DOI: https://doi.org/10.34123/jurnalasks.v17i1.739

POLITEKNIK STATISTIKA STIS

Jurnal Aplikasi Statistika & Komputasi Statistik
Vol. 17 No. 1, June 2025
e- ISSN: 2615-1367, p-ISSN:2086-4132

Available online at https://jurnal.stis.ac.id/

Separated Couples during the COVID-19 Outbreak: A Survival Support Vector Machine Analysis

Muhammad Luthfi Setiarno Putera[1*], Rafik Patrajaya[2], Setiarno[3]

[1,2]IAIN Palangka Raya, Palangka Raya, Indonesia, [3]Universitas Palangka Raya, Palangka Raya, Indonesia
**Corresponding Author: E-mail address: m.luthfi@iain-palangkaraya.ac.id*

ARTICLE INFO

Article history:
Received 16 February, 2024
Revised 22 November, 2024
Accepted 10 December, 2024
Published 24 February, 2025

Keywords:
Cox Proportional Hazard; Feature Selection; Religious Court; Separation; Survival Support Vector Machine

Abstract

Introduction/Main Objectives: The separation between spouses has been rising noticeably in recent years in Palangka Raya, particularly during the COVID-19 outbreak. **Background Problems:** An analysis of time-to-event on those separations will be undertaken quantitatively using survival analysis by comparing the results yielded by Cox proportional hazards (PH) regression and non-parametric Survival Support Vector Machine (SUR-SVM). **Novelty:** This work suggests a feature selection method that looks for influencing elements related to the c-index by employing backward elimination. **Research Methods:** This study's data came from Indonesia's Supreme Court webpage, including a database of separation verdicts from the Palangka Raya Religious Court, spanning from April 2020 to March 2021. The response variables were the time-to-separation (marriage length until separation) (t) and the censored state of the occurrence (θ). **Finding/Results:** Based on SUR-SVM, the factors contributing the most to the separation are the absence of children, unsteady employment of appellants, and finance motive as the primary reason. In terms of concordance index and Akaike Information Criterion (AIC), the SUR-SVM outperformed the Cox proportional hazard model. These values of SUR-SVM were 59.24 and 1899.78, respectively. SUR-SVM correctly classified 59.24% of separations based on the chronological order of events.

1. Introduction

The COVID-19 outbreak spreading throughout the world has paralyzed various aspects of life, such as social, economic, health, and many more. As of November 2022, there are currently more than 6 million cases of COVID-19 in Indonesia since the emergence of the first victim in March 2020 [1]. One of the impacts felt by society as a result of the pandemic is couple's hardship [2], [3]. Such problem arises due to various factors, such as negative emotions due to financial difficulties, home confinement, and even dissatisfaction with government policies [4]. Tensions and disputes which occur, notably in vulnerable couples, will be more prone to enduring conflicts and could point to divorce or marriage separation [5], [6]. Annual report of Religious Court of Palangka Raya showed that number of marriage separation there indicated an upturn since COVID-19 outbreak in 2020.

Analysis of time-to-event situations might be undertaken quantitatively using survival analysis, one of which is marriage separation. In Indonesia, separation for Muslim couples is generally administered by a Religious Court situated in the regency or city capital [5]. Such cases could reach thousands each year. Several survival studies on marriage separation were conducted by [7] and [8].

Both found that separation was motivated by variety factors, ranging from socio-economic and cultural ones [7], [8].

Survival analysis on marriage separation is commonly dealt with censored data. The censored data is a set of values whose occurrence status is unknown since the study ended before the failure occurred [9]. As [7] and [8] revealed about separation, their dataset were classified to the right-censored like most other events. Right censorship in separation occurs as the spouses return to living together eventually or separation is not noticed by the end of the research.

Separation cases could generally be modeled using Cox proportional hazards (PH) regression [10], [11]. Beyond its common use as a conventional method of survival analysis, this method has plenty shortcoming yet. As the assumption of proportional across time for the hazard of two individuals fails to be met, Cox PH regression is challenging to employ [12]. Likewise, if covariate dependencies occur, another model is needed in the form of a non-parametric model which does not require these assumptions [13]. Non-parametric methods have been used in many techniques, including Survival Support Vector Machine (SUR-SVM).

The other study by Abdel Sater discover that numerous social and economic factors drastically influence the survival time of a marriage. With respect to survival, having children decreases marriage survival and a high level of education or a high income increases marriage survival. Applications of a survival analysis model has great potential in social science settings. The model applied Support Vector Machine as well to identify the influencing factor to the duration of marriage. The difference lies on the absence of feature selection to seek for the factor that contributes more to the duration of marriage among separated spouses[8].

