Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Sentiment Analysis Using Grok AI as an Auto-Labeling Tool in The Text Processing Stage

Yoga Handoko Agustin[1], Dede Kurniadi[2], Indri Tri Julianto[3]*, Benedicto B. Balilo Jr [4]
[1)2)3)] Department of Computer Science, Institut Teknologi Garut, Indonesia,
[4] CS/IT, Bicol University, Legazpi City, Philippines
[1)]yoga.handoko@itg.ac.id, [2)]dede.kurniadi@itg.ac.id, [3)]indritrijulianto@itg.ac.id, [4)] bjbbalilo@bicol-u.edu.ph

Submitted :March 1, 2025 | **Accepted** : April 3, 2025 | **Published** : April 16, 2025

Abstract: A critical aspect of Natural Language Processing (NLP) is text processing, where text labeling represents the most significant challenge due to its resource-intensive nature when conducted manually. At this stage, automatic labeling emerges as a more practical solution, particularly with the advent of Artificial Intelligence (AI), which offers tools to address this obstacle. Grok AI, equipped with a new feature operable on Platform X, provides a promising approach. This study aims to leverage the Grok AI feature on Platform X for automatic text labeling. The research methodology involves labeling text data obtained from a public dataset. To assess the quality of the labeling results, an evaluation method employing Naive Bayes classification modeling is applied. The findings reveal that Grok AI's performance closely approximates that of human labeling. The highest accuracy achieved by Grok AI is 51.71% using the k-Nearest Neighbors (k-NN) algorithm, approaching the human labeling accuracy of 60.52% with k-NN. Furthermore, Grok AI surpasses the performance of VADER labeling, which achieves an accuracy of only 49.49% with Naive Bayes. Consequently, the Grok AI feature on Platform X presents a viable alternative for the automatic labeling of text data.

Keywords: ai, grok, labeling, x.

INTRODUCTION

Natural language processing, among other areas, has been profoundly affected by the rise of AI (Russel and Norvig 2021). NLP, a subfield of AI focused on the interaction between computers and human language, has the potential to revolutionize how we interact with technology (Nurwanda, Suarna, and Prihartono 2024). However, a significant challenge in NLP research is the manual annotation of text data, a time-consuming and resource-intensive task (Abro, Talpur, and Jumani 2023; Julianto, Kurniadi, Septiana, et al. 2023; Parlika et al. 2020; Utami and Erfina 2021). In recent years, researchers have sought to automate text labeling to mitigate this challenge (Ashari 2023; Julianto, Kurniadi, and Jr 2023; Telnoni, Suryatiningsih, and Rosely 2020). These methods aim to reduce manual effort by employing machine learning techniques to automatically assign labels to text data (Rachmat et al. 2023).

This research proposes to automated text labeling by leveraging the Grok AI feature on Platform X. Grok AI, a powerful AI-powered chatbot, can understand and respond to natural language queries, making it a valuable tool for various tasks, including text labeling (Moeis et al. 2024). The proposed approach has the potential to significantly accelerate the development of NLP models and applications (Asri et al. 2022). By reducing the time and effort required for manual labeling, researchers and developers can focus on building more sophisticated and powerful NLP systems (Xiao and Jin 2021). Additionally, the proposed method can be applied to a wide range of text-based tasks, such as sentiment analysis, text classification, and chatbot development (Basha et al. 2023).

Several prior studies have investigated text labeling for sentiment analysis, employing a range of techniques but largely overlooking the potential of AI assistants. The first study examined sentiment analysis of the Vidio application, utilizing VADER Lexicon and Inset for text labeling (Gaja, Maulana, and Komarudin 2023). It reported an accuracy of 90% with Inset and 73% with VADER Lexicon. The second study explored sentiment analysis of the PeduliLindungi application, applying VADER Lexicon for data labeling (Illia, Eugenia, and Rutba 2021), and concluded that VADER’s lexicon-based method, designed for social media, enhanced its effectiveness. The third study assessed public sentiment toward electric vehicles, using VADER Lexicon for labeling (Anwar

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

and Permana 2023), finding that 95% of opinions were positive, with 5% negative. The fourth study compared VADER Lexicon and Naïve Bayes Classifier for labeling a dataset on public sentiment toward investment applications (Amaliah and Nuryana 2022), revealing that Naïve Bayes outperformed VADER Lexicon, achieving 78% accuracy versus 67%. The fifth study analyzed customer reviews, employing VADER Lexicon for labeling (Barik and Misra 2024), and demonstrated an impressive accuracy of 98.64%. To succinctly summarize these findings and highlight research gaps, a table is provided, as presented in Figure 1.

