ISSN 2087-3336 (Print) | 2721-4729 (Online)
TEKNOSAINS: Jurnal Sains, Teknologi dan Informatika
Vol. 12, No. 1, 2025, page. 122-132
http://jurnal.sttmcileungsi.ac.id/index.php/tekno
DOI: 10.37373

Sentiment analysis study of library services using support vector machine methods

Risa Putri Arianty*, M. Riko Anshori Prasetya, Subhan Panji Cipta, Bayu Nugraha

* Universitas Sari Mulia, Indonesia, Jln. Pramuka No.02, Pemurus Luar, Kec. Banjarmasin Timur, Kota Banjarmasin, Kalimantan Selatan 70238

* pputtrii.rissaaa@gmail.com

Submitted: 19/07/2024 Revised: 13/08/2024 Accepted: 23/08/2024

ABSTRACT

Academic institutions' libraries contribute significantly to students' success by giving them access to the information resources they need. Sentiment analysis of library services is crucial in the current digital era to comprehend student perspectives and enhance service quality. The purpose of this study is to assess how XYZ University students feel about the library services thereby applying the Support Vector Machine (SVM) technique. Data obtained through a survey utilizing a Google Form is used in this study to analyze sentiment using the SVM algorithm. Preprocessing steps include data cleaning, normalization, tokenization, stopword removal, and stemming. Eighty percent of the data were used to train the SVM model, while twenty percent of the data were used to test it. The evaluation findings demonstrated that the SVM model could classify sentiment with a 90% accuracy level. This result was confirmed using an accuracy, precision, recall, and F-measure metrics confusion matrix. The sentiment analysis's findings revealed that most students had a favorable opinion of the library's services, particularly when it came to the staff's effectiveness and the facilities' cleanliness. But there are still certain areas that require work, including managing noise and adding more pertinent book selections. To increase overall library user happiness, it is advised to put in place a program that focuses on enhancing the quality of a quieter learning environment and providing a selection of books that meet the needs of students.

Keywords: Sentiment analysis; library services; student sentiment; support vector machine

1. INTRODUCTION

Because they make it easier for students to get the information resources they require, academic libraries are essential to promoting student achievement. To comprehend student perspectives and enhance service quality in the current digital era, it is critical to examine sentiment regarding library services [1]. Public libraries have introduced cutting-edge concepts, tools, and resources into their offerings, thereby augmenting their value [2]. Academic libraries are in a strategic position to take the lead in offering services connected to research data. This is regarded as one of the sectors that require the most improvement in the future [3][4].

A range of activities that entail face-to-face interaction between people and machines are included in library services to guarantee user happiness. It comprises components including personnel performance, facility cleanliness, and information resource accessibility [5]. The degree of user happiness and the efficacy of student learning are both influenced by the quality of the services offered. This implies that information specialists, such as academic librarians, must choose the most effective way to satisfy user needs. Librarians must patrons or library users in obtaining the information they require [6][7].

Students continue to voice dissatisfaction with the services offered, despite numerous attempts to enhance library services. The richness of book collections, the attitudes of the workers, and other library amenities have all been the subject of complaints. As a result, a student sentiment study of library services must be carried out [8][9].

Many techniques for sentiment analysis have been used in past studies. Ratniasih, for instance, claims that the Sastrawi Stemming method and the Support Vector Machine (SVM) algorithm can be used to categorize student happiness [10], In the meantime, Fakhri examined the sentiment of the campus service satisfaction survey using the support vector machine (SVM) [11]. Hermanto also contrasted the effectiveness of the support vector machine (SVM) and Naïve Bayes classifiers in categorizing student complaints [12].

The Support Vector Machine (SVM) approach is used in this study to categorize students' attitudes toward the library services at XYZ University. An evaluation measurement tool called a confusion matrix is used to assess the performance of these two strategies. It is anticipated that the analysis's findings will offer suggestions for enhancing library services overall.

Based on student attitude research, this study was carried out to offer improvement suggestions that can assist the library at XYZ University in raising the standard of service. By taking into account the suggestions of students regarding their needs, the library can be improved even more. The Support Vector Machine (SVM) approach will be used in this study to discover and categorize student attitudes toward library services. This study focuses on how well SVM performs sentiment analysis and offers suggestions for enhancing services in light of the analysis's findings. By applying the SVM approach, this study not only offers insight into the efficacy of the method but also makes recommendations for enhancements to library services, making it significant for conducting sentiment analysis and providing library services.

