

Analysis and Implementations of Support Vector Machine for Predicting the Number of Ship Passengers at PT Pelni Medan Branch

Muhammad Aryo^{1*}, Wilda Rina Hasibuan²

^{1,2}Universitas Muhammadiyah Sumatera Utara
Muhhammadaryo736@gmail.com^{1*}

Abstract

This study examines the application of the Support Vector Machine (SVM) algorithm to predict the number of ship passengers at PT Pelni Medan Branch. The main problem faced by the company is the difficulty in forecasting passenger numbers due to fluctuations influenced by seasonal factors, economic conditions, and special events. Until now, the company has not had a reliable data-based prediction system, so operational decisions are still made manually, which often leads to capacity imbalances. The SVM algorithm was chosen for its ability to perform both linear and non-linear predictions by utilizing a hyperplane as a class separator. Research data were obtained from historical passenger records, which were processed through preprocessing, splitting into training and testing sets, and SVM model training. The implementation was realized through a web-based application using PHP and MySQL to make prediction results easily accessible to the company. The findings show that the SVM model can classify passenger numbers into three categories: low, medium, and high. The system was tested using the Blackbox Testing method and declared valid for all tested functions. Thus, this research contributes to supporting operational decision-making and planning at PT Pelni Medan Branch, although it is still limited by the number of variables used.

Keywords: Support Vector Machine, Prediction, Number of Passengers, PT Pelni, Machine Learning, Information System

1. Introduction

Indonesia is an archipelagic country with more than 17,000 islands, making sea transportation a vital means of connecting different regions. Transportation serves as a fundamental need for society to support economic, social, and mobility activities. Over time, transportation facilities have continued to develop, and ships remain one of the most preferred modes. At PT Pelni Medan Branch, the main challenge lies in predicting passenger numbers due to fluctuations influenced by seasonal factors, economic conditions, and special events. Until now, the company has not had an accurate data-driven prediction system, so operational decisions are still made manually, creating risks of capacity imbalance. Furthermore, this study is limited by the number of variables used, resulting in predictions that only provide categorical outputs (low, medium, high) rather than exact numerical values.

To address this issue, PT Pelni Medan Branch requires a passenger prediction system. One of the approaches applied is a prediction technique based on machine learning, namely the Support Vector Machine (SVM). SVM is a method that can be used for both classification and regression tasks. This method is capable of performing linear and non-linear predictions by determining a hyperplane that maximizes the margin between data classes. The primary goal of SVM is to construct an effective computational learning model that separates data within high-dimensional feature spaces. In this way, SVM can utilize patterns influenced by seasonal factors, economic conditions, and special events, making it a suitable algorithm for identifying trends in predicting passenger numbers.

Several studies have utilized Support Vector Machine (SVM) in various research contexts. One such study was conducted by Drajat Indra Purnama and Oki Prasetia Hendarsin, who applied Support Vector Regression (SVR) to forecast the number of departing air passengers in Central Sulawesi. The findings revealed that SVR with an RBF kernel optimized through grid search was able to address nonlinear data patterns with very high accuracy, achieving a MAPE of 7.28% on the training data and 18.67% on the testing data, thereby outperforming classical methods such as ARIMA.

2. Theoretical Basis

This chapter discusses the scientific theories underlying the problem being researched, which consist of general theories and specific theories.

2.1. Previous Research

Table 1: Previous Research

No	researcher	Research Title	Research result
1	Drajat Indra Purnama, Oki Prasetya Hendarsin	Forecasting the Number of Departing Air Passengers in Central Sulawesi Using Support Vector Regression (SVR)	The study revealed that SVR with an RBF kernel optimized through grid search was able to handle nonlinear data patterns with high accuracy, achieving a MAPE of 7.28% on the training data and 18.67% on the testing data, thus outperforming classical methods such as ARIMA. [1]
2	Benni Agung Nugroho, Andika Kurnia Adi Pradana, Ellya Nurfarida	Prediction of Customer Service Arrival Time for Motor Vehicles Based on Historical Data Using Support Vector Machine	The findings reveal that employing the SVM method with a linear kernel achieved 92.5% accuracy on the training dataset, while on the testing dataset it produced an average performance of 97.33% across accuracy, precision, and recall. [2]
3	Drajat Indra Purnama, Siti Setianingsih	Support Vector Regression (SVR) Model for Forecasting Number of Passengers on Domestic Flights at Sultan Hasanudin Airport Makassar	The study revealed that the data on domestic flight passengers at Sultan Hasanudin Airport exhibited a nonlinear pattern. The best-performing model was SVR with an RBF kernel optimized through grid search. The comparison results indicated that SVR provided higher accuracy than ARIMA, as shown by lower MAPE values on both training data (7.57% vs 9.05%) and testing data (11.38% vs 12.32%). Therefore, SVR proved to be more effective in modeling and forecasting passenger numbers compared to the classical ARIMA method. [3]

