JUSIKOM PRIMA (Jurnal Sistem Informasi dan Ilmu Komputer Prima)
Vol. 7 No. 2, February 2024 E-ISSN: 2580-2879

APPLICATION OF THE ARIMA MODEL FOR PREDICTION OF MONTHLY DIVORCE RATE IN THE RELIGIOUS COURTS IN SOUTH SUMATRA

Sri Fitriati[1], Dian Palupi Rini[2], Firdaus[3]
[1,2,3]Sriwijaya University Palembang
srifitriaty@gmail.com

ABSTRACT

This research discusses predictions of divorce rates in Religious Courts in South Sumatra. This is important to do because the divorce rate has a changing trend. Sometimes high and sometimes low but always at a high trend number, so soCourt officials or social scientists in developing effective strategies for overcoming marriage problems, allocating resources, or supporting families who need counseling can be prepared in advance, especially at the Religious Courts in South Sumatra to reduce the divorce rate because the purpose of marriage is not to divorce. This research discusses the divorce rate in terms of predicting/forecasting because much research has been done on the divorce rate by examining the causes of divorce. This research uses the ARIMA (Autoregressive Integrated Moving Average) model to predict. The ARIMA model is a method that has been widely used in forecasting research to get good results. The research results are that the ARIMA model used is (1,0,2) and (2,0,2), with an error rate of only 0.48% using the MAPE method.

Keywords:predictions, numbers, divorce, arima, mape.

INTRODUCTION

In recent years, data analysis techniques have become increasingly relevant in various fields, including the justice system. One area where data analysis can provide valuable insight is in predicting divorce rates. Divorce breaks the marital bond between husband and wife, who can no longer live in harmony as husband and wife (Sitopu, HW, 2018).Insights into predicted divorce rates play an important role in understanding societal trends and planning future resources. South Sumatra, located in Indonesia, has a diverse population and various socio-economic factors that can influence divorce rates. By analyzing historical data from Religious Courts in South Sumatra, we can identify patterns and trends that explain the factors underlying changes in divorce rates over time. These insights can help policymakers and stakeholders to

develop targeted interventions and policies to address the leading causes of divorce in the region.

The divorce rate is an essential social indicator that reflects marital dynamics and family structure. Understanding and predicting divorce rates can help policymakers, court officials, and social scientists develop effective strategies for dealing with marital problems, allocating resources, and supporting needy families. The South Sumatra Religious Court is one of the courts that resolves divorce cases. The divorce rate in South Sumatra is not always the same, and the trend is not always decreasing or increasing. As per data from 2018-2022, in 2018, the number of divorce cases was 3168 cases; in 2019, it reached 3375 points; in 2020, it reached 2987 cases; in 2021, it reached 3378 topics; and in 2022, it reached 3431 cases. This type of data is a type of time series data. Time series analysis that has good results and is popularly used is the Auto-Regressive Integrated Moving Average (ARIMA) model (Albeladi, K., etc.2023), so this research aims to provide accurate predictions regarding the monthly divorce rate in Religious Courts in South Sumatra.

LITERATURE REVIEW

As evidence of the research that will be carried out, the study will be compared to similar research that has been carried out. The previous study that was raised was:

They are predicting Divorce prospects using Ensemble Learning: Support Vector Machine, Linear Model, and Neural Network: Mian Muhammad Sadiq Fareed, etc. 2022. *Support Vector Machine*(SVM), passive-aggressive classifier, and Linear Perceptron Model (MLP) are applied in the context of divorce prediction. Question-based data sets are created by field staff. Responses to questions provide essential data about whether a marriage will likely turn into divorce in the future. Cross-validation e-validation is applied in 5 layers, and the performance results of the evaluation metrics are confirmed. The accuracy score is 100%. The findings produce key indicators of divorce and the most significant factors when predicting divorce. Accuracy calculations use the Receiver Operating Characteristic (ROC) in this paper.