Compared to earlier research, this one has a few advantages, one of which is that it makes use of the SVM method to examine students' attitudes about library services. The analysis's findings demonstrate that the accuracy level approaches 90%, demonstrating the potency of SVM in assessing students' attitudes toward library services. Second, based on the research findings, this study offers suggestions for enhancing library services in addition to its main focus on sentiment analysis. Because of this, the study can help raise the standard of library services. Third, because XYZ students provided the data, the study's conclusions are more pertinent and the university library can use them immediately.

In this study, the SVM algorithm is implemented using the Scikit-learn module and the Python programming language. The benefits of this study are in raising customer satisfaction and operational efficiency in the library, which would eventually help students at XYZ University succeed academically.

2. METHOD

The support vector machine (SVM) method is used in this study's quantitative approach to examine students' attitudes toward the library services at XYZ University. The first step of the research process is gathering information from students via surveys regarding their experiences using library services. The information is then divided into three categories: positive, negative, and neutral. as demonstrated in Figure 1. Study Flow.

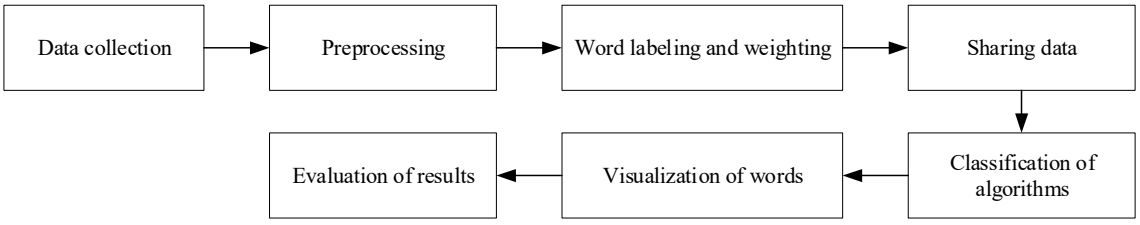

Figure 1. Research flow

There are multiple steps to this research:

Data Collection: All current students at XYZ University were given a Google Form to complete to facilitate direct observation, which was the method used in this study to collect data. The information was then gathered and saved in Microsoft Excel file format.

Preprocessing: This step is completed once all the data has been gathered and is prepared for further processing.

- Data Cleaning: At this point, superfluous words and characters are eliminated from the gathered data [13].
- Case Folding: In this step, the text is normalized by transforming every letter in a comment to either lowercase or uppercase overall [14].
- Tokenizing: At this point, lengthy texts are filtered and specific characters, such as punctuation, are eliminated [15].
- Stopword/Filtering: This stage entails eliminating terms like "and," "which," "or," "is," and other like terms that are deemed unnecessary [16].
- Stemming: Every word is examined, and the root word is found.

Labeling and Word Weighting: The goal of the labeling process is to give each piece of data in the dataset a unique label or category. Subsequently, word weighting is done to assign values or weights to all words in the comments, including neutral, negative, and positive sentences [17][18]. This process is done after applying the steaming step.

Data division: Data division consists of the following two processes: 80% Training Data and 20% Testing Data. Training information currently, a machine learning technique is being utilized to train the model with 80% of the data. Evaluating data Additionally, 20% of the data is used to evaluate how well the trained model performs when using training data.

Support Vector Machine (SVM) Classification: Next, the Support Vector Machine (SVM) technique can be used to classify the data. Word visualization: A type of data visualization that displays the frequency of occurrence of words in a text is called word visualization. The visualization's word size corresponds to the word's frequency of occurrence in the text [19].

Results evaluation: Following the conclusion of the full research series, results evaluation might be carried out. In this assessment, a confusion matrix for student sentiment analysis of library services using the SVM approach will be used to analyze the accuracy level performance.

