IT FOR SOCIETY, Vol. 09, No. 01
ISSN 2503-2224

SENTIMENT ANALYSIS OF STUDENT SATISFACTION TOWARDS DISTANCE LEARNING USING MACHINE LEARNING METHOD

M.Andres
Faculty of Computing
President University
Cikarang, Indonesia
andres@president.ac.id

Tjong Wan Sen
Faculty of Computing
President University
Cikarang, Indonesia
wansen@president.ac.id

Rusdianto Roestam
Faculty of Computing
President University
Cikarang, Indonesia
rusdianto@president.ac.id

***Abstract*— The Covid-19 pandemic forces the entire society to change their way of life. One of them is the process of face-to-face learning changing into distant learning. Various responses arise from students during the implementation of this new system, both positive and negative, indicating the level of student satisfaction. The sentiment analysis of students' comments during distance learning was conducted using machine learning algorithms and tools Rapid miner. Literature study shows that the Naive Bayes, K-NN, and Decision Tree algorithms have very high accuracy, so this research uses those methods to get high-accuracy results. The research shows the following results; Naive Bayes is 93.80% and class precision for pred. Positive 93.80% and pred. negative 100.00%. The K-NN algorithm is 92.49% and class precision for pred. positive is 92.37%, pred. negative 100%. The Decision Tree method is 90.81% with a standard deviation of (+-) 0.58 and class precision for pred. positive 90.81% and class pred. negative 0.00%.**

Keywords—distance learning, sentiment analysis, Naive Bayes, K-NN, Decision Tree, and Rapid Miner

I. INTRODUCTION

The usual way of teaching and learning in universities is face-to-face, where teachers and students are in the same classroom. The teacher delivers the lesson directly to the students with the help of a projector to display the material in PowerPoint or PDF format for easy understanding. Students can communicate directly with the teacher who is in front of them and the teacher can easily see the classroom situation and the condition of the students attending the class. In 2019, the Covid-19 pandemic began to spread from China to various countries. In March 2020, the first confirmed case of COVID-19 was reported in Depok, West Java, Indonesia, and it continued to spread, causing widespread concern and fear among people to engage in public activities, including in universities.

Due to concern and fear, the government implemented the Community Activity Restrictions abbreviated as PPKM. All activities that involve physical contact and many people gathering in one place are restricted by government regulations to control the spread of the COVID-19 virus. To continue the teaching and learning process, universities generally apply distance learning systems. The distance learning during the pandemic was established by the government or PPKM. Learning and teaching that is usually done in the classroom with face-to-face interactions must be temporarily replaced by remote or distance learning. This paper proposes a solution: implementing a personalized learning path, allowing students to choose their studies In this distance learning process, there are many challenges and obstacles for both teachers and students. This learning method is new for both teachers and students, with both positive and negative aspects, difficulties and conveniences for teachers and students. The hope is that the lessons delivered by teachers can be well received by students. To find out the effectiveness of the distance learning process for students, the author conducted this research. The aim is for educators and lecturers at President University to evaluate and take steps and solutions so that the distance learning process can be even better if distance learning is carried out in the future. Sentiment analysis is conducted using Rapid Miner, which is an open-source tool. The usual way of teaching and learning in universities is face-to-face, where teachers and students are in the teacher deliver the lesson directly to the students with the help of a projector to display the material in PowerPoint or PDF format for easy understanding. Students can communicate directly with the teacher who is in front of them and the teacher can easily see the classroom situation and the condition of the students attending the class. In 2019, the Covid-19 pandemic began to spread from China to various countries. In March 2020, the first confirmed case of COVID-

19 was reported in Depok, West Java, Indonesia, and it continued to spread, causing widespread concern and fear among people to engage in public activities, including in universities. Due to the concern and fear, the government implemented the Community Activity Restrictions or abbreviated as PPKM. All activities that involve physical contact and many people gathering in one place are restricted by government regulations to control the spread of the COVID-19 virus. To continue the teaching and learning process, universities generally apply distance learning systems. The distance learning during the pandemic was established by the government or PPKM. Learning and teaching that is usually done in the classroom with face-to-face interactions must be temporarily replaced by remote or distance learning. In this distance learning process, there are many challenges and obstacles for both teachers and students. This learning method is new for both teachers and students, with both positive and negative aspects, difficulties and conveniences for teachers and students. The hope is that the lessons delivered by teachers can be well received by students. To find out the effectiveness of the distance learning process for students, the author conducted this research. The aim is for educators and lecturers at President University to evaluate and take steps and solutions so that the distance learning process can be even better if distance learning is carried out in the future. Sentiment analysis is conducted using Rapid Miner, which is an open-source tool.