Comparison Of Fuzzy Time Series AndARIMAModel.K. Senthamarai Kannan, etc. 2019. This research compares the ARIMA model with fuzzy time series—the methodology used in the time series in Box Jenkins methodology, which is a predictable average error. We can find the average error using the formula in fuzzy time series. Comparison of actual models, ARIMA models, and Fuzzy Time series models. It is observed that the error value of the fuzzy time

series is less than the error value of the ARIMA time series. It was concluded that the fuzzy time series had better results than the ARIMA time series model and was also suitable for predicting fuel prices in India. In this paper, the calculation of prediction accuracy uses Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE) with error score values of 4%, 3%, and 2%, respectively.

*Divorce Prediction Using Machine Learning Algorithms in Ha'il Region, KSA.*Etaf Saleh Alshawarbeh, etc 2023.The 54 items in the divorce predictor scale (DPS) are used as features or attributes in the machine learning model. In addition to the DPS, a personal information form was used to collect participants' data to conduct this research in a more structured and traditional manner. Of the 148 participants, 116 were married, while 32 were divorced. With the use of Artificial Neural Network (ANN), Naïve Byes (NB), and RF algorithms, the effectiveness of DPS was tested in this research. The correlation-based feature selection method is used to identify the top 6 features from the same dataset, and the highest accuracy rate is 96.61% using MAPE (Mean Absolute % Error) calculations with ANN. The results show that DPS can predict divorce. This scale can assist counselors and family therapists in case formulation and developing family intervention plans.

METHODS

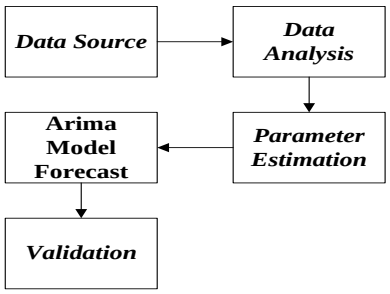

Figure 1. Research methods

Figure 1 shows the research flow used, from creating data sources, data analysis, parameter estimation, building ARIMA models, and validation.

1. Data Source

Table 1. Data on Divorce from Religious Courts in South Sumatra

	2018	2019	2020	2021	2022

January	336	443	415	395	431
February	244	301	305	393	269
March	250	286	262	321	327
April	248	263	27	257	199
May	211	183	55	191	289
June	145	228	188	416	366
July	331	348	312	197	275
August	281	258	364	296	323
September	394	388	410	294	263
October	314	312	262	233	241
November	288	245	318	271	293
December	126	120	69	114	155
AMOUNT	3168	3375	2987	3378	3431

To analyze data in predicting divorce rates in religious courts in South Sumatra in table 1, the data used is an Integer pattern taken based on data sources from first-level religious courts in South Sumatra. The number of identified cans is 60 of type Integer with various conditions.

2. Data Analysis

Time series data in ARIMA models may contain seasonal variations. Seasonal variations, or seasonality, are cycles that repeat regularly over time. The cyclical structure in a time series may or may not be seasonal. If it repeats consistently at the same frequency, it is seasonal. Otherwise, it is not seasonal and is called a cycle. Table 3.1 shows that monthly data could be more consistent, producing different figures, not in a particular season because divorce is not seasonal. Once seasonality is identified, the data can be modeled. Seasonal models can be removed from the time series. A time series in which the seasonal component has been removed is called seasonal stationary. Time series with a clear seasonal part are called non-stationary.

3. Parameter Estimation

Parameter Estimation is a branch of statistics that uses sample data to estimate distribution parameters. The technique used for parameter estimation is called an estimator. In this research, the parameter estimated is parameter d, where d is the differencing test required on the dataset. If the dataset is stationary, differencing is

unnecessary, so the d value is 0. If the dataset requires a differencing test, then the d parameter has a value of 1 or 2 (Etaf Saleh, Alshawarbeh, etc., 2023).