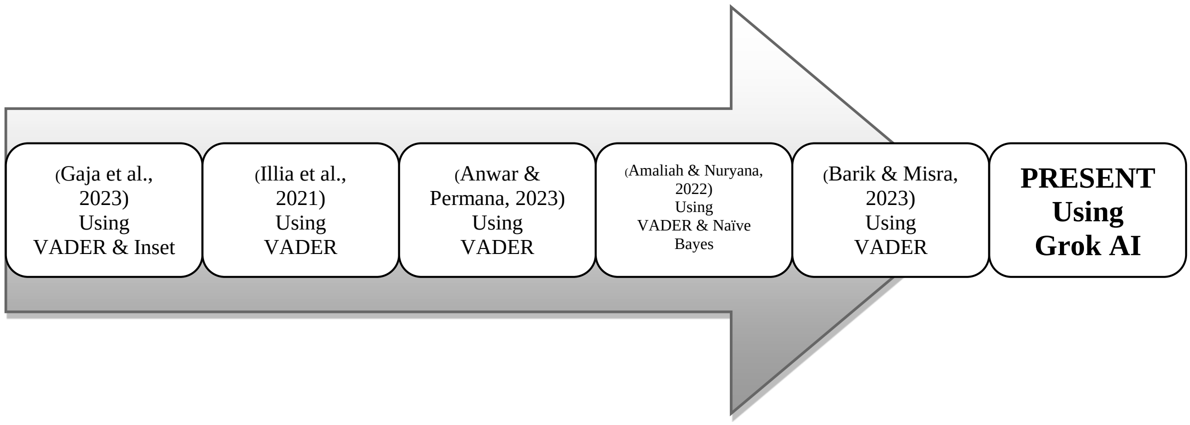

Fig 1. Research Gap

This study is significant due to the scarcity of research exploring AI assistants in text labeling processes. The research will utilize the Grok AI for text labeling, comparing their outcomes with those of human/manual labeling and labeling performed with VADER. The evaluation will involve two algorithms—K-Nearest Neighbors (K-NN) and Naïve Bayes—using K-Fold Cross Validation to assess the effectiveness of each text labeling method. The dataset employed is a publicly accessible financial sentiment dataset sourced from Kaggle.com, comprising 499 data points.

K-Nearest Neighbors (K-NN) offers several benefits, such as its robustness to noisy training data, rapid processing speed, simplicity of comprehension, and suitability for large datasets (Cholil et al. 2021; Julianto, Kurniadi, Septiana, et al. 2023). Its effectiveness was evidenced by (Andriana, Hilabi, and Hananto 2023), who rated its performance as "Excellent Classification." Naïve Bayes, on the other hand, excels in classification tasks due to its proficiency with small datasets, efficiency in parameter selection, and quick execution times (Julianto, Kurniadi, and Jr 2023; Pebdika, Herdiana, and Solihudin 2023). Research by (Insan, Hayati, and Nurdiawan 2023) rated Naïve Bayes as achieving a "Good Classification" performance. Consequently, K-NN and Naïve Bayes were selected as the algorithms for the classification process in this study.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

METHOD

The research methodology consists of four phases, depicted in the research framework presented in Figure 2.

