

## Enhancing Classification Performance Through Employing Feature Selection Algorithms

Wisam A. Qader <sup>1\*</sup> , Musa M.Ameen <sup>1</sup> , and Bilal I. Ahmed <sup>2</sup>

<sup>1</sup>Department of Computer Engineering, Faculty of Engineering, Tishk International University, Erbil, Iraq.

<sup>2</sup> Department of Information Technology, Faculty of Applied Science, Tishk International University, Erbil, Iraq.

### Article History

Received: 23.12.2024

Revised: 19.02.2025

Accepted: 24.02.2025

Published: 25.02.2025

Communicated by: Assist. Prof. Dr. Abubakar M. Ashir

\* Email Address:

[wisam.abdulaziz@tiu.edu.iq](mailto:wisam.abdulaziz@tiu.edu.iq)

\*Corresponding Author



Copyright: © 2024 by the author. Licensee Tishk International University, Erbil, Iraq. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 2.0 Generic License (CC BY-NC 2.0) <https://creativecommons.org/licenses/by-nc/2.0/>

**Abstract:** Data classification is a pivotal area of research due to its critical importance across a wide range of applications, such as healthcare, finance, and predictive analytics. This study introduces an innovative approach designed to enhance classification accuracy and produce highly precise outputs. The methodology incorporates five distinct feature selection algorithms—Quick Branch and Bound (QBB), Las Vegas Filter (LVF), Branch and Bound (B&B), Focus, and Sequential Floating Forward Search (SFFS)—to identify the most relevant features. These are followed by the application of five classification algorithms—Sequential Minimal Optimization (SMO), Stacking, Random Tree (RT), Naive Bayes (NB), and Decision Table (DT)—to categorize the data. The classifiers, selected from diverse classification families, were rigorously evaluated using seven datasets encompassing various domains. The proposed approach achieved accuracy improvements ranging from 0.843 to 0.966 for Naive Bayes, 0.837 to 0.964 for SMO, 0.806 to 0.962 for Decision Table, 0.783 to 0.941 for Random Tree, and 0.877 to 0.986 for Stacking, demonstrating a significant enhancement compared to traditional methods. Among the tested combinations, the integration of QBB as a feature selection method with the Stacking classification algorithm yielded the highest performance, achieving an accuracy of 99.2% in certain datasets. Ultimately, the combination of the best-performing feature selection algorithm and the most accurate classification algorithm was identified, showcasing the versatility and robustness of this approach for practical applications across multiple fields.

**Keywords:** Data Classification; Feature Selection; Naive Bayes; Decision Tree; Sequential Minimum Optimization; Random Tree; Stacking.

### 1. Introduction

Data classification plays a critical role in numerous applications, including healthcare, finance, and predictive analytics [1]. However, one of the primary challenges in classification tasks is handling high-dimensional feature spaces, where the presence of irrelevant and redundant features can negatively impact classification accuracy and increase computational complexity [2]. Feature selection methods help mitigate these issues by identifying the most relevant features while eliminating noise, thereby improving classification performance [3]. Despite advancements in feature selection techniques, many existing methods suffer from scalability issues and high computational costs, making them less effective for large, complex datasets [4]

There are two primary methods of categorization in use today; the first involves the manual application of rules, vocabularies, and patches/parts of pictures to the classification of texts and visual sceneries. The second option is to utilize automation, which employs a wide range of techniques from fields like machine learning and information retrieval to do classification [1].

Classification methods are generally categorized into two main approaches. The first is manual classification, which relies on predefined rules, vocabularies, and the segmentation of textual and

---

visual data into distinct components, such as patches or parts of images, to facilitate sorting. The second approach is automated classification, which leverages advanced methodologies, including machine learning algorithms and information retrieval techniques, to perform systematic and efficient categorization [1].

It is feasible to ensure a system that performs well for a specific domain by developing algorithms with high rate of accuracy [2]. Because of this, different algorithms of feature selection are employed to various classifiers to develop a powerful and new approach for the purpose of classification.