4. Arima Model Forecast

The Arima model consists of three parameters: p, q, and d with the format (p, d, q). First, the data is tested to be stationary. If it is stationary, then the value of d = 0; conversely, if it is not, then the value of d is an integer number from 1 and 2. The values for p and q have an integer number from 0 to 5. The models must be tested individually until they get the best model for the dataset.

5. Validation

For validation, the Arima model that has been determined must calculate the error rate to obtain forecasting accuracy. This research uses the MAPE (Mean Absolute Percentage Error) method to estimate the forecasting error rate.

RESULTS

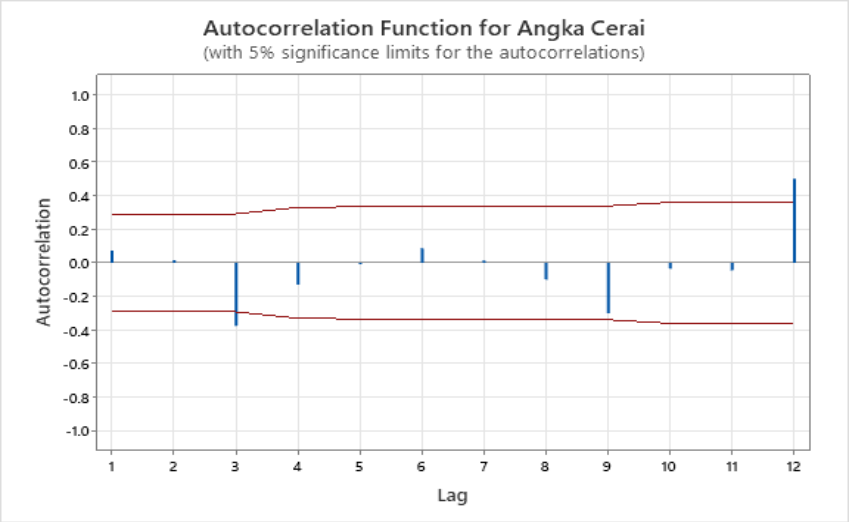

Figure 2. Autocorrelation Function (ACF) Test

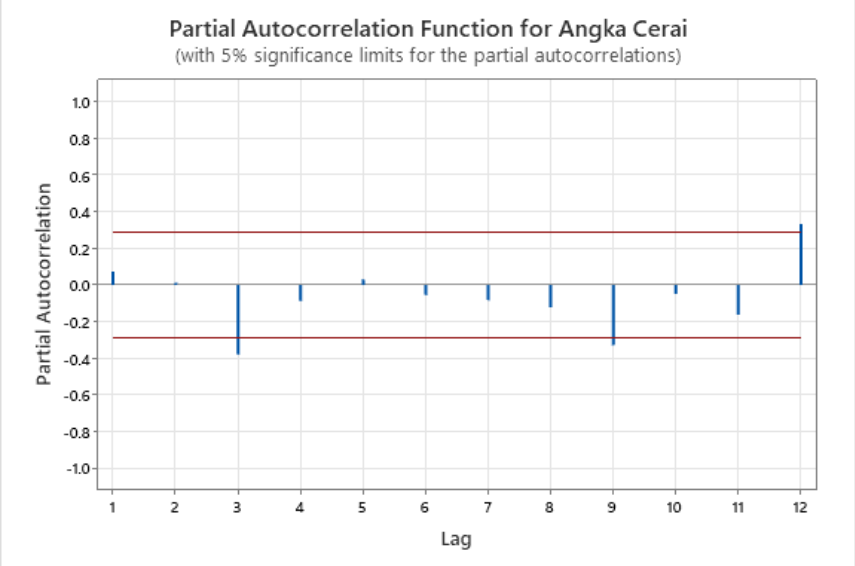

Figure 3. Partial Autocorrelation Function (PACF) Test

Figure 2 and Figure 3 explain that testing stationary data uses the ACF and PACF tests in the ARIMA model. It determines whether the ARIMA model should carry out differencing tests. The differencing test aims to make the data stationary, and the parameter d will have a value. Figure 1 and Figure 2 shows that the dataset is stationary, as evidenced by the lag number not exceeding 1 or -1 (Alsuwaylimi, A., 2022).