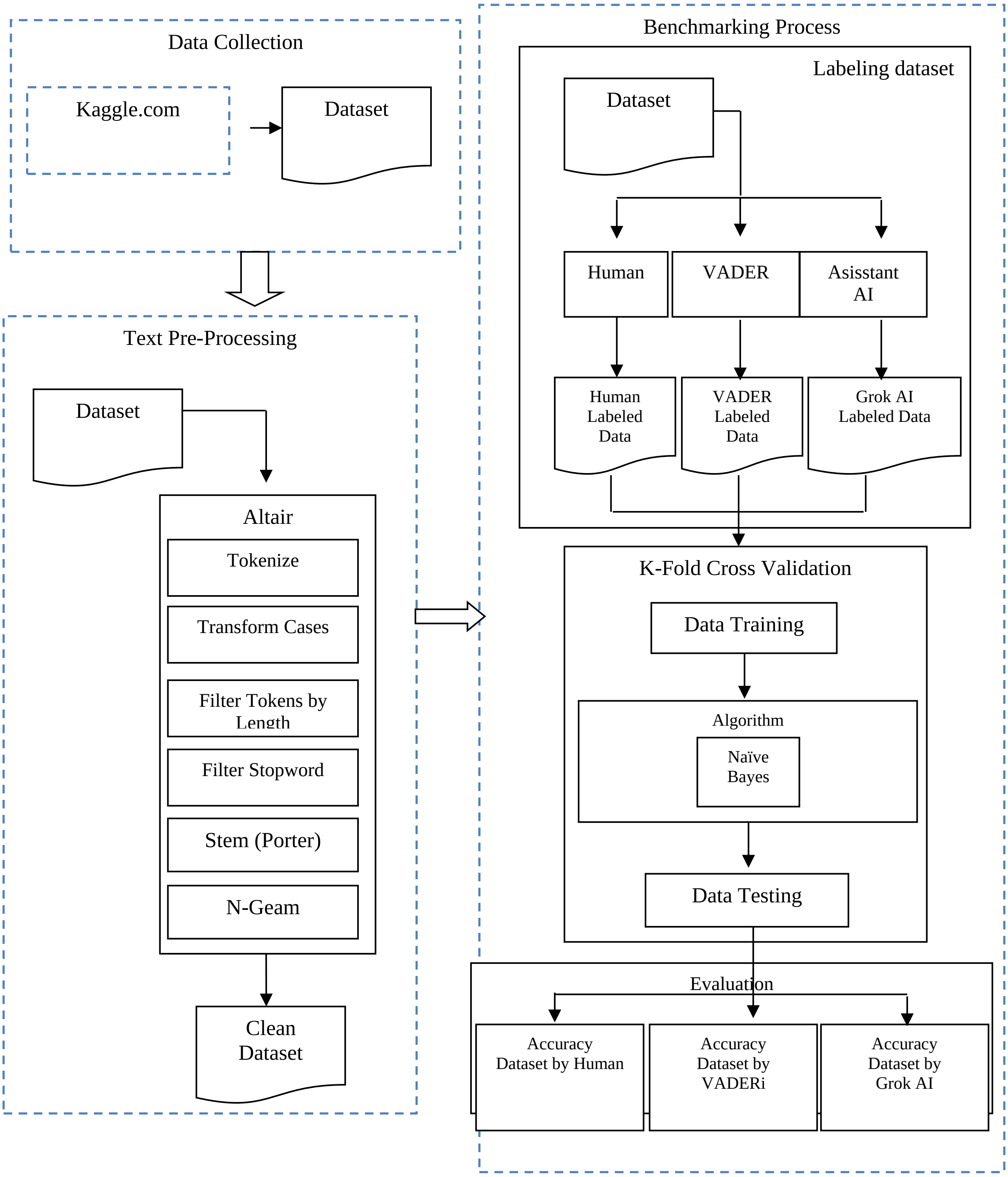

Fig 2. Research Framework

In The research methodology is divided into four key stages, as outlined below:

The first stage is Data Collection. The data, consisting of 500 manually labeled financial sentiment entries, is obtained from Kaggle.com (Kaggle, 2022).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The second stage is Text Pre-Processing. Although the dataset from Kaggle.com is already free of noisy and redundant data, additional pre-processing steps are applied to evaluate the performance of each labeling model. These steps include:

1. Tokenize: Breaking down sentences into smaller units, or tokens (Prasetya et al. 2021);
2. Transform Cases : To standardize the text, all existing capital letters will be changed to lowercase. This process is implemented to maintain consistent letter casing during the classification model's processing, thereby avoiding potential issues in the tokenization stage (Watori, Aryanti, and Junaidi 2020);
3. Filter Token by Length: Restricting words based on minimum and maximum character limits (Julianto et al. 2024; Julianto, Kurniadi, and Jr 2023);
4. Filter Stopword : Removing irrelevant or insignificant words from sentences using a predefined stopword list (Sutedi, Julianto, and Fitriani 2024) ;
5. Stem (Porter): Reducing words to their root forms (Albab, P, and Fawaiq 2023; Hidayat, Cahyana, and Julianto 2024);
6. N-Grams : involves taking an N-character portion from a string, will divide words according to their order within a sentence (Khairunnisa, Adiwijaya, and Faraby 2021).