II. RESEARCH METHODOLOGY

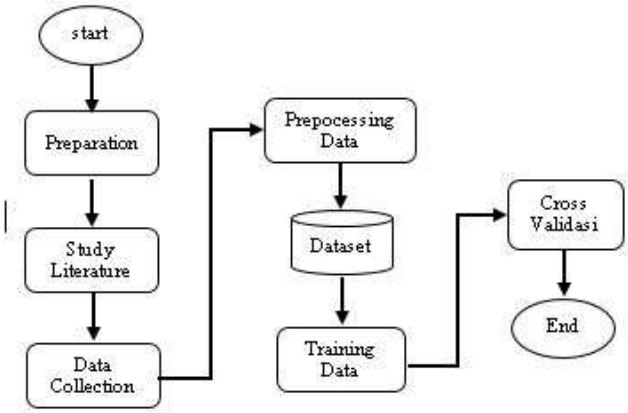


Figure 1.1 Research Methodology

Step to Research Methodology :

1. Preparation, background creation, problem identification, problem formulation, objectives, problem limitations, research benefits, and novelty.
2. Study literature is taken from journal websites such as Google Scholar, Garuda, and Indexed Sinta, and others. The literature is reviewed as a basis for searching for a research topic.
3. Data collection, namely data collection taken from student comment data on campus (PUIS).
4. Data preprocessing, namely discarding.

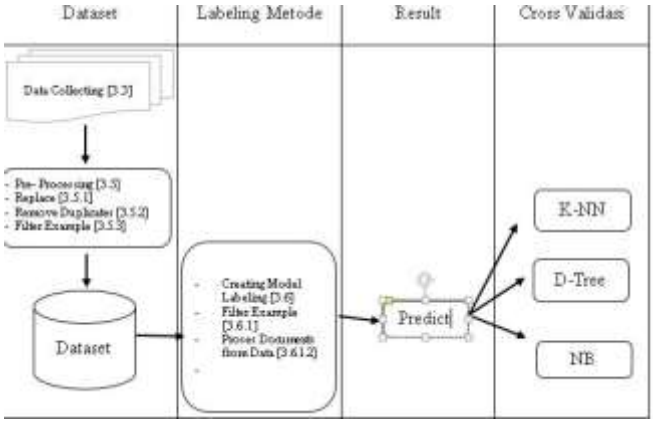


Figure 1.2

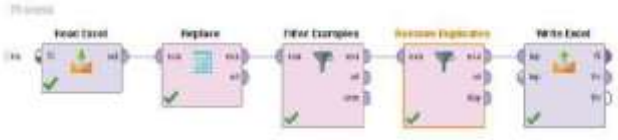
1. In this research, data in the form of student comments was used
2. Labeling uses the Naive Bayes algorithm
3. Results from sentiment analysis labeling
4. Cross Validation The validation process uses three algorithms, namely: Naive Bayes, K-NN and Decision Tree.

2.1 COLLECTING DATA

Figure 2.1 Data Collecting

1. The dataset was downloaded from the campus information system which contains data for 2020 when the distance learning process was held.
2. Data for 2020 was collected and used as training and testing data. The data contains student comments which will be processed into information that has sentiment value.

