

Data	τ	$\tau_{\alpha;db}$			**Keputusan**
		$\tau_{0,01;224}$	$\tau_{0,05;224}$	$\tau_{0,1;224}$	
Harga emas dunia (Agustus 2000-Juli 2020)	-1.7561	-3.99	-3.43	-3.13	Menerima H_0

Tabel 4. Hasil pengujian akar unit dengan uji Augmented Dickey-Fuller/ADF

Data	τ	$\tau_{\alpha;db}$			**Keputusan**
		$\tau_{0,01;224}$	$\tau_{0,05;224}$	$\tau_{0,1;224}$	
Data hasil proses *differencing* logaritma natural harga emas dunia	-5.4533	-3.99	-3.43	-3.13	Menolak H_0

Tabel 5. Hasil Penaksiran dan Pengujian Signifikansi Parameter

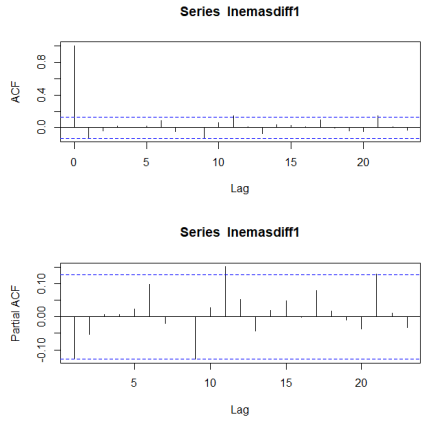
Model	Parameter	SE Parameter	t_{hitung}	t_{tabel}= $t(df = n-1; \alpha = 2{,}5\%)$	Keputusan
ARIMA (0,1,1) dengan konstanta	$\hat{\theta}_1 = -0{,}1377$	0.0662	-2,08006	-1,9810 (digunakan - t_{tabel} sebagai pembanding)	Menolak H_0
	$\mu = 0{,}0080$	0.0026	3,076923	1,9810	Menolak H_0

Pada Gambar 4 terlihat bahwa data telah memenuhi asumsi kestasioneran data. Hal ini diperkuat dengan hasil uji akar unit dengan uji *Augmented Dickey-Fuller/ADF* sebagai berikut:

Hipotesis Uji akar unit dengan uji *Augmented Dickey-Fuller/ADF*

H_0: $\rho = 0$ (terdapat akar unit)

H_0 ditolak jika τ memiliki nilai kurang dari (lebih negatif) dari $\tau_{\alpha;db}$. Nilai τ diatas memiliki nilai kurang dari (lebih negatif) dari $\tau_{\alpha;db}$ maka H_0 ditolak. Tidak terdapat akar unit yang menandakan bahwa data telah stasioner.



Gambar 5. Plot ACF dan PACF data hasil proses *differencing* logaritma natural harga emas dunia

Selanjutnya, proses identifikasi model ARIMA yang tepat untuk memodelkan data hasil proses differencing logaritma natural harga emas dunia sebagai berikut:

Pada Gambar 5(ACF) terlihat bahwa plot ACF signifikan pada lag ke-11 dan lag ke-21 dan Gambar 5(PACF) juga terlihat bahwa plot PACF signifikan pada lag ke-11 dan lag ke-21 sehingga model yang dicobakan pada data harga emas dunia adalah ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(1,1,1), ARIMA(0,1,1) dan ARIMA(0,1,2). Berdasarkan hasil estimasi parameter dan pengujian signifikansi parameter, diperoleh model dengan semua parameter yang signifikan disajikan pada tabel 5 dengan keterangan:

- df atau derajat bebas pada t_{tabel} bernilai 238 sebab n pada model ini adalah 239. Banyak data harga emas dunia yang dipergunakan adalah 240 namun karena adanya proses *differencing* berakibat banyak data (n) menjadi 239 data.
- t_{tabel} untuk $t(df = 238; \alpha = 2{,}5\%)$ tidak tertera pada tabel, namun dapat diperoleh melalui proses interpolasi.