The third stage is Dataset Labeling. This phase involves labeling the dataset using four distinct approaches:

1. Human Labeling: Manual annotation by individuals;
2. VADER Lexicon Labeling: Annotation using the VADER Lexicon tool;
3. Grok AI Labeling: Annotation performed by the Grok AI from X;

The fourth stage is Model Validation. During this phase, the effectiveness of each labeling outcome will be evaluated using the K-Nearest Neighbors (K-NN) and Naïve Bayes algorithms, with model validation performed through K-Fold Cross Validation. The K-Nearest Neighbors algorithm is commonly applied in classification tasks. It operates by categorizing data into classes based on their proximity or similarity to the training data (Pratama, Umaidah, and Voutama 2021). The steps of this algorithm include:

1. Specifying the value of k;
2. Computing the distance between the data to be classified and the labeled data;
3. Identifying the k smallest distances;
4. Assigning a class to the data based on a distance metric.
   The proximity calculation, utilizing a distance matrix, can be expressed with the following formula:

$$dist(X_{1,}X_2) = \sqrt{\sum_{i=1}^{n}(x_{1i} - x_{21})} \qquad (1)$$

The Naïve Bayes Algorithm, rooted in Bayes' theorem, is a classification technique that predicts future probabilities based on historical patterns (Insan et al. 2023). Its formula is represented as:

$$P(X|Y) = \frac{P(X|Y)P(X)}{P(Y)} \qquad (2)$$

The final part of this stage is Evaluation. Here, the accuracy of each model will be measured to identify which one delivers the highest performance.

RESULT

The initial stage yields a dataset sourced from Kaggle.com, comprising 499 entries of manually labeled financial sentiment data. This dataset is displayed in tabular format, as illustrated in Table 1.

Table 1. Dataset

No	Sentences	Sentiment
1	The GeoSolutions technology will leverage Benefon 's GPS solutions by providing Location Based Search Technology	Positive
2	ESI on lows	Negative
3	For the last quarter of 2010	Positive
4	According to the Finnish-Russian Chamber of Commerce	Neutral
5	The Swedish buyout firm has sold its remaining 22.4 percent stake	Neutral
….		
499	You 're not alone	Neutral

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The second stage produces a dataset prepared for performance assessment, having been processed through Text Processing steps such as Tokenize, Transform Cases, Filter Token By Length, Filter Stopword, Stem (Porter) and N-Grams. These pre-processing tasks were executed using Altair. The modeling process is depicted in Figure 3.

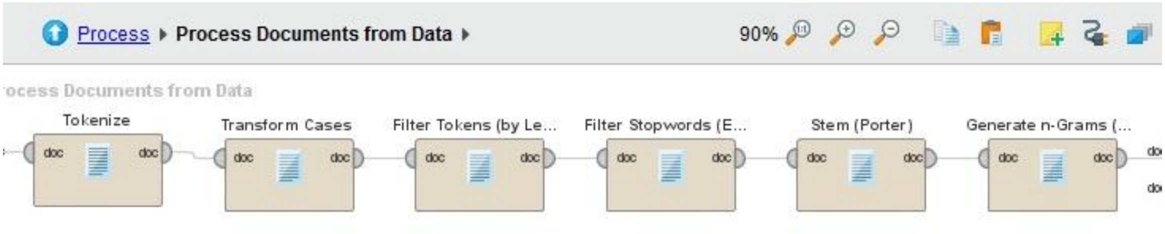

Fig 3. Text Processing

The third phase generates a dataset labeled by humans, VADER, and Grok AI. The human labeling was already finalized, as the dataset sourced from Kaggle came pre-labeled manually. VADER-based labeling was carried out using Altair, which incorporates the VADER operator. The modeling process is illustrated in Figure 4.

.

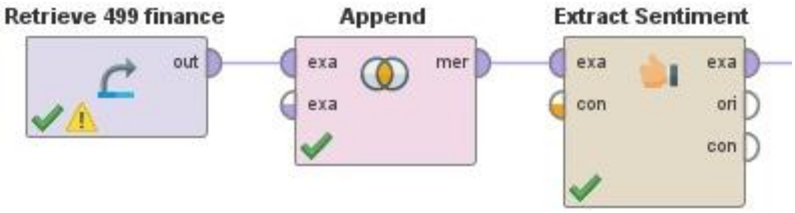

Fig 4. Labeling using VADER

This model serves as a specialized operator provided by the Altair framework, incorporating a feature that enables the use of VADER for enhanced sentiment analysis. The results derived from this detailed modeling process, which leverages both Altair's functionality and VADER's capabilities, are visually represented in Figure 5 for clarity and further interpretation.