A large number of couples registering for marital separation amid the COVID-19 outbreak strengthens the need to analyze the characteristics of separated Muslim couples and the influencing factors behind it. Such motives would be explained by comparing the result of two models, Cox PH and SUR-SVM. This work also suggested a feature selection method that looks for influencing elements related to the c-index by employing backward elimination. By doing so, it will enable to determine the socio-economic factors motivating the separation among Muslim couples in early period of pandemic.

2. Material and Methods

2.1. Reference Review

SUR-SVM is a machine learning method which does not require proportional hazard assumption to be fulfilled [11]. Besides, SUR-SVM is able to be conducted on high-dimensional data [14]. Because separations were observed in numerous regions and were specifically linked to a number of socioeconomic factors, it is therefore appropriate for these matters. Plenty of study demonstrated the superiority of SUR-SVM over alternative methods. Studies applying several kinds of datasets discovered that SUR-SVM performed considerably better than Cox PH [15]. Furthermore, recent work on glioma dataset suggested that SUR-SVM topped Cox PH in terms of prediction [16]. However, our motives to conduct this study based on the less number of application of SUR-SVM in social matters or particularly law matters.

2.2. Kaplan-Meier Curve and Survival Model Assumption

To assess the survival rates of marital separation, one might utilize the Kaplan-Meier curve which is related to the use of log-rank test. Another common use of the log-rank test is to see whether survival curves for different categories in a variable differ from one another [17]. The following is the hypothesis

H_0 : Survival curve categories do not significantly differ from one another

H_1 : Survival curve categories differ significantly from one another

where test statistic is expressed as

$$\chi^2 = \sum_{l=1}^{L} \frac{\left(O_l - E_l\right)^2}{E_l} \tag{1}$$

where $O_l - E_l = \sum_{k=1}^{n}(m_{lk} - e_{lk})$ and $e_{lk} = \left(\dfrac{n_{lk}}{\sum_{l=1}^{L}\sum_{k=1}^{n} n_{lk}} \right)\left(\sum_{l=1}^{L}\sum_{k=1}^{n} m_{lk} \right)$.

Description :

O_l : number of case in the l-th category

E_l : expected number of cases in the l-th category

m_{lk} : number of cases in the l-th category bearing event at time t_k

n_{lk} : number of at-risk cases bearing an instantaneous event in the l-th category before time t_k

e_{lk} : expected value in the p-th category at time t_k

l : number of categories in a variable

If $\chi^2 > \chi^2_{\alpha,(L-1)}$, then H_0 is rejected, indicating at least one difference in the survival curve for a variable [11].

Besides testing on differences among categories, the Cox PH model should meet the proportional hazard assumption, suggesting that the hazard ratio is independent of time [18]. Such test is relied on Schoenfeld error calculated by

$$SR_{mk} = z_{mk} - E\left(z_{mk} \mid R\left(t_{(mk)}\right)\right) \tag{2}$$

and the conditional probability in Equation (2) is acquired from

$$E\left(z_{mk} \mid R\left(t_{(mk)}\right)\right) = \frac{\sum_{i \in R\left(t_{(mk)}\right)} z_{mk} \exp(\beta x_i)}{\sum_{i \in R\left(t_{(mk)}\right)} \exp(\beta x_i)} \tag{3}$$

where

SR_{mk} : Schoenfeld error of the m-th predictor for cases experiencing an event at time $t_{(k)}$

Z_{mk} : the value of the m-th predictor for cases experiencing an event at time $t_{(k)}$

Following that step is to generate a rank variable f_r corresponding to its survival time. A value of 1 is assigned to the scenario where the event occurs for the first time, and so forth. Next, examine the association between the Schoenfeld error and the ranking variable f_r using the hypothesis, as follows:

H_0 : $\rho = 0$

H_1 : $\rho \neq 0$

and below is the statistic test

$$t_{\text{test}} = \frac{corr_{f_r,SR_{mk}}\sqrt{n-2}}{\sqrt{1-\left(corr_{f_r,SR_{mk}}\right)^2}} \tag{4}$$

with $corr_{f_r,SR_{mk}} = \dfrac{\operatorname{cov}(f_r, SR_{mk})}{\sqrt{\operatorname{var}(f_r)\operatorname{var}(SR_{mk})}}$. The decision is to reject H_0 if $|t_{\text{test}}| > t_{(\alpha/2,n-2)}$. Since there is a clear association between the survival time rank variable and Schoenfeld error, the proportional hazard assumption is not valid [19].