$$I = \frac{range\ _{t\ value}}{range\ _{df}}\ (current\ df - lowest\ df)$$

Hasil interpolasi = $t\ \min - I$

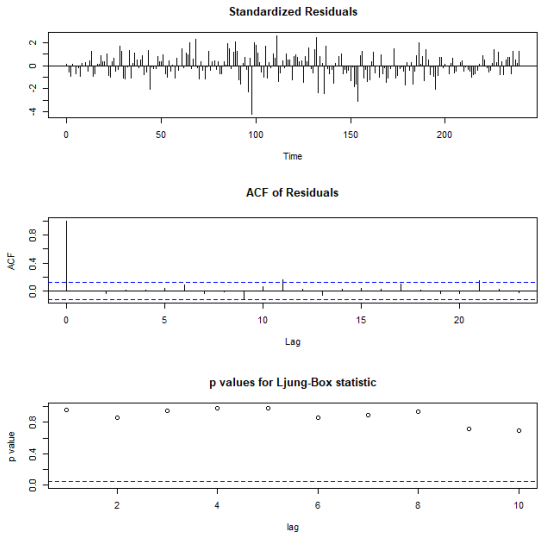
Sehingga untuk memperoleh t_{tabel} pada uji statistik diatas :

$I = \frac{t_{df=100;0,025} - t_{df=1000;0,025}}{1000-100}\ (238 - 100)$

$I = \frac{1,984-1,962}{900} \times 138 = 0,003373333$

Hasil interpolasi = 1,984 - 0,003373333 = 1,980626667 ~ 1,9810

Berdasarkan Tabel 5, selanjutnya dilakukan pemeriksaan diagnostik yang meliputi uji *white noise* dan distribusi normal pada galat.



Gambar 6. Plot ACF galat dan *p-values* untuk Ljung Box *statistic*

Pada Gambar 6, *plot* ACF menunjukkan bahwa residual model sudah memenuhi model *white noise* karena bernilai di bawah 10% (untuk *lag* lebih besar dari 1). Sedangkan nilai *p-valu*e Ljung-Box juga diatas garis batas 5%, yang menandakan hipotesis nol residual tidak mengandung korelasi serial diterima.

Pada Tabel 6 terlihat bahwa Model ARIMA(0,1,1) lulus uji pada pemeriksaan diagnostik sebab galat yang dihasilkan pada model ARIMA(0,1,1) memenuhi asumsi *white noise* dan uji normalitas data. Artinya galat dari model ARIMA(0,1,1) merupakan model yang baik dari data harga emas dunia. Selanjutnya tidak dilakukan proses pemilihan model terbaik berdasarkan hasil perhitungan *AIC* maupun SBC dikarenakan hanya ada satu model yang lulus uji pemeriksaan diagnostik. Pengukuran ketepatan model dilakukan dengan perhitungan *MAPE* dihasilkan untuk model ARIMA(0,1,1) adalah 3,70%. Perhitungan MAPE menggunakan 240 data aktual yang dipergunakan dalam membangun model serta 240 data hasil prediksi yang diperoleh dari model (data tertera pada Lampiran 1).

Tabel 6. Pemeriksaan Diagnostik

Tabel 6.	**Pemeriksaan Diagnostik**	
Model	*White noise*	Galat berdistribusi Normal
ARIMA (0,1,1) dengan konstanta	Memenuhi	Normal

Prediksi Model ARIMA

Berdasarkan Tabel 3, model ARIMA(0,1,1) dengan konstanta jika dituliskan dalam bentuk persamaan diperoleh model sebagai berikut:

$$(1 - B)\hat{Y}_t = 0,0081 - 0,1367\varepsilon_{t-1}$$

dimana $\hat{Y}_t = \ln \hat{Z}_t$

sehingga perhitungan nilai prediksi untuk Z_t adalah

$$\hat{Z}_t = \exp(\hat{Y}_t)$$

Keterangan:

$\hat{Z}_t$: prediksi harga emas dunia

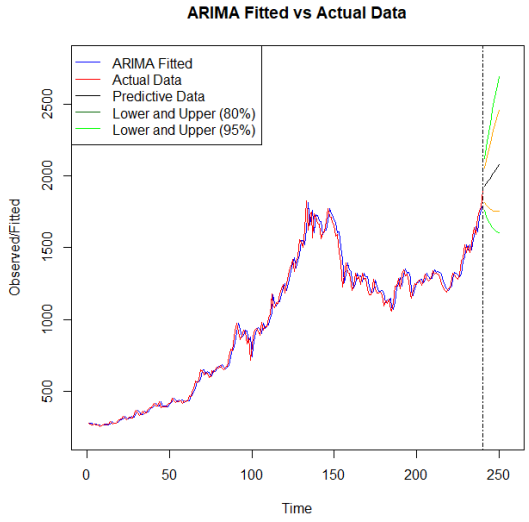
Pada Lampiran 1, ditampilkan hasil perhitungan nilai prediksi menggunakan

Tabel 7. Harga Emas Dunia (interval keyakinan 95%)

Waktu	Harga emas dunia (US$/Troy ons)		
	Prediksi	Prediksi Terendah	Prediksi Tertinggi
1 Agustus 2020	1930,046	1758,346	2118,512
1 September 2020	1945,651	1720,322	2200,491
1 Oktober 2020	1961,381	1693,192	2272,050
1 November 2020	1977,240	1672,146	2338,002
1 Desember 2020	1993,227	1655,097	2400,434
1 Januari 2021	2009,343	1640,923	2460,478

model ARIMA(0,1,1) untuk data *in sample* maupun data *out sample*.

Berikut adalah plot data harga emas dunia, nilai penyuaian *in sample*, serta nilai prediksi terendah dan prediksi tertinggi dengan interval keyakinan 95% dengan model ARIMA(0,1,1) di atas. Pada tabel 7 diperlihatkan hasil prediksi harga emas dunia per 1 Agustus 2020 hingga prediksi per 1 Januari 2021 mengalami kenaikan dengan rata-rata kenaikan sebesar 15,8594 US$/Troy ons setiap bulannya.



Gambar 6. Plot data harga emas dunia, runtun tersuai dan prediksi harga emas dunia

KESIMPULAN DAN SARAN

Kesimpulan

1. Model ARIMA (0,1,1) dengan konstanta adalah model ARIMA terbaik untuk memodelkan data aktual harga emas dunia periode Agustus 2000-Juli 2020 dimana per Januari 2020 dunia telah dilanda pandemi Covid-19 sehingga model ini dapat dipergunakan untuk prediksi harga emas dunia di masa pandemi Covid-19. Adapun MAPE dari model ini adalah sebesar 3,70%.
2. Model ARIMA (0,1,1) dengan konstanta menghasilkan prediksi harga emas dunia per Agustus 2020 hingga Januari 2021 berturut-turut sebesar 1930,046; 1945,651; 1961,381; 1977,240; 1993,227; 2009,343. Rata-rata kenaikan harga emas dunia per bulannya selama periode ini (Agustus 2020 hingga Januari 2021) diperkirakan 15,8594 US$/Troy ons emas

Saran

Berdasarkan hasil analisis menggunakan model prediksi ARIMA, saran yang diajukan peneliti untuk penelitian selanjutnya adalah:

1. Berdasarkan hasil penelitian ini diharapkan ada penelitian lanjutan yang mempergunakan model lainnya selain ARIMA yang dapat dipergunakan sebagai pembanding kesesuaian model untuk harga emas dunia.
2. Berdasarkan hasil prediksi harga emas dunia yang cenderung mengalami kenaikan (periode Agustus 2020 hingga Januari 2021), emas merupakan investasi yang aman dan menguntungkan di masa pandemi Covid-19.