2.2 PPRE-PROCESSING



2.3 PRE PROCESSING

a. Replace
b. Remove Duplicates
c. Filter Example

2.3 LABELLING PROCCESS

Figure 2.4 labeling process

2.4 Cross Validation

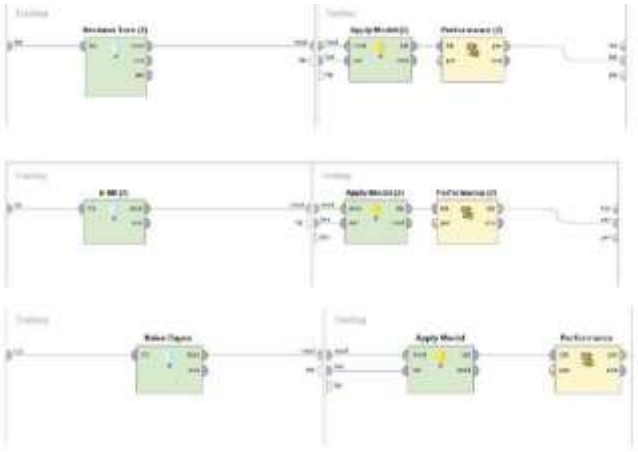


Figure 2.5 Cross Validation Algorithm

III. RESULT AND DISCUSS

Based on the results of the research, the following results were obtained. Naive Bayes is 93.80% with a standard deviation (+/-) of 1.50% and class precision for pred. Positive 93.80% and pred. negative 100.00%. Recal class for true positive 100.00%, true negative 34.69%. The K-NN algorithm method is 92.49% with standard deviation (+/-) - 1.97% and class precision for pred.positive is 92.37%, pred.negative is 100%. Recall class for true positive 100% and true negative 18.37%. Decision Tree method 90.81% with standard deviation (+-) 0.58 and class precision for pred. positive 90.81% and pred.negative class 0.00%. Recall class for true positive 100.00% and true negative 0.00%.

IV. CONCLUSIONS

Based on the results of the research with Methodology Algoritmn Naive Bayes, K-NN and Decision Tree. Naive Bayes 93,8% accurate, K-NN 92,49% and Decision Tree 90.81%. Standar Devisition Naive Bayes 1,50%, K-NN 1.97% and Decision Tree 0,58%.

REFERENCES

[1] Amra, I. A., & Maghari, A. Y. (2017). Student Performance Prediction Using KNN and Naïve Bayesian. 8th International Conference on Information Technology (ICIT). Amman: IEEE.

[2] An, Y., Sun, S., & Wang, S. (2017). Naïve Bayes Classifier for Music Emotion Classification Based on Lyrics. *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS).* Wuhan: IEEE.

[3] Anggraini, S., Akbar, M., Wijaya, A., Syaputra, H., & Sobri, M. (2021). Klasifikasi Gejala Penyakit Coronavirus Disease 19 (COVID-19) Menggunakan Machine Learning. Journal of Software Engineering Ampera.

[4] Arista, A. (2022). Comparison Decision Tree and Logistic Regression Machine Learning Classification Algorithms to determine Covid-19. *Sinkron : Jurnal dan Penelitian Teknik Informatika*.

[5] Azizah, E. N., Wibawa, A. P., & Pujianto, U. (2019). *Perbandingan Algoritma Naive Bayes dan C4.5 Untuk Prediksi Kelulusan Mahasiswa Dalam Ruang Belajar Virtual.* Malang: Universitas Negeri Malang.

[6] Barus, S. P. (2022). Penerapan Model Decision Tree pada Machine Learning untuk Memprediksi Calon Potensial Mahasiswa Baru. *Jurnal Ikraith Informatika*

[7] Budi, S. (2017). Text Mining Untuk Analisis Sentimen Review Film Menggunakan Algoritma K-Means; Text Mining For Movie Review Sentiment Analysis Using K-Means Algorithm. *Techno.COM*.

[8] Chandani, V., Wahono, R. S., & Purwanto. (2015). Komparasi Algoritma Klasifikasi Machine Learning Dan Feature Selection pada Analisis Sentimen Review Film. *Journal of Intelligent Systems*.

[9] Firmansyah, & Yulianto, A. (2021). Machine Learning with Decision Tree for Predict Invoice Payment, Case Study: Gramedia Jakarta. *Journal of Informatics and Telecommunication Engineering*.

[10] Fitrianah, D., Dwiasnati, S., H, H. H., & Baihaqi, K. A. (2021). Penerapan Metode Machine Learning untuk Prediksi Nasabah Potensial menggunakan Algoritma Klasifikasi Naïve Bayes. *Faktor Exacta*.