2.3. Survival Function and Hazard Function

The survival function $S(t)$ represents the likelihood that an item will endure or avoid an occurrence or failure until a specific point in time. Given that T is the length of time till an event happens,

$$S(t) = P(T > t) = \int_t^{\infty} f(u)\,du = 1 - F(t) \quad (5)$$

the function $S(t)$ is calculated by applying equation (5). The hazard function $h(t)$ calculates the probability at which the event occurs during any given time point, which is given in Equation (6)

$$h(t) = -\frac{\partial \log S(t)}{\partial t} \quad (6)$$

As time goes on, the probability of occurrence of events will be higher [20]. Besides those two, there is cumulative hazard function. Such function could be written as

$$H(t) = \int_0^t h(u)\,du \text{ or } -H(t) = \ln S(t). \quad (7)$$

The function $H(t)$ could be interpreted as cumulative amount of hazard up to time t.

2.4. Cox Proportional Hazard Model

Cox Proportional Hazard model is a way to understand how predictors influence the survival function of the event, such as marriage separation. Let Z denotes the predictors. [20] declared that Cox regression can be expressed as

$$h(t \mid Z = z) = h_0(t)\exp\left(z^{\mathrm{T}}\beta\right) \quad (8)$$

with

$h(t \mid Z = z)$	: hazard function
B	: vector of coefficients of each predictor
Z	: vector of predictor
$h_0(t)$	: baseline hazard function

2.5. Survival Support Vector Machine

Survival Support Vector Machine (SUR-SVM) is machine learning model which employs a prognostic index, unlike Cox model which relies on hazard function. The prognostic could be expressed as the probability of the couple to reconcile due to mediation or others. Refer to [11], SUR-SVM has a utility function u

$$\mathbf{u}(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\varphi}(\mathbf{x}) \quad (9)$$

where $\mathbf{w}$ is vector of parameter and $\boldsymbol{\varphi}(\mathbf{x})$ is the transformation of predictor x. Besides, SUR-SVM has an objective function, expressed in

$$\min_{\mathbf{w},\xi} \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{\gamma}{2}\sum_{j}\sum_{k,j<k} v_{jk}\xi_{jk}; \gamma \geq 0 \quad (10)$$

and the constraint function is expressed in

$$\begin{aligned} &\mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}_k) - \mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}_j) \geq 1 - \xi_{jk}; \forall j < k \\ &\xi_{jk} \geq 0; \forall j < k. \end{aligned} \quad (11)$$

Indicator v_{ij} is a comparison between i-th case and j-th case which fulfills

$$v_{jk} = \begin{cases} 1, \left(t_j < t_k, \delta_j = 1\right) \\ 0, \left(t_j < t_k, \delta_j = 0\right) \end{cases} \quad (12)$$

with ξ_{jk} in Equation (11) is the value of violations due to an error in ordering the occurence time-to-event [21].

2.6. Feature Selection and Goodness of Survival Model

Feature selection is a technique for identifying features that contain specific relevant predictors in order to produce a more favorable model. One of these techniques is backward elimination. The elimination is accomplished by modeling all predictors with Cox PH and SUR-SVM. After that, remove one of the least significant factors and regress the remaining predictors using both models. Continue the elimination process until every predictor gets an opportunity to be eliminated. The less contributing predictor could be determined by having a higher c-index [17]. The collection of predictors with the highest c-index is then used to re-model the Cox PH and SUR-SVM models. The concordance index, often known as the c-index, measures the order of the prognostic function and the observed survival time for both censored and uncensored data. A model with a higher c-index value represents a stronger survival model [11]. The Akaike Information Criterion (AIC) is another metric used to assess the goodness of a survival model. Model with lower AIC metric indicates better stability of survival model [22].

2.7. Data and Variables

This study's data came from Indonesia's Supreme Court webpage, including a database of separation verdicts from the Palangka Raya Religious Court which could be accessed in https://putusan3.mahkamahagung.go.id/pengadilan/profil/pengadilan/pa-palangkaraya.html. Spanning from April 2020 to March 2021, the time frame under observation was 1 year. There were 319 decisions regarding separation released within observed period. The analysis for regression aim took several variables into account. The response variables were the time-to-separation (marriage length until separate) (t) and censored state of the occurence (θ). Summary of variables is shown in Table 1.

Table 1. Variable summary

Variable	Explanation	Scale
Survival time (t)	Marriage length (time-to-separation), in years	Ratio
Status (θ)	State of the occurence	Nominal
	0: not denounced to separation (censored)	
	1: denounced to separation	
$z1$	Complainant's age at marriage	Ratio
$z2$	Complainant's level of school	Ordinal
$z3$	Complainant's job status	Nominal
$z4$	Appellant's age at marriage	Ratio
$z5$	Appellant's level of school	Ordinal
$z6$	Appellant's job status	Nominal
$z7$	Number of kids	Ratio
$z8$	Motive of separation	Nominal

Refer to Table 1, variable z_2 and z_5 comprise identical four categories, those are elementary school, junior high school, secondary school, and higher school. Variable z_3 and z_6 comprise identical four categories, including unskilled employees, skilled employees, semi-professionals, and working professionals. Variable z_8 consists of five categories, those are leaving duties, disputes, economic pressure, moral crises, and physical weaknesses.

