SENTIMENT ANALYSIS OF PUBLIC OPINION REGARDING FUEL OIL ON TWITTER BY COMPARING CLASSIFICATION ALGORITHMS

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer pada
Program Studi Sistem Informasi

Oleh:

GITA WIDARMA
11950311555

FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2023

LEMBAR PERSETUJUAN

SENTIMENT ANALYSIS OF PUBLIC OPINION REGARDING FUEL OIL ON TWITTER BY COMPARING CLASSIFICATION ALGORITHMS

TUGAS AKHIR

Oleh:

GITA WIDARMA
11950311555

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 20 Juni 2023

Ketua Program Studi	**Pembimbing**
Eki Saputra, S.Kom., M.Kom.	**Dr. Rice Novita, S.Kom., M.Kom.**
NIP. 198307162011011008	**NIK. 130510011**

LEMBAR PENGESAHAN

SENTIMENT ANALYSIS OF PUBLIC OPINION REGARDING FUEL OIL ON TWITTER BY COMPARING CLASSIFICATION ALGORITHMS

TUGAS AKHIR

Oleh:

GITA WIDARMA
11950311555

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 15 Juni 2023

Pekanbaru, 15 Juni 2023
Mengesahkan,

Dekan

Dr. Hartono, M.Pd.
NIP. 196403011992031003

Ketua Program Studi

Eki Saputra, S.Kom., M.Kom.
NIP. 198307162011011008

DEWAN PENGUJI:

Ketua : **Arif Marsal, Lc., MA.**

Sekretaris : **Dr. Rice Novita, S.Kom., M.Kom.**

Anggota 1 : **Mustakim, ST., M.Kom.**

Anggota 2 : **Nesdi Evrilyan Rozanda, S.Kom., M.Sc.**

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada peneliti. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin peneliti dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman dan tanggal pinjam pada *form* peminjaman.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis di- acu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 15 Juni 2023
Yang membuat pernyataan,

GITA WIDARMA
NIM. 11950311555

LEMBAR PERSEMBAHAN

Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Alhamdulillahi Rabbil 'Alamin, segala puji bagi Allah *Subhanahu Wa Ta'ala* sebagai bentuk rasa syukur atas segala nikmat yang telah diberikan tanpa ada kekurangan sedikitpun. Shalawat beserta salam tak lupa pula kita ucapkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan *"Allahumma Sholli'ala Sayyidina Muhammad Wa'ala Ali Sayyidina Muhammad"*. Semoga kita semua selalu senantiasa mendapat syafaat-Nya di dunia maupun di akhirat, *aamiin ya rabbal'aalamiin*.

Saya persembahkan hadiah istimewa karya kecil ini sebagai salah satu bentuk bakti, rasa terima kasih, dan hormat kepada Ayah dan Ibu tercinta. Terima kasih yang tak terhingga karena telah merawat dan membesarkan saya dengan setulus hati dan penuh perjuangan hingga saya bisa sampai tahap saat ini. Berkat doa dan kasih sayangmu, anakmu telah berhasil memperoleh gelar sarjana seperti yang engkau harapkan. Tiada apapun di dunia ini yang dapat membalas semua jasa-jasa dan pengorbananmu. Untuk itu saya anakmu ini selalu mendoakan yang terbaik untuk Ayah dan Ibu agar bahagia dunia dan akhirat serta diberikan tempat istimewa di sisi-Nya kelak. Dan pastinya saya juga berterima kasih yang tak terhingga kepada saudara kandung tercinta saya yaitu Abang dan Adik yang telah memberikan saya pelajaran dan pemahaman mengenai indahnya kehidupan yang damai sebagai saudara.

Kepada Bapak dan Ibu Dosen Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau yang telah memberikan ilmu pengetahuan, pengalaman, dan kebaikan selama perkuliahan, saya ucapkan terima kasih banyak dan semoga menjadi amal jariyah. *Aamiin.*

Untuk sahabat terdekat yang tidak bisa saya sebutkan satu persatu dan pastinya juga teman-teman seperjuangan, terima kasih berkat kalian masa perkuliahan menjadi lebih bermakna semoga dimasa mendatang kita bisa bertemu lagi dalam keadaan yang lebih baik.

Wassalamu'alaikum Warahmatullahi Wabarakaatuh.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Alhamdulillahi Rabbil 'Alamin, bersyukur kehadirat Allah *Subhanahu Wa Ta'ala* atas segalarahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan Tugas Akhir ini. Tidak lupa sholawat beriringan salam selalu tercurahkan untuk Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan melantunkan *Allahumma Sholli'ala Sayyidina Muhammad Wa'ala Alihi Muhammad*. Tugas Akhir ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Sarjana Komputer di Program Studi Sistem Informasi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Banyak pemangku kepentingan telah berperan dalam mendukung dan membimbing peneliti pada proses penelitian dan penulisan Tugas Akhir ini. Maka dari itu, ungkapan terima kasih juga peneliti ucapkan kepada:

1. Bapak Prof. Dr. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi.
4. Ibu Siti Monalisa, S.T., M.Kom sebagai Sekretaris Program Studi Sistem Informasi.
5. Ibu Dr. Rice Novita, S.Kom., M.Kom sebagai Dosen Pembimbing Tugas Akhir ini yang telah banyak meluangkan waktu dan memberikan masukan, nasehat, serta motivasinya baik itu dalam penyelesaian Tugas Akhirini.
6. Bapak Arif Marsal, LC., MA sebagai Ketua Sidang peneliti yang telah banyak memberikan arahan, masukan, nasihat serta motivasi dalam penyelesaian Tugas Akhir ini juga dalam perkuliahan.
7. Bapak Mustakim, S.T., M.Kom sebagai Penguji I peneliti yang telah banyak memberikan arahan, masukan, nasihat serta motivasi dalam penyelesaian Tugas Akhir ini juga dalam perkuliahan.
8. Bapak Nesdi Evrilyan Rozanda, S.Kom., M.Sc sebagai Penguji II peneliti yang telah banyak memberikan arahan, nasihat, masukan serta motivasi dalam penyelesaian Tugas Akhir ini juga dalam perkuliahan.
9. Bapak Anofrizen, S.Kom., M.Kom sebagai Dosen Pembimbing Akademik peneliti yang telah banyak memberikan arahan, masukan, dan motivasi selama perkuliahan mulai dari semester 1 hingga semester 8 ini.
10. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom sebagai Kepala Laboratorium Program Studi Sistem Informasi.

11. Seluruh Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah banyak memberikan ilmunya kepada peneliti. Semoga ilmu yang diberikan dapat peneliti amalkan dan menjadi amal jariyah.
12. Seluruh Pegawai dan Staf Fakultas Sains dan Teknologi yang telah membantu dan mempermudah proses administrasi selama perkuliahan ini.
13. Kedua orang tua peneliti, yaitu Ayah Syamsidar dan Ibu Widiastuti tercinta yang tanpa lelah selalu memberikan semangat, motivasi, *support*, serta doa terbaiknya dan selalu menjadi motivasi peneliti dalam menyelesaikan Strata 1 (S1) ini.
14. Abang Andri Widarma, Adik Imam Al-Buqhari, Kakak Ipar Nova Ramadhani, dan Keponakan Al-Shaki Zehan Widarma tercinta. Terimakasih telah memberikan perhatian, semangat, *support* serta doa kepada peneliti.
15. Sahabat-sahabat tercinta peneliti terutama dari Kost Revolusioner, yaitu Aji Nursalim, Haryanto, Indra Saputra Purba, Jefri Gustian, Muhammad Risky Agustin, Syafri Sirait, dan Yogi Saputra.
16. Keluarga besar Sistem Informasi 19 khususnya teman-teman dari *Premium Class*, KKN Desa Dusun Pusaka 2022, serta teman seperbimbingan Tugas Akhir ini.
17. Semua pihak yang namanya tidak dapat disebutkan satu persatu, yang telah banyak membantu dalam pelaksanaan serta penyelesaian Tugas Akhir ini.

