International Journal of Engineering, Science and InformationTechnology
Volume 5, No. 1 (2025) pp. 19-27
ISSN 2775-2674 (online)
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v5i1.623
Research Paper, Short Communication, Review, Technical Paper

News Popularity Prediction in West Sumatera Using Autoregressive Integrated Moving Average

Ansharulhaq Aminsyah[1]*, Nurdin[2], Zara Yunizar[1]

[1]*Department of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia*
[2]*Department of Information Technology, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia*

**Corresponding author Email: ansharulhaq.200170191@mhs.unimal.ac.id*

The manuscript was received on 17 June 2024, revised on 18 August 2024, and accepted on 28 November 2024, date of publication 1 January 2025

Abstract

The increasing public interest in reading online news is undoubtedly a challenge for news portals as online news providers. Therefore, this research was conducted to predict news popularity in West Sumatra through the FajarSumbar.com news portal using the Autoregressive Integrated Moving Average (ARIMA) model. This research aims to develop a forecasting model that can assist in estimating the popularity of each news category so that news portals can devise more effective content strategies. The data used in this study includes the number of monthly news impressions from March 2021 to June 2024, which are grouped into various categories such as Religion & Culture, Industrial Economics, Criminal Law, etc. Using the ARIMA method, which can handle time series data and overcome data non-stationarity problems through differencing and the use of grid search in optimization to find the best parameters based on the lowest evaluation metric. The results show that the ARIMA model can provide reasonably accurate predictions, although the level of accuracy varies between categories. The Mean Absolute Percentage Error (MAPE) values obtained are as follows: Religion & Culture 26%, Industrial Economy 29%, Criminal Law 29%, Health 40%, Sports 38%, Tourism & Entertainment 26%, Education 27%, Government Politics 31%, Social Environment 27%, and Technology 51%. The Technology and Health news categories show higher error rates than others, while Religion & Culture and Tourism & Entertainment have better accuracy rates. Thus, the ARIMA model can be used to predict future trends in news popularity, helping editors plan content strategies that are more relevant and interesting to readers. However, improvements are needed for news categories that have high variability.

Keywords: *ARIMA, Forecasting, News Popularity, MAE, MAPE.*

1. Introduction

Information technology is developing rapidly in the digitalization era, affecting various aspects of life, including mass media. Mass media, as a means of communication and information, is now more easily accessible to the public through print, electronic, or cyber media such as online news portals [1]. News portals are digital platforms that focus on the publication and dissemination of the latest news and are increasingly popular following the development of digital transformation [2]. This ease of access to information is also driven by the rapid growth of social media such as Instagram, Facebook, and Twitter, which has increased people's interest in online news.

The increase in online news makes news portals face the challenge of predicting reader interest and how their news can reach more people. Predicting the popularity of news is vital in helping the media identify news likely to be in demand in the future to devise better distribution strategies. Prediction or forecasting is the process of estimating future events based on historical data, the results of which can be used by companies or organizations to make strategic decisions [3].

Previous research conducted by Zhang and Lin used the Random Forest algorithm to predict news popularity. It showed that this algorithm has a high level of accuracy with an ROC value of 0.81 [4]. In addition, another study was also conducted by Syulistyo and Agustin using the Multi-Layer Perceptron algorithm and found an accuracy of 76% in predicting news popularity, compared to Random Forest, which has an accuracy of 70% [5].

This research uses the Autoregressive Integrated Moving Average (ARIMA) method to predict news popularity. ARIMA is an accurate short-term forecasting model that only uses historical data without the need for independent variables. This model was first developed by George Box and Gwilym Jenkins and is used to analyze time series data [6]. This research aims to test the performance of the

ARIMA model in predicting news popularity on the FajarSumbar.com online news portal, which is one of the primary sources of information for the people of West Sumatra.

The urgency of this research is to support the media in anticipating the type of content readers are interested in in the evolving digital era. Media such as FajarSumbar.com can make strategic decisions regarding content preparation, release, and distribution through news popularity predictions. The results of this prediction also help increase the reach of readers and strengthen their engagement with the news presented. In the current rapid information flow situation, content production efficiency is needed. By predicting news topics that have the potential to attract more readers, the media can allocate resources more optimally. In addition, these predictions allow media outlets to provide news more in line with readers' preferences, improving users' experience in accessing their news portals.

This research is expected to contribute to the media industry in West Sumatra and its surroundings in the context of digital transformation. It can provide insight and also a basis for the editorial staff as decision-makers in planning the next steps that are better in the presentation or release of news in the future and provide insight into reader preferences and interests, which journalists and media can use to present more interesting and relevant content or news.