2.8. Analysis Method

All categorical variables are converted to dummies using the number of categories (n_l) – 1 before being regressed. What is needed to do survival analysis on the separation dataset are explaining descriptive statistics, analyzing the curve of Kaplan-Meier and log-rank test, testing proportional hazard assumption, and analyzing survival and hazard curve. Afterwards, the data is being modeled using the Cox PH and Survival Support Vector Machine. Then, the step is selecting significant predictors from each survival method using feature selection. To seek for the better model, this work also provide goodness-of-fit based on c-index and AIC.

3. Results and Discussion

The number of separation cases filing by the Muslim couple in Palangka Raya Religious Court during April 2020 – March 2021 was 319 cases. Of all 319 occurrences, 64 were censored and 255 were uncensored. Marriage length (t) averages 10.26 years, with a median of exactly 9 years. The marriage lasted the longest 37 years, while the shortest was 0 years. A discrepancy between the mean and median revealed that the distribution of the marriage length of separated couples is asymmetric. The survival probability of the marriage is shown in Figure 1.

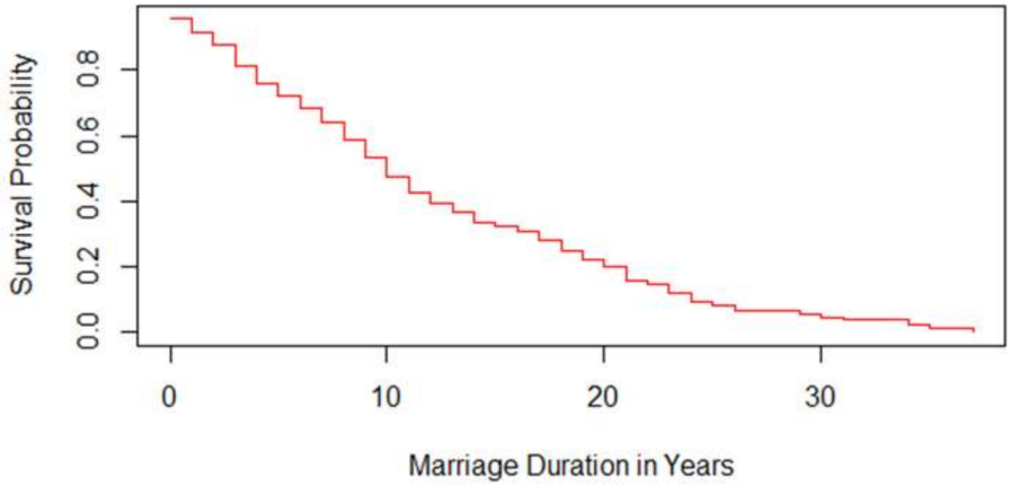

Figure 1. Survival probability of marriage of separated couples

Refer to Figure 1, the marriage length of separated couples indicated to decrease drastically at early phase of marriage within 0 – 5 years and 5 – 10 years. The survival probability was indicated steady after 25 years of marriage. It is coherent with the work of [23] noted that couple with longer marriage life was more likely to maintain their harmonious household life rather than newly-wed couple. The results of the log-rank test for each predictor are shown in Table 2, computed using Equation (1), where significant predictor is in bold.

Table 2. Summary of log-rank test

Variable	Log-rank Value	d.f	p-value
$z1$	19.1	1	0.001
$z2$	14.8	3	0.002
$z3$	12.0	3	0.007
$z4$	3.2	1	0.070
$z5$	4.0	3	0.300
$z6$	0.8	3	0.900
$z7$	65.7	1	0.001
$z8$	3.4	4	0.500

At $\alpha = 5\%$, Table 2 showed that four significant predictors, e.g. complainant's age at marriage ($z1$), complainant's education ($z2$), complainant's job status ($z3$), and number of kids ($z7$). It implied those four variables have different survival curves between groups (categories). Henceforth, complainant's age at marriage and number of kids could cause significant differences to the survival probability of marriage. It is relevant to the work of [11] which declared that some categories many socio-economic determinant have were able to differ the survival curve between categories.

There were eight variables to identify whether they all had a substantial impact on how long the problematic couples survive. The test result for the proportional hazard assumption, which was determined using Equations (2), (3), and (4), is shown in Table 3. The significant predictor at $\alpha = 5\%$ is in bold.