3. RESULTS AND DISCUSSION

Data collection

A questionnaire used in this study to gauge student satisfaction with the library services at XYZ University was used to gather data. Every student who was enrolled at the time of the survey completed it via Google Forms. The purpose of this questionnaire was to gauge how satisfied students were with the facilities, staff services, cleanliness, and other aspects of the library's offerings. To gather students' perspectives regarding the caliber of services rendered by library staff, the questions posed to them focused on library services. 144 respondents' answers to this questionnaire were successfully gathered through distribution. Table 1 four instances of gathered data are provided.

Table 1. Sample data

No.	Principal	Semester	UNISM Library Opinion
1.	Faculty of Science & Technology	7	The officers are rude
2.	Faculty of Humanities	3	Clean and quite complete
3.	Health Faculty	1	Cleanliness can be improved again
4.	Faculty of Humanities	5	Service must be further improved

Preprocessing

After collection, the data is cleaned and prepped for future analysis through preprocessing steps. The preprocessing actions performed are listed below:

- Data Cleaning: At this point, superfluous words and characters, such as periods, commas, hashtags, links, usernames, and other symbols, are eliminated from the gathered data. The goal of this data-cleaning procedure is to eliminate superfluous information from the library review text to speed up the analysis process. The "pandas" library is used in this procedure to manipulate data, while regular expressions, or re, are used to manipulate text. This is an illustration of Table *2*. regarding data cleansing.

Table 2. Data cleaning

Input	Output
Clean, nice and neat	Clean nice and neat

- Case Folding: Using Python's "(lower)" function, every text is automatically converted to lowercase at this point to accomplish word normalization. This procedure guarantees that the format of every text is consistent. Table *3* is an example of a case folding.

Table 3. Case folding

Input	Output
The officers are rude	the officers are rude

- Tokenizing: At this point, the information that has passed through multiple earlier stages will be divided into smaller pieces known as tokens. Tokenization is done with the help of the Natural Language Toolkit (Nltk) package. Text is divided into tokens using the nltk word_tokenize function. Table 4 provides one example. Regarding tokenization.

Table 4. Tokenizing

Input	Output
library services are quite good	[library, services, are, quite, good]

- Stopword/Filtering: At this point, terms like "and," "the," "in," "is," and "at" are eliminated because they don't add much to the text's meaning in English or aren't important. Example Table 5: Stopword/filtering Examples.

Table 5. Stopword/filtering

Input	Output
the service at this library is friendly	[service, library, friendly]

- Stepping: Every word is examined at this point to see which prefixes and suffixes need to be eliminated to return it to its original form. Literary libraries are used in the stemming process. The words are processed to transform them into their most basic form when the required libraries have been imported. Table 6 is an example of stemming.

Table 6. Stemming

Input	Output
[service, good, friendly]	Service good friendly

Labeling and weighting of words

A Lexion-based technique employing compound score was used to label 144 data points in this study's labeling process. This strategy entails using a word dictionary that has been evaluated using a sentiment score. The word's polarity—positive, neutral, or negative—is indicated by this score. Table 7 shows this labeling method.

Table 7. Label

Text	Score	Sentiment
Officers rude	-0.4588	Negative
Good	0.4404	Positive
Independent	0.0000	Neutral

Example of a manual calculation of the Compound Score:

a. Tokenization and Initial Score on the text of one of the sentiments, namely "good" is broken down into words: ["good"]. By getting a sentiment score = 1.9 (this value is taken from the VADER dictionary).
b. Calculating Compound Score: VADER uses the following normalization formula to calculate Compound Score:

$$\text{Compound Score} = \frac{Total\ Score}{\sqrt{total\ score^2 + \propto}} \quad (1)$$

Where:

Total score : Total sentiment score
∝ (alpha) : 15 (constant in VADER formula)

Using this formula:

$$Compoung\ score = \frac{1.9}{\sqrt{(1.9^2) + 15}} = \frac{1.9}{\left(\sqrt{3.61 + 15}\right)} = \frac{1.9}{\sqrt{18.61}} = \frac{1.9}{4.32} = 0.4404$$

Table 7 necessitates manual computation using word weighting and sentiment analysis methods. Using VADER (Valence Aware Dictionary and Sentiment Reasoner), which evaluates each word in the text for its score, sentiment is identified. The Compound Score, which is used to define the sentiment label—positive if more than 0, negative if less than 0, and neutral if equal to 0—is obtained by combining and normalizing the sentiment scores of each word.

