

Theoretical Analysis of Naive Bayes and Selective Naive Bayes Machine Learning Algorithms

Jaison Victor Joseph Francis Xavier & Sri Krishna Pathuri

Abstract

This paper provides a comprehensive theoretical analysis of the extensively employed machine-learning algorithm Naive Bayes and its variant, the Selective Naive Bayes algorithm. The complexity, effectiveness, and scalability of both schemes are analyzed in detail to reveal their theoretical foundations and identify potential development avenues [1] [2]. The analysis takes a comprehensive approach, encompassing relevant aspects such as problem identification, literature review, formulation of a theoretical framework, analysis of properties, and comparative analysis.

Despite being a variant of the Naive Bayes algorithm, the Selective Naive Bayes algorithm exhibits distinct characteristics and, by some metrics, increased efficiency [3]. In addition, it is essential to compare their efficacy in diverse applications, such as spam detection, medical diagnosis, and sentiment analysis [4] [5] [6].

The purpose of this paper is to provide valuable methodological recommendations for algorithm improvement by engaging with prior research and critically evaluating the utility and shortcomings of these algorithms. Even though the Naive Bayes algorithm and its variant utilize relatively simple mathematical logic, their capacity for complex classification tasks [7] is notable. Therefore, any advancements in our understanding of these algorithms have significant ramifications, potentially affecting a variety of machine learning research and application domains.

In accordance with extant exploratory studies [11][12], it is expected to investigate the strengths of these algorithms and their suitability for various tasks. The simplicity and efficiency of Naive Bayes classification make it an excellent choice for large datasets or real-time applications [13]. However, when dealing with complex relationships and interdependencies within the data, its simplistic approach can become a limitation [14]. Selective Naive Bayes, on the other hand, extends the Naive Bayes algorithm by choosing a subset of features to consider, thereby potentially enhancing performance [15]. However, the optimal feature selection procedure may pose additional computational challenges and contribute to overfitting due to an increase in complexity and computational burden. This research is poised to investigate these theoretical strengths and limitations in greater depth, define the contexts in which each algorithm excels, and propose methods for improving their robustness and generalizability [16][17].

Finally, the clarity, coherence, and relevance of the research paper's results are evaluated, with a particular emphasis on their practical implications. The anticipated findings are expected to advance our theoretical comprehension of the Naive Bayes and Selective Naive Bayes algorithms, identify subtleties that may contain key improvement potential, and provide a comprehensive analysis of their properties and performance [8] [9] [10]. The study's practical implications could extend to any field where machine learning is used as a research or analytical instrument, thereby having a broad impact.

Index Terms

Naive Bayes, Selective Naive Bayes, Machine Learning, Algorithm Analysis, Theoretical Framework, Property Analysis, Performance Improvement

I. INTRODUCTION

ALGORITHMS for machine learning are techniques used to address a variety of problems, facilitate decision-making, and predict outcomes. With accelerated technological advancement and data proliferation, the significance of machine learning in a variety of disciplines, such as healthcare, finance, information technology, and business analytics, among others, has grown [1].

Naive Bayes (NB) and Selective Naive Bayes (SNB) are two machine learning classification principles widely used in the aforementioned fields. Naive Bayes is favored due to its simplicity, dependability, scalability, and effective implementation across a wide range of contexts, including spam detection, medical diagnosis, and customer satisfaction [3][4][5]. However, the Naive Bayes classifier implies the strong independence of features, which may not always be the case in practical applications, resulting in subpar performance [6].

The variant Selective Naive Bayes algorithm was developed to circumvent the excessively simplistic approach of the Naive Bayes algorithm. Instead of considering all features, SNB chooses a subset, potentially enhancing performance [7]. The significance of determining what this "improvement" entails and whether its optimality can be applied to a variety of applications [8] is crucial. Consequently, while both variants are founded on similar theoretical principles, they have distinct advantages and disadvantages and may be better adapted to particular problems or data types.

<https://github.com/jaison304/Machine-Learning-Project/upload/main>

Despite the straightforward application and practical success of these algorithms, theoretical voids remain when examining their properties. Prior research on Naive Bayes has primarily focused on empirical comparisons and analyses of its efficacy in various contexts. The lack of a comprehensive comprehension of the theoretical properties of Naive Bayes and, more importantly, the Selective Naive Bayes algorithm [9] has been left by these investigations.