2.2. Predictions

Prediction is a systematic process of estimating future events based on information obtained from the past, with the aim of minimizing errors (the difference between actual outcomes and expected results). Prediction does not necessarily provide definite answers but rather serves as an effort to obtain results that are as close as possible to what may occur. [4] In a business context, prediction or forecasting refers to the activity of estimating future sales, including projected sales volume, market potential, and market share. Moreover, prediction plays an important role in helping companies plan their inventory needs, as it provides optimal insights that enable organizations to minimize the risk of planning errors to the lowest possible level. [5]

2.3. Support Vector Machine

Support Vector Machine (SVM) is a machine learning algorithm used for classification and regression tasks by identifying the optimal hyperplane that separates data with the maximum margin. By applying the principle of Structural Risk Minimization, SVM can efficiently and accurately handle high-dimensional data. The fundamental concept of SVM is to find a hyperplane that optimally separates two classes, meaning the boundary provides the widest possible margin between them. [6]

3. Research Methodology

This section outlines the methods, steps, and techniques employed in conducting the research to ensure the thesis objectives are achieved systematically.

3.1. Research Type

This study employs a quantitative approach by statistically processing numerical data obtained from observations, documentation, and interviews to train a prediction model based on Support Vector Machine (SVM). The focus is not on developing a new algorithm but on applying SVM to predict ship passenger numbers at PT Pelni Medan Branch using operational data (year, month, ship name, and boarding passengers). The outcomes include both an analysis of SVM effectiveness and a web-based prototype system built with PHP and MySQL, capable of predicting passenger categories (low, medium, high). Thus, this applied research provides practical outputs in the form of a prediction system that supports the company's decision-making process.

3.2. Research Flow

The research flow implemented in this study describes the systematic stages carried out by the researcher, starting from data collection to the evaluation of results. In general, the research flow can be illustrated as follow.

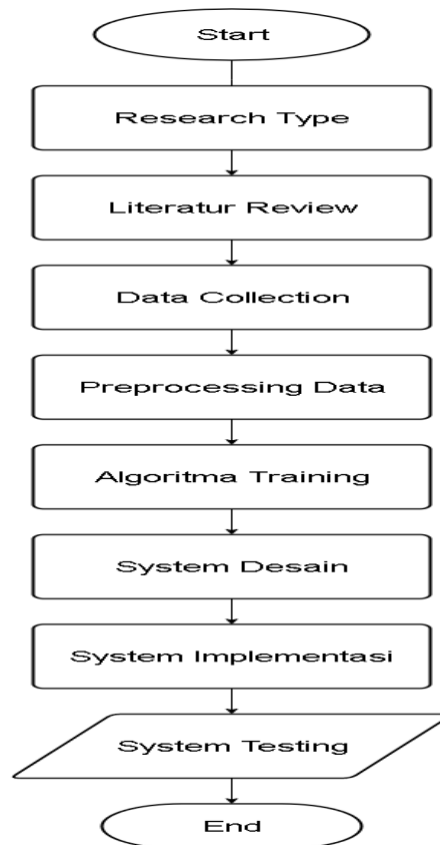


Fig. 1: Research Flow Chart

3.3. Literature Review

Gathering theories, concepts, and previous studies as a foundation and reference for developing the research framework.

3.4. Data Collection

Collecting and reviewing historical data on the number of ship passengers at PT Pelni Medan Branch to serve as the basis for analysis in this study.

3.5. Data Preprocessing

Conducting initial data processing, including data cleaning, normalization, creation of supporting variables, and splitting the dataset for training and testing purposes.

3.6. Algorithm Training

Implementing the Support Vector Machine (SVM) method to construct and develop predictive and classification models using the data that has undergone prior processing, in order to generate more accurate and systematic prediction outcomes.

3.7. System Design

Developing and designing a comprehensive system layout, which includes the system architecture, operational process flow, and user interface design, serving as a crucial foundation for the implementation and operation of the developed model.