Word weighting is accomplished in Python using TfidfVectorizer from the scikit-learn module. Term Frequency-Inverse Document Frequency (TF-IDF) is a text processing technique that aids in determining a word's significance about sentiment. Table 8 provides an example of word weighting, which looks like this:

Table 8. Word weighting

	Term	idf	tf_0	tf_1	tf_2	tf_3	tf_4	tf_5	tf_6	tf_7
0	accessed	3.871201	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	active	4.276666	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	add	2.890372	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	added	3.871201	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Data Sharing

The dataset is split into two sets, training, and testing, as part of the data separation procedure. Testing data is used to evaluate the model's performance, whereas training data is used to create and train the classification model. Eighty percent of the data in this study were utilized for testing, while the remaining twenty percent were used for training. This 80:20 divide, with 80% of the data for training and 20% for testing, is a widely utilized technique, according to the journal Gustientiedina and Handayani [20][21].

There were 144 respondents in the research dataset, which was split into two primary sets: training data and testing data. To teach the model to identify patterns and generate precise predictions, up to 80% of the dataset—115 respondents—was used as training data. The model's performance was assessed and its capacity to predict previously untested new data was confirmed using the testing data from the remaining 20% of respondents, or 29 people. To guarantee that the final model can generalize effectively and does not overfit the training set, this dataset partition is crucial.

Support vector machine (SVM) algorithm classification

The Support Vector Machine (SVM) approach for classification entails locating the best hyperplane to divide two groups into distinct sets of data. By acting as a dividing line, this hyperplane maximizes the margin that separates the two classes. In this study, Python and the scikit-learn library were used as tools for categorization modeling. The 144 students at XYZ University who completed questionnaires provided the dataset, which was composed of labeled student comments into three categories: good, negative, and neutral. The student comments and their classification labels are included in the dataset, which is in the format of a text file or CSV (Comma Separated Values). The following matrix displays the findings of the SVM classification in this investigation. Figure 2. Matrix support vector machine.

The Support Vector Machine (SVM) technique was used to classify data on testing data, and the findings are shown in Figure 2. Matrix support vector machine as a confusion matrix. The model's ability to forecast the testing data class based on the actual label is demonstrated by this confusion matrix. Three classes—Negative, Neutral, and Positive—make up this matrix. Overall, this model performs satisfactorily, as evidenced by the assessment matrix, which demonstrates the model's excellent data classification abilities.

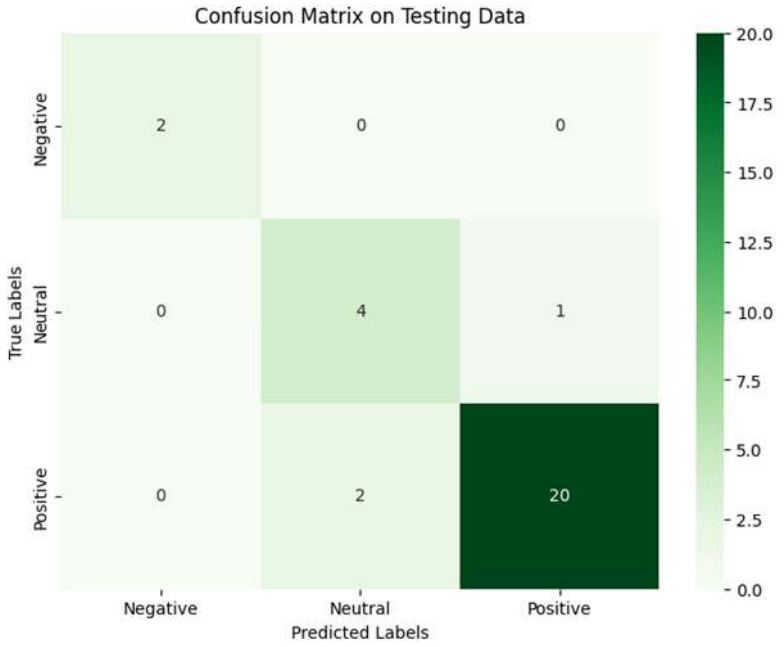

Figure 2. Matrix support vector machine

Visualization of words

Sentiment analysis revealed that most students had nice things to say about XYZ University's library offerings. This demonstrates how happy students are with the library's offerings. Figure 3. Word visualization: Word cloud generated from student evaluations of library services.

