e-ISSN: 2614-8404
p-ISSN:2776-3234

# Implementation of Data Mining on Tourist Visits Patterns on Lombok Island Tourism Objects

**Saikin[1*)], Sofiansyah Fadli[2], Maulana Ashari[3]**
[1,3]Program Studi Sistem Informasi, STMIK Lombok
[2]Program Studi Teknik Informatika, STMIK Lombok
Email: [1]eken.apache@gmail.com, [2]sofiansyah182@gmail.com, [3]aarydarkmaul@gmail.com

***Abstract*** −Foreign tourists entering Indonesia in 2017 and 2018 have increased. From the data obtained on the website of the Ministry of Tourism (Kemenpar) the number of foreign tourists in 2017 was 14,039,799, while in 2018 there were 15,806.1, with a comparison of the number of tourists from the two years, the percentage increase in tourists was 12.58%. The data analysis approach using a classification model is a data analysis approach by studying the data and making predictions with the new data. in the classification model, there are many algorithms that can be applied in data analysis, one of which is the Decision Tree algorithm. This study aims to analyze the pattern of tourist visits based on the objects visited by the number of tourists visiting certain tourist objects. From the modeling using the Decesion Tree C4.5 Algorithm and the scenario of splitting the data into three parts, the highest accuracy value was obtained for splitting data of 80:20 for train and testing data and max depth 7, which obtained an accuracy of 94% for train data and 92% for data. testing. Modeling with the Boostrap Aggregating Method, the accuracy score obtained on training data is 93% and testing data is 92. percent. 3 accuracy results from using bagging reduce the accuracy of the C4.5 algorithm on the data training side from 94% to 93 percent, while the accuracy of testing data is still the same, namely 92%.

***Keywords – C4.5 Algorithm, Data Mining, Decision Tree, Boostrap Aggregatting Method.***

## I. INTRODUCTION

Foreign tourists entering Indonesia in 2017 and 2018 have increased. From the data obtained on the website of the Ministry of Tourism (Kemenpar) the number of foreign tourists in 2017 was 14,039,799, while in 2018 there were 15,806.1, with a comparison of the number of tourists from the two years, the percentage increase in tourists was 12.58%. . The number of incoming tourists is calculated from various entrances scattered in several areas that have tourist attractions, including the Lombok area in the Province of West Nusa Tenggara (NTB). Although from statistical data obtained from the tourism service website in the area in the range of 2017 to 2018, there was a decrease in the number of tourists, but in that year the number of tourists entered was more than two million tourists.

The number of tourist visits certainly increases the activity of the economic industry in various sectors, such as hospitality, culinary, transportation services and also tour and travel services. Companies engaged in these fields provide travel services for tourists, who will visit certain tourist objects. In providing services the company will adjust to the needs and interests of tourists, but the obstacle is the difficulty of predicting the needs or interests of tourists, an obstacle faced by companies engaged in tour and travel services. Understanding the needs and interests of tourists in choosing a tourist attraction to be visited by using existing data in the past to predict the interest of tourists in visiting the selected tourist attraction. Analysis of past data to predict future needs will assist management in making decisions, especially those closely related to tourist visiting patterns.

The data analysis approach using a classification model is a data analysis approach by studying the data and making predictions with new data. in the classification model, there are many algorithms that can be applied in data analysis, one of which is the Decision Tree algorithm. Decision trees are one of the most popular classification methods because they are easy to interpret by humans. The concept of a decision tree is to convert data into a decision tree and decision rules, [1]. Decision trees are widely used to solve decision-making cases such as medicine (diagnosing patient disease), computer science (data structure), psychology (decision-making theory) and so on [2].

Research conducted by [3] entitled Predicting the Number of Foreign Tourist Visits to Bali Using Support Vector Regression with Genetic Algorithms. This study aims to predict tourist visits, the results of this study. Based on the MAPE value test, the obtained value is 2.513% with the best parameters, namely the lamda range 1 - 10, the complexity range 1 - 100, the epsilon range 0.00001 - 0.001, the gamma range 0.00001 - 0.001, sigma range 0.01 – 3.5, SVR 1250, generation GA 90, population 70, crossover rate 0.6, mutation rate 0.4, number of features 2 and number of prediction period 1 month. And succeeded in modeling the data on foreign tourist visits to Bali according to short-term predictions. while the research conducted by [4] Decision Tree in Analyzing Lake Poso Tourist Visitor Data for Decision Making. The results obtained from this study are the number of visitors more than 28,984 having the statement "Many" which is dominated by local tourists while the value with the description "Less" is for foreign tourists. This is one of the important points in determining the right strategy to develop tourism in Poso Lake.

