IOP Conf. Series: Materials Science and Engineering **732** (2020) 012079 doi:10.1088/1757-899X/732/1/012079

GIS for coffee shops classification and routing using Naive Bayes method

E Rohadi[1], A Amalia[2,*], J D Bagaskara[1], B Harijanto[1] and S Adhisuwignjo[2]

[1] Information Technology Department, State Polytechnic of Malang, Jl. Soekarno Hatta No. 9 Malang 65141, Indonesia
[2] Electrical Engineering Department, State Polytechnic of Malang, Jl. Soekarno Hatta No. 9 Malang 65141, Indonesia

*amalia@polinema.ac.id

Abstract. In recent years, the number of coffee shops has grown rapidly in Malang whose locations spread in various places. This condition makes the consumers having trouble to find the place that meets with their needs of the price and comfort level. In this works, the Geographical Information System of Coffee Shop Business Classification in Malang based on criteria is proposed. This system can classify coffee shop data according to the consumer desires using the Naïve Bayes method. Users simply provide a choice of price criteria and desired level of comfort on this website-based system. The classification results are used to make it easier for users to obtain information, both the map of locations and the route to reach the coffee shops that meet the criteria expected by the user. Based on the testing that has been done, 100% of users stated that they could find a coffee shop according to the desired criteria. As a result, the system promises as the application in determining the selection of coffee shops corresponds to the consumer criteria.

1. Introduction

Nowadays the coffee shops have become a growing business in many cities such as Malang, East Java. The number of coffee consumers in Indonesia increases every year. The coffee consumption reached 276 million kilograms in 2016, and coffee plantations production grows significantly as the coffee production reached 10,141 tons by 2017. Both the growth of coffee consumers and the ease of getting supplies of raw materials for coffee beans in Malang became a significant factor in the abundance of coffee shop businesses as an emerging business in Malang.

The coffee shop visitors have various backgrounds ranging from a student of many universities and also workers due to having a small discussion, doing assignments, client meetings and so on reasons. The scattered location, different opening times, prices and levels of comfort, has become difficulties for the consumers to deliver their purposes. Based on that problem, in this research, the application of Geographic Information Systems (GIS) classification of coffee shop business in Malang based on criteria and its management input and output of spatial data referencing to geographical and factual data of the coffee shops proposed [1,2]. The criteria of proposed application have clarified of the coffee shops based on price and comfort level using the Naïve Bayes method.

The system is based on a website in order to make it easier for users to access. The proposed system has been addressed to provide input in the form price and comfort level criteria. Price criteria have been

classified into coffee shops with sub-criteria of low prices, medium prices, and high prices. The criteria for the comfort level of coffee shops have been classified into coffee shops with ordinary, sufficient and comfortable sub criteria. Furthermore, the application displays the location, along with a review of the coffee shops that matches the expected criteria by the user. This system is expected to be able to classify coffee shops based on price and comfort level so that users can search for several coffee shops according to the desired criteria. The proposed GIS for coffee shops classification and routing use the Naïve Bayes method in order to classify the data [3-6]. The purpose of developing this system is to provide an application that facilitates users in getting information and location of coffee shops.

2. The method and descriptions of system

2.1. Method

Naive Bayes algorithm is one of the most effective and efficient inductive learning algorithms for data mining along with machine learning. This algorithm belongs to the wrapper approach and considered a simple classifier based on the classical statistical theory "Bayes theorem." The Bayesian algorithm is branded "naïve" because it is founded on Bayes Rule, which has a strong supposition that the features are conditionally independent from each other with regard to the class. In the literature, the NB algorithm has proven its effectiveness in various domains such as text classification, improving search engine quality, image processing, fault prediction, and medical diagnoses [6].