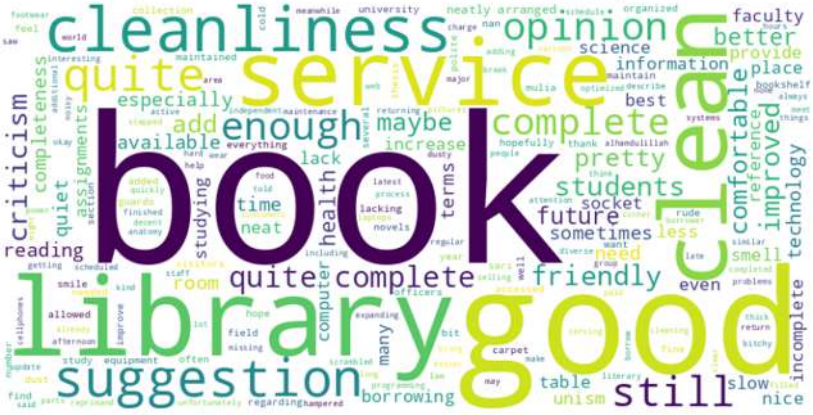

Figure 3. Word visualization

Evaluation of results

Python, Google Colab, and Microsoft Excel are the tools utilized in this work to facilitate the evaluation of categorization outcomes using a facilitated Vector Machine (SVM). Moreover, a confusion matrix is used in the assessment of SVM classification outcomes. A table that lists the proportion of test data that are correctly and incorrectly identified is called a confusion matrix. An explanation of the assessment metrics derived from a confusion matrix is provided below.

- Accuracy: Measures how accurate the model is in predicting both classes (positive, negative, and neutral). Accuracy calculation formula:

$$\boldsymbol{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

Where:

- TP (True Positive): Number of correct positive predictions.
- TN (True Negative): Number of correct negative predictions.
- FP (False Positive): Number of incorrect positive predictions.
- FN (False Negative): Number of incorrect negative predictions.

- Precision: Indicates the proportion of true positive predictions made by the model, or how well it recognizes the positive class. Accurate computation formula:

$$\boldsymbol{Precision} = \frac{TP}{TP+FP} \tag{3}$$

- Recall: This metric indicates how well the model can identify the positive class as a whole, or how many positive classes it has properly recognized. Remember the calculating formula:

$$\boldsymbol{Recall} = \frac{TP}{TP+FN} \tag{4}$$

- F-Measure: A measure that, depending on preference, assigns equal or differing priority to memory and precision. The formula for calculating F-Measure:

$$\boldsymbol{F-Measure} = \frac{2\times Recall\times Precision}{Recall+Precision} \tag{5}$$

Table 9 shows the confusion matrix of the support vector machine (SVM):

Table 9. Matrix support vector machine

	Actual Negative (TP)	Actual Neutral (TN)	Actual Positive (TN)
Predicted Negative	2	0	0
Predicted Neutral	0	4	1
Predicted Positive	0	2	20

Table 9 indicates that the evaluation technique can be manually computed using the Support Vector Machine (SVM) algorithm using 29 data samples.

- Accuracy:

$$\textbf{Accuracy} = \frac{2+4+20}{2+0+0++0+4+1+0+2+20} = \frac{26}{29} \approx 0.90$$

This shows that the Support Vector Machine (SVM) algorithm has an accuracy of 90%

- Precision for positive class:

$$\boldsymbol{Precision} = \frac{20}{20+1+0} \approx 0.95$$

- Recall for positive class:

$$\boldsymbol{Recall} = \frac{20}{0+2+20} = 0.91$$

- F-Measure for positive class:

$$\boldsymbol{F-Measure} = \frac{2\times 0.95\times 0.91}{0.95+0.91} \approx 0.93$$

Next, to compute this evaluation technique, instructions from imported libraries, including Sklearn, Matplotlib, and Seaborn, are run with the assistance of Google Colab. This results in the following SVM algorithm classification evaluation calculation: Figure 4 concering the assessment findings:

```
Accuracy: 0.90
Classification Report:
              precision    recall  f1-score   support

    Negative       1.00      1.00      1.00         2
     Neutral       0.67      0.80      0.73         5
    Positive       0.95      0.91      0.93        22
```

Figure 4. Evaluation results

Based on these performance measures, the evaluation of the classification results using the confusion matrix thus demonstrates that the Support Vector Machine (SVM) algorithm utilized is capable of classifying data very well.