DAFTAR PUSTAKA

Abdullah, L. (2012). ARIMA Model for Gold Bullion Coin Selling Prices Forecasting. *International Journal of Advances in Applied Sciences*, *1*(4). https://doi.org/10.11591/ijaas.v1i4.1495

Baker, S., Bloom, N., Davis, S., & Terry, S. (2020). COVID-Induced Economic Uncertainty. *National Bureau of Economic Research*. https://doi.org/10.3386/w26983

Bandyopadhyay, G. (2016). Gold Price Forecasting Using ARIMA Model. *Journal of Advanced Management Science*, *March*, 117–121. https://doi.org/10.12720/joams.4.2.117-121

Fauziah, A., & Surya, M. E. (2016). Peluang investasi emas jangka panjang melalui produk pembiayaan BSM cicil emas (studi pada bank syariah mandiri K.C. Purwokerto). *Islamadina: Jurnal Pemikiran Islam*, *16*(1), 57–73.

Gujarati, D. N. (2004). *Basic Econometric* (4th ed.). The McGraw−Hill Companies.

Ji, Q., Zhang, D., & Zhao, Y. (2020). Searching for safe-haven assets

during the COVID-19 pandemic. *International Review of Financial Analysis*, *71*(April), 101526. https://doi.org/10.1016/j.irfa.2020.101526

Rahmawati, Wahyuningsih, S., & Syaripuddin. (2019). Peramalan laju produksi minyak bumi provinsi kalimantan timur menggunakan metode dca dan arima. *Journal of Statistical Application and Computational Statistics*, *11 No 1*, 73–86.

Rosadi, D. (2006). *Pengantar Analisa Runtun Waktu*. Universitas Gadjah Mada.

Rosadi, D. (2011). *Analisis Ekonometrika&Runtun Waktu Terapan dengan R*. Penerbit ANDi.

WHO. (2020a). *Coronavirus*. https://www.who.int/health-topics/coronavirus#tab=tab_1

WHO. (2020b). Coronavirus disease (COVID-19): Situation Report – 188. *WHO*, *July*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

Widarjono, A. (2007). *Ekonometrika: Teori dan Aplikasi untuk Ekonomi dan Bisnis*. Ekonisia.

Yulianti, N., & Silvy, M. (2013). Sikap Pengelola Keuangan Dan Perilaku Perencanaan Investasi Keluarga Di Surabaya. *Journal of Business and Banking*, *3*(1), 57–68.

LAMPIRAN

Lampiran 1. Perbandingan Data Aktual dengan Data Prediksi Harga Emas Dunia *In Sample*