2. Literature Review

2.1. Previous Research

Based on research conducted by Sulaeman Nurman Muhammad and Sudarmin Nusrang in applying the ARIMA method to analyze and forecast rice production in Maros Regency, South Sulawesi Province. Using rice production data sourced from the Maros Regency Statistics Agency for the period 2001 - 2018 shows that the ARIMA (0,2,1) model is suitable for forecasting the amount of rice production in Maros Regency. The forecasting results show that the rice production in Maros Regency has increased yearly with an average increase of 3807.1 tonnes [7].

The ARIMA method has also been applied by I Gede Iwan Sudipa et al., and this method effectively predicts the trend of BBCA, BBRI, and BMRI stock data. The data period used is from January 2018 to June 2023. The results showed that the ARIMA model of each stock produced a low MAPE error, with a MAPE value of 4% for BBCA, 5% for BBRI, and 7% for BMRI. The MAPE value generated by each model is included in the MAPE value with a high level of precision because it is below 10%[8].

Following the research topic, similar research has been conducted by Zhang and Lin regarding the prediction and Evaluation of decision tree algorithms such as Random Forest, C4.5, and CART. Using the online news popularity dataset obtained from the UCI Machine Learning Repository, a summary of data about articles published by Mashable within 2 years. The results of this study show that Random Forest is the best, with an ROC value of 0.81, followed by other algorithms, 0.78 for C4.5 and 0.64 for CART[4].

Similar research has also been conducted by Arie Rachmad Syulistyo et al., who analyzed the performance of Neural Network algorithms and other artificial intelligence techniques in predicting the popularity of news articles that can help the media find out whether their news will become popular. The test results show that the accuracy of the Multi-Layer Perceptron method is 76%, and Random Forest gives an accuracy of 70% [5].

Research conducted by Gusti Tasya Meilania and Lola Malihah aims to determine the characteristics of job seeker data, form a forecasting model, and forecast the number of job seekers registered in Banjar Regency using the Double Exponential Smoothing and ARIMA models. Based on the analysis conducted, the ARIMA (1,0,1) model is the best model that can be used to forecast the number of job seekers in Banjar Regency because it has the smallest MSE value where the MSE of ARIMA (1,0,1) is 7352.05 while the MSE of DES is 9070.67. From this model, the results show that in short-term forecasting, the number of job seekers in Banjar Regency is expected to increase for the following 3 periods [9].

Research by Miaomiao Wang et al. examines the application of the ARIMA model in predicting future outbreaks of chickenpox in his country, namely China. The data is historical data on the monthly incidence of chickenpox outbreaks reported in China from 2005 to 2018. Based on the results of parameter tests (all $P<0.05$) and Ljung-Box tests (all $P>0.05$), the ARIMA $(1, 1, 1)\times(0, 1, 1)12$ model was determined as the optimal model based on the coefficient of determination R2 (0.271) and standardized BIC (14.970). The values produced by the ARIMA $(1, 1, 1)\times(0, 1, 1)12$ model are very close to the observed values in 2019, with an average relative error between the actual and predicted values of 15.2%. In conclusion, the ARIMA model can be used to predict future trends in chickenpox outbreaks. This provides a scientific benchmark for chickenpox prevention and control strategies [10].

2.2. Data Mining

Data Mining is a process that uses one or more computer learning techniques to analyze and automate knowledge [11]. Through various approaches, data mining aims to extract patterns or trends within the previously unseen data, thereby providing valuable insights for decision-making[12]. Data mining combines traditional analytical methods with sophisticated algorithms to process large volumes of data, using statistical, mathematical, artificial intelligence, and machine learning techniques to extract and identify information. Several methods are often used in data mining, such as classification, clustering, association, and forecasting [13].

2.3. Forecasting

Forecasting estimates a future value based on existing data and information about past events [14]. Prediction is a process for forecasting something that may come based on past and present data information to minimize errors [15]. The results of the forecasting process can be used to estimate what will happen in the future to make the right plans and strategies to deal with future changes or situations. All forecasting has the same concept: using past data to estimate or project future data [16]. There is a wide range of forecasting methods, and the most commonly used include time series analysis, such as smoothing, linear regression, moving average (MA), autoregressive (AR), autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA); causal methods such as multiple regression analysis; and qualitative methods such as sales force polling, and executive prediction.