Semoga segala doa dan dorongan yang telah diberikan selama ini menjadi amal kebajikan dan mendapat balasan setimpal dari Allah *Subhanahu Wa Ta'ala*. Peneliti menyadari bahwa penulisan Tugas Akhir ini masih banyak terdapat kekurangan dan jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangat diharapkan untuk kesempurnaan Tugas Akhir ini dan semoga Laporan Tugas Akhir ini bermanfaat bagi kita semua. Akhir kata peneliti ucapkan terima kasih.

Wassalamu'alaikum Warahmatullahi Wabarakaatuh.

Pekanbaru, 20 Juni 2023

Penulis,

GITA WIDARMA

NIM. 1195031155

Gmail GITA WIDARMA <11950311555@students.uin-suska.ac.id>

[ISITIA 2023] Your paper #1570900516 ('Comparison of Machine Learning Classification Algorithms and Feature Selection Particle Swarm Optimization on Public Opinion Regarding Fuel Oil')

ISITIA 2023 (isitia@its.ac.id) <isitia=its.ac.id@edas.info> Sen, 15 Mei 17.46
Balas-Ke: ISITIA 2023 <isitia@its.ac.id>
Ke: Gita Widarma <11950311555@students.uin-suska.ac.id>, Rice Novita <rice.novita@uin-suska.ac.id>, Mustakim Mustakim <mustakim@uin-suska.ac.id>, Nesdi Evrilyan Rozanda <nesdi.rozanda@uin-suska.ac.id>
Cc: Muhammad Attamimi <attamimi@ee.its.ac.id>, Vita Lystianingrum <vita@ee.its.ac.id>, Prasetiyono Hari Mukti <prasetiyono.hm@gmail.com>, Dimas Fajar Uman Putra <dimasfup@ee.its.ac.id>

Dear Mr. Gita Widarma,

Congratulations!

We are pleased to inform you that your manuscript:
Paper ID: 1570900516
Paper Title: Comparison of Machine Learning Classification Algorithms and Feature Selection Particle Swarm Optimization on Public Opinion Regarding Fuel Oil
Authors: Gita Widarma, Rice Novita, Mustakim Mustakim and Nesdi Evrilyan Rozanda
has now been ACCEPTED for presentation in 2023 International Seminar on Intelligent Technology and Its Applications (ISITIA).

The reviews are included below, and can also be found at https://edas.info/showPaper.php?m=1570900516 using your EDAS login name 11950311555@students.uin-suska.ac.id.

Please find the following details on your final manuscript preparation, registration, as well as video submission.

For the final manuscript:
1. You must revise your manuscript according to the comments and suggestions given by the reviewers (see below).
2. Ensure that you have added all the author names, both in your manuscript as well as in EDAS in the same sequence.
3. You are required to ensure that the similarity rating of your manuscript is less than 30%. Submissions with a similarity rating of 30% or above will be dropped and no longer considered for presentation in ISITIA 2023.
4. Please ensure that your manuscript has been checked for grammatical and typographical errors, and that your manuscript is IEEE pdf eXpress compliant.
5a. A formatting checklist is given here to help you prepare your manuscript before uploading to IEEE pdf eXpress. You should also upload the confirmed checklist during the registration process.
5b. The details on how to use IEEE pdf eXpress to generate your PDF file can be found here: https://elib.its.ac.id/conf/isitia/main/home/submissions/.
6. You are required to transfer copyright of the manuscript to IEEE if you wish your manuscript to be published in IEEExplorer. This can be done electronically through EDAS.
7. IEEE reserves the right to exclude a paper from distribution after the conference (e.g. removal from IEEE Xplore) if the paper is not presented at the conference.

To register, please go to the following address: https://elib.its.ac.id/conf/isitia/main/registration/

Video upload details:
A recorded live online presentation is required for IEEExplore submission, however, we also ask a pre-recorded video for backup as suggested by the IEEE. The instruction on the video requirement can be found in: https://elib.its.ac.id/conf/isitia/main/2021/05/10/pre-recorded-presentation-video-instruction/

Thank you very much.

Sincerely yours,

Sentiment Analysis of Public Opinion Regarding Fuel Oil on Twitter by Comparing Classification Algorithms

1st Gita Widarma
Student of Information Systems
Universitas Islam Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
11950311555@students.uin-suska.ac.id

2nd Rice Novita
Department of Information Systems
Universitas Islam Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
rice.novita@uin-suska.ac.id

3rd Mustakim
Department of Information Systems
Universitas Islam Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
mustakim@uin-suska.ac.id

4th Nesdi Evrilyan Rozanda
Department of Information Systems
Universitas Islam Sultan Syarif Kasim Riau
Pekanbaru, Indonesia
nesdi.rozanda@uin-suska.ac.id

Abstract- The increase in the price of fuel oil (BBM) has led to various responses or opinions from the public, one of which is conveying their aspirations through social media, namely Twitter. In a tweet, it only consists of a few fragments of words which certainly contain useful information. Text mining is one way to overcome this problem because text mining has analysis that focuses on analyzing an opinion or comment that is extracted into sentiment information whether the information is positive or negative. In this study, three classification algorithms are used most often in data mining, namely Decision Tree, K-Nearest Neighbor, and Naïve Bayes. The purpose of this study is to look at the results of sentiment from public opinion about rising fuel prices on Twitter and to compare the performance of the Decision Tree, K-Nearest Neighbor, and Naïve Bayes Algorithms. This study also used the Particle Swarm Optimization (PSO) Feature Selection to improve the results of the accuracy of the algorithm used. The data obtained is 5000 tweet data consisting of 2165 negative sentiment tweet data and 2835 positive sentiment tweet data. The data is tested based on different amounts of data, namely 1000 data, 3000 data, and 5000 data. The result of the best algorithm when tested using 1000 data is the classic K-Nearest Neighbor algorithm with a value of 75.60%. While the best accuracy value in the test (3000 Data) is the PSO-based K-Nearest Neighbor algorithm with a value of 78.21%. Meanwhile, in testing (5000 Data) the classic K-Nearest Neighbor algorithm is the best with an accuracy value of 74.02%.

Keywords: Sentiment, Fuel Oil, Decision Tree, K-Nearest Neighbor, Naïve Bayes, Particle Swarm Optimization (PSO)

I. INTRODUCTION

Indonesia is a country that has abundant natural wealth, all of which can be proven by the large reserves of oil, gas, nickel, coal, and others. Even with this abundant natural wealth, there are still many Indonesian people who live on the poverty line[1][2]. Based on data from the BP Statistical Review 2022, Indonesia has succeeded in producing 692 thousand barrels of oil per day and contributing 0.8% of world oil production which also ranks 24th as the largest oil-producing country in the world. Petroleum is dubbed as the mother of various commodities because petroleum has a large and important role in human survival[3].

PT. Pertamina is a state-owned company engaged in the energy sector and the largest dividend contributor in Indonesia. Currently, one of the products provided by the company has experienced a significant increase in price, namely fuel oil (BBM). Pertalite is the fuel most used by drivers and will reach 23 million kilos in 2021. This will certainly have an impact on the burden on people's lives in Indonesia because petroleum can cause inflation or drive up prices widely[4][5].

This situation has the potential to generate various responses or opinions from the general public, one of which is conveying their aspirations via social media Twitter. Twitter is one of the most widely used applications by the people of Indonesia because Twitter is considered easy to use and is known to be fast in spreading information[6]. In Indonesia, there are 19.5 million Twitter users out of a total of 500 million global users. Based on the results of crawling data that was carried out on 27 December 2022 - 27 January 2023, there were 109,470 data that discussed "BBM". With so many tweets or comments discussing the fuel issue, the authors are interested in conducting research on sentiment analysis on the effects of rising fuel prices. Public opinion on a tweet usually only consists of fragments of words whose meanings are difficult to understand. Thus of course, there is a lot of valuable data, so the role of science is very necessary[7][8].