This study aims to cover this vacuum by revealing the theoretical foundations of these algorithms and identifying their properties that are not well understood [10][11]. It juxtaposes the two algorithms in a comparative analytics arrangement in order to evaluate their efficacy critically and identify potential development avenues [12]. This invariant analysis offers insightful insights into the performance characteristics, capabilities, and limitations of both algorithms when dealing with data of differing nature and complexity.

The purpose of this paper is to shed light on aspects of these algorithms that are frequently overlooked in the literature, thereby contributing to our collective understanding of these tools [13]. It is a crucial step in attaining a comprehensive comprehension of machine learning algorithms, ultimately facilitating the creation of sophisticated, versatile, and robust systems [14]. This paper's theoretical analysis seeks to facilitate the optimization of these machine learning algorithms, leading to enhanced performance in machine learning applications [15].

In this context, this paper will examine in detail the following topics: problem identification, literature review, development of a theoretical framework, property analysis, comparative analysis, theoretical insights, and result evaluation [16]. It is anticipated that the study's findings will advance the theoretical comprehension of these algorithms and open up new avenues in extant Machine Learning and Data Mining methodologies.

The emphasis is on delivering a clear and coherent analysis that demonstrates not only theoretical but also practical relevance. By meticulously evaluating the properties of Naive Bayes and Selective Naive Bayes, this study aims to uncover blueprints for their enhancement, thereby facilitating significant advances in machine learning applications [17][18][19][20].

This research is important because it contributes to a broader understanding of the factors influencing the efficacy of these algorithms, identifies their fundamental properties, and establishes a connection between their performance and these properties [7]. In addition, it will investigate any discrepancies between the theoretically ideal conditions for optimal performance of these algorithms and the conditions observed in real-world use cases. By combining a solid theoretical foundation with a comprehension of practical contexts, the research creates a bridge between theory and practice, thereby enhancing the applicability of the findings. It identifies any fundamental elements of the algorithms that may have been neglected in their practical application and highlights any enhancements that could considerably increase their effectiveness [8][9][10].

II. MOTIVATION

The motivation for conducting this extensive theoretical research on the Naive Bayes and Selective Naive Bayes algorithms stems from the obvious lack of theoretical investigation surrounding these widely used machine-learning tools. The theoretical properties that regulate the functioning, strengths, limitations, and nuances of these algorithms [1][2] remain unexplored despite the multitude of empirical studies conducted. This knowledge gap necessitates an exhaustive comprehension of their underlying theoretical principles, which serves as the primary impetus for this research.

Significantly, the applicability of comprehensive theoretical understanding extends beyond the confines of academic learning and into real-world situations, offering potential benefits across industries. In the healthcare industry, for instance, an increase in the efficacy of these algorithms can impact the precision of diagnoses, potentially saving lives. In the financial industry, it can improve the accuracy of predictive instruments used for market analysis and forecasting, thereby facilitating improved decision-making [8][9]. Similarly, developments in these algorithms could bolster IT's cybersecurity measures, thereby protecting vital data and systems [10].

Simultaneously, this research can make a substantial contribution to the foundational knowledge of the constantly evolving field of machine learning. By providing new perspectives on algorithm analysis, [11] could foster innovation and inspire novel research trends. Inherent in these motivations is the aspiration to challenge the accepted norms of existing methodologies, thereby challenging the existing limits of comprehension while simultaneously contributing to the expansion of knowledge [12][13].

The objective of separating these theoretical needles from the haystack is to provide significant insights into the machines that accelerate the adoption and application of machine learning across a variety of industries. As a result, with these motivations as the impetus, this research aims to make significant advancements in the vast field of machine learning, paving the way for future researchers to investigate, analyze, and create with a deeper understanding of their tools.

III. MAIN CONTRIBUTION

- 1) The primary purpose of this research paper is to offer a thorough theoretical analysis of the Naive Bayes and Selective Naive Bayes algorithms.