This application uses the Naïve Bayes method to classify coffee shop data based on price and comfort levels criteria by classifying coffee shops into A-class, B-class, and C-class. The Bayesian classifier in this research is used to represent the probability of price and comfort level criteria. The results of the classification are used to facilitate users to obtain the information and location maps from coffee shops that meet the criteria expected by the user. The advantage of using Naive Bayes is that this method requires only a small amount of training data to determine the estimated parameters required in the classification process [7,8].

2.2. System design

Figure 1. System design.

Figure 1 shows the design of the GIS for coffee shops classification and routing by Using Naive Bayes Method. This system consists of several parts which are:

- The administrators can create, update and delete of the data related to the coffee shop.
- The database to store all the data regarding to this system.
- Naïve Bayes method for data processing.

- Users can choose the criteria they want such as price and comfort level, users can see a list of coffee shops that match with the user's criteria and view coffee shops on maps with different colour icons for each class.

2.3. System flowchart

Figure 2 shows the system flowchart used to implement the system. The user must enter the criteria needed to filter the coffee shop data. Therefore, the system will process the input data from the user and display the filtering results of the coffee shop according to user criteria.

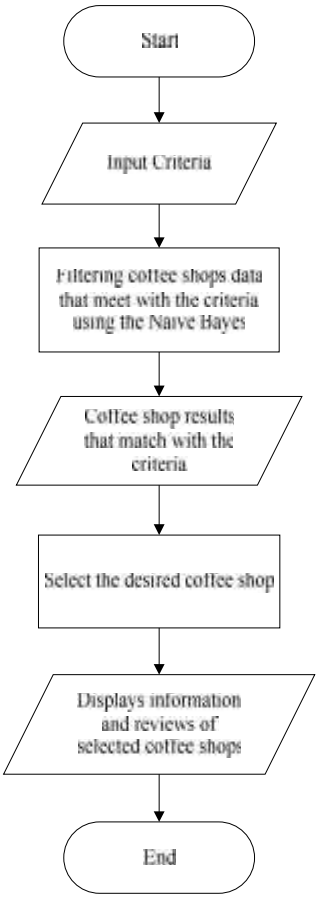

Figure 2. System flowchart.

3. Results and discussion

3.1. The implementation of Naïve Bayes method

For the purpose of this study, data was collected from 34 coffee shops including the information of coffee shops such us longitude and latitude of each coffee shops, then price and comfort level. This application uses the Naïve Bayes method to classify coffee shop data based on price and comfort levels criteria by classifying coffee shops into A-class, B-class, and C-class. The A-class coffee shops are a coffee shop with the results of calculation methods that get the best assessment points, B-class gets a moderate assessment points, and C-class is the coffee shop with the lowest assessment points. The Naïve Bayes method can classify data with small amounts of data training to estimate the parameters needed for classification [7,9].

The results of the classification are used to facilitate users to obtain the information and location maps from coffee shops that meet the criteria expected by the user. Users only need to provide input in the form of a choice of price and comfort level criteria on this website based application. The price

criteria classified into cheap, medium and expensive while, the comfort level criteria classified into ordinary, sufficient and comfortable to facilitate users in choosing the appropriate price and comfort level options. The system will calculate the data using the Naïve Bayes method to present coffee shop in accordance with the criteria entered by the user, and then the system will display the coffee shop information along with a map route to the selected coffee shop. The coffee shop data classified by the price and comfort level criteria by weighting each sub-criteria as shown in table 1. The value of weighting obtained from the total number of respondents' choices in each sub-criteria.

Table 1. Price criteria weighting.

Price Criteria	Comfort Level Criteria	Weight
Cheap	Ordinary	3
Medium	Sufficient	2
Expensive	Comfortable	1

The highest weighting value indicated that it received the highest number of respondents. Table 2 shows 5 out of 34 coffee shop data with prices and comfort levels. The value of the price obtained from the average price acquired from each coffee shop, while the value of the comfort level obtained from the average level of comfort that comes from the respondent's questionnaire at each coffee shop. Trials by comparing the results of manual calculations using Excel with the calculations made by the system have been done and provide the same results. The final probability value obtained from the calculation of the likelihood value which is then performed to get the maximum value of each probability as shown in table 3.