Table 3. Summary of proportional hazard assumption test

Variable	Correlation (ρ)	Chi-square	p-value
$z1$	0.048	0.507	0.476
$z2$	**-0.176**	**7.471**	**0.006**
$z3$	**0.145**	**5.496**	**0.019**
$z4$	-0.007	0.014	0.904
$z5$	-0.020	0.103	0.748
$z6$	-0.022	0.136	0.712
$z7$	**0.374**	**40.273**	**0.001**
$z8$	-0.003	0.002	0.961

Table 3 displays complainant's education ($z2$), complainant's job status ($z3$), and number of kids ($z7$), all in bold, with test results that contradict H_0 of Equation (4). As a result, the proportional assumption is not met. It shows a strong association between the Schoenfeld error and the survival time ranking. In particular, SUR-SVM, a substitute means which dispenses with the proportional hazard assumption, is required because all of the variables involved are necessary to represent survival time. Having been calculated by using Equations (5) and (6), Figure 2 displays the cumulative survival function and the cumulative hazard function.

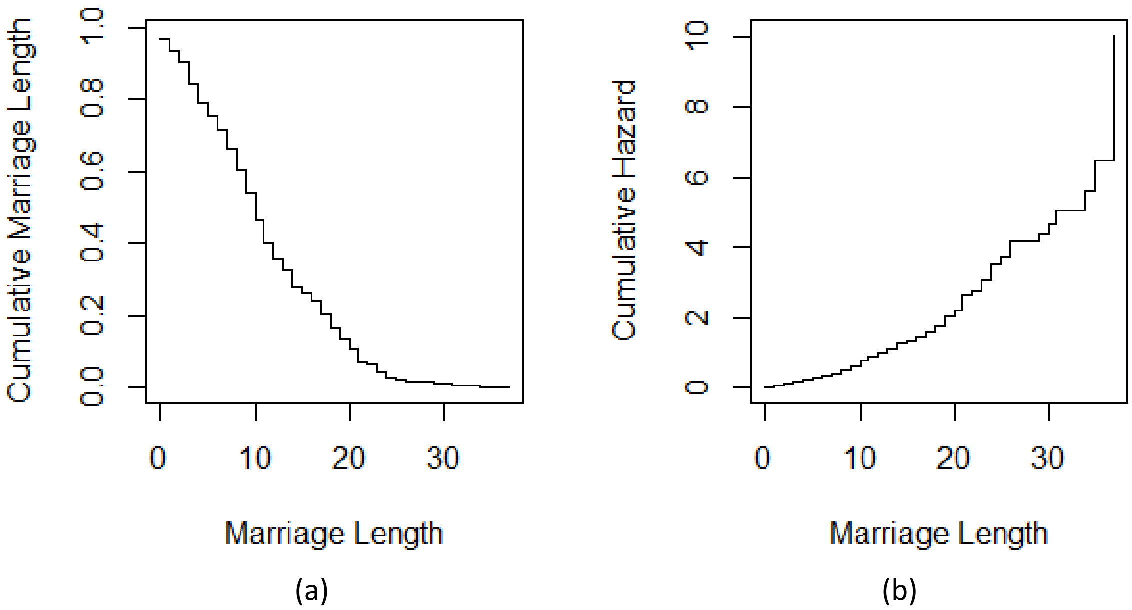

Figure 2. Cumulative survival and cumulative hazard of marriage length

Figure 2(a) shows that spouses with 0 - 10 years of marriage are more likely to be separate than they with longer marriages. The likelihood that a separated spouse will survive drops off significantly between the ages of 0 and 10 years, beyond that, it appears to be more stable. Figure 2(b) shows a trend that rises from left to right, similar to a staircase. That increase suggested that the likelihood of a pair divorcing increased with the length of the marriage. It is similar to the study of [24] which revealed that prolonged disputes and negative emotions within household life might contribute to elevated separation rate. Table 4 shows parameter estimation from the Cox proportional hazard model applied to separated couples' marital length data. The numbers which come after the dot (.) symbol in a variable, for instance $z3.1$, $z3.2$, etc., correspond to the category in a categorical variable.