Lombok Island is one of the areas that is a favorite of tourists visiting both local and foreign tourists. To find out the interests of tourists in choosing tourist objects to be visited, it is necessary to analyze tourist visit data to make it easier to provide services for tourists. Based on previous research that predicts tourist visits with the Support Vector

e-ISSN: 2614-8404
p-ISSN:2776-3234

Machine (SVM) method and analyzes the number of tourist visits using the decision tree method. This study tries to analyze the pattern of tourist visits based on the objects visited by the number of certain tourist objects. The method used in this research is the decesion tree C4.5 method and the bootrap aggreagating (bagging) method. Tools used to process jupyter notebook traveler data. The results of this study will also provide recommendations to the company to provide services or offer tourist objects provided to tourists.

## II. RESEARCH METHODOLOGY

### A. Literature Review

Modeling the number of foreign tourist arrivals in Batam using arima and time series regression. [5] The purpose of this study is to model with arima and time series regression and use it on serial data and trend and seasonal data. 1. The best suitable ARIMA model is ARIMA (0,1,1) (0,1,1)12, after testing the significance and residual assumptions.2. The appropriate time series regression model is one involving the time variable (t), the month dummy variable (January to December) and observations to (t-2) and (t-3). Tests of residual assumptions can also be met. 3. After calculating RMSE and MAPE, the best model is given by Time Series Regression.

Prediction of the Number of Foreign Tourist Visits to Bali Using Support Vector Regression with Genetic Algorithms. [3] The purpose of this study is to predict the number of foreign tourist visits to Bali using SVR with Genetic Algorithm optimization. The test results show the MAPE value obtained is 2.513% with the best parameters, namely lamda range 1-10, complexity range 1-100, epsilon range 0.00001 - 0.001, gamma range 0.00001 - 0.001, sigma range 0.01 - 3, 5, SVR 1250, generation GA 90, population 70, crossover rate 0.6, mutation rate 0.4, number of features 2 and number of prediction period 1 month. Based on the test results, the GA-SVR method on data on foreign tourist visits to Bali is suitable for short-term predictions.

Decision Tree in Analyzing Lake Poso Tourist Visitor Data for Decision Making. [4] The purpose of this study is to analyze data on tourist visits to Poso Lake by using the Decesion Tree algorithm for decision making. The results of this study, visitors to Lake Poso tourism with the number of visitors more than 28,984 have a lot of information and are dominated by local tourists, while foreign tourists with more than 417 visitors and less than 1,874 still have less information.

[6] Village Classification in Gianyar Regency: Extraction and Classification of Village Potential. The purpose of this study is to identify potential natural attractions, socio-cultural potential of the community, as well as supporting facilities and infrastructure that are useful in building or developing villages in Gianyar Regency as DTW. The results of potential identification are then used as information in clustering villages as tourist attractions. The results of potential identification are then used as information in clustering villages as tourist attractions.The conclusion of this study is that there are 6 potential attributes of the village that can be used to develop the villages of Gianyar Regency as a tourist attraction. These six attributes are: (a) Village atmosphere; (b) Uniqueness of flora & fauna; (c) village community arts; (d) The existence of temples as attractions; (e) Accessibility; and (f) Travel comfort; Three village clusters were formed based on their potential, namely a village cluster that has developed as a DTW consisting of 13 villages (Cluster I), a developing village cluster of 24 villages (Cluster II), and an undeveloped village cluster consisting of 33 villages (Cluster III); Cluster I has advantages in the uniqueness of flora & fauna, comfort, and artistic potential of the village community. Cluster II excels in the attributes of village atmosphere and temples as an attraction, while cluster III excels in accessibility to and between villages within the same cluster.