Table 2. VADER Modeling Result

No	Sentences	Sentiment	Score
1	The GeoSolutions technology will leverage Benefon 's GPS solutions by providing Location Based Search Technology	Positive	0.67
2	ESI on lows	Negative	-0.69
3	For the last quarter of 2010	Positive	0.66
4	According to the Finnish-Russian Chamber of Commerce	Neutral	0
5	The Swedish buyout firm has sold its remaining 22.4 percent stake	Neutral	0
….			
499	You 're not alone	Neutral	0

The outcomes of the labeling process are reflected in the Score attribute, which serves as the basis for classification. According to the established criteria, a score greater than or equal to 0.05 is categorized as positive, a score less than or equal to -0.05 is deemed negative, and a score of exactly 0 is considered neutral, as outlined by (Illia et al. 2021). Analysis of the modeling results reveals a distribution of 239 instances labeled as positive, 86 as negative, and 174 as neutral, providing a comprehensive breakdown of the dataset's sentiment composition. Subsequently, the next phase of labeling was conducted using Grok AI, with the process executed directly on the Grok AI platform for efficiency and accuracy. A detailed illustration of this labeling procedure is provided in Figure 6, offering a clear visual representation of the methodology and its outcomes.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

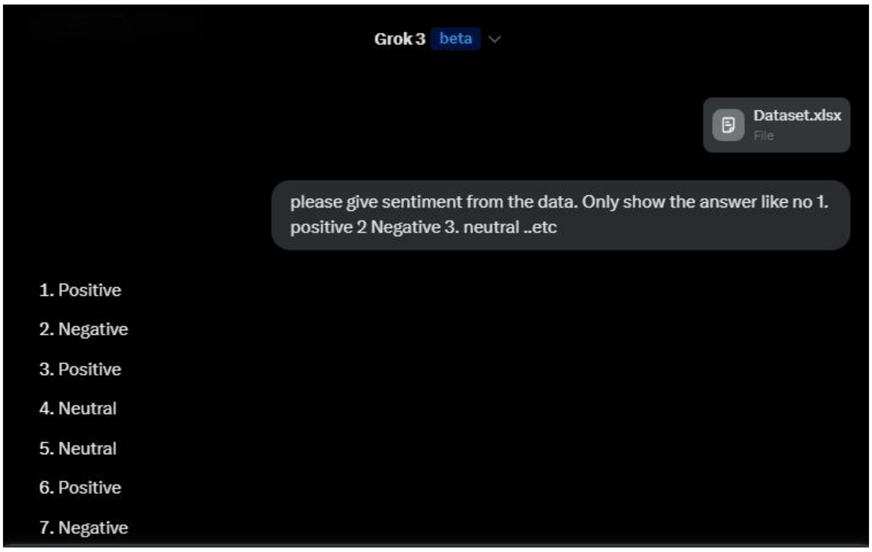

Fig 6. Labeling using Grok AI

The labeling procedure yielded a total of 151 instances classified as positive, 125 identified as negative, and 223 designated as neutral as seen in Figure 7, reflecting a balanced yet diverse distribution of sentiment across the dataset. During this process, the specific instruction given was, "Please give sentiment from the data. Only Show the answer like no 1. Positive, 2 Negative, 3. Neutral etc", ensuring a straightforward and systematic approach to sentiment categorization. This clear directive facilitated consistent labeling, allowing for an accurate assessment of the emotional tone within the text, which is critical for subsequent analysis and interpretation of the results.

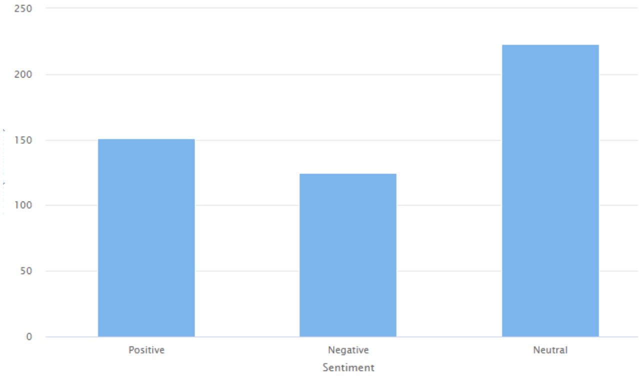

Fig 7. Sentiment Visualization

In the fourth phase, a comprehensive evaluation is conducted to generate performance metrics tailored to each dataset, providing valuable insights into the effectiveness of the applied methods. The modeling efforts were executed using Altair, a robust platform that facilitated the implementation of two distinct algorithms: Naïve Bayes and K-Nearest Neighbors (K-NN), chosen for their suitability in classification tasks. To ensure the reliability and generalizability of the results, validation was carried out through the K-Fold Cross Validation technique, which systematically assesses model performance by partitioning the data into multiple subsets for training and testing. This meticulous process, encompassing both the modeling and validation steps, is visually detailed in Figure 8, offering a clear and structured representation of the methodology and its outcomes for further analysis and discussion.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