Data aktual harga emas dunia bulan Agustus 2000 - Juli 2020

[1]	278,3	273,6	264,9	270,1	272,0	265,6	266,8	257,9	264,0	265,3	270,6	266,2
[13]	274,4	292,4	279,5	273,9	278,7	282,1	296,7	302,6	308,9	326,5	313,5	303,2
[25]	312,4	323,9	318,0	316,8	347,6	368,3	350,2	335,9	339,1	364,5	346,0	354,0
[37]	375,7	385,4	384,5	396,8	415,7	402,2	396,4	427,3	387,0	394,0	392,6	391,0
[49]	410,4	418,7	428,5	451,3	437,5	421,8	436,5	428,7	435,0	416,3	435,9	429,9
[61]	433,8	469,0	465,1	494,6	517,1	570,8	561,6	581,8	651,8	642,5	613,5	634,2
[73]	625,9	598,6	604,1	646,9	635,2	652,0	669,4	663,0	680,5	661,0	648,1	666,9
[85]	673,0	742,8	792,0	782,2	834,9	922,7	972,1	916,2	862,8	887,3	926,2	913,9
[97]	829,3	874,2	716,8	816,2	883,6	927,3	941,5	922,6	890,7	978,8	927,1	953,7
[109]	951,7	1.008,0	1.039,7	1.181,1	1.095,2	1.083,0	1.118,3	1.113,3	1.180,1	1.212,2	1.245,5	1.181,7
[121]	1.248,3	1.307,8	1.357,1	1.385,0	1.421,1	1.333,8	1.409,3	1.438,9	1.556,0	1.535,9	1.502,3	1.628,3
[133]	1.828,5	1.620,4	1.724,2	1.745,5	1.565,8	1.737,8	1.709,9	1.669,3	1.663,4	1.562,6	1.603,5	1.610,5
[145]	1.684,6	1.771,1	1.717,5	1.710,9	1.674,8	1.660,6	1.577,7	1.594,8	1.472,2	1.392,6	1.223,8	1.312,4
[157]	1.396,1	1.326,5	1.323,6	1.250,6	1.201,9	1.240,1	1.321,4	1.283,4	1.295,6	1.245,6	1.321,8	1.281,3
[169]	1.285,8	1.210,5	1.171,1	1.175,2	1.183,9	1.278,5	1.212,6	1.183,1	1.183,5	1.189,4	1.172,1	1.094,9
[181]	1.131,6	1.115,5	1.141,5	1.065,8	1.060,3	1.117,3	1.233,9	1.234,2	1.289,2	1.214,8	1.318,4	1.349,0
[193]	1.306,9	1.318,8	1.271,5	1.170,8	1.150,0	1.208,6	1.252,6	1.247,3	1.266,1	1.272,0	1.240,7	1.266,6
[205]	1.316,2	1.281,5	1.267,0	1.277,9	1.306,3	1.339,0	1.315,5	1.322,8	1.316,2	1.300,1	1.251,3	1.223,7
[217]	1.200,3	1.191,5	1.212,3	1.220,2	1.284,7	1.319,7	1.312,8	1.293,0	1.282,8	1.310,2	1.409,7	1.426,1
[229]	1.519,1	1.465,7	1.511,4	1.465,6	1.520,0	1.582,9	1.642,5	1.592,4	1.713,4	1.755,1	1.781,2	1.900,3

Data peramalan harga emas dunia bulan Agustus 2000 - Juli 2020

[1]	276,7	280,5	276,7	268,6	272,1	274,2	268,9	269,2	261,5	265,8	267,5	272,3
[13]	269,2	275,9	292,4	283,5	277,4	280,8	284,2	297,3	304,3	310,7	326,9	317,9
[25]	307,6	314,3	325,2	321,5	320,0	346,4	368,1	355,5	341,3	342,1	364,2	351,3
[37]	356,5	376,0	387,2	388,0	398,8	416,7	407,4	401,1	427,0	395,4	397,4	396,4
[49]	394,9	411,5	421,1	430,9	452,0	443,0	428,1	438,8	433,5	438,3	422,6	437,5
[61]	434,4	437,4	468,3	469,3	495,0	518,1	567,8	567,0	584,4	647,2	648,3	623,2
[73]	637,8	632,6	608,0	609,5	646,8	641,9	655,8	672,9	669,7	684,5	669,5	656,3
[85]	670,8	678,1	739,4	790,9	789,7	835,2	917,5	972,2	931,2	878,9	893,3	929,0
[97]	923,3	848,4	877,6	743,0	812,2	880,4	928,1	947,2	933,4	903,7	975,9	941,2
[109]	959,6	960,5	1.009,4	1.043,8	1.170,5	1.114,2	1.096,0	1.124,2	1.123,8	1.181,6	1.217,7	1.251,6
[121]	1.200,7	1.251,6	1.310,4	1.361,4	1.392,9	1.428,6	1.357,3	1.413,3	1.446,9	1.552,9	1.550,6	1.521,0
[133]	1.626,1	1.813,7	1.659,0	1.728,9	1.757,2	1.603,7	1.732,5	1.726,8	1.690,6	1.680,5	1.591,1	1.614,7
[145]	1.624,1	1.689,6	1.773,8	1.739,0	1.728,6	1.695,7	1.678,8	1.604,1	1.608,9	1.502,3	1.418,6	1.259,0
[157]	1.315,4	1.395,9	1.346,6	1.337,4	1.272,4	1.221,1	1.247,4	1.321,5	1.299,0	1.306,5	1.263,9	1.324,3
[169]	1.297,5	1.297,8	1.232,0	1.188,8	1.186,5	1.193,8	1.276,7	1.231,1	1.199,2	1.195,2	1.199,8	1.185,3
[181]	1.115,9	1.138,5	1.127,7	1.148,8	1.085,5	1.072,3	1.119,9	1.227,3	1.243,2	1.293,1	1.235,2	1.317,1
[193]	1.355,4	1.324,1	1.330,1	1.289,7	1.196,1	1.165,5	1.212,3	1.257,0	1.258,7	1.275,3	1.282,7	1.256,4
[205]	1.275,4	1.321,1	1.297,2	1.281,4	1.288,7	1.314,4	1.346,3	1.330,3	1.334,5	1.329,3	1.314,6	1.270,0
[217]	1.239,9	1.215,4	1.204,4	1.221,0	1.230,1	1.287,3	1.325,8	1.325,2	1.307,8	1.296,6	1.318,9	1.408,1
[229]	1.435,1	1.519,4	1.484,8	1.519,9	1.484,8	1.527,3	1.587,8	1.648,0	1.612,8	1.712,9	1.763,3	1.793,1