[7] Tourism Market Segmentation in Yogyakarta: Classification of Lifestyles of Domestic Tourists, the purpose of this research, is to identify the underlying dimensions of the lifestyles of local tourists to classify tourists who visit according to their typology of lifestyle, and to illustrate how to understand the various segments of local tourists. The results of this research are Metropolis Culture Aspiring Domestic tourists in Jogja in this cluster like shopping activities and usually buy something to take home when visiting Jogja. Happy to spend time off to do recreational activities and interested in visiting Jogja because of getting information from the media. Self-quality Explore Where domestic tourists in Jogja in this cluster have the opinion that if it is very thick with its customs and local wisdom, its culture must be preserved. In addition, tourists are also very optimistic about the future of tourism conditions in Jogja and believe that there will be more other tourists visiting the Jogja Aspiring Vacationer This cluster represents a segment of domestic tourists who have the opinion that education is very important and educational background does not affect one's income.

### B. Theoretical Foundation

#### 1. Data Mining

Data mining is a discipline that studies methods and extracts knowledge or finds patterns from data. Data itself is a recorded fact and has no meaning. And knowledge is a pattern, rule or model that emerges from the data, so data mining is called Knowledge Discovery in Database (KDD). [8].

Data mining is the analysis of the collection review is the analysis of the review of the data set to find unexpected relationships and summarize the data in a different way than before, which is understandable and useful for the data owner. [1] Data mining is an interactive and iterative process that involves several processes used, including knowing the type of data application, data selection, data cleaning, data integration, data reduction and transformation, data mining algorithms in selecting results interpretation development techniques and using the knowledge of the resulting process determined. [9].



e-ISSN: 2614-8404
p-ISSN:2776-3234

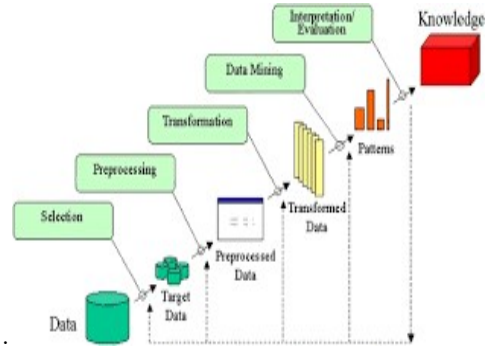


Figure 1. Knowledge Data Discovery (KDD) [10]

In the KDD process can be divided into the following [10]:

a) Selection
b) Preprocessing
c) Transformation
d) Data mining
e) Interpretation / evaluation.

In data mining, before extracting data, the dataset used must first enter the pre-processing phase. In the pre-processing phase, the data is first normalized so that later the implementation of the algorithm can run well. Then the dataset used must be large or large so that the level of the resulting pattern is getting better. Pre-processing is very useful for analyzing multi-variate datasets, the target is determined and then cleaned first. Data cleaning removes noise and missing data [11].

2. Decision Tree

A decision tree is a predictive model using a tree structure or hierarchical structure. Apart from being relatively fast in construction, the results of the models built are also easy to understand, so this Decision Tree is the most popular classification method used. A decision tree is a flow-chart like tree structure, where each internal node shows a test on an attribute, each branch shows the results of the test and the leaf node shows the classes or class distribution. [12].

$$Entropy(S) = \sum_{i=1}^{n} - p_i * log_2 p_i \quad (1)$$

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

$$SplitEntropy_A(S) = -\sum_{i+1}^{n} \frac{|Si|}{|S|} * log_2 \frac{|Si|}{|S|} \quad (3)$$

$$GainRatio(A) = \frac{Gain(A)}{SplitEntropy(A)} \quad (4)$$

where:

S =Case set
A =Attribute
n =Number of partitions on attribute A
|Si| =Number of cases in the i-th partition
|S| =Number of cases in S
Ph =Proportion from S_i to against S

3. C4.5 Algorithm

The C4.5 algorithm is a group of Decision Tree algorithms. This algorithm has input in the form of training samples and samples. Training samples in the form of sample data that will be used to build a tree that has been tested for truth. While samples are data fields that will later be used as parameters in classifying data [12].