Text mining is a technique for analyzing a data source and classifying it by looking for interesting patterns or variations of the data. Text mining also has an analysis that focuses on analyzing an assessment of one's opinions, attitudes, and emotions regarding a topic, be it an individual, organization, or product, namely sentiment analysis[9][10]. Sentiment analysis or opinion mining is a method for analyzing an opinion or comment that is extracted into sentiment information whether the information is positive or negative[11][12].

Research on sentiment analysis that focuses on fuel has been carried out a lot, one of which is research conducted by [13]. This study, discusses the impact of rising fuel prices on the economy of the lower middle class in Indonesia and tests the random forest algorithm.

Several machine learning methods that are often used in the application of sentiment analysis, such as K-Nearest Neighbor, Decision Tree, Naïve Bayes, Support Vector Machine, Random Forest, and others. Some previous studies that conducted research on the classification of sentiment, including research conducted by [14] also did a comparison of the K-Nearest Neighbor, Decision Tree, and Naïve Bayes algorithms which used 760 data that predicted student thesis graduation at Budi Luhur University. The results of this study show the accuracy performance of the K-Nearest Neighbor algorithm obtained the highest score of 80.39%, while Naïve Bayes obtained an accuracy value of 73.16% and the Decision Tree algorithm had an accuracy value of 75.00%.

While research conducted by [15] uses data from Twitter that discusses e-commerce, namely Bukalapak and Tokopedia by comparing the use of the Decision Tree, K-Nearest Neighbor, and Naïve Bayes algorithms. The purpose of this study is to find the best accuracy of the algorithm used. The highest accuracy results in this study were the Decision Tree algorithm of 80.00%, while K-Nearset Neighbor obtained a value of 78.00% and Naïve Bayes of 77.00%.

Research discussing sentiment analysis was also carried out by [16] with the research title *"Comparative analysis of Naïve Bayes, K Nearest Neighbor and C.45 method in weather forecast"* using 1442 data tweets and 8 attributes. The results of this study are the K-Nearest Neighbor algorithm with parameters k=7 and fold=5 to obtain the highest accuracy value of 71.58%, while the Decision Tree algorithm with fold=20 obtains an accuracy of 69.83% and the lowest is Naïve Bayes 68.77%.

Meanwhile, research using the PSO optimization feature has been carried out a lot, PSO is very well known for optimizing the accuracy of the combination of algorithms used. Research with titles *"Data Mining Optimization Uses C4.5 Classification and Particle Swarm Optimization (PSO) in the location selection of Student Boardinghouses"*[17]. The purpose of this research is to implement the C.45 algorithm and see the effect of using Particle Swarm Optimization (PSO). The results of testing the Decision Tree algorithm obtained an accuracy of 97.53% and increased by 0.25% after being optimized using PSO to 97.78%.

PSO feature selection is also used in a study entitled "Performance Comparison of K-Nearest Neighbor and Decision Tree C4.5 by Utilizing Particle Swarm Optimization for Prediction of Liver Disease"[18]. With the research results where the PSO-based Decision Tree C.45 algorithm has an accuracy level of 91.26% and an AUC value of 0.935 which is better than the PSO-based K-Nearest Neighbor whose accuracy is only 89.11% and an AUC value of 0.923. That the PSO-based Decision Tree is considered to be more precise in predicting disease.

Meanwhile, research testing the PSO-based Naïve Bayes algorithm has also been carried out[19]. In this study, the classification of texts regarding reviews of news articles is carried out. The purpose of this research is to see the performance of using Particle Swarm Optimization (PSO) to improve the accuracy of the Support Vector Machine and Naïve Bayes algorithms. The results of testing the Naïve Bayes algorithm obtained an accuracy value of 89.50% and an AUC value of 0.500. Meanwhile, Naïve Bayes PSO obtained an accuracy of 92.00% with an AUC of 0.550. SVM has an accuracy value of 87.50% with an AUC of 0.979, while the PSO SVM has an accuracy value of 90.50% and an AUC of 0.975.

Researchers use these three algorithms because the algorithms used in this study are 3 of the top 10 most influential algorithms in Data Mining[20]. This study also tested the number of different data dimensions, namely in test case I with 1000 data, test case II with 3000 data, and 5000 data in test case III. The use of different data dimensions aims to see whether the amount of data affects the accuracy of the PSO-based Decision Tree, K-NN, Naive Bayes, algorithm and the Decision Tree, K-NN, Naïve Bayes algorithm without PSO.

II. RESEARCH METHODOLOGY

The method in this study was adopted from the Cross-Industry Standard Process for Data Mining (CRISP-DM) model. CRISP-DM is a method for solving common problems in data mining which consists of six stages, namely business understanding, data understanding, data preparation, modeling, evaluation, and deployment. In this study, it only reached the evaluation stage and did not reach the deployment stage, namely the implementation stage on a tool. The stages of CRISP-DM can be seen in Figure 1 below.

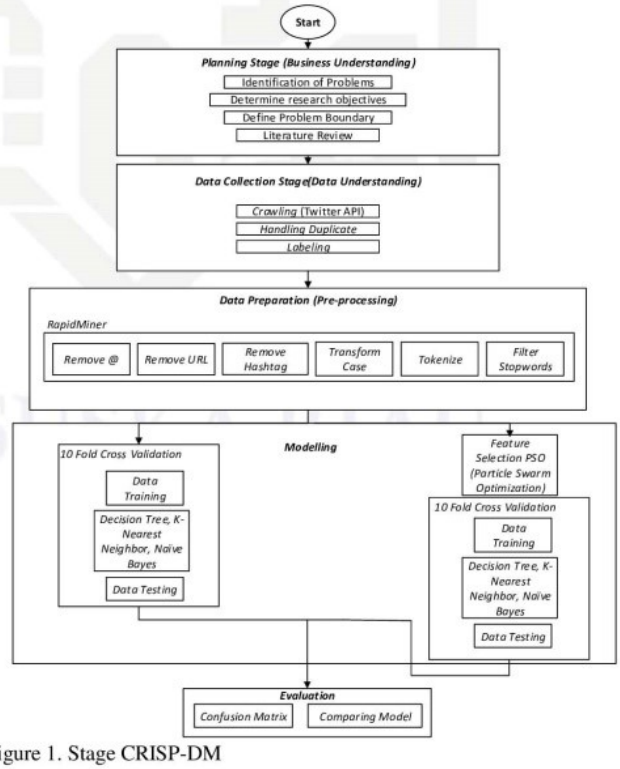

Figure 1. Stage CRISP-DM

The framework steps adopted from the Cross-Industry Standard Process for Data Mining (CRISP-DM) are as follows [21][22].

1. *Business Understanding*

At this stage what is done is to understand how the process will be passed by the researcher to obtain the expected results. Data collection is the first step so that this research can be carried out, then conducting an assessment of related research and determining research objectives[23].

2. *Data Understanding*

The next stage is understanding the data, meaning how we know where and how to obtain data. This stage begins with collecting data, describing the data statistically and visually, and evaluating the quality of the data[24].

3. *Data Preparation (Pre-processing)*

This stage is carried out before the dataset enters the modeling. This process consists of several stages, namely: (1) transform cases aims to align words into lowercase letters in a tweet, (2) remove RT to remove RT words that are often found in a tweet, (3) remove @ to remove mentions in a tweet, (4) remove http to remove links that are often found in a tweet, (5) tokenize to separate words in sentences into separate ones and remove punctuation, (6) Stopword filter aims to remove non-standard words and words liaison, (7) Filter Tokens (by length) to delete words whose number of letters does not match the predetermined number[25].

4. *Modeling*

At this stage, it begins with determining and modeling the algorithms to be tested using datasets that have gone through the preprocessing stages to obtain Accuracy, Precision, Recall, and AUC results from each algorithm.

5. *Evaluation*

This stage evaluates the results of the modeling by comparing the Accuracy, Precision, Recall, and AUC values between models based on PSO feature selection and not. [19].