- 2) The purpose of this study is to delve deeper into the important properties of these algorithms, such as complexity, effectiveness, and scalability, and to provide a better understanding of their theoretical foundations.
- 3) Identify and analyze potential enhancement areas in these algorithms, thereby contributing to the development of more sophisticated and effective machine learning tools.
- 4) Provide a comparative analysis of Naive Bayes and Selective Naive Bayes in order to comprehend their advantages and disadvantages, as well as the situations in which each variant offers the best performance.
- 5) Evaluate the results of the theoretical analysis for clarity, coherence, and applicability, ensuring that the findings can be used to improve the performance of machine learning applications in various disciplines.
- 6) Contributions are anticipated to include the provision of a more comprehensive understanding, the identification of key challenges and limitations, and the provision of methodological suggestions for improvement.
- 7) Contribute to the broader field of machine learning in academia and industry by highlighting the practical implications and potential impact of these theoretical findings.

IV. RELATED WORKS

Multiple studies have examined the properties, effectiveness, and applications of Naive Bayes and its variant, the Selective Naive Bayes, in various disciplines.

Nosiel et al. Despite its simplicity, the study revealed that Naive Bayes could perform classification tasks with a high degree of accuracy.

Li and Liu [2] examined Naive Bayes in greater detail as part of their causality-based attribute weighting research. They utilized a genetic algorithm to determine weightings and analyzed how various weightings affected the effectiveness of the Naive Bayes method in their experiments. Their discovery enhanced understanding of the sensitivity of the Naive Bayes classifier to feature weighting.

In their study, Xue, Wei, and Guo [3][6] demonstrated the applicability of Naive Bayes by developing a real-time Naive Bayes classifier accelerator on FPGA. The study demonstrated how modifying the algorithm's properties can improve its efficacy, especially in real-time applications.

In Zamroni et al.'s [12] study, which utilized Naive Bayes for the classification of cardiac disease, medical contexts were observed. Simultaneously, the investigation of Damanik et al. [14] utilized the algorithm to predict mortality due to Covid-19. These studies highlighted the algorithm's sensitivity to variations in its application context, thereby enhancing our knowledge of the Naive Bayes algorithm.

In their investigation [7], Ke et al. also demonstrated the optimization potential of Naive Bayes. Using optimized Naive Bayes models, the researchers predicted the occurrence of rock-bursts, demonstrating the algorithm's adaptability when augmented with optimization techniques.

Both Syrien et al. [9] and Rahmadani et al. [10] investigated the efficacy of the Naive Bayes algorithm in the context of sentiment analysis. With a focus on the algorithm's potential in social media contexts, these studies provided valuable insights into the algorithm's utility in handling unstructured data.

In their study [11], Religia and Maulana presented Genetic Algorithm optimization on Naive Bayes for airline customer satisfaction classification. This study not only emphasized the algorithm's adaptability, but also the significance of feature selection and optimization for enhancing its efficacy.

Kushwah and Kulkarni [18] used Naive Bayes in the field of data mining to identify competitors from large unstructured datasets. This study demonstrated the efficiency and effectiveness of the Naive Bayes algorithm in managing large data sets and performing complex classification tasks.

In addition, Asnimar et al.'s [17] research reveals that the Naive Bayes algorithm has a substantial application in the education sector. They used the algorithm and its Decision Tree counterpart to classify potential scholarship recipients, demonstrating that Naive Bayes can be a valuable decision-making aid in the field of education.

Interestingly, in the field of consumer behavior, Kharisma et al. [16] contrasted various model combinations of the Naive Bayes algorithm using Term Weighting Techniques in a sentiment analysis study. The outcomes demonstrated the algorithm's efficacy in processing complex data and its adaptability to adjusting to increase efficiency.

Similarly, Palupi [4][19] classified the impoverished in Indonesia using the Naive Bayes algorithm. Through their research, they highlighted the algorithm's proficiency in dealing with diverse datasets, highlighting its significance in socioeconomic studies.

Fadil et al. [5] used Naive Bayes to construct a waste classifier, while Yusuf et al. [8] utilized an Autoregressive Exogenous Neural Network approximated by a Naive Bayes filter to examine fire behavior. These studies demonstrated the algorithm's adaptability and extensive applicability across a variety of disciplines, each of which has distinct requirements.

In an intriguing application-related study, Zaeni et al. [15] classified stride length using the Naive Bayes algorithm. These studies contribute to the diversification of the Naive Bayes classifier's applications, thereby strengthening its scalability.