Analysis

A 90% accuracy rate for the Support Vector Machine (SVM) technique indicated that students had nice things to say about the library's services. Positive remarks included "quiet atmosphere," "very good service," and "friendly and helpful staff." This is how sentiment is distributed: Favorable Attitude (75.69%): Regarding library services, the vast majority of students left extremely excellent ratings. With remarks like "the service in the library is very good, I always find the books I'm looking for," "the atmosphere in the library is quiet," and "the staff is friendly and helpful," this result demonstrates that the majority of students who use the library are content with the services offered. This high degree of satisfaction suggests that the library has been successful in fulfilling students' requirements and expectations. Figure 5. An illustration of a favorable opinion of library services.

Figure 5. Example of positive sentiment toward library services

Negative Sentiment (6.25%): A tiny percentage of students expressed dissatisfaction, with the most common grievances being that "the book collection is incomplete" and "the library is too noisy and it is difficult to concentrate here." Even though there aren't many, the library should be aware of these unfavorable comments so that it can keep improving. Figure 6: An illustration of a negative attitude toward library services.

Figure 6. Example of negative sentiment toward library services

Neutral Sentiment: According to remarks like "limited experience" and "service was good enough but not always consistent," some students held a neutral opinion (18.06%). This opinion offers further information for bettering services.

The Support Vector Machine (SVM) algorithm performs exceptionally well in classifying student sentiment and managing the complexity and variety of comment data with 90% accuracy. The majority of students had extremely positive things to say about the library, as evidenced by the student sentiment diagram graph generated in Figure 7. Comments like "the service in the library is very good," "the atmosphere is calm," and "the staff are friendly and helpful" were common. With a percentage of 75.69%, positive sentiment predominates, compared to just 6.25% for negative sentiment and 18.06% for neutral sentiment.

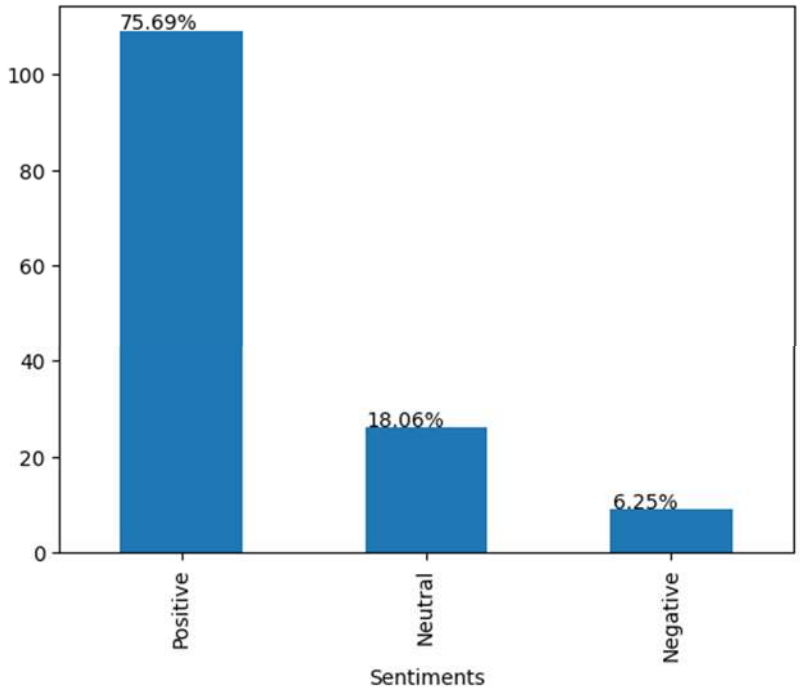

Figure 7. Student sentiment diagram

Figure 7's high percentage of positive emotion indicates that most students have a favorable opinion of the library's services. Reduced indifference and negative feelings also give libraries important information about how to keep raising the caliber and consistency of their offerings. Such sentiment research is highly helpful in comprehending consumer happiness and creating plans for effective service enhancement.