2.4. Autoregressive Integrated Moving Average

ARIMA is one of the time series models discovered and developed in 1970 by George Box and Gwilym Jenkins, so this method can also be called Box-Jenkins. This ARIMA model creates a short-term forecasting model that completely ignores independent variables when making forecasts using historical data. According to Singh et al., the ARIMA model gives weight to past and error values to correct model predictions, making it better than regression or other exponential methods [17].
ARIMA (p,d,q) model can be written as follows:

$$t\phi_p(B)\,(1-B)^d\,Z_t = \theta_q(B)e_t \quad \text{.....(1)}$$

Description:

2.5. News

News is a collection of text, speech, and images humans need to add to and update their insights or information [18]. Usually presented in print, broadcast, internet, or by word of mouth to a third person or crowd. In the form of a record or description of a story related to an event or incident that has just occurred, a news report, or a notification. News is categorized into several types: Straight News, Depth News, Investigation News, and Opinion News. The choice of which type to publish depends on factors including the topic's relevance, the readers' interest, and the news's actuality. News can be said to be good if the writing in the media, both print and electronic, consists of 5W + 1H elements (What, Why, Where, Who, When, and How) and contains news values or journalistic values [19].

3. Research Method

The system plan for predicting the popularity of news using ARIMA involves several stages. Firstly, historical data on the number of new visitors is collected, and various kinds of news topics will be used, namely religion & culture, industrial economy, criminal law, politics, etc., with scraping techniques through the fajarsumbar.com media admin page. The data then becomes preprocessed data. Next, hyperparameter tuning is performed for both models. ARIMA models were built and trained. The models are evaluated using test data, with Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) as evaluation metrics. The research flow is depicted in Figure 1

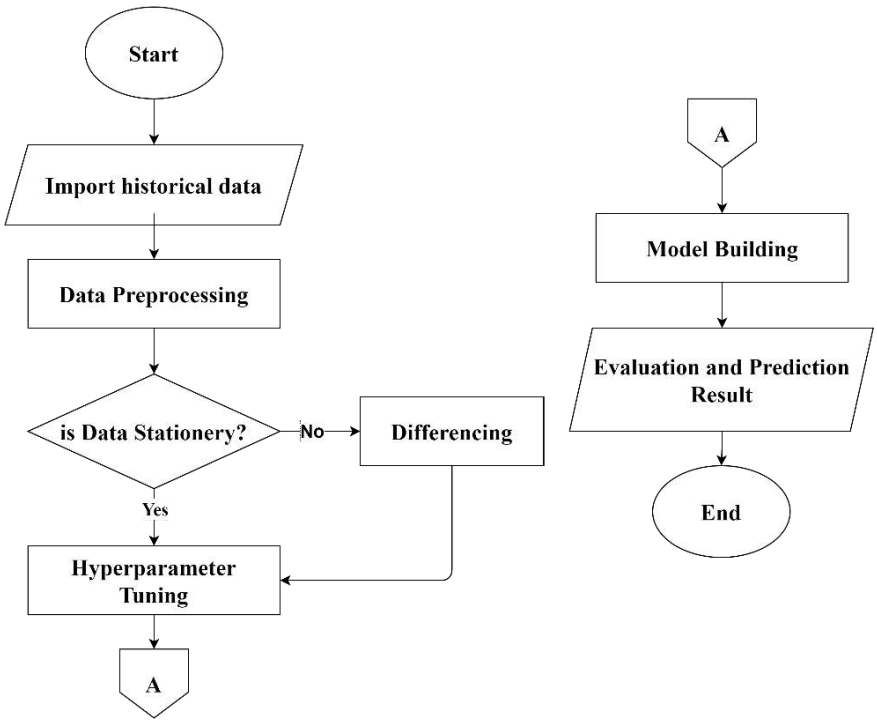

Fig 1. Research Flow Diagram

3.1. Library Import

Before model building, importing libraries needs to be done to facilitate model building; in this case, the libraries used are as follows.

1. Streamlit Library: Streamlit is an open-source library used to create interactive, simple, and fast web applications, especially applications related to data science and machine learning. This library makes it easy for users to build interfaces and interactions with users, visualize data, and display predictive models.
2. Pandas Library: is a library that provides data analysis structures and tools. Using pandas, users can easily manipulate and analyze data in tabular format. This library is used to process and manage historical data used in this research and to perform analysis and visualization of model results.
3. Numpy Library: Numpy is an essential library for Python numerical computing. It supports multidimensional arrays and provides various efficient mathematical and statistical functions. This analysis used numpy to perform statistical calculations and data manipulations needed during the model training and evaluation process.
4. Scikit-Learn library: a library that provides various functions to measure machine learning performance. This research uses sci-kit-learn to evaluate the model's prediction accuracy by measuring the model error.
5. Plotly Library: Plotly is a library for creating interactive graphs and data visualizations. It supports creating various visualizations, such as line graphs and subplots. This thesis uses a plot to display prediction results and historical data in an informative and interactive visual form.
6. Library Statsmodel: Statsmodels adalah library yang menyediakan beragam alat statistik dan econometrika untuk analisis time series atau deret waktu. Dalam penelitian ini, statsmodels digunakan untuk membangun model peramalan jumlah tayangan berita menggunakan metode Autoregressive Integrated Moving Average (ARIMA).