Petunjuk Penulisan
JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

Naskah dikirim dalam bentuk *softcopy* ke alamat email pppm@stis.ac.id disertai dengan daftar riwayat hidup ringkas penulis. Format naskah mengacu pada Petunjuk Penulisan Naskah berikut:

Naskah dibuat menggunakan *Microsot Office Word* 2010. Seluruh bagian dalam naskah diketik dengan huruf *Times New Roman*, ukuran 12, spasi 1,5, ukuran kertas A4 dan marjin 2 cmuntuk semua sisi, serta jumlah halaman 15-20. Untuk kepentingan penyuntingan naskah, seluruh bagian naskah (termasuk tabel, gambar dan persamaan matematika) dibuat dalam format yang dapat disunting oleh editor.

Gaya penulisan naskah untuk Jurnal Aplikasi Statistika dan Komputasi Statistik ditulis dalam Bahasa Indonesia dengan gaya naratif. Pembabakan dibuat sederhana dan sedapat mungkin menghindari pembabakan bertingkat. Tabel dan gambar harus mencantumkan sumber jika dari data sekunder. Tabel, gambar dan persamaan matematika diberi nomor secara berurut sesuai dengan kemunculannya. Semua kutipan dan referensi dalam naskah harus tercantum dalam daftar pustaka, dan sebaliknya sumber bacaan yang tercantum dalam daftar pustaka harus ada dalam naskah. Format sumber: Nama Penulis dan Tahun. Nomor dan judul table diletakkan di bagian atas table dan dicetak tebal, sedangkan nomor dan judul gambar diletakkan di bagian bawah gambar dan dicetak tebal.

Bagian naskah berisi:

Judul. Judul tidak melebihi 12 kata dalam Bahasa Indonesia.

Data Penulis. Berisi nama lengkap semua penulis tanpa gelar, asal institusi, dan alamat email.

Abstrak. Ditulis dalam Bahasa Inggris dan Bahasa Indonesia, maksimum 100 kata untuk masing-masing abstrak dan berisikan tiga hal yaitu topik yang dibahas, metodologi yang dipergunakan dan hasil yang didapatkan.

Kata Kunci. Berisi kata atau frasa (maksimum 5 subjek) yang sering dipergunakan dalam naskah dan dianggap mewakili dan atau terkait dengan topik yang dibahas.