Several pertinent tools were employed in this study to assist the research and assure the validity and quality of the analysis. The primary programming language used to create the SVM algorithm, analyze data, and evaluate metrics was Python. Python programs can be written and run for free on the cloud computing platform Google Colab. Before data is entered into the SVM model, it is processed, analyzed, and visualized using Microsoft Excel. Libraries may better analyze user satisfaction, create strategies for service development, make long-term plans, and maximize resources with the aid of sentiment analysis. High user satisfaction ratings are another crucial assessment factor that helps make sure libraries stay current and sensitive to the requirements of students.

4. CONCLUSION

Using the Support Vector Machine (SVM) technique, this study examines how XYZ University's students feel about its library services. Results based on calculations done by hand and with Python tools indicate that SVM can classify student sentiment with an accuracy rate of 90%. Around 75.69% of the student answers were positive, with the majority focusing on the staff's performance, the facilities' cleanliness, and the availability of information sources. About 6.25% of respondents complained about noise and the absence of book collections, whilst 18.06% of comments were neutral. Adding and upgrading book collections, controlling noise, and giving library employees more training are all suggestions for enhancing library services. Additionally, this study creates avenues for future research, including the assessment of digital library services and the application of more advanced sentiment analysis techniques. All things considered, this survey offers insightful information on how students see library services and lays a solid basis for future service enhancements. By following these suggestions, the library at XYZ University will be able to maintain its current level of service excellence and assist with the academic pursuits of its students.

REFERENCE

[1] K. A. Barfi, S. K. Parbie, C. K. Filson, M. V. Teye, K. Kodua-Ntim, and E. Ayensu, "*Assessing the quality of services at an academic library*," Heliyon, vol. 9, no. 12, Dec. 2023, doi: 10.1016/j.heliyon.2023.e22449.

[2] J. Kociubuk and R. Willett, "'*Youth service librarians…are not just providers of books': Caregivers' perceptions of the value of public library services during the* COVID-19 *pandemic*," Libr Inf Sci Res, vol. 45, no. 3, Jul. 2023, doi: 10.1016/j.lisr.2023.101256.

[3] S. Sediyaningsih, M. P. Ristiyono, K. Launggu, and P. Ochieng Juma, "*De-contextual communication: Factors influencing usage intentions of metaverse technology in digital library services*," Heliyon, vol. 9, no. 10, Oct. 2023, doi: 10.1016/j.heliyon.2023.e20388.

[4] Ade Ajie Ferizal, Mohamad Anas Sobarnas, and Djoko nursanto, "Sistem Informasi Perpustakaan Berbasis Web di SMK Fatahillah Cileungsi," INFOTECH : Jurnal Informatika & Teknologi, vol. 2, no. 2, pp. 104–113, Dec. 2021, doi: 10.37373/infotech.v2i2.178.

[5] E. R. Lidinillah, T. Rohana, and A. R. Juwita, "Analisis sentimen twitter terhadap *steam* menggunakan algoritma *logistic regression* dan *support vector machine*," TEKNOSAINS : Jurnal Sains, Teknologi dan Informatika, vol. 10, no. 2, pp. 154–164, Jul. 2023, doi: 10.37373/tekno.v10i2.440.

[6] R. Martin-Melon, T. Hernández-Pérez, and S. Martínez-Cardama, "*Research data services* (RDS) *in Spanish academic libraries*," *Journal of Academic Librarianship*, vol. 49, no. 4, Jul. 2023, doi: 10.1016/j.acalib.2023.102732.

[7] A. Surya et al., "Analisis Kepuasan Pelanggan Pada Sekolah Tinggi Teknologi Muhammadiyah Cileungsi-Bogor," 2020. [*Online*]. *Available*: http://jurnal.sttmcileungsi.ac.id/index.php/tekno

[8] Y. Riady, M. Sofwan, M. Mailizar, T. M. Alqahtani, L. N. Yaqin, and A. Habibi, "*How can we assess the success of information technologies in digital libraries? Empirical evidence from Indonesia*," *International Journal of Information Management Data Insights*, vol. 3, no. 2, Nov. 2023, doi: 10.1016/j.jjimei.2023.100192.