The extensive body of research conducted within the context of the Naive Bayes algorithm revealed the algorithm's effectiveness in a variety of domains. While referencing the theoretical foundations, the majority of the publications [12][13][14] focused on application-related studies. Consequently, the primary objective of this study is to examine and compare the fundamental theoretical properties of Naive Bayes and Selective Naive Bayes. In light of the extensive range of applications and the demonstrated efficacy of these algorithms, this paper's effort to elucidate their theoretical aspects becomes especially pertinent and opportune in order to maximize their potential.

Consequently, this paper seeks to consolidate the numerous experiments and theories associated with the Naive Bayes and Selective Naive Bayes algorithms, refining the understanding of their complexities and exploring their performances under a variety of scenarios, thereby enhancing the existing body of literature from a novel angle.

V. PROPOSED FRAMEWORK

In its analysis of the Naive Bayes and Selective Naive Bayes algorithms, this study employs an exhaustive methodology. The proposed framework begins with a problem identification phase in which the emphasis is placed on identifying the specific aspects of these algorithms that require a deeper understanding and explanation [1][2]. The investigation focuses on the complexity, effectiveness, scalability, and fundamental properties of the algorithms. Potential improvement areas for these algorithms are identified.

Once the problem has been identified, we conduct a comprehensive literature review [3][4][5] of studies that have analyzed the performance of Naive Bayes and Selective Bayes in various contexts. This review sheds light on the biases of previous studies and identifies the research voids this study should investigate. The purpose of this stage is to collect as much information as feasible regarding the issues identified in the previous step.

A theoretical framework is then developed [6][7] after gaining a comprehension of the current research landscape and identifying the voids. This framework enables a clearer comprehension of the theoretical foundations of the Naive Bayes and Selective Naive Bayes algorithms [8] with the aid of research discussing the theory underlying the Naive Bayes and Selective Naive Bayes algorithms. This step lays the groundwork for subsequent analysis by establishing a direct link between the theoretical properties of the algorithms and their functional outcomes [9][10].

With a solid theoretical framework in place, the complexity, effectiveness, and scalability of the algorithms are analyzed. This analysis dissects each property, examining how it interacts with other properties and contributes to the algorithm's overall efficacy [11][12]. This is an essential step in comprehending the strengths of the algorithms, their applications, and potential advancement areas.

Following the property analysis is a comparative analysis in which Naive Bayes and Selective Naive Bayes are directly contrasted to determine their relative strengths and the contexts in which each performs optimally [13][14]. This analysis puts into practice the comprehension of the properties garnered in the preceding steps. It highlights the differences between the outcomes of the two algorithms, identifies areas where one may be superior to the other, and presents opportunities to enhance the less optimal algorithm [15][16].

The final phase of analysis is to derive theoretical conclusions from the analyses. This entails deciphering the comparative and property analyses, comprehending what they imply for the algorithms, and identifying potential improvement areas. [17][18].

Following the analyses, an evaluation is conducted to ensure the results' clarity, coherence, and applicability [19]. This method verifies that the theoretical analyses can be comprehended and applied to improve the performance of Naive Bayes, Selective Naive Bayes, and other machine learning algorithms.

The results of the comparative and property analyses provide a comprehensive comprehension of the intricate workings of the Naive Bayes and Selective Naive Bayes algorithms. These findings pave the way for identifying algorithmic enhancement opportunities [15, 16]. For instance, if a particular property appears to limit the efficacy of an algorithm in a specific context, the algorithm can be modified to compensate for these limitations. Similarly, if a property that improves efficiency in a particular use case is identified, it may be possible to apply this understanding to enhance performance in other scenarios.

The research assesses whether the theoretically ideal conditions for the algorithms correspond to the conditions encountered in real-world applications. This analysis is essential for comprehending the divide between theory and practice and for utilizing the insights gained from the practical application of these algorithms to better comprehend their theoretical foundations [16][17]. This exercise is aided by the vast data on this topic provided by a variety of previous studies.

This evaluation also contributes to the comprehension of the practical implications of the research findings. While emphasis is placed on the theoretical foundations of algorithms, the ultimate objective is to improve their functionality and efficacy. Therefore, it is essential to comprehend the practical implications of these analyses in order to make relevant and actionable recommendations for enhancement. These findings are applicable not only to machine learning, but also to any field where these algorithms are applied, thereby contributing to a vast multitude of disciplines. [18][19].