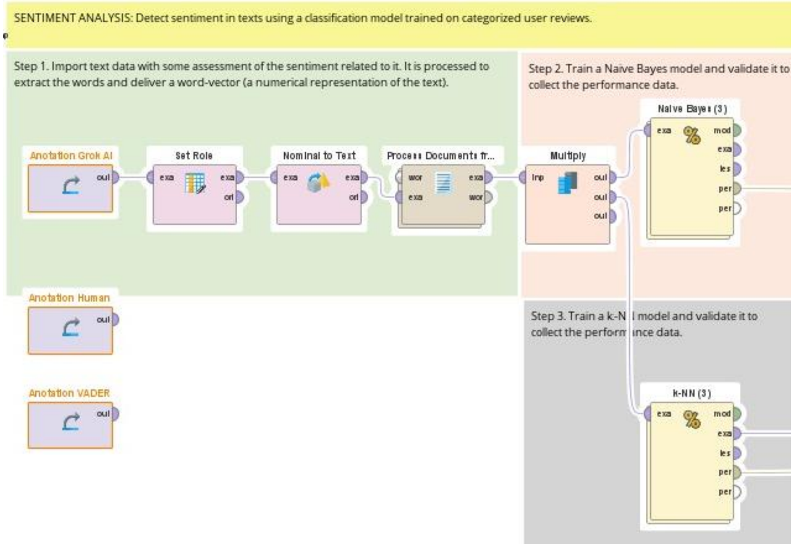

Fig 8. Performance Testing Model

The performance evaluation for each dataset is conducted using a consistent methodology, employing identical modeling techniques and operators to ensure uniformity across the tests, with the sole variable being the substitution of the dataset itself. This standardized approach allows for a fair and reliable comparison of how the models perform across different data inputs, isolating the impact of the dataset on the outcomes. The detailed results of these model performance tests, which highlight key metrics and insights derived from the analysis, are systematically compiled and presented in Table 3, providing a clear and concise overview for review and interpretation.

Table 3. Result Performance Test

No	Dataset	K-NN		Naïve Bayes	
		Accuracy	Kappa	Accuracy	Kappa
1	Labeling (Human)	60.52%	0.205	53.51%	-0.012
2	Labeling (VADER)	48.51%	0.207	49.49%	0.060
3	Labeling (Grok AI)	51.71%	0.205	48.50%	0.091

DISCUSSION

Grok AI leverage sophisticated natural language processing (NLP) techniques to transform textual input into a structured, machine-readable format, enabling deeper analysis. Utilizing this processed data, the AI Assistants compare the text against patterns they have previously been trained on, drawing from a repository of learned linguistic behaviors and structures. The outcome of this comparison serves as the foundation for classifying the text's sentiment, determining whether it conveys a positive, negative, or neutral tone. The key distinction between the two AI Assistants lies in their approaches to pattern recognition, which encompasses a blend of language models, sentiment lexicons, and contextual understanding. Each assistant adopts a unique strategy for recognizing these patterns, significantly influencing the precision with which they assign sentiment labels.

Factors such as the complexity of the language model, the comprehensiveness of the sentiment dictionary, and the ability to interpret contextual nuances all contribute to these differences in performance. Ultimately, the AI Assistant that most closely aligns with or replicates human-labeled sentiment is deemed the superior option, as it better mirrors human judgment. In this regard, Grok AI demonstrated a notably higher accuracy rate of 51.71% when evaluated using the K-Nearest Neighbors (K-NN) algorithm, a result likely attributable to Grok's more refined and advanced pattern recognition capabilities, which enable it to better discern subtle linguistic and emotional cues.