With the stationary dataset, the differencing test is unnecessary, so the parameter d is 0. The parameters p and q are from the formula (p,d,q). parameters p and q have integer values 0 to 5. Then model(0,0,1), model(1,0,0), model(1,0,1), model(1,0,2), model(1,0,3), model(1,0,4), model(1,0,5), model(0,0,2), model(2,0,0), model(2,0,1), model(2,0,2), model(2,0,3), model(2,0,4), model(2,0,5), model(0,0,3), model(3 ,0,0), model(3,0,1), model(3,0,2), model(3,0,3), model(3,0,4), model(3,0,5) , model(0,0,4), model(4,0,0), model(4,0,1), model(4,0,2), model(4,0,3), model(4, 0,4), model(4,0,5), model(0,0,5), model(5,0,0), model(5,0,1), model(5,0,2), model(5,0,3), model(5,0,4), model(5,0,5).

Of the 35 models created, they were tested one by one until the best models were obtained, namely models (1,0,2) and (2,0,2) because they had the slightest error, namely 0.48%. So, the accuracy obtained is 99.52%.

Table 2. Prediction of Divorce Rates in 2023

Model	Actual

(1,0,2)	(2,0,2)	2023	Month
354	354	381	January
281	281	306	February
241	241	242	March
171	171	117	April
350	350	354	May
210	210	227	June
270	270	278	July
270	270	297	August
204	204	208	September
270	270	88	October
270	270		November
270	270		December

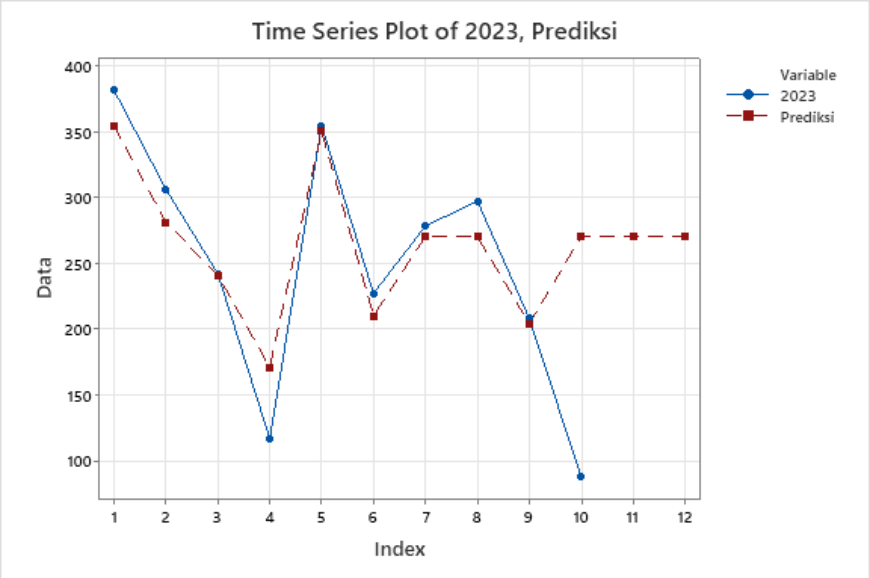

Figure 4 Predicted and Actual Graphics for 2023

Model (1,0,2) and model (2,0,2) were implemented to predict the monthly divorce rate in the South Sumatra Religious Courts, which can be seen in Table 2 and Figure 4. There are predicted numbers and actual numbers, which are close in Table 2, and Figure 4 does not show actual figures in November and December 2023 because this research was conducted in November 2023.

DISCUSSION

This research contributes in 2 (two) ways:

1. Development research in which this research takes the side of predicting/forecasting the divorce rate, where research on divorce has been widely discussed, but another study discusses divorce in terms of the causes of divorce and predictions of divorce using machine learning algorithms, not data mining like this research.
2. The South Sumatra Religious Court exists because the institution can help prepare counseling strategies or the administrative burden of registering divorce, which is a changing trend.