By strictly adhering to this framework, a profound understanding of the Naive Bayes and Selective Naive Bayes algorithms is achieved. In addition, a holistic view of how these properties interact reveals which properties must be prioritized

and which must be optimized. Therefore, it affords the opportunity to attempt to maximize the effectiveness of these algorithms based on the context in which they are employed.

Flowing through these steps—beginning with identification, moving through analysis, deriving insights from analysis, and concluding with a rigorous evaluation—offers a systematic method for comprehending the Naive Bayes and Selective Naive Bayes algorithms [20]. This systematic approach ensures the thoroughness and rigor of the research, rendering the findings trustworthy and the drawn conclusions valid.

Documentation of the process, findings, evaluations, and insights concludes the entire process of gaining an understanding of these algorithms via the proposed framework. By detailing the entire procedure, the research contributes to the existing corpus of literature and lays the groundwork for future investigation into these algorithms. It serves as a springboard for future research on each of the aspects examined in this paper. In addition, the detailed insights provided here can serve as a valuable resource for anyone working with or considering working with these algorithms, thereby making a significant contribution to the field.

VI. DATA DESCRIPTION

The research paper significantly relies on previous academic and scientific literature. This study is predominately based on theoretical data extracted from multiple academic sources that discuss and research the Naive Bayes and Selective Naive Bayes algorithms at length. These sources include peer-reviewed journals, conference papers, and other academic publications that have examined these algorithms in the past and contributed valuable findings to the academic community [1]-[20].

The used literature is extensive and comprises a multiplicity of studies that provide a comprehensive understanding of the Naive Bayes and Selective Bayes algorithms. This corpus investigates not only the conventional applications of these algorithms, but also their applicability and efficacy in nontraditional contexts. Consequently, this description of the collected data assists in orienting the research study toward a detailed exploration of the chosen algorithms, laying the groundwork for addressing the problem statement discussed previously. The subsequent sections of this research paper aim to conduct a more thorough analysis of these data and derive insightful conclusions that contribute to a broader comprehension of these commonly used machine learning algorithms.

The selected literature spans multiple methodologies and contexts, offering therapeutics for a variety of use-cases ranging from spam detection to heart disease prediction [1][12]. Also included are studies that employ Naive Bayes in a variety of disciplines, including education [17], information security [1], and sentiment analysis [9][10][16]. The breadth of the selected literature strengthens the robustness and dependability of the research's findings and conclusions. In addition, it assures a comprehensive comprehension of the investigated algorithms, which is essential for the purpose of this theoretical analysis.

VII. COMPARISON

As we delve deeper into the specifics of properties, it is important to note Naive Bayes' versatility in managing large datasets with nominal and numeric attributes [1][9]. This resiliency is largely attributable to its conditional independence assumption, which enables it to effectively manage irrelevant characteristics.

In contrast, the Selective Naive Bayes algorithm typically outperforms Naive Bayes in situations involving feature interdependencies. It is distinguished by the addition of a feature selection phase, which ensures a more precise and focussed classification [3][6]. This inclusion frequently results in the Selective Naive Bayes method outperforming the standard Naive Bayes method in terms of classification precision [5][11].

Alternatively, while the feature selection phase of Selective Naive Bayes improves accuracy, it also introduces complexity to the algorithm [6]. Considerations for entropy thresholds and cutoff points, which are frequently adjusted using techniques such as Genetic Algorithms [2], add to its complexity. Consequently, while the Selective Naive Bayes algorithm excels in high-dimensional problems, its complexity may prevent its use in less demanding applications.

In terms of data scalability, while both algorithms efficiently manage growing datasets, the Selective Naive Bayes variant excels when the feature set is large [3][13]. Here, its feature selection phase is advantageous because it filters out irrelevant features, thereby improving processing speed and precision [12][14]. This characteristic is especially advantageous when working with high-dimensional datasets and real-time applications.

The Naive Bayes algorithm has exhibited impressive efficacy across a variety of applications, including sentiment analysis [9][10][16] and healthcare predictions [12][14]. Popular due to its simplicity, sturdiness, and capacity to manage high-dimensional spaces. However, when relationships exist between predictors, the Selective Naive Bayes method can improve classification accuracy [4][6][15].