Table 4. Cox model parameter estimates

Variable	Coefficient (β)	Hazard Ratio	p-value	Variable	Coefficient (β)	Hazard Ratio	p-value
*z*1	0.001	1.00	0.93	*z*5.3 (higher school)	-0.051	0.95	0.85
*z*2.1 (junior high school)	0.131	1.14	0.58	*z*6.1 (unskilled employees)	0.313	1.37	0.27
***z*2.2 (secondary school)**	**0.667**	**1.95**	**0.01**	*z*6.2 (skilled employees)	0.436	1.55	0.12
***z*2.3 (higher school)**	**0.830**	**2.29**	**0.01**	*z*6.3 (semi-professionals)	0.107	1.11	0.72
*z*3.1 (unskilled employees)	0.106	1.11	0.71	***z*7***	**-0.643**	**0.52**	**0.01**
*z*3.2 (skilled employees)	0.196	1.22	0.47	*z*8.1 (leaving duties)	-0.468	0.63	0.66
*z*3.3 (semi-professionals)	0.449	1.57	0.11	*z*8.2 (disputes)	0.331	1.39	0.75
*z*4	0.025	1.02	0.05	*z*8.3 (economy pressures)	0.136	1.14	0.89
*z*5.1 (junior high school)	0.187	1.21	0.40	*z*8.4 (moral crisis)	-0.112	0.89	0.91
*z*5.2 (secondary school)	0.023	1.02	0.91				
Likelihood ratio test		131.7		d.f. = 19	p-value = <0,01		

As shown in Table 4, the likelihood ratio test yields a test statistic of 131.7 and a p-value less than 0.01 in order to assess the importance of the parameters jointly. The choice to reject H_0 in cases where at least one variable significantly affects the separation rate was suggested by the p-value. According Table 4, the partial test on parameter found that three terms representing two variables were significant to the separation rate, shown in bold. These variables are the complainant's level of schooling, which includes secondary school and higher school, as well as the number of kids. The best Cox model based on Equation (7) is

$$h(t|Z=z)=h_0(t)\exp(0.667z_{2.2}+0.830z_{2.3}-0.643z_7)$$

A couple with more kids may have a lower hazard ratio, as indicated by the negative sign (-) for the number of kids. Table 4's hazard ratio, particularly for the major predictors, can be understood as a gauge of how these factors affect the rate of separation. For example, the number of kids hazard ratio (represented by the * symbol) is 0.52. Increasing by one kid would presumably result in a 0.52-fold decrease in the separation rate. Therefore, it is possible that couples who have more kids may have longer marriages.

Conversely, the hazard ratio is roughly 2.29 for categorical factors such as the complainant's degree from a higher school. Such figure implies that complainants who graduated higher school have separation rate 2.29 times higher than they who graduated elementary school as the reference. Hence, it can be said that the complainant’s level of school contributed statistically to the separation. The c-index of the SUR-SVM model was 58.83, and the c-index of the Cox model was 23.22, according to an algorithm that used the Kernel Radial Basis Function. After feature selection, the c-index of the SUR-SVM and Cox-PH models is shown in Table 5.

Table 5. Eliminated variables' contribution to all models

Eliminated Variable	C-index of Cox PH	C-index of SUR-SVM	Eliminated Variable	C-index of Cox PH	C-index of SUR-SVM
$z1$	23.23	57.56	$z5$	23.25	58.73
$z2$	24.31	58.63	$z6$	22.62	58.30
$z3$	22.73	58.65	$z7$	22.48	56.76
$z4$	23.64	57.66	$z8$	23.48	58.00

As seen in Table 5, the number of children ($z7$) had the largest decrease in the SUR-SVM c-index. Table 5 produces a difference of 2.07 between 58.83 and 56.76. It is thought to be the primary factor affecting the order in which the prognostic index and survival time are related. Besides number of kids, others contributing to the c-index of SUR-SVM are complainant's age at marriage ($z1$), appellant's age at marriage ($z4$), reason of separation ($z8$), and appellant's job status ($z6$). Other predictors were not considered since they had slight margin, not more than 0.5.

Additionally, Table 5 demonstrates that the number of kids had the most notable decline in the Cox model c-index. Since there is notable difference between the final c-index and the overall c-index, the number of kids has the most impact on the length of marriage. Afterwards, the selected features for SUR-SVM model are complainant's age at marriage ($z1$), appellant's age at marriage ($z4$), appellant's job status ($z6$), number of kids ($z7$), and reason of separation ($z8$), while the selected feature for Cox-PH model is complainant's job status ($z3$), appellant's job status ($z6$), and number of kids ($z7$). Table 6 shows an overview of both survival models' performance after applying the selected features.