III. RESULT AND DISCUSSION

1. *Business Understanding*

This study uses datasets originating from public comments about rising fuel prices on social media Twitter. This study aims to analyze public opinion sentiment on the effect of rising fuel prices on social media Twitter and look at the performance of each algorithm when tested using different amounts of data and the effect of using the PSO feature selection. Data collection has also been through a review of related literature.

2. *Data Understanding*

For data collection, namely crawling using the RapidMiner application, the query used is "BBM" based on the "latest and most popular" tweets with the parameters that match those mentioned in the research method. The results of the data obtained are then exported into a Microsoft Excel file. The crawling results obtained 109,470 tweet data. Before the data is manually labeled by linguists into positive and negative sentiments, the data will first go through the cleansing stage or delete duplicate data. Positive comments are categorized based on the presence of words that contain positive meanings, such as: good, good, continue, enthusiasm, cool, smart, smart, and so on. Meanwhile, negative opinions are collected based on words that have bad meanings, and are rude or insulting, such as: stupid, stressed, stupid, shrewd, stupid, slow, stupid, stupid, tadpole, expensive, and so on. The amount of data obtained is 5000 tweet data consisting of 2165 negative sentiment tweet data and 2835 positive sentiment tweet data.

TABLE II. CRAWLING RESULT

No.	Id.	Tweet
1	1572837473398680000	Hari ini @jokowi sibuk prioritaskan proyek abal2 "IKN" seolah-olah tidak sadar Indonesia sedang krisis bayar utang (cicilan pokok 400T, bunga 405T. 1/3 APBN) sehingga rakyat dikorbankan dgn kenaikan harga BBM, pajak, listrik. Situ jangan egois, sono belajar dari Pres Habibie?? https://t.co/ekOWOdIdl0
2	1572738093408740000	Pemerintah Deteksi 1,3 Juta Penerima BLT BBM Salah Sasaran https://t.co/QpPYZpcyar
…	…	…
5000	1573209178000460000	BLT BBM salah satu upaya untuk menjaga daya beli masyarakat di Jambi #BLTBBMBantuMasyarakat #BansosBBMRingankanRakyat

TABLE III. LABELING RESULT

No.	Tweet	Sentiment
1	Hari ini @jokowi sibuk prioritaskan proyek abal2 "IKN" seolah-olah tidak sadar Indonesia sedang krisis bayar utang (cicilan pokok 400T, bunga 405T. 1/3 APBN) sehingga rakyat dikorbankan dgn kenaikan harga BBM, pajak, listrik. Situ jangan egois, sono belajar dari Pres Habibie?? https://t.co/ekOWOdIdl0	Negatives
2	Pemerintah Deteksi 1,3 Juta Penerima BLT BBM Salah Sasaran https://t.co/QpPYZpcyar	Negatives
…	…	…
5000	@tanyarlfes waktu sebelum bbm naik rm padang depan rumah ku serba 11k, ehh skrg jadi serba 12k?	Negatives

3. *Data Preparation (Pre-processing)*

In the pre-processing stage, there are 7 stages with 5000 data used. An example of data results that go through the pre-processing stage can be seen in Table IV below.

TABLE IV. *OUTPUT* FROM *PRE-PROCESSING*

Process	*Output*
Genuine Sentiment	RT @L_Vanthoe25: Biar BBM naik, Nyawa Rakyat Papua tetap Murah Meriah. #RipHAM. #TolakImpunitas #SavePapua https://t.co/9fAH5ehR25

Remove RT	@L_Vanthoe25: Biar BBM naik, Nyawa Rakyat Papua tetap Murah Meriah. #RipHAM. #TolakImpunitas #SavePapua https://t.co/9fAH5ehR25
Remove @	Biar BBM naik, Nyawa Rakyat Papua tetap Murah Meriah. #RipHAM. #TolakImpunitas #SavePapua https://t.co/9fAH5ehR25
Remove URL	Biar BBM naik, Nyawa Rakyat Papua tetap Murah Meriah. #RipHAM. #TolakImpunitas #SavePapua
Remove Hashtag	Biar BBM naik, Nyawa Rakyat Papua tetap Murah Meriah.
Transform Cases	biar bbm naik, nyawa rakyat papua tetap murah meriah.
Tokenize	biar, bbm naik, nyawa, rakyat, papua, tetap, murah, meriah
Filter Stopwords	bbm naik nyawa rakyat papua tetap murah meriah

4. *Modeling*

The modeling stage uses 10-fold cross-validation, which is the method of dividing the dataset into 10 parts, 1 of which is used as testing data, while the other part is training data. Then the datasets that have been labeled and have gone through the preprocessing process will be tested using the Decision Tree, K-NN algorithm, Naïve Bayes, and the use of feature selection PSO.

4.1 Testing the Decision Tree Algorithm, K-NN, and Naïve Bayes

The test will use three different data dimensions, namely, 1000 data, 3000 data, and 5000 data. The initial step for testing the algorithm is to use Excel Read operators to read datasets that have been labeled with positive sentiment and negative sentiment in .xls file format. The main process in testing the algorithm used can be seen in figure2 below.

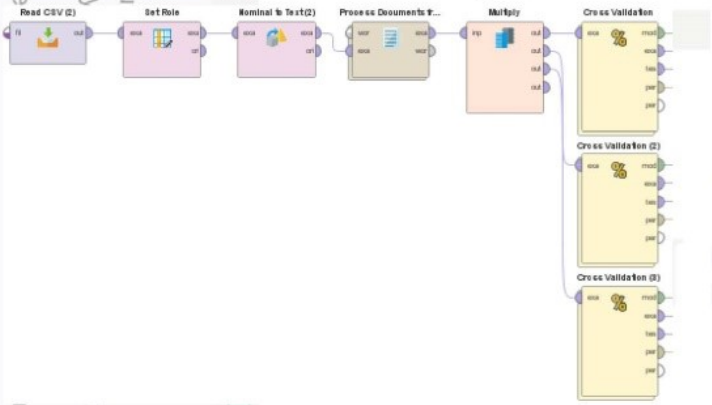

Figure 2. Main Process Modeling on Rapidminer

The next stage is Set Role to determine the fields in the class of the dataset, then enter the Nominal to Text stage to change all nominal attributes to text. Furthermore, the Process Document operator is the stage of data preparation (pre-processing) which consists of the stages of Transform Cases, Tokenize, Filter Stopwords, and Filter Tokens (by Length) which can be seen in Figure 3 below.

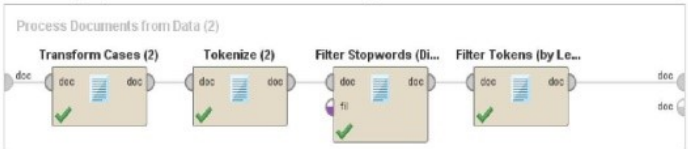

Figure 3. Processes in the Operator Process Documents

Before entering the Cross Validation stage the data will enter the Multiply stage which functions to be able to use three Cross Validation operators (number of folds = 10) at the same time in which there are Decision Tree with criterion parameter = gain_ratio, maximal depth = 10, confidence = 0.1, minimum gain = 0.01, minimum leaf size = 2, minimum size for split = 4, number of prepruning alternatives = 3. Meanwhile, the K-Nearest Neighbor parameter is k = 5 , and Naïve Bayes which only focuses on cross validation values. For details can be seen in figure 4 (Decision Tree), figure 5 (K-Nearest Neighbor), figure 6 (Naïve Bayes). Cross Validation in it there are two parts, namely training, and testing. The training section contains the implementation of the algorithms used, while the testing section consists of the Apply Model operator to implement the algorithm and Performance operators to measure the level of accuracy, precision, recall, and AUC value of each algorithm used. This test method repeats the test 10 times.