3.2. Data Import

The dataset or historical data is the number of new visitors on the fajarsumbar.com news portal. The dataset used in this study consists of several news topics, such as religious & cultural news topics, industrial economics, criminal law, politics, etc., from March 2021 to June 2024. This data includes the movement of news visitors for each topic with a monthly frequency. This historical data forms the basis of the analyses and predictions conducted in this study.

3.3. Data Preprocessing

Data Preprocessing is a critical stage in preparing data so that it is ready to be used for forecasting models. In this case, the stages of preprocessing are as follows:

1. Data Visualization: Visualising the data provides an overview of the movement in the number of news impressions across different categories. The following figure displays a line graph showing the monthly trend of news impressions in some of the main categories analyzed.
2. Handling Outliers: Outliers are data points that deviate significantly from the mean and have high Z-Score values. Z-Score can identify data that deviates from normal limits and handle outliers according to the needs of the analysis. Once the outliers are identified, the next step is determining whether to remove, adjust or retain them based on further study.
3. Stationary Check: Stationary check is a crucial step to ensure that the data used in the model is stationary, i.e., its statistics, such as mean or average, variance, and covariance, are constant over time. Stationarity is necessary because ARIMA models can only work well on stationary data. One of the stationary test methods is the ADF (Augmented Dickey-Fuller) Test, which provides a statistical value and p-value. If the p-value is smaller than the significance level (0.05), then the null hypothesis can be rejected, and the data can be concluded to be stationary.
4. Splitting Data: The last stage in preprocessing is separating the data into test and training data. Usually, the data is divided into a certain ratio, such as 80% (32 data) for training data and the remaining 20% (8) for test data. This division is essential to ensure that the model can be assessed or evaluated using data not used during training, thus providing a more accurate picture of the model's performance when applied to new data.

3.4. Hyperparameter Tunning

Optimization and hyperparameter tuning are essential steps to ensure the performance of the ARIMA model. In an ARIMA model, p, d, and q parameters are determined through ACF and PACF analysis with a value range of 0 to 3 [20]. These parameters control the p (autoregressive), d (differencing), and q (moving average) parts of the ARIMA model. It is crucial to correctly identify the parameters to ensure the model can effectively capture the data patterns. Grid search techniques find combinations that minimize forecasting error, explicitly targeting the lowest Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Grid search will test various combinations of parameters to identify the optimal settings to minimize MAE. By fine-tuning these parameters, the model is adapted to capture complex patterns in the data, ensuring accurate and reliable forecasts. This careful optimization process strengthens the predictive power of the model.

3.5. Building ARIMA Model

The model-building stage is crucial in transforming the processed data into actionable insights. The forecasting model is developed and refined at this stage using the optimized hyperparameters determined in the previous steps. The model is designed to predict trends, such as the popularity of news articles. One of the most suitable models for time series forecasting in this context is the Autoregressive Integrated Moving Average (ARIMA) model. ARIMA is a statistical technique that combines three main components: autoregression, differencing, and moving average. The autoregression component considers past values to predict future values, differencing helps make the data stationary by removing trends, and the moving average component captures the dependency between an observation and

the residual error of the moving average model applied to the delayed observation. By combining these elements, ARIMA can effectively model time series data, providing reliable predictions about the popularity of news articles.

3.6. Results and Model Evaluation

After training, the ARIMA model is evaluated using test data. Testing the results is intended to obtain accurate data from the built system, ensuring that the system runs according to the research objectives [21]. The Evaluation is done by calculating several metrics, including Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

MAE is a performance accuracy evaluation method that calculates the average absolute difference between actual data and predicted results [22]. The MAPE formula is presented in the equation below:

.....(2)

This metric helps assess how accurately the model predicts news popularity levels by comparing the expected results with actual data. Visualizing predictions versus actual data clearly explains the model's performance, highlighting areas where predictions may have failed.

4. Result And Discussions

4.1. Prediction Results Analysis

The previous table shows that the ARIMA model predicts news popularity with varying evaluation metric results. The following analysis further explores the pattern of data movement and the prediction results among news topics; figures 2, 3, and 4 below provide a visualization of news topic trends.