At the learning stage, the C4.5 algorithm has 2 working principles, [12]:

a. Decision tree creation. The purpose of the decision tree induction algorithm is to construct a tree data structure that can be used to predict the class of a new case or record that does not yet have a class. C4.5 constructs a decision tree using the divide and conquer method. At first, only root nodes were created by applying the divide and conquer algorithm. This algorithm chooses the best case solution by calculating and comparing the gain ratio, then the nodes formed at the next level, divide and conquer algorithm will be applied again until the leaves are formed.
b. Making rules (rule set). The rules formed from the decision tree will form a condition in the form of if-then. These rules are obtained by tracing the decision tree from root to leaf. Each node and branching conditions will form a condition or an if, while for the values contained in the leaf will form a result or a then.

4. Bagging

Bagging is a voting method in which base-learners are differentiated from their training through a slightly different training set. Generating a sample L that is slightly different from the given sample is done by bootstrapping, when given a training set X of size N, then N random samples of X are shown with replacement [13]. Bagging

e-ISSN: 2614-8404
p-ISSN:2776-3234

was invented by [14] which stands for "bootstrap aggregating". [15] in his book states that bagging is a technique of the ensemble method by manipulating training data, training data is duplicated d times with sampling with replacement, which will produce d new training data, then from d training data The classifiers will be built called bagged classifiers [15].

According to [14] the stages of bagging can be considered as follows:

1. Bootstrap stages
   a. Take n samples randomly from the training data.
   b. After the sample is taken, arrange the best tree based on the training data.
   c. Repeat steps a-b b times to obtain B classification trees.
2. Aggregating Stages

The aggregating stage is identical to the majority vote, namely making predictions / guesses from a combination of B fruit classification trees.

C. Methodology

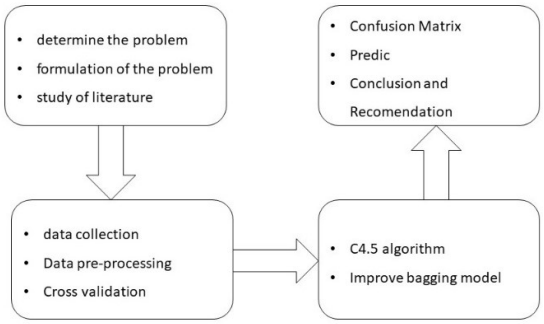


Figure 2. Research Flow

D. Data Collection

The data used in this study is secondary data, obtained from the office of the Lombok Ceria Holiday tour and travel company. The data obtained is in the form of company daily activity data, namely data on tourist attraction visits, from 2014 to 2015. The data consists of 11 (eleven) columns namely date, name, participants (number of participants), hotels, programs (tourism objects to be visited). ), restaurants, transportation, guides (tour guides), remarks, expansions and invoices. Feature deletion is also carried out on features that have no effect on data modeling, such as feature names, feature deletion will be carried out.

E. Data Pre-processing

Data cleaning aims to clean up the missing (Nan value) and or empty (null values). In the data obtained, there are several data features that are missing or null values. Several techniques are used to handle missing data, the first method is to fill in the blank or missing data with the highest value (max) of the feature, and the second method is done to the missing data by deleting features whose missing values are too high. Removed features such as restaurants, remarks, expansions and invoices. The figure table below shows the percentage of missing data for each feature.

| | TOTAL DATA MISSING | persen_missing |
|---|---|---|
| tanggal | 0 | 0.000000 |
| nama | 0 | 0.000000 |
| Peserta | 11 | 6.043956 |
| Hotel | 11 | 6.043956 |
| Program | 0 | 0.000000 |
| Resto | 95 | 52.197802 |
| transport | 31 | 17.032967 |
| Guide | 58 | 31.868132 |
| Remark | 150 | 82.417582 |
| Expenses | 176 | 96.703297 |
| Invoice | 135 | 74.175824 |

Figure 3. The percentage of missing data on each feature.

F. Feature Breakdown

After doing data cleaning, the next step is to transport data. Before carrying out data transportation, data is separated on the date feature, in that feature there are three values namely year, month and date, so the separation is done. Data transformation is not carried out on all data features, but feature transformation is carried out on only a few features such as hotel features, program features, transport features. display data transformation as Figure below:

| peserta | tanggal | bulan | tahun | hotel | program | Transport |
|---|---|---|---|---|---|---|
| rendah | akhir | triwulan3 | 2014 | 0.166667 | 0.714286 | 0.166667 |
| rendah | akhir | triwulan3 | 2014 | 0.166667 | 0.714286 | 0.166667 |
| rendah | akhir | triwulan3 | 2014 | 0.166667 | 0.285714 | 0.166667 |
| rendah | akhir | triwulan3 | 2014 | 0.166667 | 0.285714 | 0.166667 |
| rendah | akhir | triwulan3 | 2014 | 0.333333 | 0.714286 | 0.500000 |

Figure 4. Data Transformation.