Table 6. Performance metric of survival model

Predictors Selected	**Indices**	**Cox-PH**	**SUR-SVM**
All	C-index	23.22	58.83
	AIC	2361.13	2318.22
Feature selection	C-index	25.41	59.24
	AIC	2004.35	1899.78

According to Table 6, SUR-SVM's post-feature selection c-index is 0.41 higher than SUR-SVM's c-index across all predictors, which generated results between 59.24 and 58.83. In other side, the Cox model's c-index after feature selection is 2.19 times higher than the Cox model's c-index across all predictors. As evidenced by the greater c-index and lower AIC of SUR-SVM, it suggests that SUR-SVM is superior than the Cox model. Thus, 59.24% time-to-separation and a suitable prognostic sequence could be obtained using the enhanced SUR-SVM model. It suggested that SUR-SVM correctly classified 59.24% of separation situations as occurring in the correct order.

4. Conclusion

SUR-SVM underperforms semi-parametric Cox proportional hazard model in terms of quantifying the relationship between the predictors and survival duration of dissolved Muslim couples. The result demonstrated that the feature selection applied in both models was effective in optimizing the survival model by eliminating less contributed predictors. Various socio-economic factors substantially influenced the duration of marriage among separated Muslim couples in Palangka Raya, including the number of kids, the appellant's employment, and the reason for separation. Future studies could compare the length of marriages befor COVID-19 and its aftermath by applying machine learning techniques. The challenge is lied on the exploring best machine learning method along with best feature selection technique.

Ethics approval

The ethical guidelines were followed for conducting this investigation. Every individual participant in the study gave their informed permission.

Acknowledgments

We would like to thank the Religious Court of Palangka Raya for their invaluable support in this study. Our research findings have been substantially enhanced by the Religious Court of Palangka Raya's publication statistics and databases. All stakeholders' participation and assistance have been crucial to the accomplishment of this study.

Competing interests

All the authors declare that there are no conflicts of interest.

Funding

This study received no external funding.

Underlying data

Derived data supporting the findings of this study are available from the corresponding author on request.

Credit Authorships

Muhammad Luthfi Setiarno Putera: Methodology, Software Data Processing and Analysis, Data Visualization, Writing. **Rafik Patrajaya:** Writing, Layout and Editing. **Setiarno:** Conceptualization, Supervision, Validation.

References

[1] M. A. Widiawaty, K. C. Lam, M. Dede, and N. H. Asnawi, "Spatial differentiation and determinants of COVID-19 in Indonesia," *BMC Public Health*, vol. 22, no. 1, p. 1030, May 2022, doi: 10.1186/s12889-022-13316-4.

[2] P. R. Pietromonaco and N. C. Overall, "Implications of social isolation, separation, and loss during the COVID-19 pandemic for couples' relationships," *Current Opinion in Psychology*, vol. 43, pp. 189–194, Feb. 2022, doi: 10.1016/j.copsyc.2021.07.014.

[3] H. Prime, M. Wade, and D. T. Browne, "Risk and resilience in family well-being during the COVID-19 pandemic," *American Psychologist*, vol. 75, no. 5, pp. 631–643, 2020, doi: 10.1037/amp0000660.

[4] M. D. Puspitasari and M. Gayatri, "COVID-19 and Marital Dissolution in West Java, Indonesia," *The Family Journal*, no. 10664807221124246, 2022, doi: https://doi.org/10.1177/10664807221124246.

[5] I. Rais, "The impact of COVID-19 pandemic on divorce rates among Indonesian Muslim societies," *Indonesian Journal of Islam and Muslim Societies*, vol. 11, no. 2, 2021, doi: 10.18326/ijims.v11i2.271-297.

[6] A. Thadathil and S. Sriram, "Divorce, Families and Adolescents in India: A Review of Research," *Journal of Divorce & Remarriage*, vol. 61, no. 1, pp. 1–21, 2020, doi: https://doi.org/10.1080/10502556.2019.1586226.

[7] S. Norouzi *et al.*, "Marriage survival in new married couples: A competing risks survival analysis," *PLoS ONE*, vol. 17, no. 8, p. e0272908, 2022, doi: https://doi.org/10.1371/journal.pone.0272908.

[8] R. Abdel-Sater, "Marriage Dissolution in the United States: A Survival Analysis Approach," *Journal of Divorce & Remarriage*, vol. 63, no. 4, pp. 235–261, 2022, doi: https://doi.org/10.1080/10502556.2022.2042788.

[9] M. Nemati, J. Ansary, and N. Nemati, "Machine-Learning Approaches in COVID-19 Survival Analysis and Discharge-Time Likelihood Prediction Using Clinical Data," *Patterns*, vol. 1, no. 5, p. 100074, Aug. 2020, doi: 10.1016/j.patter.2020.100074.

[10] V. Berg, A. Miettinen, M. Jokela, and A. Rotkirch, "Shorter birth intervals between siblings are associated with increased risk of parental divorce," *PLOS ONE*, vol. 15, no. 1, p. e0228237, Jan. 2020, doi: 10.1371/journal.pone.0228237.