CONCLUSION

In conclusion, the findings of this study demonstrate that employing alternative labeling approaches through AI assistants such as Grok AI from X for sentiment analysis is a viable option, as evidenced by performance test outcomes that reveal no substantial disparities when compared to human labeling. Among the methods evaluated, human labeling paired with the K-Nearest Neighbor (K-NN) algorithm emerged as the most effective, delivering a commendable accuracy of 53.71% and a Kappa statistic of 0.205, indicating a moderate level of agreement beyond chance. The primary objective of this research was to assess and compare the efficacy of various labeling

*name of corresponding author

techniques, shedding light on their relative strengths and limitations. These results pave the way for future investigations, which could focus on refining and boosting the performance metrics achieved here, potentially by optimizing the algorithms or incorporating a broader range of AI assistants. Such efforts could further enhance the reliability and precision of automated sentiment analysis, bridging the gap between machine-driven and human judgment in increasingly sophisticated ways.

.

REFERENCES

Abro, Abdul Ahad, Mir Sajjad Hussain Talpur, and Awais Khan Jumani. 2023. "Natural Language Processing Challenges and Issues : A Literature Review." *Journal of Science* 36(4):1522–36.

Albab, M. Ulil, Yohana Karuniawati P, and Mohammad Nur Fawaiq. 2023. "Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic." *Jurnal TRANSFORMATIKA* 20(2):1–10.

Amaliah, Fitrah, and I. Kadek Dwi Nuryana. 2022. "Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter." *Journal of Informatics and Computer Science* 3(3):384–93.

Andriana, Herda, Shofa Shofia Hilabi, and Agustia Hananto. 2023. "Penerapan Metode K-Nearest Neighbor Pada Sentimen Analisis Pengguna Twitter Terhadap KTT G20 Di Indonesia." *JURIKOM (Jurnal Riset Komputer)* 10(1):60–67.

Anwar, Muchamad Taufiq, and DennyPermana Arief Permana. 2023. "Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER." *Jurnal Teknik Informatika Dan Sistem Informasi* 10(1):783–92.

Ashari, Ilham Firman. 2023. "Analysis Sentiments In Facebook Down Case Using Vader And Naive Bayes Classification Method." *Multitek Indonesia: Jurnal Ilmiah* 16(2):75–89.

Asri, Yessy, Widya Nita Suliyanti, Dwina Kuswardani, and Muhamad Fajri. 2022. "Pelabelan Otomatis Lexicon Vader Dan Klasifikasi Naive Bayes Dalam Menganalisis Sentimen Data Ulasan PLN Mobile." *PETIR: Jurnal Pengkajian Dan Penerapan Teknik Informatika* 15(2):264–75.

Barik, Kousik, and Sanjay Misra. 2024. "Analysis of Customer Reviews with an Improved VADER Lexicon Classifier." *Journal of Big Data* 11(10):1–29.

Basha, M. John, S. Vijayakumar J. Jayashankari, Ahmed Hussein, and Alawadi Pulatova. 2023. "Advancements in Natural Language Processing for Text Understanding." Pp. 1–9 in *E3S Web of Conferences 399*. Vol. 04031.

Cholil, Saifur Rohman, Titis Handayani, Rastri Prathivi, and Tria Ardianita. 2021. "Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa." *IJCIT (Indonesian Journal on Computer and Information Technology)* 6(2):118–27.

Gaja, Muhammad Yusuf Rismanda, Iqbal Maulana, and Oman Komarudin. 2023. "Analisis Sentimen Opini Pengguna Aplikasi Vidio Pada Ulasan Playstore Menggunakan Algoritma Naive Bayes." *JATI (Jurnal Mahasiswa Teknik Informatika)* 7(4):2767–74.

Hidayat, Taupik, Rinda Cahyana, and Indri Tri Julianto. 2024. "Analisis Sentimen Layanan Sistem Informasi Akademik Mahasiswa Menggunakan Algoritma Naive Bayes." *Algoritma* 21(1):119–30.

Illia, F., M. P. Eugenia, and S. A. Rutba. 2021. "Sentiment Analysis on PeduliLindungi Application Using TextBlob and VADER Library." Pp. 278–88 in *The 1 International Conference on Data Science and Official Statistics (ICDSOS)*. Jakarta: Politeknik Statistika.

Insan, Moh Khoirul, Umi Hayati, and Odi Nurdiawan. 2023. "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes." *JATI (Jurnal Mahasiswa Teknik Informatika)* 7(1):478–83.

Julianto, Indri Tri, Dede Kurniadi, and Benedicto B. Balilo Jr. 2023. "Enhancing Sentiment Analysis With Chatbots : A Comparative Study Of Text Pre-Processing." *JUTIF* 4(6):1419–30.

Julianto, Indri Tri, Dede Kurniadi, Benedicto B. Balilo Jr, and Fauza Rohman. 2024. "The Role Of Feature Selection In Enhancing The Accuracy Of AI Assistant Auto-Labeling." *Jurteksi* 11(1):85–92.