Pendahuluan. Memuat latar belakang, studi sebelumnya yang relevan, permasalahan ataupun hipotesis yang akan diuji dalam penelitian, ruang lingkup penelitian, serta tujuan dari penelitian.

Metodologi terdiri atas:

a. **Tinjauan Referensi.** Bagian ini menguraikan landasan konseptual dari tulisan dan berisi alasan teoritis mengapa pertanyaan penelitian dalam artikel diajukan. Di samping itu penulis dapat mengutip studi yang relevan sebelumnya untuk melengkapi justifikasi mengenai kerangka pikir penelitian.
b. **Metode Analisis.** Bagian ini berisi informasi teoritis dan teknis yang cukup memadai untuk pembaca dapat mereproduksi penelitian dengan baik termasuk di dalamnya uraian mengenai jenis dan sumber data serta variabel yang digunakan. Dalam hal keperluan verifikasi hasil, editor dan mitra bestari (*reviewer*) berhak meminta data mentah (*raw data*) yang digunakan penulis.

Hasil dan Pembahasan. Tuliskan hasil yang didapat berdasarkan metode yang digunakan disertai analisis terhadap variabel-variabelnya . Dapat disajikan berupa tabel, gambar, hasil pengujian hipotesis dengan disertai uraian analitis yang mengangkat poin-poin penting berdasarkan konsepsi teoritisnya.

Kesimpulan dan Saran. Bagian ini memuat kesimpulan dari hasil dan implikasinya secara akademis, dan saran yang dapat diberikan berdasarkan temuan dari pembahasan. Bagian ini juga memuat keterbatasan penelitian dan kemungkinan penelitian lanjutan yang dapat dilakukan dengan penggunaan/pengembangan variabel, metode analisis ataupun cakupan wilayah penelitian lainnya.

Daftar Pustaka. Daftar pustaka disusun berdasarkan urutan abjad dengan ketentuan sebagai berikut:

Publikasi Buku

1. Penulis satu orang
   Enders, Walter. 2010. *Applied Econometric Time Series, Third Edition*. New Jersey: Wiley.
2. Penulis dua orang
   Pyndick, Robert. S. dan Rubinfeld, Daniel L. 2009. *Microeconomics, Seventh Edition*. New Jersey: Pearson Education.
3. Penulis tiga orang
   Fotheringham, A. S., Brunsdon, C, dan Charlton, M. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. West Sussex: John Wiley & Sons.

Artikel dalam jurnal

Romer, P. 1993. Idea Gaps and Object Gaps in Economic Development. *Journal of Monetary Economics*, Vol. 32 (3), 543–573.

Artikel online

Woodward, Douglas P. 1992. Locational Determinants of Japanese Manufacturing Start-Ups in the United States. *Southern Economic Journal*, Vol. 58 (3), 690-708. http://www.jstor.org/discover/10.2307/1059836 (Diakses 1 Sepetember, 2014).

Buku yang ditulis oleh lembaga atau organisasi

BPS. 2009. *Analisis dan Penghitungan Tingkat Kemiskinan 2008*. Jakarta: BPS.

Kertas kerja (*working papers*)

Edwards, S. 1990. Capital Flows, Foreign Direct Investment, and Debt-Equity Swaps in Developing Countries. *NBER Working Paper*, 3497.

Makalah yang direpresentasikan

Zhang, Kevin H. 2006. Foreign Direct Investment and Economic Growth in China: A Panel Data Study for 1992-2004. *Conference of WTO, China, and Asian Economies*. Beijing.

Karya yang tidak dipublikasikan

Hartono, Djoni. 2002. Analisis Dampak Kebijakan Harga Energi terhadap Perekonomian dan Distribusi Pendapatan di DKI Jakarta: Aplikasi Model Komputasi Keseimabangan Umum (Computable General Equilibrium Model. *Tesis*. Jakarta.

Artikel di koran, majalah, dan periodik sejenis

Reuters. (2014, September 17). Where is Inflation?. *Newsweek*.