[11] M. L. S. Putera and S. Setiarno, "Machine learning survival analysis on couple time-to-divorce data," *Desimal: Jurnal Matematika*, vol.5, no. 3, Art. no. 3, Dec. 2022, doi: 10.24042/djm.v5i3.13742.

[12] V. Van Belle, K. Pelckmans, S. Van Huffel, and J. A. K. Suykens, "Support vector methods for survival analysis: a comparison between ranking and regression approaches," *Artificial Intelligence in Medicine*, vol. 53, no. 2, pp. 107–118, Oct. 2011, doi: 10.1016/j.artmed.2011.06.006.

[13] Md. S. Satu, K. Howlader, M. P. Hosen, N. Chowdhury, and M. A. Moni, "Identifying the Stability of Couple Relationship Applying Different Machine Learning Techniques," in *2020 11th International Conference on Electrical and Computer Engineering (ICECE)*, Dec. 2020, pp. 246–249. doi: 10.1109/ICECE51571.2020.9393131.

[14] S. Goli, H. Mahjub, J. Faradmal, and A.-R. Soltanian, "Performance Evaluation of Support Vector Regression Models for Survival Analysis: A Simulation Study," *ijacsa*, vol. 7, no. 6, 2016, doi: 10.14569/IJACSA.2016.070650.

[15] S. Lee and H. Lim, "Review of statistical methods for survival analysis using genomic data," *Genomics Inform*, vol. 17, no. 4, p. e41, Dec. 2019, doi: 10.5808/GI.2019.17.4.e41.

[16] R. Zhao, Y. Zhuge, K. Camphausen, and A. V. Krauze, "Machine learning based survival prediction in Glioma using large-scale registry data," *Health Informatics Journal*, vol. 28, no. 4, p. 14604582221135427, 2022, doi: https://doi.org/10.1177/14604582221135427.

[17] P. Wang, Y. Li, and C. K. Reddy, "Machine Learning for Survival Analysis: A Survey," *ACM Comput. Surv.*, vol. 51, no. 6, p. 110:1-110:36, Feb. 2019, doi: 10.1145/3214306.

[18] Z. Zhang, J. Reinikainen, K. A. Adeleke, M. E. Pieterse, and C. G. M. Groothuis-Oudshoorn, "Time-varying covariates and coefficients in Cox regression models," *Ann Transl Med*, vol. 6, no. 7, p. 121, Apr. 2018, doi: 10.21037/atm.2018.02.12.

[19] A. Sheng and S. K. Ghosh, "Effects of Proportional Hazard Assumption on Variable Selection Methods for Censored Data," *Statistics in Biopharmaceutical Research*, vol. 12, no. 2, pp. 199–209, Apr. 2020, doi: 10.1080/19466315.2019.1694578.

[20] D. G. Kleinbaum and M. Klein, "Kaplan-Meier Survival Curves and the Log-Rank Test," in *Survival Analysis : A Self-Learning Text, Third Edition*, in Statistics for Biology and Health. , Springer International Publishing, 2012, pp. 55–96. Accessed: Jan. 27, 2024. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4419-6646-9_2

[21] D. D. Prastyo, H. A. Khoiri, S. W. Purnami, Suhartono, S.-F. Fam, and N. Suhermi, "Survival Support Vector Machines: A Simulation Study and Its Health-Related Application," in *Supervised and Unsupervised Learning for Data Science*, M. W. Berry, A. Mohamed, and B. W. Yap, Eds., in Unsupervised and Semi-Supervised Learning. , Cham: Springer International Publishing, 2020, pp. 85–100. doi: 10.1007/978-3-030-22475-2_5.

[22] K. Burke, M. C. Jones, and A. Noufaily, "A Flexible Parametric Modelling Framework for Survival Analysis," *Journal of the Royal Statistical Society Series C: Applied Statistics*, vol. 69, no. 2, pp. 429–457, 2020, doi: https://doi.org/10.1111/rssc.12398.

[23] P. Chi, Q. Wu, H. Cao, N. Zhou, and X. Lin, "Relationship-oriented values and marital and life satisfaction among Chinese couples," *Journal of Social and Personal Relationships*, vol. 37, no. 8–9, pp. 2578–2596, Aug. 2020, doi: 10.1177/0265407520928588.

[24] D. A. Widyastari, P. Isarabhakdi, P. Vapattanawong, and M. Völker, "Marital Dissolution in Postmodern Java, Indonesia: Does Early Marriage Increase the Likelihood to Divorce?," *Journal of Divorce & Remarriage*, vol. 61, no. 8, pp. 556–573, Nov. 2020, doi: 10.1080/10502556.2020.1799308.