Table 2. The list of coffee shops.

No	Coffee Shops	Price	Comfort Level	Price conversion	Comfort Level Conversion
1	Lataroma	13333,3	3,86667	Low	Ordinary
2	Ini Kopi	13400	4	Low	Ordinary
3	Noch coffee	22266,7	4,06667	Medium	Sufficient
4	Handlebar	18200	4,13333	Medium	Sufficient
5	DW coffee shop	22333,3	4,53333	Medium	Comfortable

Table 3. The probability of price and comfort level criteria.

Parameters	Price conversion	Probability	Value
Price	Low	A	0,9
	Low	B	0,73
	Low	C	0
	Medium	A	0,1
	Medium	B	0,28
	Medium	C	0,67
	Expensive	A	0
	Expensive	B	0
	Expensive	C	0,3
Comfort Level	Ordinary	A	0
	Ordinary	B	0,72
	Ordinary	C	0,84
	Sufficient	A	0,3
	Sufficient	B	0,23
	Sufficient	C	0,17
	Comfortable	A	0,7
	Comfortable	B	0
	Comfortable	C	0

3.2. The system testing

In this section, the results of system implementation and system testing discussed. Figure 3 shows the home page of this system. In this home page, users can choose options from price and comfort level criteria according to user needs. Hereafter, the system processes and displays the results of coffee shop recommendations in accordance with the criteria entered by the user as shown in figure 4. This system also displays a Google maps along with the route to the selected coffee shop by displaying the user's location point and the location of the coffee shop. Google Maps was developed using the Google API and JavaScript programming language on this system.

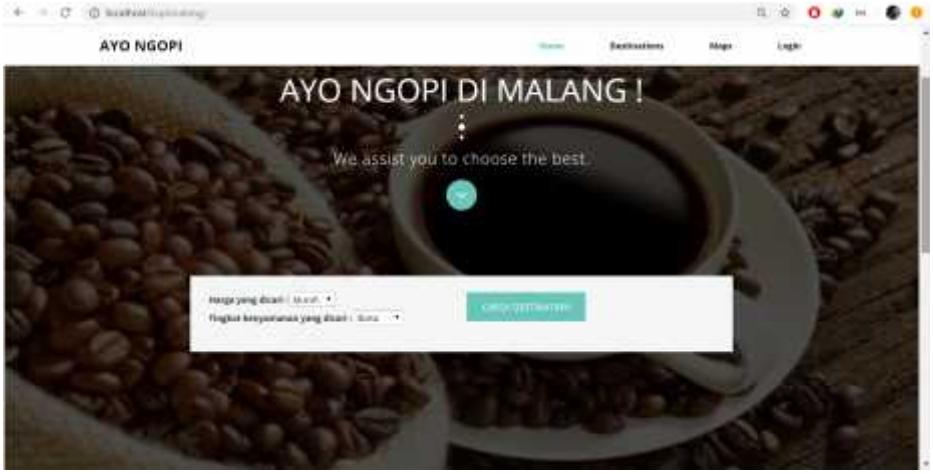

Figure 3. System home page.