The Naive Bayes and Selective Naive Bayes models exhibit functional distinctions owing to their fundamental differences. The summary below compares them in terms of their foundations, complexity and efficiency, effectiveness, and scalability.

Table 1: Comparative Analysis of Naive Bayes and Selective Naive Bayes

Parameters	Naive Bayes	Selective Naive Bayes
Foundational Assumption	Assumes feature independence	Relaxes feature independence through feature selection
Complexity and Efficiency	Efficient, may struggle with extensive feature-sets	Added complexity through feature selection phase, typically more efficient
Efficacy	Robust and proficient but might dip under violated assumption of feature independence	Tends to outperform Naive Bayes when inter-feature relationships exist
Scalability	Demonstrates scalability, could be strained by many features	Better scalability through feature selection, particularly with high-dimensional data

Understanding complexity is a crucial factor to consider when conducting a comparative analysis. In this regard, the Naive Bayes algorithm excels due to its simplicity and independence assumptions [2][18]. The Naive Bayes algorithm's foundational concepts make it an ideal choice for machine learning novices and basic application contexts [1][10] because they are simple to comprehend and implement.

However, while Selective Naive Bayes is a variant of Naive Bayes, it transcends simplicity. The addition of a phase for feature selection inherently adds a layer of complexity [3][15]. Nevertheless, Selective Naive Bayes offers a trade-off between complexity and enhanced performance, which is particularly advantageous when dealing with high-dimensional data [3][6]. Thus, the decision depends on factors such as the classification accuracy requirement and the dimensionality of the problem.

A comprehension of efficacy provides additional substance to our comparison. With its simple mechanism of conditional independence-based probabilities, Naive Bayes provides efficient learning and prediction [2][10]. Nevertheless, Selective Naive Bayes, with its feature selection phase, frequently provides a higher level of efficacy, particularly when the dataset contains a large number of features [6][14]. In this case, there is a trade-off between the effectiveness of learning and the complication introduced by the feature selection procedure.

Similarly, it is crucial to consider how susceptible these algorithms are to data variations and disturbance. Despite its relative simplicity, the Naive Bayes algorithm demonstrates remarkable robustness to irrelevant features [1][8]. However, when the feature independence assumption is violated substantially, it may perform poorly [2][11]. Due to its feature selection phase, the Selective Naive Bayes algorithm may produce superior results in such situations [3][5].

In addition, when scalability considerations, such as growing data volumes or extensive feature sets, enter the picture, the situation becomes more complex. While the Naive Bayes algorithm is generally efficient, its processing speed is slowed by increasing complexity due to a large number of features [9][18]. In contrast, due to its design, the Selective Naive Bayes algorithm can effectively manage larger data volumes and a greater number of features. This is primarily attributable to the feature selection phase, which enables the algorithm to scale better with increasing dimensionality [3][6][13].

Collectively, these comparative characteristics demonstrate that the Naive Bayes and the Selective Naive Bayes models, despite having similar foundations, exhibit distinct properties and behaviors under various conditions. In addition, they emphasize the multifaceted nature of these algorithms and the need to delve deeper into their theoretical comprehension in order to optimize their application in a variety of contexts [17][19][20].