3.8. System Implementation

Developing a system or prototype that effectively integrates the constructed algorithmic model with the company's data, enabling the system to be directly utilized for supporting decision-making, analysis, and predictive operations within the company's operational context.

3.9. System Testing

Conducting comprehensive testing on the accuracy of the model and the performance of the developed system, with the aim of ensuring that the results meet the expected quality standards and effectively fulfill the users' needs and expectations.

4. Results and Discussion

The results of the analysis and implementation of the Support Vector Machine method for predicting ship passenger numbers based on the operational data of PT Pelni Medan Branch, developed using PHP as the programming language and MySQL as the database server, are presented as follows:

1. Form Login

The login form resulting from the analysis and implementation of the Support Vector Machine for predicting ship passenger numbers based on the operational data of PT Pelni Medan Branch is shown in Fig. 2

Fig. 2: Form Login

2. Form Home

The Home Form from the analysis and implementation of the Support Vector Machine for predicting ship passenger numbers based on the operational data of PT Pelni Medan Branch displays the main menu page, which consists of Home, Passenger Data, SVM, and Exit, as shown in Fig. 3

Fig. 3: Form Home

3. Form Passenger Data

The Passenger Data Form from the analysis and implementation of the Support Vector Machine for predicting ship passenger numbers based on the operational data of PT Pelni Medan Branch displays the passenger data management page, as shown in Fig. 4

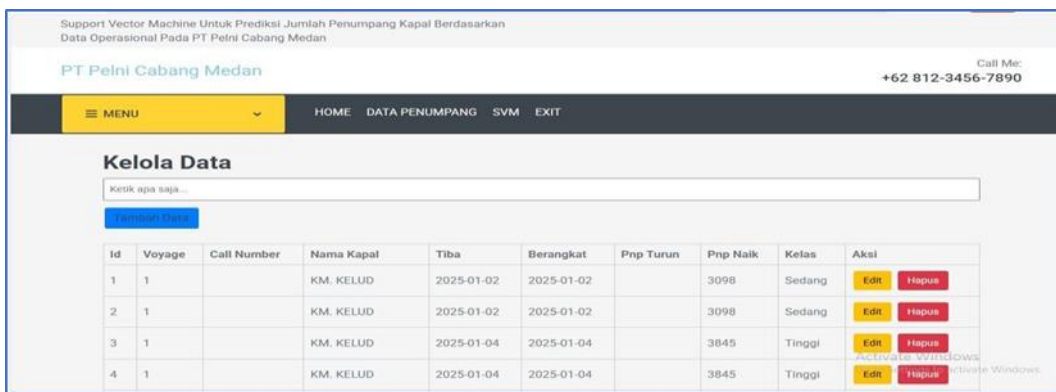


Fig. 4: Form Passenger Data

4. Form SVM

The SVM Form from the analysis and implementation of the Support Vector Machine for predicting ship passenger numbers based on the operational data of PT Pelni displays the class prediction page. This page includes several input variables such as year, month, ship name, and embarked passengers, which are used to predict the total number of passengers, as shown in Fig. 5

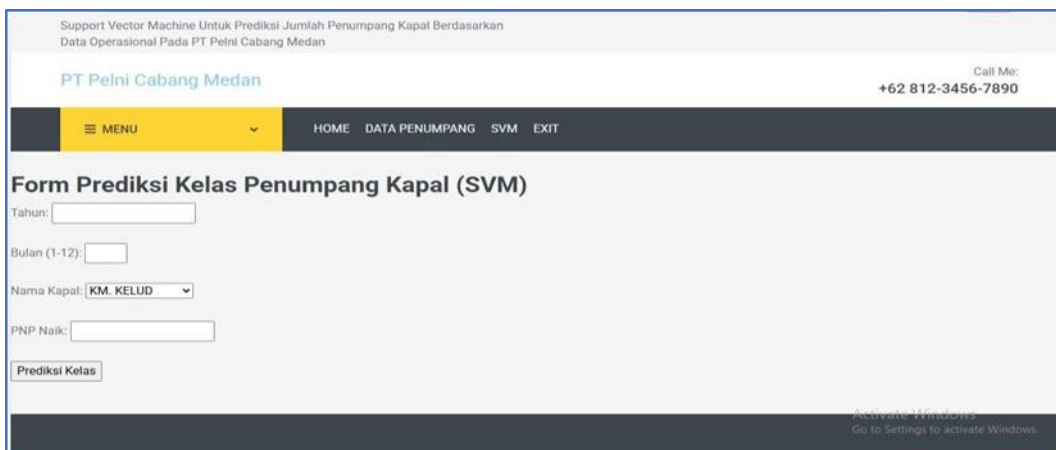


Fig. 5: Form SVM

In the next stage, after all the variables have been entered and the class prediction process is executed, the prediction results are displayed as shown in Figure 6.