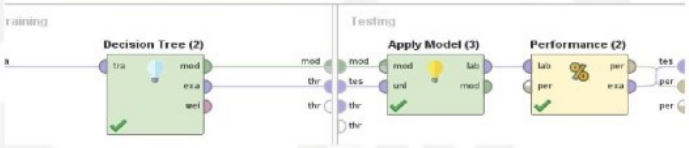

Figure 4. Process *Cross Validation Decision Tree*

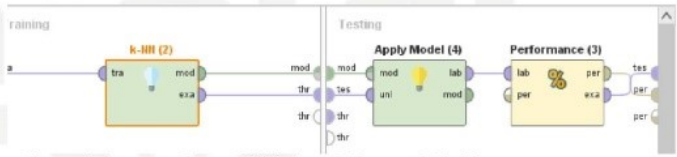

Figure 5. Process *Cross Validation K-Nearest Neighbor*

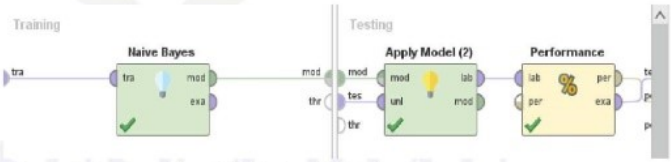

Figure 6. Process *Cross Validation Naïve Bayes*

4.2 Testing the Decision Tree Algorithm, K-NN, and Naïve Bayes + *PSO*

After modeling the Decision Tree, K-NN, and Naïve Bayes algorithms without using the PSO feature selection, the next step is to model the comparison of the three algorithms by adding the Particle Swarm Optimization (PSO) feature selection which also uses three different dimensions, namely 1000 data, 3000 data, and 5000 data. The purpose of adding the PSO is to optimize the level of accuracy of the algorithm used. The stages are not much different from the previous one, only adding the PSO operator which consists of Cross

Validation of the algorithm used after the Multiply operator. The parameters used for the three tests of the algorithm used are Cross-Validation with k-fold = 10. While the parameters used in PSO are Population Size = 5 , Maximum Number = 30, and Inertia Weight value = 1. For the main process in using PSO This can be seen in Figure 7 below.

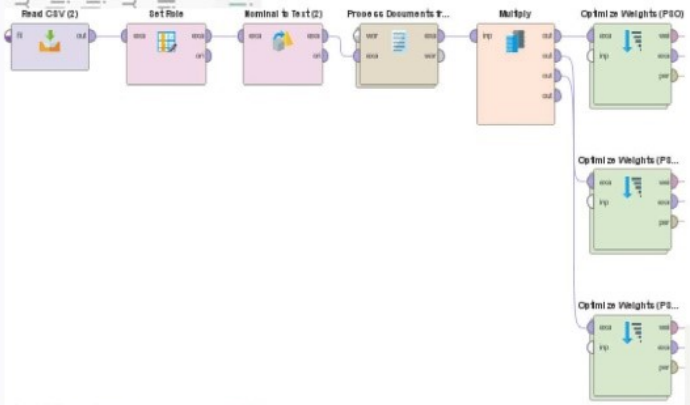

Figure 7. Main Process Modeling on PSO-based Rapidminer

5. *Evaluation*

After the classification process is carried out, the next stage is an evaluation to see the Accuracy, Precision, Recall, and AUC values for three different data dimensions and the modeling that has been done before. Based on Table V the experimental results of 1000 data, in the Decision Tree algorithm, it can be seen that the addition of the PSO feature selection has a positive impact on Accuracy, Precision, Recall, and AUC whose initial values are 70.80%, 70.08%, 98.20%, and 0.571 to 71.70%, 70.59%, 98.80% and 0.581. Whereas in the K-NN algorithm with the addition of PSO, it has a negative impact because the Accuracy, Precision, Recall, and AUC values have decreased from the initial values of 75.60%, 78.01%, 88.46%, and 0.735 to 73.10%, 76.65%, 86.07%, and 0.704. Meanwhile, in the Naïve Bayes algorithm, the addition of PSO only has a positive effect on Accuracy, Precision, and AUC values where the initial values are 69.20%, 80.06%, 0.498 which increase to 69.90%, 82.50%, 0.599 and have a negative impact on the Recall value because decreased from the initial value of 71.97% to 69.73%.

TABLE V. CLASSIFICATION MODEL EVALUATION (1000 DATA)

Algoritma	Accuracy	Precision	Recall	AUC
Decision Tree	70.80%	70.08%	98.20%	0.571
K-NN	**75.60%**	78.01%	88.46%	**0.735**
Naïve Bayes	69.20%	80.06%	71.97%	0.498
Decision Tree + PSO	71.70%	70.59%	**98.80%**	0.581
K-NN + PSO	73.10%	76.65%	86.07%	0.704
Naïve Bayes + PSO	69.90%	**82.50%**	69.73%	0.599

The results of testing 3000 data can be seen in Table VI. In the Decision Tree, the addition of the PSO feature selection has a positive impact on the Accuracy, Precision, and AUC values where it has increased from the initial values of 74.71%, 74.10%, and 0.549 to 75.47%, 74.72%, and 0.563. In contrast to the Recall value has decreased from the initial value of 99.82% to 99.72%. Whereas in the K-NN algorithm with the addition of the PSO feature selection it only has a positive impact on the Accuracy and Recall values, the initial values are 78.07% and 88.95% to 78.21% and 92.93%. In contrast to the Precision and AUC values which decreased from the initial values of 82.15% and 0.781 to 80.09% and 0.754. Meanwhile, in the Naive Bayes algorithm with the PSO feature selection, the values that experienced an increase were only Precision and AUC, whose initial values were 85.01% and 0.522 to 86.18% and 0.594. The values for Accuracy and Recall have decreased from 71.51% and 73.46% to 70.72% and 70.74% after being optimized using PSO.

TABLE VI. CLASSIFICATION MODEL EVALUATION (3000 DATA)

Algoritma	Accuracy	Precision	Recall	AUC
Decision Tree	74.71 %	74.10%	**99.82%**	0.549
K-NN	78.07%	82.15%	88.95%	**0.781**
Naïve Bayes	71.51%	85.01%	73.46%	0.522
Decision Tree + PSO	75.47%	74.72%	99.72%	0.563
K-NN + PSO	**78.21%**	80.09%	92.93%	0.754
Naïve Bayes + PSO	70.74%	**86.18%**	70.74%	0.594

The test results of 5000 data can be seen in Table VII. In the Decision Tree the addition of the PSO feature selection has a positive impact on the Accuracy, Precision, and Recall values where it has increased from the initial values of 61.78%, 91.65%, and 12.93% to 61.90%, 92.57%, and 13.03%. The AUC value has no effect, which is still 0.500 even after PSO optimization. Whereas in the K-NN algorithm with the addition of the PSO feature selection it has a negative impact on the overall Accuracy, Precision, Recall, and AUC values whose initial values are 74.02%, 71.19%, 67.35%, and 0.805 to 68.58%, 70.88%, 46.70%, and 0.760. Meanwhile, in the Naive Bayes algorithm with the PSO feature selection, the values that experienced an increase were only Accuracy, Precision, and AUC, whose initial values were 71.30%, 63.90%, and 0.602 to 73.74%, 68.31%, and 0.658. The value of Recall has decreased from 77.59% to 73.72% after being optimized using PSO.