REFERENCES

- [1] Nosiel, N., Andriyanto, S., & Said Hasibuan, M. (2021). Application of Nave Bayes Algorithm for SMS Spam Classification Using Orange. *International Journal of Advanced Science and Computer Applications*.
- [2] Li, M., & Liu, K. (2019). Causality-Based Attribute Weighting via Information Flow and Genetic Algorithm for Naive Bayes Classifier. *IEEE Access*, 7, 150630-150641.
- [3] Xue, Z., Wei, J., & Guo, W. (2020). A Real-Time Naive Bayes Classifier Accelerator on FPGA. *IEEE Access*, 8, 40755-40766.
- [4] Palupi, E.S. (2022). CLASSIFICATION OF THE POOR IN INDONESIA USING NAIVE BAYES ALGORITHM AND NAIVE BAYES ALGORITHM BASED ON PSO. *Jurnal Riset Informatika*.
- [5] Fadil, I., Helmiawan, M.A., Supriadi, F., Saeppani, A., Sofiyan, Y., & Guntara, A. (2022). Waste Classifier using Naive Bayes Algorithm. *2022 10th International Conference on Cyber and IT Service Management (CITSM)*, 1-5.
- [6] Xue, Z., Wei, J., & Guo, W. (2020). A Real-Time Naive Bayes Classifier Accelerator on FPGA. *IEEE Access*, 8, 40755-40766.
- [7] Ke, B., Khandelwal, M., Asteris, P.G., Skentou, A.D., Mamou, A., & Armaghani, D.J. (2021). Rock-Burst Occurrence Prediction Based on Optimized Naïve Bayes Models. *IEEE Access*, 9, 91347-91360.
- [8] Yusuf, S.A., Alshdadi, A.A., Alghamdi, R., Alassafi, M.O., & Garrity, D.J. (2020). An Autoregressive Exogenous Neural Network to Model Fire Behavior via a Naïve Bayes Filter. *IEEE Access*, 8, 98281-98294.
- [9] Syrien, A., Hanumanthappa, M., & Sundaravadivazhagan, B. (2018). Sentiment Analysis of Tweets using Nave Bayes Algorithm through R Programming. *International Journal of Computer Sciences and Engineering*.
- [10] Rahmadani, P.S., Tampubolon, F.C., Jannah, A.N., Hutabarat, N.L., & Simarmata, A.M. (2022). Tiktok Social Media Sentiment Analysis Using the Nave Bayes Classifier Algorithm. *Sinkron*.
- [11] Religia, Y., & Maulana, D. (2021). Genetic Algorithm Optimization on Nave Bayes for Airline Customer Satisfaction Classification. *JISA(Jurnal Informatika dan Sains)*.
- [12] Zamroni, A., Novitawati, N., Rahayu, D.N., Doharma, R., & Taufik, A. (2022). HEART DISEASE CLASSIFICATION MODEL IN LEVEL II HOSPITAL MOH. RIDWAN MEUREKSA USING NAVE BAYES. *Infotech: Journal of Technology Information*.
- [13] Benmahamed, Y., Kemari, Y., Teguar, M., & Boubakeur, A. (2018). Diagnosis of Power Transformer Oil Using KNN and Nave Bayes Classifiers. *2018 IEEE 2nd International Conference on Dielectrics (ICD)*, 1-4.
- [14] Damanik, A.R., Hartama, D., & Damanik, I.S. (2022). Prediksi Tingkat Kematian di Indonesia Akibat Covid-19 Menggunakan Algoritma Naïve Bayes Prediction of Mortality Rate in Indonesia due to Covid-19 Using the Naïve Bayes Algorithm.
- [15] Zaeni, I.A., Primadi, W., Osman, M.K., Anzani, D.R., Lestari, D., & Handayani, A.N. (2022). The Naïve Bayes Algorithm for the Stride Length Classification. *2022 International Conference on Electrical and Information Technology (IEIT)*, 90-94.
- [16] Kharisma, R.S., Muttafi'ah, & Dahlan, A. (2021). Comparison of Naïve Bayes Algorithm Model Combinations with Term Weighting Techniques in Sentiment Analysis. *2021 4th International Conference on Information and Communications Technology (ICOIACT)*, 160-163.
- [17] Asnimar, Achmad, A., Yuyun, Iskandar, A.O., & Mansyur (2022). Classification Of Prospective Scholarship Recipients Kartu Indonesia Pintar (KIP) With Decision Tree Algorithm And Naïve Bayes. *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*.
- [18] Kushwah, A., & Kulkarni, Y.C. (2018). Identifying Competitors from Large Unstructured Dataset Using Nave Bayes Classifier and Apriori Algorithm. *International Journal of Computer Sciences and Engineering*.
- [19] Palupi, E.S. (2022). CLASSIFICATION OF THE POOR IN SUMATERA AND JAVA ISLAND USING NAIVE BAYES ALGORITHM AND NAIVE BAYES ALGORITHM BASED ON.
- [20] Okikiola, F.M., Adewale, O.S., & Obe, O.O. (2023). A DIABETES PREDICTION CLASSIFIER MODEL USING NAIVE BAYES ALGORITHM. *FUDMA JOURNAL OF SCIENCES*.