G. Data Binning

Data binning aims to group data features based on certain criteria into smaller ones. In the processed data, the features of the binned data are the features of the participants grouped into low, medium and high. The date features are grouped into early, mid and late groups, while the number features are grouped into quarter1, quarter2, quarter3 and quarter4 groups.

| peserta | tanggal | bulan | |
|---|---|---|---|
| rendah | akhir | trwiulan4 | 89 |
| | awal | triwulan1 | 88 |
| | pertengahan | triwulan3 | 23 |
| | akhir | triwulan3 | 20 |
| | pertengahan | triwulan2 | 16 |
| sedang | akhir | trwiulan4 | 12 |
| rendah | awal | triwulan2 | 8 |
| tinggi | akhir | trwiulan4 | 6 |
| sedang | awal | triwulan1 | 3 |

Figure 5. Data Binning

e-ISSN: 2614-8404
p-ISSN:2776-3234

H. Target Labels

Determination of the target label from the classification is determined by conducting data clusters, where the target label is divided into two classes, namely the high level 0 visitor class marked with the number 0 (zero), and the 1 or low level class marked with the number 1 (one).

| kelas_kunjungan | |
|---|---|
| 1 | 139 |
| 0 | 126 |

Figure 6. Classification target label

## III. RESULTS AND DISCUSSION

Classification of data is divided in two ways, namely by classification with the decesion tree algorithm C4.5, the second way by using boostrap aggregating (Bagging). data classification with the C4.5 algorithm uses three data splitting scenarios, namely :

1) Data training and testing 50%.
2) Data training 70% and data testing 30%.
3) Data training 80% and data testing 20%.

A. Classification With C4.5 Algorithm.

In the classification with the C4.5 algorithm, three scenarios and one to seven max depths were tried, from the results obtained, the highest value and range of accuracy values in the training data and testing data were in the 80% vs 20% splitting scenario, and in the max depth experiment. 7. The accuracy value obtained is 0.94 for training data and 0.92 for testing data.

Table 1. C4.5 Classification accuracy value

| kenario splitting | | Max Depth | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 50% : 50% | Train | 0,89 | 0,90 | 0,93 | 0,94 | 0.95 | 0,96 | 0,96 |
| | Test | 0,86 | 0,86 | 0,88 | 0,87 | 0,86 | 0,86 | 0,86 |
| 70 % : 30% | Train | 0,88 | 0,88 | 0,91 | 0,93 | 0,94 | 0,94 | 0,95 |
| | Test | 0,86 | 0,86 | 0,89 | 0,88 | 0,89 | 0,89 | 0,89 |
| 80 % : 20% | Train | 0,88 | 0,88 | 0,91 | 0,91 | 0,93 | 0,93 | 0,94 |
| | Test | 0,86 | 0,86 | 0,88 | 0,88 | 0,92 | 0,90 | 0,92 |

B. Confusion Matrix Algoritma C4.5

The results of the classification with the C4.5 algorithm were tested using a confusion matrix, on training data and testing data. in the training data the values that are correctly predicted (True Positive) are 94, the values that are correctly predicted are sala (True Negative) are 104, while the false positive and True negative data are each 7. While the confusion matrix on the data testing data that is predicted to be true positive is 22 and predicted is 27, while on true negative 1 and false positive is 3.