TABLE VII. CLASSIFICATION MODEL EVALUATION (5000 DATA)

Algoritma	Accuracy	Precision	Recall	AUC
Decision Tree	61.78 %	91.65%	12.93%	0.500
K-NN	**74.02%**	71.19%	67.35%	**0.805**
Naïve Bayes	71.30%	63.90%	**77.59%**	0.602
Decision Tree + PSO	61.90%	**92.57%**	13.03%	0.500
K-NN + PSO	68.58%	70.88%	46.70%	0.760
Naïve Bayes + PSO	73.74%	68.31%	73.72%	0.658

IV. CONCLUSIONS

Based on the results and analysis processes that have been carried out in this study, some final conclusions can be drawn as follows:

1. Comment data obtained from Twitter regarding public opinion regarding BBM during the period December 27 2022 - January 27 2023, the tendency for comments submitted to contain positive and negative comments. Of the 5000 test data, there are 2835 positive sentiment data and 2165 negative sentiment data. Here it can be concluded that Twitter social media users in Indonesia have a positive attitude toward the policy of increasing fuel prices during this period time.
2. After going through the testing phase using the classical algorithm model and the Particle Swarm Optimization (PSO) based algorithm, several values of different accuracy levels were obtained. The use of Particle Swarm Optimization (PSO) is proven to increase the level of accuracy of several models that have been done. In the application of (1000 Data), the best accuracy is the classic K-Nearest Neighbor algorithm with a value of 75.60%. While the best accuracy value in the test (3000 Data), is the PSO-based K-Nearest Neighbor algorithm with a value of 78.21%. Meanwhile, in testing (5000 Data), the classic K-Nearest Neighbor algorithm is the best with an accuracy value of 74.02%.
3. The amount of data affects the level of accuracy of each algorithm used, both classical and Particle Swarm Optimization (PSO) based algorithms.

REFERENCES

[1] S. U. A. et all Wardani Widia, Suriana, "Dampak kenaikan Bahan Bakar Minyak (BBM) Terhadap Inflasi dan Implikasinya Terhadap Makroekonomi di Indonesia," *J-Las*, vol. 2, no. 1, pp. 209–217, 2022.

[2] C. E. Stanford, "Coal Resources, Production and Use in Indonesia," *Coal Handb. Towar. Clean. Prod.*, vol. 2, pp. 200–219, Oct. 2013, doi: 10.1533/9781782421177.2.200.

[3] D. Yuliani, S. Saryono, D. Apriani, and M. Ro, "Dampak Kenaikan Harga Bahan Bakar Minyak (BBM) Terhadap Sembilan Bahan Pokok (Sembako) Di Kecamatan Tambun Selatan Dalam Masa Pandemi," vol. 2, no. 2, pp. 320–326, 2022.

[4] T. Dartanto, "Reducing fuel subsidies and the implication on fiscal balance and poverty in Indonesia: A simulation analysis," *Energy Policy*, vol. 58, pp. 117–134, 2013, doi: 10.1016/j.enpol.2013.02.040.

[5] G. R. Hrp and dan N. Aslami, "Analisis Dampak Kebijakan Perubahan Publik Harga BBM terhadap Perekonomian Rakyat Indonesia," *J. Ilmu Komputer, Ekon. dan Manaj.*, vol. 2, no. 1, pp. 1464–1474, 2022.

[6] I. Santoso, Windu Gata, and Atik Budi Paryanti, "Penggunaan Feature Selection di Algoritma Support Vector Machine untuk Sentimen Analisis Komisi Pemilihan Umum," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 3, pp. 364–370, 2019, doi: 10.29207/resti.v3i3.1084.

[7] M. A. H. Fitri Veny Amilia, Rachmadita Andreswari, "Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm," *Procedia Comput. Sci.*, vol. 161, pp. 765–772, 2019, doi: 10.1016/j.procs.2019.11.181.

[8] W. G. Hernawati, "Sentimen Analisis Operasi Tangkap Tangan Kpk Menurut Masyarakat Menggunakan Algoritma Support Vechtor Machine, Naïve Bayes, Berbasis Particle Swarm Optimizition," *Fakt. Exacta*, vol. 12, no. 3, p. 230, 2019, doi: 10.30998/faktorexacta.v12i3.4992.

[9] M. Alzate, M. Arce-Urriza, and J. Cebollada, "Mining the text of online consumer reviews to analyze brand image and brand positioning," *J. Retail. Consum. Serv.*, vol. 67, no. March, p. 102989, 2022, doi: 10.1016/j.jretconser.2022.102989.

[10] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. e-ISSN*, vol. 2548, no. 12, p. 964X, 2017.

[11] C. Y. Man, D. A. Palmer, and J. Qian, "The Belt and Road Initiative on Twitter: An annotated dataset," *Data Br.*, vol. 45, p. 108711, 2022, doi: 10.1016/j.dib.2022.108711.

[12] M. Rezki, D. N. Kholifah, M. Faisal, P. Priyono, and R. Suryadithia, "Analisis Review Pengguna Google Meet dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes," *J. Infortech*, vol. 2, no. 2, pp. 264–270, 2020, doi: 10.31294/infortech.v2i2.9286.

[13] K. Nurbagja, N. Saputra, A. Riyadi, and M. N. Tentua, "Sentiment Analysis of the Increase in Fuel Prices Using Random Forest Classifier Method The percentage of the poor population in Indonesia from 1990 to 2003," vol. 5, no. 1, pp. 132–144, 2023, doi: 10.12928/biste.v5i1.7414.

[14] A. Solichin, "Comparison of decision tree, Naïve Bayes and K-nearest neighbors for predicting thesis graduation," *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, pp. 217–222, 2019, doi: 10.23919/EECSI48112.2019.8977081.

[15] A. Bayhaqy, S. Sfenrianto, K. Nainggolan, and E. R. Kaburuan, "Sentiment Analysis about E-Commerce from Tweets Using Decision Tree, K-Nearest Neighbor, and Naïve Bayes," *2018 Int. Conf. Orange Technol. ICOT 2018*, no. November 2019, 2018, doi: 10.1109/ICOT.2018.8705796.

[16] Y. Findawati, I. R. I. Astutik, A. S. Fitroni, I. Indrawati, and N. Yuniasih, "Comparative analysis of Naïve Bayes, K Nearest Neighbor and C.45 method in weather forecast," *J. Phys. Conf. Ser.*, vol. 1402, no. 6, 2019, doi: 10.1088/1742-6596/1402/6/066046.

[17] A. Waluyo, H. Jatnika, M. R. S. Permatasari, T. Tuslaela, I. Purnamasari, and A. P. Windarto, "Data Mining Optimization uses C4.5 Classification and Particle Swarm Optimization (PSO) in the location selection of Student Boardinghouses," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, 2020, doi: 10.1088/1757-899X/874/1/012024.

[18] Z. Fadilah and Murnawan, "Performance Comparison of K-Nearest Neighbor and Decision Tree C4.5 by Utilizing Particle Swarm Optimization for Prediction of Liver Disease," *Int. J. Open Inf. Technol.*, vol. 9, no. 10, pp. 9–15, 2021.

[19] N. K. Wardhani *et al.*, "Sentiment analysis article news coordinator minister of maritime affairs using algorithm naive bayes and support vector machine with particle swarm optimization," *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 24, pp. 8365–8378, 2018.

[20] X. Wu *et al.*, *Top 10 algorithms in data mining*, vol. 14, no. 1. 2008. doi: 10.1007/s10115-007-0114-2.

[21] A. Azevedo and M. F. Santos, "KDD , SEMMA AND CRISP-DM : A PARALLEL OVERVIEW Ana Azevedo and M . F . Santos," *IADIS Eur. Conf. Data Min.*, pp. 182–185, 2008, [Online]. Available: http://recipp.ipp.pt/handle/10400.22/136%0Ahttp://recipp.ipp.pt/bitstream/10400.22/136/3/KDD-CRISP-SEMMA.pdf

[22] K. Sigit, A. P. Dewi, G. Windu, Nurmalasari, T. Muhamad, and N. Kadinar, "Comparison of Classification Methods on Sentiment Analysis of Political Figure Electability Based on Public Comments on Online News Media Sites," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 662, no. 4, 2019, doi: 10.1088/1757-899X/662/4/042003.

[23] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.

[24] F. Martinez-Plumed *et al.*, "CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 8, pp. 3048–3061, 2021, doi: 10.1109/TKDE.2019.2962680.

[25] I. Taleb, R. Dssouli, and M. A. Serhani, "Big Data Pre-processing: A Quality Framework," *Proc. - 2015 IEEE Int. Congr. Big Data, BigData Congr. 2015*, no. March 2016, pp. 191–198, 2015, doi: 10.1109/BigDataCongress.2015.35.