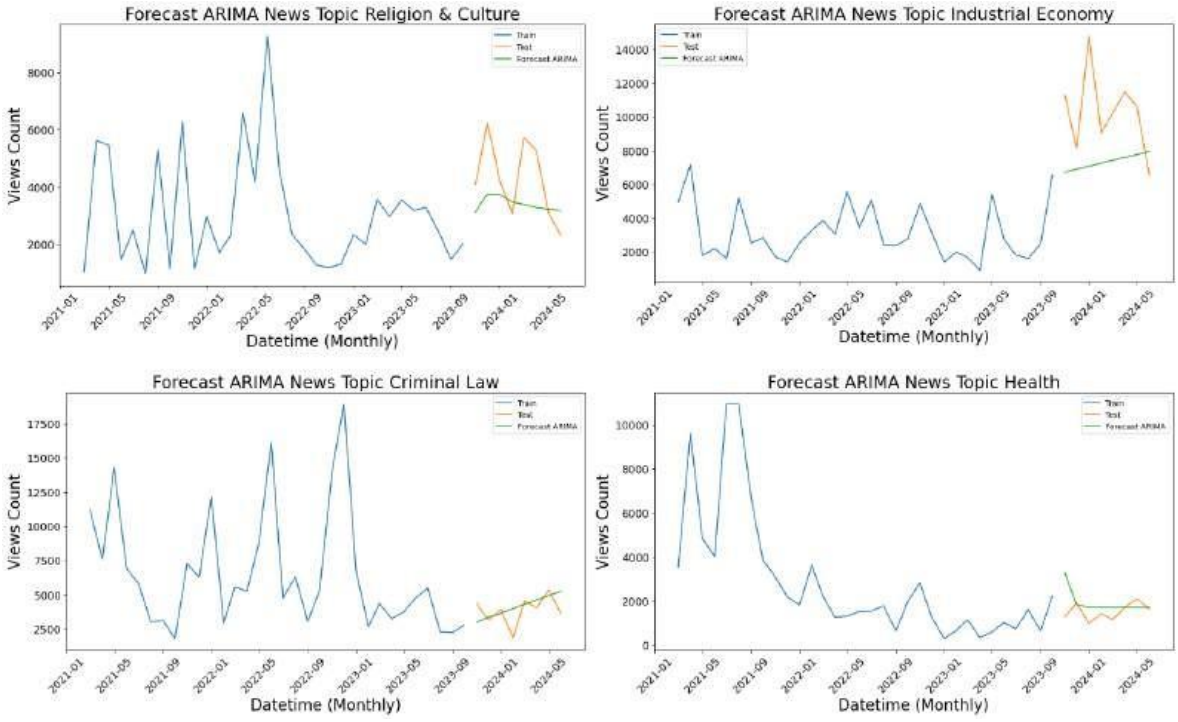

Fig 2. Graph Visualization of News Topics

Based on the 'Religion & Culture' Topic graph above it shows significant fluctuations in news popularity in the training data, especially with a high peak in popularity in mid-2022. However, the test data shows a more steady decline, with the ARIMA predictions tending to stabilize and flatten relatively in the future, suggesting a more conservative prediction for future popularity trends. Furthermore, the 'Industrial Economy' topic shows considerable fluctuations, with several peaks and dips in some periods. The test data indicates significant spikes, but the ARIMA predictions for this topic are more moderate, with a slight upward trend in the future. However, the

popularity of the news in the test data tends to drop dramatically. On the topic of 'Criminal Law,' the popularity of news related to this topic peaks in mid-2022, but the test data shows a sharp decline. The ARIMA forecast predicts a slightly increasing trend but remains conservative, indicating that the popularity of the news may stabilize and increase slightly in the future. 'Health' topics, then, showed less fluctuation than other topics. The ARIMA predictions show a stable flat trend, with little change in popularity in the future. This suggests that ARIMA predictions expect the popularity of news on this topic to remain stable without significant fluctuations. Further news topics can be seen in Figure 3 below:

Fig 3. Graph Visualization of News Topics

Based on Figure 3 above, the 'Sports' Topic graph shows that the training data for this topic is relatively stable, with minor fluctuations until early 2023. However, the test data shows a significant spike in mid-2024, and the ARIMA prediction shows a trend that will likely increase. The model predicts an increase in news popularity on sports topics, although not as high as the spike seen in the test data. Later on in the 'Education' Topic, the graph shows considerable fluctuations in the popularity of the news, with an upward trend from 2022 until mid-2023. The test data shows a sharp spike in mid-2024, but the ARIMA prediction predicts more moderate stability and a flat trend in the future. The ARIMA model does not expect a significant spike, as the test data shows.
Furthermore, on the 'Tourism & Entertainment' topic, the popularity of news in training data shows some significant fluctuations, especially in early 2021. The test data shows a declining trend, and the ARIMA predictions predict a relatively flat stability in the future without any significant changes in news popularity on this topic. Topic 'Politics': The graph shows considerable fluctuations in the training data for politics, especially with some sharp peaks and dips. The test data also shows a significant spike in mid-2024. However, the ARIMA predictions forecast a slightly increasing but more moderate trend compared to the spikes seen in the test data. For other news topics, please see Figure 4 below:

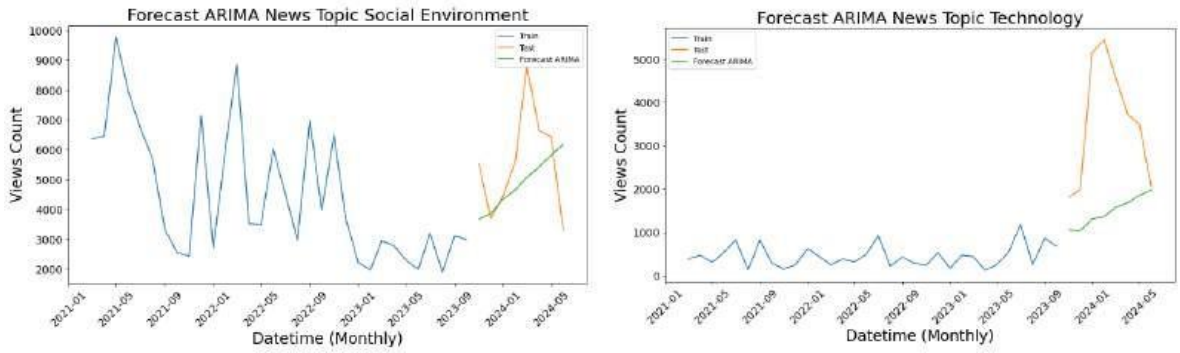

Fig 4. Graph Visualization of News Topics

Based on Figure 4 above, the 'Sosial Lingkungan' Topic shows considerable fluctuations in news popularity, with the highest peak occurring in mid-2021. After that, popularity tends to decline until early 2023. The test data shows a significant spike at the end of 2023, and the ARIMA prediction predicts a steady increase in the future. The model predicts an upward trend, although not as large as the spike seen in the test data. Finally, the 'Teknologi' Topic shows a relatively stable trend with small fluctuations until early 2023. However, the test data shows a significant spike at the end of 2023. The ARIMA predictions for this topic show a moderately increasing trend in the future, but still lower than the sharp spike seen in the test data.

4.2. Evaluation Prediction Result

In this research, problem analysis, database analysis, and implementation of the ARIMA method to predict news popularity are essential to ensure the success of this research. The following table shows the evaluation value of ARIMA prediction in predicting news popularity with various news topics, covering the period from March 2021 to June 2024 with monthly frequency. The results are summarized in Table 1.

Table 1. Prediction Result

No	New Topic	MAE	MAPE
1	Religion & Culture	1204	26%
2	Industrial Economy	3272	29%
3	Criminal Law	861	29%
4	Health	510	40%
5	Sports	3950	38%
6	Tourism & Entertainment	588	26%
7	Education	2420	27%
8	Politics	2653	31%
9	Social Environment	1445	27%
10	Technology	2044	51%

Based on the prediction results in the table above, the ARIMA algorithm shows satisfactory results in predicting the popularity of news topics. On the topic of 'Religion & Culture,' ARIMA produced an MAE of 1204 and a MAPE of 26%, while on the topic of 'Industrial Economy,' an MAE of 3272 and a MAPE of 29% were recorded. For 'Criminal Law,' an MAE of 861 and a MAPE of 29% were obtained. Some other topics, such as 'Health' had an MAE of 510 with a MAPE of 40%, and 'Sports' showed the highest MAE of 3950 with a MAPE of 38%. The topics' Tourism & Entertainment' and 'Education' obtained MAEs of 588 and 2420, with MAPEs of 26% and 27% respectively. The topic 'Politics' recorded an MAE of 2653 and a MAPE of 31%, while 'Social Environment' produced an MAE of 1445 with a MAPE of 27%. Finally, the topic 'Technology' recorded an MAE of 2044 and the highest MAPE of 51%. These results show that the error rate of ARIMA predictions varies by news topic, with MAPE ranging from 26% to 54%. While the ARIMA model provided pretty good predictions for most topics, some topics, such as 'Health' and 'Technology,' showed higher error rates, suggesting that the model's accuracy may differ depending on the topic analyzed. Next is the overall forecasting result, which can be seen in Figure 5 below:

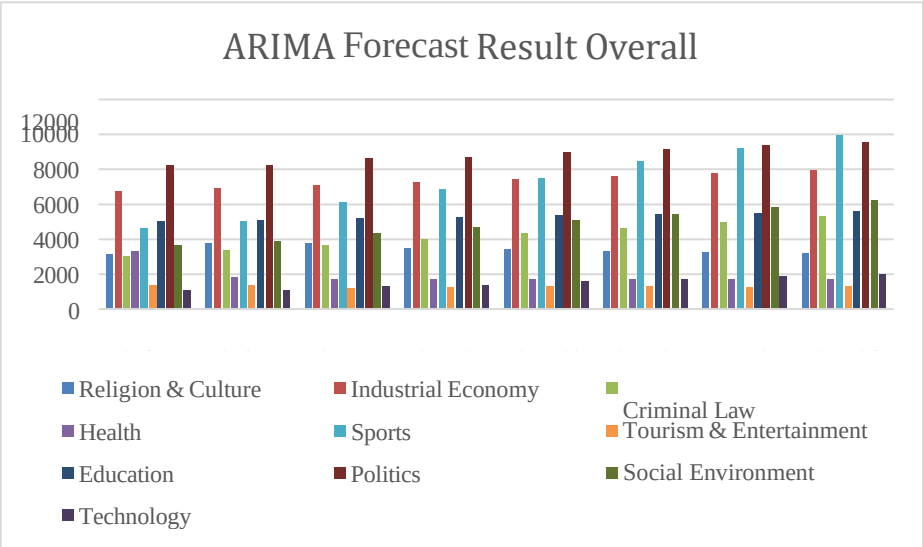

Fig 5. Forecast Result Overall

Based on Figure 5 above, it can be seen that the popularity trend of each news category varies in the period from November 2023 to June 2024. The Industrial Economy and Sports categories consistently show the highest popularity throughout the prediction period. These two categories continue to have significant spikes, especially in early 2024. This suggests that economic and sports news remained the main

focus of readers during this period. Religion & Culture, Politics, and Criminal Law also show significant trends, albeit with more variable fluctuations. The popularity of news in these categories tended to increase steadily, with spikes in certain months, indicating a continued interest in politics, religion, and law news.
Meanwhile, the categories of Health, Education, Tourism & Entertainment, and Social Environment show relatively stable trends. Health and education news tended to stabilize, with some small increases towards the middle of 2024. Technology topics remain at the lowest level among the other categories, indicating relatively less interest from readers in these topics. Overall, this graph shows that while there are variations in the level of popularity of each category, such issues as the economy, sports, and politics remain the most popular with readers. In contrast, other categories show a more stable but moderate popularity.

5. Conclusion

This research shows that the Autoregressive Integrated Moving Average (ARIMA) model can be used to predict news popularity on the FajarSumbar.com news portal with different levels of accuracy for each news topic. Evaluation results using Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) show that the ARIMA model is entirely accurate in predicting popularity with the results of prediction evaluation metrics from each news category producing varying Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) values, namely in the Religion & Culture news category has MAE and MAPE of 1204 and 26%, Industrial Economy by 3272 and 29%, Criminal Law by 861 and 29%, Health by 510 and 40%, Sports by 3950 and 38%, Tourism & Entertainment by 588 and 26%, Education by 2419 and 27%, Politics by 2653 and 31%, Social Environment by 1445 and 27% and Technology news category with MAE by 2044 and MAPE 51%. This ARIMA model can be considered successful in predicting the popularity of news on the fajarsumbar.com news portal with varying degrees of accuracy. Overall, news media editors can use the predictions generated to develop more effective content strategies based on reader preferences. Although ARIMA provides good results, combining ARIMA models with other prediction models, such as Random Forest or Neural Network, can improve accuracy in news categories with more complex patterns.