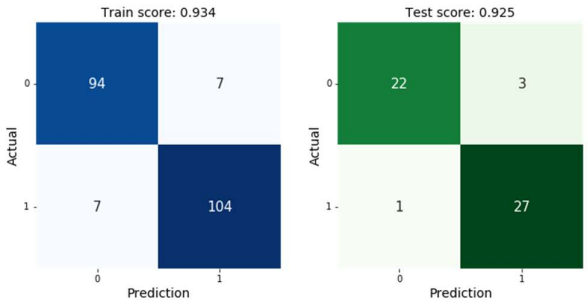


Figure 7. Confusion matrix classification

The test results using the confusion matrix, the accuracy value is 0.92 percent and the precision value is 0.96 percent and the recall value is 0.9 percent and the f1-score value is 0.93 percent.

```
================================================

accuracy_score of C4.5  :  0.9245283018867925

precision_score of C4.5 :  0.9642857142857143

recall_score of C4.5    :  0.9

F1-score  of C4.5       :  0.9310344827586207

================================================
```

Figure 8. Classification accuracy value

C. Bootsrap Aggregating (Bagging)

Testing the model using the bagging method, the value of n_estimator or the number of decision trees built seven times. Then in aggregating so that the accuracy value obtained on the training data and testing data is 0.93% and testing is 0.92%.

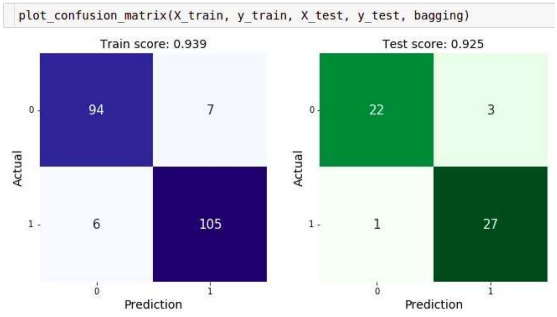


Figure 9. Confusion Matrix Model Bagging

e-ISSN: 2614-8404
p-ISSN:2776-3234

D. Modeling Results

Judging from the modeling results on the month of visit data which is divided into four quarters. In class 0 visits, the highest number of visits occurred in the 1st quarter, and in class 1 visits, the highest number of visits occurred in January.

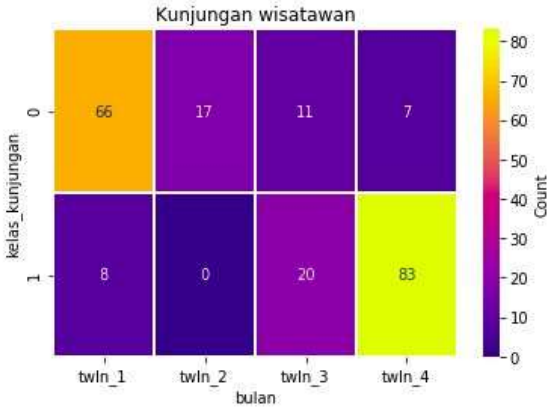


Figure 10. Quarterly visiting class

In the class of visits to tourist objects, the modeling results show that the highest number of level 0 visitors tested Gili Nanggu attractions, and the highest number of level 1 visitors visited Mandalika Kuta, Gili Nanggu and Pink Beach attractions.

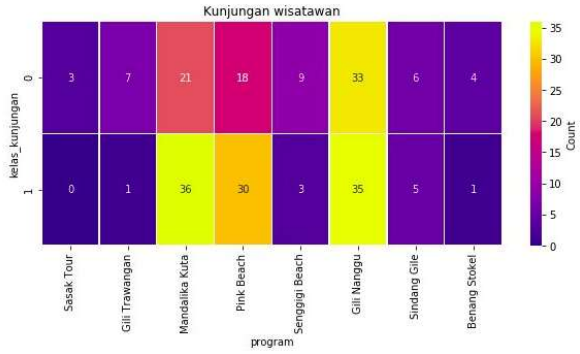


Figure 11. Tourist Attraction Visiting Class.

## IV. CONCLUSION

A. Conclusion

1. From the modeling using the Decision Tree C4.5 algorithm and the scenario of splitting the data into three parts, the highest accuracy value obtained is 80:20 data splitting for train and testing data and max depth 7, which gets 94% accuracy for train data and 92% for data testing.
2. Modeling with the Bootstrap Aggregating Method, the accuracy score obtained on training data is 93% and testing data is 92. percent.
3. The accuracy results from using bagging reduce the accuracy of the C4.5 algorithm on the data training side from 94% to 93 percent, while the accuracy of testing data is still the same at 92%.
4. The prediction results for the highest class of visits are at level 0 in the first quarter where in the first quarter it is in January, February and March. while at the level occurred in the 4th quarter, namely October, November and December.
5. The attractions with the most visitors at level 0 of the visitor class are Gili Nanggu attractions, while at level 1 tourists tend to choose Mandalika Kuta, Gili Nanggu and Pink Beach attractions.