[11] Fitriyyah, S. N., Safriadi, N., & Pratama, E. E. (2019). Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes. *Jurnal Edukasi dan Penelitian Informatika*.

[12] Ipmawati, J., Kusrini, & Luthfi, E. T. (2017). Komparasi Teknik Klasifikasi Teks Mining Pada Analisis Sentimen. *Indonesian Journal on Networking and Security*, 28.

[13] Izati, N. A., Warsito, B., & Widiharih, T. (2019). Prediksi Harga Emas Menggunakan Feed Forward Neural Network Dengan Metode Extreme Learning Machine. *JURNAL GAUSSIAN*.

[14] Khoeri, I., & Mulyana, D. I. (2021). Implementasi Machine Learning dengan Decision Tree Algoritma C4.5 dalam Penerimaan Karyawan Baru pada PT. Gitareksa Dinamika Jakarta . *Jurnal Sosial dan Teknologi (SOSTECH)* , 615.

[15] Kurniawan, R., & Apriliani, A. (2020). Analisis Sentimen Masyarakat Terhadap Virus Corona Berdasarkan Opini Dari Twitter Berbasis Web Scrapper. *Jurnal Informatika Sains dan Teknologi*.

[16] Lestari, A. R., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tentang Opini Pilkada Dki 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes dan Pembobotan Emoji. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.

[17] Najar, A. M., Sudarsana, I. W., Albab, M. U., & Andhika, S. (2022). Machine Learning untuk Identifikasi Jenis Kanker Darah (Leukemia). *Vygotsky: Jurnal Pendidikan Matematika dan Matematika*.

[18] Nasution, M. R., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *Jurnal Informatika*.

[19] Oktaviana, N. E., Sari, Y. A., & Indriati. (2019). Analisis Sentimen Terhadap Kebijakan Kuliah Daring Selama Pandemi Menggunakan Pendekatan Lexicon Based Features Dan Support Vector Machine. *Jurnal Teknologi Informasi dan Ilmu Komputer*.

[20] Pudjajana, A. M., & Manongga, D. (2018). Sentimen Analisis Tweet Pornografi Kaum Homoseksual Indonesia Di Twitter Dengan Naive Bayes. *Jurnal SIMETRIS*.
[21] Puspita, R., & Widodo, A. (2020). Perbandingan Metode KNN, Decision Tree, dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS. *Jurnal Informatika Universitas Pamulang*.
[22] Rahman, M., Ahmed, M. T., Nur, S., & Islam, A. Z. (2022). The prediction of Coronavirus Disease 2019 Outbreak on Bangladesh Perspective Using Machine Learning: A Comparative Study. *International Journal of Electrical and Computer Engineering (IJECE)*.
[23] Ruhyana, N. (2019). Analisis Sentimen Terhadap Penerapan Sistem Plat Nomor Ganjil/Genap Pada Twitter Dengan Metode Klasifikasi Naive Bayes. *Jurnal IKRA-ITH Informatika*.
[24] Samsir, Ambiyar, V, U., E, F., & W, R. (2021). Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes. *Jurnal Media Informatika Budidarma*.
[25] Sari, D. N., Adelia, F., Rosdiana, F., Butar, B. B., & Hariyanto, M. (2020). Analisa Sentimen Terhadap Review Produk Kecantikan Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika Universitas Muhammadiyah Tangerang*.
[26] Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online JD.ID Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi. *SIMETRIS*.
[27] Simamora, R. J., Tibyani, & Sutrisno. (2019). Peramalan Curah Hujan Menggunakan Metode Extreme Learning Machine. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.
[28] Sipayung, E. M., Maharani, H., & Zefanya, I. (2016). Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Informasi*.
[29] Wardani, N. R., & Erfina, A. (2021). Analisis Sentimen Masyarakat Terhadap layanan Konsultasi Dokter Menggunakan Algoritma Naive Bayes. *SISMATIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika)* (p. 11). Sukabumi: Universitas Nusa Putra.
[30] Wiyono, S., & Abidin, T. (2018). Perbandingan Algoritma Machine Learning SVM dan Decision Tree untuk Prediksi Keaktifan Mahasiswa. *Jurnal Dan Penelitian Teknik Informatika*.