References

[1] D. Kusuma Habibie and M. Administrasi Publik Universitas Gadjah Mada Jl Sardjito, “DWI FUNGSI MEDIA MASSA,” 2018.
[2] S. Wahyuni, “PERBANDINGAN EFEKTIVITAS TEMU KEMBALI INFORMASI PADA PORTAL BERITA ONLINE DI YOGYAKARTA (HARIAN JOGJA DAN TRIBUN JOGJA).”
[3] M. Y. Darsyah, “Peramalan Pola Data Musiman Dengan Model Winter’s & ARIMA,” 2015.
[4] Y. Zhang and K. Lin, "Predicting and Evaluating the Online News Popularity based on Random Forest," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Aug. 2021. doi: 10.1088/1742-6596/1994/1/012040.
[5] A. Rachmad Syulistyo and V. Meliana Agustin, "Predicting News Article Popularity with Multi Layer Perceptron Algorithm," 2022. [Online]. Available: https://www.kaggle.com/waseemakramkhan/the-tribune-news-articles.
[6] S. P. Fauzani and D. Rahmi, “Penerapan Metode ARIMA Dalam Peramalan Harga Produksi Karet di Provinsi Riau,” *Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT)*, vol. 2, no. 4, pp. 269–277, 2023.
[7] S. Nurman, M. Nusrang, and Sudarmin, "Analysis of Rice Production Forecast in Maros District Using the Box-Jenkins Method with the ARIMA Model," *ARRUS Journal of Mathematics and Applied Science*, vol. 2, no. 1, pp. 36–48, Feb. 2022, doi: 10.35877/mathscience731.
[8] I. G. I. Sudipa, R. Riana, I. N. T. A. Putra, C. P. Yanti, and M. D. W. Aristana, “Trend Forecasting of the Top 3 Indonesian Bank Stocks Using the ARIMA Method,” *SinkrOn*, vol. 8, no. 3, pp. 1883–1893, Jul. 2023, doi: 10.33395/sinkron.v8i3.12773.
[9] L. M. Malihah and G. T. Meilania, “PERBANDINGAN MODEL PERAMALAN JUMLAH PENCARI KERJA MENGGUNAKAN ARIMA DAN DOUBLE EXPONENTIAL SMOOTHING,” *Jurnal Litbang Sukowati : Media Penelitian dan Pengembangan*, vol. 7, no. 2, pp. 169–178, Nov. 2023, doi: 10.32630/sukowati.v7i2.441.
[10] M. Wang *et al.*, "Methods and Applications An Autoregressive Integrated Moving Average Model for Predicting Varicella Outbreaks-China," 2019. [Online]. Available: https://weekly.chinacdc.cn/
[11] M. Maulita, “PENDEKATAN DATA MINING UNTUK ANALISA CURAH HUJAN MENGGUNAKAN METODE REGRESI LINEAR BERGANDA (STUDI KASUS: KABUPATEN ACEH UTARA),” 2023. [Online]. Available: http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/indexMayaMaulita|http://jom.fti.budiluhur.ac.id/index.php/IDEALIS/index|
[12] B. D. Samudera, N. Nurdin, and H. A. K. Aidilof, "Sentiment Analysis of User Reviews on BSI Mobile and Action Mobile Applications on the Google Play Store Using Multinomial Naive Bayes Algorithm," *International Journal of Engineering, Science and Information Technology*, vol. 4, no. 4, pp. 101–112, Oct. 2024, doi: 10.52088/ijesty.v4i4.581.
[13] M. Fikry, D. Hamdhana, and M. Qamal, "Data Mining for Processing of Research and Community Service by Lecturer Using Decision Tree Method".
[14] S. Wardah, “ANALISIS PERAMALAN PENJUALAN PRODUK KERIPIK PISANG KEMASAN BUNGKUS (Studi Kasus : Home Industry Arwana Food Tembilahan),” 2016.
[15] N. Nurdin, F. Fajriana, M. Maryana, and A. Zanati, “Information System for Predicting Fisheries Outcomes Using Regression Algorithm Multiple Linear,” *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 5, no. 2, pp. 247–258, Jan. 2022, doi: 10.31289/jite.v5i2.6023.
[16] N. Luh, W. Sri, R. Ginantra, I. Bagus, and G. Anandita, “Penerapan Metode Single Exponential Smoothing Dalam Peramalan Penjualan Barang,” 2019. [Online]. Available: http://tunasbangsa.ac.id/ejurnal/index.php/jsakti
[17] R. K. Singh *et al.*, "Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated

moving average (ARIMA) model," *JMIR Public Health Surveill*, vol. 6, no. 2, Apr. 2020, doi: 10.2196/19115.
[18] U. Rahardja, N. Lutfiani, and R. Rahmawati, “Persepsi Mahasiswa Terhadap Berita Pada Website APTISI Student Perception to the News on The APTISI Website,” 2018. [Online]. Available: http://aptisi.or.id/,
[19] J. Pendidikan and D. Konseling, “Dasar Dasar Penulisan Berita.”
[20] A. Prasetyo, N. Nurdin, and H. A. K. Aidilof, "Comparison of Triple Exponential Smoothing and ARIMA in Predicting Cryptocurrency Prices," *International Journal of Engineering, Science and Information Technology*, vol. 4, no. 4, pp. 63–71, Oct. 2024, doi: 10.52088/ijesty.v4i4.577.
[21] M. Faisal, Nurdin, Fajriana, and Z. Fitri, "Information and Communication Technology Competencies Clustering for students for Vocational High School Students Using K-Means Clustering Algorithm", doi: 10.52088/ijesty.v1i4.318.
[22] A. A. Suryanto, A. Muqtadir, and S. Artikel, “PENERAPAN METODE MEAN ABSOLUTE ERROR (MEA) DALAM ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PRODUKSI PADI Info Artikel : ABSTRAK,” no. 1, p. 11, 2019.
[23] A. N. M. F. Faisal, A. Rahman, M. T. M. Habib, A. H. Siddique, M. Hasan, and M. M. Khan, "Neural networks based multivariate time series forecasting of solar radiation using meteorological data of different cities of Bangladesh," *Results in Engineering*, vol. 13, Mar. 2022, doi: 10.1016/j.rineng.2022.100365.