Fig. 6: Form SVM

5. Conclusion

The results of this study indicate that the application of SVM can classify the number of ship passengers into three categories, namely low, medium, and high. The predictions produced are categorical in nature and therefore do not provide exact numerical values of passengers. Furthermore, the model relies on a simple linear approach based on the maximum and minimum values in each class, making it closer to a heuristic method rather than a full mathematical implementation of SVM, which is also due to the limited availability of data.

Acknowledgement

I would like to express my deepest gratitude to my parents who have continuously supported me throughout the completion of this research. I am also sincerely thankful to my supervisor, Mrs. Wilda Rina HSB, S.T., M.Kom, who has patiently and wholeheartedly guided me until this study was successfully completed.

References

- [1] Drajat Indra Purnama, Oki Prasetia Hendarsin "Peramalan Jumlah Penumpang Berangkat Melalui Transportasi Udara di Sulawesi Tengah Menggunakan Support Vector Regression (SVR)," *JAMBURA JOURNAL OF MATHEMATICS*, Vols. Vol. 2, No. 2, pp. 49-59, July 2020, DOI: <https://doi.org/10.34312/jjom.v2i2.4458>.
- [2] Benni Agung Nugroho, Andika Kurnia Adi Pradana, Ellya Nurfarida "Prediksi Waktu Kedatangan Pelanggan Servis Kendaraan Bermotor Berdasarkan Data Historis menggunakan Support Vector Machine," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, Vols. Vol. 7, No. 1, April 2021.
- [3] S. S. Drajat Indra Purnama, Sarika Afrizal " Model Support Vector Regression (SVR) untuk Peramalan Jumlah Penumpang Penerbangan Domestik di Bandara Sultan Hasanudin," *JMSK (Jurnal Matematika, Statistika & Komputasi)*, Vols. Vol. 16, No. 3, 391-403, May , 2020, Doi: 10.20956/jmsk.v%vi%i.9176 .
- [4] Devina Larassati, Ati Zaidiah, Sarika Afrizal "SISTEM PREDIKSI PENYAKIT JANTUNG KORONER MENGGUNAKAN METODE NAÏVE BAYES," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, Vols. Volume 07, Nomor 02, Juni 2022 : 533–546.
- [5] Lusi Suryadi, Ngajiyanto, Novia Eka Pratiwi, Novia Eka Pratiwi, Pakartika Riswanto "PENERAPAN DATA MINING PREDIKSI PENJUALAN MEBEL TERLARIS MENGGUNAKAN METODE K-NEAREST NEIGHBOR(K-NN)," *JUSIM (Jurnal Sistem Informasi Musirawas)*, Vols. Vol. 7, No. 2, Desember 2022.
- [6] S. Eva Mahdyta Kiswana, Sukmadiningtyas "KOMPARASI METODE PREDIKSI RESTOCK DENGAN PENDEKATAN K-NEAREST NEIGHBOR (K-NN) DAN SUPPORT VECTOR MACHINE (SVM)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vols. Vol. 9 No. 2, April 2025.
- [7] S. P. Mohanty, U. Choppali, and E. Kougianos, "Everything you wanted to know about smart cities," *IEEE Consum. Electron. Mag.*, vol. 5, no. 3, pp. 60–70, 2016, doi: 10.1109/MCE.2016.2556879.
- [8] W. A. Jabbar, W. K. Saad, and M. Ismail, "MEQSA-OLSRv2: A multicriteria-based hybrid multipath protocol for energy-efficient and QoS-aware data routing in MANET-WSN convergence scenarios of IoT," *IEEE Access*, 2018, doi: 10.1109/ACCESS.2018.2882853.
- [9] D. Niyigena, C. Habineza, and T. S. Ustun, "Computer-based smart energy management system for rural health centers," 2016, doi: 10.1109/IRSEC.2015.7455005.
- [10] F.-Z. Younsi, A. Bounnekar, D. Hamdadou, and O. Boussaid, "SEIR-SW, Simulation Model of Influenza Spread Based on the Small World Network," *Tsinghua Sci. Technol.*, vol. 20, no. 5, pp. 460–473, 2015.