CONCLUSION

From the discussion in the previous chapter, conclusions can be drawn, namely: The use of the ARIMA (Autoregressive Integred Moving Average) model in calculating predictions of divorce rates produces a minor error, namely 0.48%, from the (1,0,2) and (2,0,2) models of the 35 models that have been tested. The best ARIMA model prediction data has 2 (two) models, namely (1,0,2) and (2,0,2), which can be used both to get predictions of the monthly divorce rate in 2023.

LIMITATIONS

Table 2 and Figure 4 show that in October 2023 the actual and predicted figures have a massive difference because model testing may not be optimal. The creation of 2 (two) models where there should only be 1 (one) best model. This could result in an imbalance in the divorce rates in 2021 and 2022, as shown in Table 1. So, in future research, the ARIMA model can be compared with the seasonal time-series model. For example, the Seasonal ARIMA or the Single Exponential Smoothing model can be used.

REFERENCES

Albeladi, K., etc.(2023). Time Series Forecasting using LSTM and ARIMA. International Journal of Advanced Computer Science and Applications, 14(1).313-320.

Alsuwaylimi, AA(2022). Comparison of ARIMA, ANN, and Hybrid ARIMA-ANN Models for Time Series Forecasting. Information Sciences Letters An International Journal. 2(12).1003-1016.

Alshawarbeh, E.S., Shafqat, A., Alraqad, T., Moumen, A., Saber, H., & Shafqat, R. (2023). Divorce Prediction Using Machine Learning Algorithms in Ha'il Region, KSA. 1-13.

Atashi, V., Gorji, H.T., Shahabi, S.M., Kardan, R., & Lim, Y.H. (2022). Water level forecasting using deep learning time-series analysis: A case study of the Red River of the north. Water, 14(12), 1-18.

de O. Santos Júnior DS, de Oliveira JFL, etc. (2019). An intelligent hybridization of ARIMA with machine learning models for time series forecasting.17-21.

Flores, N., & Silva, S. (2021). Machine learning model to predict the divorce of a married couple. 3C Tecnologia, 83-95.

Kannan, K.S., SulaigaBeevi, M., & Fathima, SSA (2019). A comparison of fuzzy time series and arima models. International Journal of Scientific & Technological Research, 8(8), 1872-1876.

Mehmood, M.S., Zafar, Z., Sajjad, M., Hussain, S., Zhai, S., & Qin, Y. (2022). Time series analysis and surface urban Heat Island intensity forecasting using the Arima model in Punjab, Pakistan. Land, 12(1), 2-20.

Navas Thorakkattle, M., etc. (2022). Forecasting the trends of covid-19 and the causal impact of vaccines using Bayesian structural time series and Arima. Annals of Data Science, 9(5), 1025-1047.

Nassiri, H., etc. (2023).Forecasting time trend of road traffic crashes in Iran using the macro-scale traffic flow characteristics. Heliyon, 9(3). 1-17

Sadiq Fareed, MM, etc. (2022). Predicting divorce prospects using ensemble learning: Support vector machine, linear model, and neural network. Computational Intelligence and Neuroscience. 2022. 1-15

Siami-Namini S., Tavakoli N., Siami Namin A. (2019). A Comparison of ARIMA and LSTM in Forecasting Time Series.1394-1401.

Sitopu, HW (2018). Forecasting the Number of Divorces in Medan City in 2018-2020

Torabipour, T., Siadat, S., & Taghavi, H. (2022). A Data Mining Approach using CNN and LSTM to Predict Divorce before Marriage. International Journal of Web Research, 5(1), 88-95.

Vaibhava Lakshmi R., Radha S. (2021). Time series forecasting for the Adobe software company's stock prices using the ARIMA (BOX-JENKIN') model. 81-90.