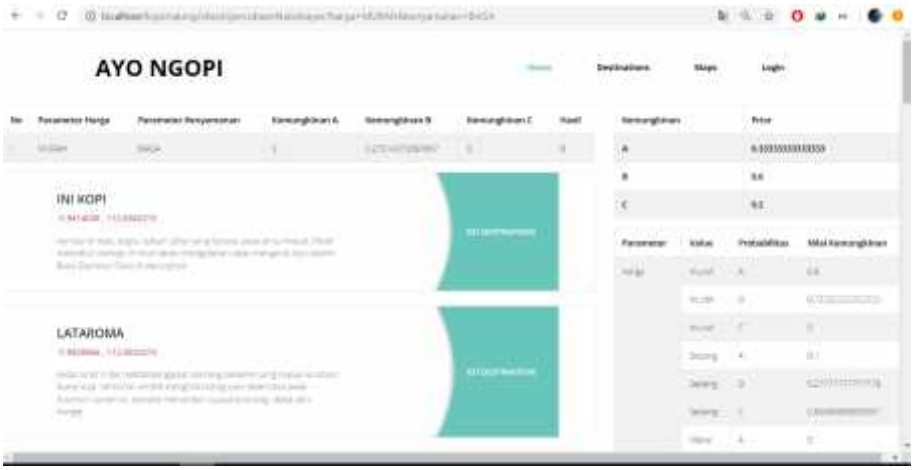

Figure 4. Coffee shop recommendation page.

System testing by 20 coffee shop visitors is carried out to ensure that the system is running as expected. At the time of testing, the author explained the features and how to use the application first. After completing the system testing, the questionnaire distributed to users and the following results were obtained:

- 100% of users or 20 out of 20 users claim that this system is easy to use.
- 90% of users or 18 out of 20 users state that this system has an attractive design.
- 100% of users or 20 out of 20 users state that users can find coffee shops according to criteria.
- 85% of users or 17 out of 20 users stated that the information provided was very informative.

- 100% of users or 20 out of 20 users stated that the location routes to the coffee shop provided was very helpful.

4. Conclusion

GIS for coffee shops classification and routing by using Naive Bayes method has been presented. The Naïve Bayes method can classify coffee shops based on criteria so that users can search for coffee shops according to their preferences. The implementation of Google Maps can be developed using JavaScript so that this system can display directions from the user's location to the desired coffee shop. Based on testing that has been done 100% of users stated that they could find a coffee shop in accordance with the desired criteria and information from the location of the coffee shop they got was very useful. As a result, this system can be used as an application in determining the selection of coffee shops according to user criteria.

Acknowledgment

The Authors would like to thank The State Polytechnic of Malang for the support to attend this conference.

References

[1] Zou L and Brooks S 2019 A dynamic approach for presenting local and global information in geospatial network visualizations *GeoInformatica* pp 1-25
[2] McKenney M, Frye R, Dellamano M, Anderson K and Harris J 2017 Multi-core parallelism for plane sweep algorithms as a foundation for GIS operations *GeoInformatica* **21** pp 151-174
[3] Abdulkareem S A, Mustafa Y T, Augustijn E W and Filatova T 2019 Bayesian networks for spatial learning: a workflow on using limited survey data for intelligent learning in spatial agent-based models *Geoinformatica* **23** pp 243-268
[4] Li P, Banerjee S and McBean A M 2011 Mining boundary effects in a really referenced spatial data using the Bayesian information criterion *Geoinformatica* **15** pp 435-454
[5] Afeni B O, Aruleba T I and Oloyede I A 2017 Hypertension Prediction System Using Naive Bayes Classifier *Journal of Advances in Mathematics and Computer Science* pp 1-11
[6] Taha A M, Mustapha A and Chen S D 2013 Naive Bayes-guided bat algorithm for feature selection *The Scientific World Journal* **2013** pp 1-9
[7] Harahap F, Harahap A Y N, Ekadiansyah E, Sari R N, Adawiyah R and Harahap C B 2018 Implementation of Naïve Bayes Classification Method for Predicting Purchase *6th Int. Conf. on Cyber and IT Service Management (CITSM)* pp 1-5
[8] Ordonez A, Paje R E and Naz R 2018 SMS Classification Method for Disaster Response Using Naïve Bayes Algorithm *Int. Symposium on Computer, Consumer and Control (IS3C)* pp 233-236
[9] Anuradha C and Velmurugan T 2016 Feature Selection Techniques To Analyse Student Acadamic Performance Using Naïve Bayes Classifier *3rd Int. Conf. on Small & Medium Business* pp 345-350