Julianto, Indri Tri, Dede Kurniadi, Yosep Septiana, and Ade Sutedi. 2023. "Alternative Text Pre-Processing Using Chat GPT Open AI." *Janapati* 12(1):67–77.

Khairunnisa, Syifa, Adiwijaya Adiwijaya, and Said Al Faraby. 2021. "Pengaruh Text Preprocessing Terhadap Analisis Sentimen Komentar Masyarakat Pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)." *Jurnal Media Informatika Budidarma* 5(2):406–14.

Moeis, Dikwan, Nasir Usman, Andi Harmin, Muhammad Faisal, Ida Mulyadi, and Musdalifa Thamrin. 2024. "Microsoft Copilot Training for Monitoring Student Learning : A Case Study Vocational High School Makassar - Indonesia." *Jurnal Pengabdian Masyarakat* 4(3):1911–22.

Nurwanda, Nana Suarna, and Willy Prihartono. 2024. "Penerapan NLP (Natural Language Processing) Dalam Analisis Sentimen Pengguna Telegram Di Playstore." *JATI (Jurnal Mahasiswa Teknik Informatika)* 8(2):1841–46.

Parlika, Rizky, Sunu Ilham Pradika, Amir Muhammad Hakim, and Kholilul Rachman N. M. 2020. "Analisis

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14632
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Sentimen Twitter Terhadap Bitcoin Dan Cryptocurrency Berbasis Python TextBlob." *Jurnal Ilmiah Teknologi Informasi Dan Robotika* 2(2):33–37.

Pebdika, Angga, Ruli Herdiana, and Dodi Solihudin. 2023. "Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima PIP." *JATI (Jurnal Mahasiswa Teknik Informatika)* 7(1):452–58.

Prasetya, Adam, Ferdiansyah Ferdiansyah, Yesi Novaria Kunang, Edi Surya Negara, and Winoto Chandra. 2021. "Sentiment Analisis Terhadap Cryptocurrency Berdasarkan Comment Dan Reply Pada Platform Twitter." *Journal of Information Systems and Informatics* 3(2):268–77.

Pratama, Aditiya Yoga, Yuyun Umaidah, and Apriade Voutama. 2021. "Analisis Sentimen Media Sosial Twitter Dengan Algoritma K-Nearest Neighbor Dan Seleksi Fitur Chi-Square (Kasus Omnibus Law Cipta Kerja)." *Sains Komputer & Informatika* 5(2):897–910.

Rachmat, Budi Kurniawan, Achmad Suwarisman, Iis Afriyanti, Aditya Wahyudi, and Dedi Dwi Saputra. 2023. "Analisis Sentimen Complain Dan Bukan Complain Pada Twitter Telkomsel Dengan SMOTE Dan Naïve Bayes." *Jurnal Teknologi Informasi Dan Komunikasi)* 7(1):107–13.

Russel, S. J., and P. Norvig. 2021. *Artificial Intelligence: A Modern Approach (4th Ed.)*. London: pearson.

Sutedi, Ade, Indri Tri Julianto, and Leni Fitriani. 2024. "Segmentasi Wilayah Terdampak Bencana Berdasarkan Fitur Geo-Posisi." *JTIIK* 11(4):797–804.

Telnoni, Patrick Adolf, Suryatiningsih, and Ely Rosely. 2020. "Pelabelan Data Dengan Latent Dirichlet Allocation Dan K-Means Clustering Pada Data Twitter Menggunakan Bahasa Indonesia Data Labeling Using Latent Dirichlet Allocation and K-Means Clustering on Indonesian-Based Twitter." *Jurnal Elektro Telekomunikasi Terapan (JETT)* 7(2):885–92.

Utami, Dian Siti, and Adhitia Erfina. 2021. "Analisis Sentimen Pinjaman Online Di Twitter Menggunakan Algoritma Support Vector Machine (SVM)." *SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika)* 1(1):299–305.

Watori, Jananto, Riska Aryanti, and Agus Junaidi. 2020. "Penggunaan Algoritma Klasifikasi Terhadap Analisa Sentimen Pemindahan Ibukota Dengan Pelabelan Otomatis." *Jurnal Informatika* 7(1):85–90.

Xiao, Yu, and Zhezhi Jin. 2021. "Summary of Research Methods on Pre-Training Models of Natural Language Processing." *Open Access Library Journal* 8(7).