[9] T. H. Wicaksono, Imam Yuadi, and Ira Puspitasari, "*Opinion mining toward work from office policies on post-pandemic* covid-19 *by using supervised learning*," TEKNOSAINS : Jurnal Sains, Teknologi dan Informatika, vol. 11, no. 1, pp. 1–7, Jan. 2024, doi: 10.37373/tekno.v11i1.525.

[10] N. L. Ratniasih, N. W. N. Jayanti, and ..., "Klasifikasi Kepuasan Mahasiswa Menggunakan Algoritma *Support Vector Machine* dan Metode *Stemming* Sastrawi," Prosiding CORISINDO …, pp. 373–378, 2023, [*Online*]. *Available*: https://stmikpontianak.org/ojs/index.php/corisindo/article/view/357%0Ahttps://stmikpontianak.org/ojs/index.php/corisindo/article/download/357/90

[11] I. Nur Fakhri and R. Febrian Umbara, "Analisis Sentimen pada Kuisioner Kepuasan Terhadap Layanan dan Fasilitas Kampus Universitas Dengan Menggunakan Klasifikasi *Support Vector Machine* (SVM)".

[12] H. Hermanto, A. Mustopa, and A. Y. Kuntoro, "Algoritma Klasifikasi *Naive Bayes* Dan *Support Vector Machine* Dalam Layanan Komplain Mahasiswa," JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer), vol. 5, no. 2, pp. 211–220, 2020, doi: 10.33480/jitk.v5i2.1181.

[13] M. Riko Anshori Prasetya and A. Mudi Priyatno, "Penanganan Imputasi *Missing Values* pada Data *Time Series* dengan Menggunakan Metode *Data Mining*", doi: 10.37034/jidt.v5i1.324.

[14] A. Tawakuli, B. Havers, V. Gulisano, D. Kaiser, and T. Engel, "*Survey:Time-series data preprocessing: A survey and an empirical analysis," Journal of Engineering Research* (Kuwait), 2024, doi: 10.1016/j.jer.2024.02.018.

[15] A. Carla and H. Soetanto, "Implementasi *Text Mining* untuk Analisis Sentimen Pada Pengguna PLN *Mobile* Mengunakan Metode *Naïve Bayes*," 2024.

[16] N. Nofiyani and W. Wulandari, "*Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining*," Jurnal Media Informatika Budidarma, vol. 6, no. 3, p. 1621, Jul. 2022, doi: 10.30865/mib.v6i3.4332.

[17] Friska Aditia Indriyani, Ahmad Fauzi, and Sutan Faisal, "Analisis Sentimen Aplikasi Tiktok Menggunakan Algoritma *Naïve Bayes* dan *Support Vector Machine*," TEKNOSAINS : Jurnal Sains, Teknologi dan Informatika, vol. 10, no. 2, pp. 176–184, Jul. 2023, doi: 10.37373/tekno.v10i2.419.

[18] J. Dalle, D. Hastuti, and M. R. A. Prasetya, "*The Use of an Application Running on the ant Colony Algorithm in Determining the Nearest Path Between Two Points," Journal of Advances in Information Technology*, vol. 12, no. 3, pp. 206–213, Aug. 2021, doi: 10.12720/jait.12.3.206-213.

[19] S. Y. Syn, D. Sinn, and S. Kim, "*Innovative Public Library Services During the COVID-19 pandemic: Application and Revision of Social Innovation Typology*," Libr Inf Sci Res, vol. 45, no. 3, Jul. 2023, doi: 10.1016/j.lisr.2023.101248.

[20] G. Gustientiedina, M. Siddik, and Y. Deselinta, "Penerapan *Naïve Bayes* untuk Memprediksi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademis," Jurnal Infomedia, vol. 4, no. 2, p. 89, 2020, doi: 10.30811/jim.v4i2.1892.

[21] A. S. Handayani, S. Soim, T. E. Agusdi, Rumiasih, and A. Nurdin, "Klasifikasi Kualitas Udara Dengan Metode *Support Vector Machine*," JIRE (Jurnal Informatika & Rekayasa Elektronika), vol. 3, no. 2, pp. 187–199, 2020.