LAMPIRAN A

POSTER KEGIATAN

LAMPIRAN B

SURAT DARI JURNAL

B.1 Copyright Receipt

IEEE COPYRIGHT AND CONSENT FORM

To ensure uniformity of treatment among all contributors, other forms may not be substituted for this form, nor may any wording of the form be changed. This form is intended for original material submitted to the IEEE and must accompany any such material in order to be published by the IEEE. Please read the form carefully and keep a copy for your files.

Sentiment Analysis of Public Opinion Regarding Fuel Oil on Twitter by Comparing Classification Algorithms

Mr. Gita Widarma, Dr. Rice Novita, Mr. Mustakim Mustakim and Mr. Needi Evriyan Rozanda

2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)

U.S. GOVERNMENT EMPLOYEE CERTIFICATION AND TRANSFER OF INTERNATIONAL COPYRIGHT

This will certify that all authors of the Work are U.S. Government employees and prepared the Work on a subject within the scope of their official duties. As such, the Work is not subject to U.S. copyright protection. In order to enable IEEE to claim and protect copyright rights in international jurisdictions, the undersigned hereby assigns to the IEEE all such international rights under copyright that may exist in and to the above Work, and any revised or expanded derivative works submitted to the IEEE by the undersigned based on the Work.

GENERAL TERMS

1. The undersigned represents that he/she has the power and authority to make and execute this form.
2. The undersigned agrees to indemnify and hold harmless the IEEE from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.
3. The undersigned agrees that publication with IEEE is subject to the policies and procedures of the IEEE PSPB Operations Manual.
4. In the event the above work is not accepted and published by the IEEE or is withdrawn by the author(s) before acceptance by the IEEE, the foregoing copyright transfer shall be null and void. In this case, IEEE will retain a copy of the manuscript for internal administrative/record-keeping purposes.
5. For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
6. The author hereby warrants that the Work and Presentation (collectively, the "Materials") are original and that he/she is the author of the Materials. To the extent the Materials incorporate text passages, figures, data or other material from the works of others, the author has obtained any necessary permissions. Where necessary, the author has obtained all third party permissions and consents to grant the license above and has provided copies of such permissions and consents to IEEE

You have indicated that you DO wish to have video/audio recordings made of your conference presentation under terms and conditions set forth in "Consent and Release."

CONSENT AND RELEASE

1. In the event the author makes a presentation based upon the Work at a conference hosted or sponsored in whole or in part by the IEEE, the author, in consideration for his/her participation in the conference, hereby grants the IEEE the unlimited, worldwide, irrevocable permission to use, distribute, publish, license, exhibit, record, digitize, broadcast, reproduce and archive, in any format or medium, whether now known or hereafter developed: (a) his/her presentation and comments at the conference; (b) any written materials or multimedia files used in connection with his/her presentation; and (c) any recorded interviews of him/her (collectively, the "Presentation"). The permission granted includes the transcription and reproduction of the Presentation for inclusion in products sold or distributed by IEEE and live or recorded broadcast of the Presentation during or after the conference.
2. In connection with the permission granted in Section 1, the author hereby grants IEEE the unlimited, worldwide, irrevocable right to use his/her name, picture, likeness, voice and biographical information as part of the advertisement, distribution and sale of products

incorporating the Work or Presentation, and releases IEEE from any claim based on right of privacy or publicity.

BY TYPING IN YOUR FULL NAME BELOW AND CLICKING THE SUBMIT BUTTON, YOU CERTIFY THAT SUCH ACTION CONSTITUTES YOUR ELECTRONIC SIGNATURE TO THIS FORM IN ACCORDANCE WITH UNITED STATES LAW, WHICH AUTHORIZES ELECTRONIC SIGNATURE BY AUTHENTICATED REQUEST FROM A USER OVER THE INTERNET AS A VALID SUBSTITUTE FOR A WRITTEN SIGNATURE.

Gita Widarma — **Signature**

27-05-2023 — **Date (dd-mm-yyyy)**

Information for Authors

AUTHOR RESPONSIBILITIES

The IEEE distributes its technical publications throughout the world and wants to ensure that the material submitted to its publications is properly available to the readership of those publications. Authors must ensure that their Work meets the requirements as stated in section 8.2.1 of the IEEE PSPB Operations Manual, including provisions covering originality, authorship, author responsibilities and author misconduct. More information on IEEE's publishing policies may be found at http://www.ieee.org/publications_standards/publications/rights/authorrightsresponsibilities.html Authors are advised especially of IEEE PSPB Operations Manual section 8.2.1.B12: "It is the responsibility of the authors, not the IEEE, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it." Authors are also advised of IEEE PSPB Operations Manual section 8.1.1B: "Statements and opinions given in work published by the IEEE are the expression of the authors."

RETAINED RIGHTS/TERMS AND CONDITIONS

- Authors/employers retain all proprietary rights in any process, procedure, or article of manufacture described in the Work.
- Authors/employers may reproduce or authorize others to reproduce the Work, material extracted verbatim from the Work, or derivative works for the author's personal use or for company use, provided that the source and the IEEE copyright notice are indicated, the copies are not used in any way that implies IEEE endorsement of a product or service of any employer, and the copies themselves are not offered for sale.
- Although authors are permitted to re-use all or portions of the Work in other works, this does not include granting third-party requests for reprinting, republishing, or other types of re-use.The IEEE Intellectual Property Rights office must handle all such third-party requests.
- Authors whose work was performed under a grant from a government funding agency are free to fulfill any deposit mandates from that funding agency.

AUTHOR ONLINE USE

- **Personal Servers**. Authors and/or their employers shall have the right to post the accepted version of IEEE-copyrighted articles on their own personal servers or the servers of their institutions or employers without permission from IEEE, provided that the posted version includes a prominently displayed IEEE copyright notice and, when published, a full citation to the original IEEE publication, including a link to the article abstract in IEEE Xplore. Authors shall not post the final, published versions of their papers.
- **Classroom or Internal Training Use.** An author is expressly permitted to post any portion of the accepted version of his/her own IEEE-copyrighted articles on the author's personal web site or the servers of the author's institution or company in connection with the author's teaching, training, or work responsibilities, provided that the appropriate copyright, credit, and reuse notices appear prominently with the posted material. Examples of permitted uses are lecture materials, course packs, e-reserves, conference presentations, or in-house training courses.
- **Electronic Preprints.** Before submitting an article to an IEEE publication, authors frequently post their manuscripts to their own web site, their employer's site, or to another server that invites constructive comment from colleagues. Upon submission of an article to IEEE, an author is required to transfer copyright in the article to IEEE, and the author must update any previously posted version of the article with a prominently displayed IEEE copyright notice. Upon publication of an article by the IEEE, the author must replace any previously posted electronic versions of the article with either (1) the full citation to the IEEE work with a Digital Object Identifier (DOI) or link to the article abstract in IEEE

B.2 Email Accepted

Gmail

GITA WIDARMA <11950311555@students.uin-suska.ac.id>

[ISITIA 2023] Your paper #1570900516 ('Comparison of Machine Learning Classification Algorithms and Feature Selection Particle Swarm Optimization on Public Opinion Regarding Fuel Oil')

ISITIA 2023 (isitia@its.ac.id) <isitia=its.ac.id@edas.info> Sen, 15 Mei 17.46
Balas-Ke: ISITIA 2023 <isitia@its.ac.id>
Ke: Gita Widarma <11950311555@students.uin-suska.ac.id>, Rice Novita <rice.novita@uin-suska.ac.id>, Mustakim Mustakim <mustakim@uin-suska.ac.id>, Nesdi Evrilyan Rozanda <nesdi.rozanda@uin-suska.ac.id>
Cc: Muhammad Attamimi <attamimi@ee.its.ac.id>, Vita Lystianingrum <vita@ee.its.ac.id>, Prasetiyono Hari Mukti <prasetiyono.hm@gmail.com>, Dimas Fajar Uman Putra <dimasfup@ee.its.ac.id>

Dear Mr. Gita Widarma,

Congratulations!