B. Recommendation

In this study using data from 2014 to 2015 for further research it is recommended to use more recent data and use the ARIMA method due to the nature of the data obtained, namely time series data.

## REFERENCES

[1] Larose, Daniel T, "Discovering Knowledge in Data : An Introduction to Data Mining", John Willey & Sons, Inc. 2005.

[2] Prasetyo, Eko, "Data Mining", Yogyakarta: Andi Offset. 2014.

[3] Listiya Surtiningsih, Muhammad Tanzil Furqon, Sigit Adinugroho, "Prediksi Jumlah Kunjungan Wisatawan Mancanegara Ke Bali Menggunakan Support Vector Regression dengan Algoritma Genetika" Jurnal Pengembangan Teknlogi Informasi dan Ilmu Komputer, Vol 2 No 8, 2018.

[4] Fredryc Joshua Pa'o, Hendry Hendry, "Decision Tree dalam Menganalisis Data Pengunjung Wisata Danau Poso untuk Pengambilan Keputusan", Jurnal Sistem Komputer dan Informatika (JSON), Vol 2, No 3, 2021.

[5] Ely Kurniawati, One Yantri, "Pemodelan Jumlah Kunjungan Wisatawan Mancanegara Di Batam Dengan Menggunakan Arima Dan Regresi Time Series", Jurnal Dimensi, Vol 7, No 3,2018.

[6] Mirah P Handayani, Putu Suciptawati, Trisna Darmayanti, Eka N Kencana, "Klasifikasi Desa/Kelurahan di Kabupaten Gianyar: Ekstraksi dan Klasifikasi Potensi Wisata", Jurnal Master Pariwisata (JUMPA) Volume 07, Nomor 02, 2021.

[7] Agnessia Mega Cahyani Andri Saputri, Devilia Sari, "Segmentasi Pasar Turisme Di Yogyakarta: Klasifikasi Gaya Hidup Wisatawan Domestik", eProceedings of Management, Vol 6, No 2, 2019

[8] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth, "From Data Mining to Knowledge Discovery in Databases", AI Magazine Volume 17 Number 3. 1996

[9] Liao, "Recent Advances in Data Mining of Enterprise Data: Algorithms and Application", Singapore: World Scientific Publishing, 2007.

[10] Yuli Mardi, "Data Mining: Klasifikasi Menggunakan Algoritma C4.5", Jurnal Edik Informatika Penelitian Bidang Komputer Sains dan Pendidikan Informatika, V2.i2 (213-219).

[11] Ahmad Rofiqul Muslikh, Heru Agus Santoso, Aris Marjuni, "Klasifikasi Data Time Series Arus Lalu Lintas Jangka Pendek Menggunakan Algoritma Adaboost Dengan Random Forest", Vol. 14, No. 1, 2018.

[12] Sunjana, "Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5, Seminar

e-ISSN: 2614-8404
p-ISSN:2776-3234

Nasional Aplikasi Teknologi Informasi (SNATI)", Yogyakarta, 2010.

[13] Clancey, W.J, "Communication, Simulation, and Intelligent Agents: Implications of Personal Intelligent Machines for Medical Education", In Proceedings of the Eighth International Joint Conference on Artificial Intelligence, 556-560. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence, 1983.

[14] Leo Beiman, "Machine Learning, Statistics Department", University of Californiaa. Berkeley, CA 94720, 1996.

[15] Amin, N.A.S., Istadi, I, "Different Tools on Multiobjective Optimization of a Hybrid Artificial Neural Network – Genetic Algorithm for Plasma Chemical Reactor Modelling", In Olympia Roeva (Editor) Real-World Applications of Genetic Algorithms. Croatia: InTech Publisher, 2012.

[16] Saikin Saikin, Sofiansyah Fadli, Maulana Ashari, "Optimization of Support Vector Machine Method Using Feature Selection to Improve Classification Results," JISA (Jurnal Informatika dan Sains), Vol 4, No 1, 2021.