We are pleased to inform you that your manuscript:
Paper ID: 1570900516
Paper Title: Comparison of Machine Learning Classification Algorithms and Feature Selection Particle Swarm Optimization on Public Opinion Regarding Fuel Oil
Authors: Gita Widarma, Rice Novita, Mustakim Mustakim and Nesdi Evrilyan Rozanda
has now been ACCEPTED for presentation in 2023 International Seminar on Intelligent Technology and Its Applications (ISITIA).

The reviews are included below, and can also be found at https://edas.info/showPaper.php?m=1570900516 using your EDAS login name 11950311555@students.uin-suska.ac.id.

Please find the following details on your final manuscript preparation, registration, as well as video submission.

For the final manuscript:
1. You must revise your manuscript according to the comments and suggestions given by the reviewers (see below).
2. Ensure that you have added all the author names, both in your manuscript as well as in EDAS in the same sequence.
3. You are required to ensure that the similarity rating of your manuscript is less than 30%. Submissions with a similarity rating of 30% or above will be dropped and no longer considered for presentation in ISITIA 2023.
4. Please ensure that your manuscript has been checked for grammatical and typographical errors, and that your manuscript is IEEE pdf eXpress compliant.
5a. A formatting checklist is given here to help you prepare your manuscript before uploading to IEEE pdf eXpress. You should also upload the confirmed checklist during the registration process.
5b. The details on how to use IEEE pdf eXpress to generate your PDF file can be found here: https://elib.its.ac.id/conf/isitia/main/home/submissions/.
6. You are required to transfer copyright of the manuscript to IEEE if you wish your manuscript to be published in IEEExplorer. This can be done electronically through EDAS.
7. IEEE reserves the right to exclude a paper from distribution after the conference (e.g. removal from IEEE Xplore) if the paper is not presented at the conference.

To register, please go to the following address: https://elib.its.ac.id/conf/isitia/main/registration/

Video upload details:
A recorded live online presentation is required for IEEExplore submission, however, we also ask a pre-recorded video for backup as suggested by the IEEE. The instruction on the video requirement can be found in:
https://elib.its.ac.id/conf/isitia/main/2021/05/10/pre-recorded-presentation-video-instruction/

Thank you very much.

Sincerely yours,

B.3 Email Review dan Comments

Muhammad Attamimi, B.Eng., M.Eng., Ph.D. (General chair)
2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)
Conference Website : http://isitia.its.ac.id/
Email: isitia@its.ac.id

Reviews/Comments:
======= Review 1 =======

> *** Comments to Author: (Please write details and comments for the authors regarding the manuscript)

The title is misleading. The content is about sentiment analysis, but the title suggest something more general. Please revise the title mentioning the important part of the paper, i.g. Twitter data, Sentiment Analysis, Oil Price, Indonesia.
Why use PSO?
Is the research already done or not yet? if it's done, the author should not use future tense "will".
Please improve the English, e.g. "Before the data will be labeled manually positive and negative sentiments" >> The data has been labelled manually with positive and negative sentiment.
What is the result of PSO feature selection? please mention not only the accuracy result, but also the selected features and how the algorithm works.
No detail needed in the conclusion. Please include the summary of the research, from the background, the aim and the result.

======= Review 2 =======

> *** Comments to Author: (Please write details and comments for the authors regarding the manuscript)

Please explain background problem, gap analysis and the contribution in this work. The current abstract contains implementation steps instead of highlighting the contribution.

The authors should not use acronym without explanation. All acronyms must be defined before use.

Please discuss references or related works of previous researches focusing on Public Opinion Regarding Fuel Oil. The methods use in this article are standard machine learning algorithms so instead of discussing their usages in works of weather forecast, liver disease, etc it would be better to discuss on text mining such as sentiment analysis on other public concern. Try use keywords: public concern twitter in Google Scholar

Fig. 1 has Indonesian words.

This work uses Rapid MIner, instead of using Fig 2-7, it would be better to give more detailed process in Fig 1.

Give more discussions about the findings and their significance about Public Opinion Regarding Fuel Oil instead about talking numbers (accuracy, etc)

======= Review 3 =======

> *** Comments to Author: (Please write details and comments for the authors regarding the manuscript)

Overall, this paper talks brings a somewhat interesting topic, but has poor presentation and lack of discussion. Revisions in many sections are required.
1. Problem statement in this paper is not well described.
2. The labeling process and how to maintain that the data labels are not bias are not well explained.
3. Analysis and discussion on the comparison results are very lacking! Section Evaluation here mostly just shows/mentions about the numbers.
4. All the figures captured from Rapidminer are not readable. Figures from screenshots are not recommended; Try re-draw them to improves figures' readability.
5. For every table showing quantitative experimental results, highlight the highest score in each column.
6. Please, fix how to write the numbers in this paper, especially for more-than-thousands numbers.
7. The English should be much refined!

LAMPIRAN C

SURAT PERNYATAAN PELABELAN

C.1 Pakar Bahasa

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama	: Dina Hartanti Hermawan, S. Pd
Tempat/ Tanggal Lahir	: Binio/ 29 juli 1996
Pekerjaan	: Guru
Alamat	: Emplasmen Kebun PKS Sei Intan, Kecamatan Pagaran Tapah Rokan Hulu

Menyatakan dengan sesungguhnya bahwa benar memvalidasi untuk pelabelan sentimen yang bersifat Positif dan Negatif dalam tweet Twitter terhadap kebijakan pemerintah mengenai naiknya harga BBM sebanyak 5000 data dalam Tugas Akhir dari :

Nama	: Gita Widarma
NIM	: 11950311555
Jurusan	: Sistem informasi
Judul Tugas Akhir	: Sentiment Analysis of Public Opinion Regarding Fuel Oil on Twitter by Comparing Classification Algorithms

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa paksaan dari pihak manapun. Atas perhatiannya saya ucapkan terima kasih.

Rokan Hulu, Mei 2023

Yang membuat Pernyataan

Dina Hartanti Hermawan, S. Pd

DAFTAR RIWAYAT HIDUP

Gita Widarma lahir di Kota Tebing Tinggi, pada tanggal 30 Desember 2001. Peneliti merupakan anak dari Bapak Syamsidar dan Ibu Widiastuti. Peneliti adalah anak kedua dari tiga bersaudara. Pada tahun 2006 peneliti memulai pendidikan dengan masuk TK Dharma Wanita Persatuan di Desa Tanjung Pauh dan lulus pada 25 Juni 2007. Lalu melanjutkan pendidikan Sekolah Dasar di SD Negeri 004 Tanjung Pauh. Peneliti menyelesaikan pendidikan Sekolah Dasar pada 8 Juni 2013. Setelah menyelesaikan pendidikan Sekolah Dasar peneliti melanjutkan pendidikan tingkat SLTP di SMP Negeri 5 Singingi Hilir yang selesai pada 11 Juni 2016. Peneliti melanjutkan pendidikan tingkat SLTA di MAN 3 Kampar. Setelah menyelesaikan pendidikan di MAN 3 Kampar pada 13 Mei 2019, peneliti pun melanjutkan pendidikan dengan menjadi mahasiswa Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau dan peneliti menyelesaikan kuliah Strata satu (S1) tersebut pada tahun 2023.

SURAT PERNYATAAN PUBLIKASI JURNAL

Saya yang bertanda tangan dibawah ini :

Nama	: Gita Widarma
NIM	: 11950311555
Jurusan	: Sistem Informasi
Judul Tugas Akhir	: *Sentiment Analysis of Public Opinion Regarding Fuel Oil on Twitter by Comparing Classification Algorithms*

Dengan ini menyatakan bahwa akan melengkapi seluruh kelengkapan administrasi Tugas Akhir Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau **berupa bukti pelaksanaan *conference* secara lengkap**. Demikian surat pernyataan ini di buat untuk dipergunakan sebagaimana mestinya.

Pekanbaru, 19 Juni 2023

Yang membuat pernyataan

Gita Widarma

NIM. 11950311555