To address these limitations, this study aims to enhance classification accuracy by integrating advanced feature selection techniques with machine learning classifiers, systematically evaluating their impact on classification performance using diverse datasets. These algorithms are known for their effectiveness in dimensionality reduction and classification accuracy enhancement were utilized, including Sequential Floating Forward Search (SFFS), Branch and Bound (B&B), Quick Branch and Bound (QBB), Focus, and the Las Vegas Filter (LVF). These algorithms offer a mix of exhaustive, heuristic, and randomized approaches, which makes them highly suitable for handling diverse and complex datasets. Similarly, classifiers like Naive Bayes, Random Tree, and Stacking were chosen for their robustness in classification tasks, offering varied learning paradigms that make them ideal for comparison.

This article is structured as follows: Previous work done in relation to the proposed approach is described in Section 2. The used methods and materials in this research are explained in Section 3. Results and discussions are in Section 4. And the final section discusses the conclusion and future works which is section 5.

## 2. Literature Review

Selecting the most useful characteristics and zeroing down on the most pertinent ones according to a set of established relevance rules is the primary function of feature selection algorithms [3]. Classification is a fundamental part of Artificial Intelligence (AI), but it is also one of the most challenging tasks since it requires great performance or accuracy [4]. Combining techniques of feature selection with classification algorithms of machine learning can result in an output with a high accuracy rate [5]. There is evidence that it requires less time by using automated tools to classify documents than manual techniques; therefore, the value of classification algorithms can be clearly underlined in this situation [6].

There is a clear difference between performance of machine learning algorithms depending on the type of classification task, according to several research. For example, SVM delivers quite great results on disease classification, such as diabetes and cancer [7]. The accuracy of algorithms is also determined by the complexity and type of data to be classified [8]. In order to classify data from medical databases, Pediatric Medical Institution in Brazil combined and used algorithms of feature selection and machine learning classifications [9].

Recent advancements in feature selection algorithms, such as evolutionary-based methods, and deep learning-based classifiers have shown promise, but often suffer from high computational costs and overfitting in small datasets. Traditional methods like Branch and Bound or Sequential Floating Forward Search remain reliable for their simplicity and lower resource requirements. Despite the progress, challenges like scalability and generalizability remain, especially in large datasets with high-dimensional features. The proposed method aims to address these drawbacks by balancing computational efficiency with high accuracy, making it relevant for current AI applications.

In this study, multiple feature selection algorithms are adopted to various classification algorithms in order to present a high performance approach for classifying the cases.

---

### 3. Methods and Materials

Feature selection plays a crucial role in improving classification accuracy by eliminating redundant and irrelevant features. However, existing feature selection techniques often struggle with computational efficiency and scalability in high-dimensional datasets

LVF, SFFS, B&B, QBB, and Focus are just five of the feature selection algorithms used to simplify data. There are also numerous data classification algorithms, in which four of them chosen from five different classification taxon families. For classification purposes, those methods are extremely useful. SMO is selected from the Functions, NB is from the Bayes, RT is from the Trees family, and DT is from the Rules family.

The selected feature selection algorithms (LVF, SFFS, B&B, QBB, and Focus) were chosen due to their ability to efficiently reduce dimensionality while preserving classification performance. These algorithms offer a mix of heuristic, exhaustive, and randomized approaches, making them suitable for handling diverse datasets. Similarly, the classification algorithms (NB, SMO, DT, RT, and Stacking) were selected based on their effectiveness in different classification paradigms, ensuring a comprehensive evaluation of performance across multiple datasets

Six times each of the four classifiers are tested, once without use of any algorithms of feature selection and five times after using each of the five algorithms of feature selection, resulting in a significant increase in the accuracy of the results. Then Stacking algorithms is used to combine the predictions of all the above-mentioned classifiers and producing a very robust model which resulting in a single boosted performance model with a very good accuracy. The Stacking algorithm was subjected to six distinct evaluations to comprehensively assess its performance. The first evaluation was conducted without employing any feature selection algorithms, serving as a baseline for comparison. The subsequent five evaluations were carried out after applying each of the five feature selection algorithms to the datasets, allowing for a detailed analysis of the impact of feature selection on the algorithm's classification accuracy and overall effectiveness.

The increase in accuracies can be observed clearly when comparing the algorithms that showed their performances before using any of the algorithms of feature selection to the algorithms that showed their performances after using algorithms of feature selection.

Finally, combination of the Quick Branch and Bound feature selection algorithm with the Stacking classification algorithm demonstrated the highest accuracy among all evaluated methods, highlighting its superior performance in feature selection and classification. This finding underscores the effectiveness of integrating advanced feature selection techniques with robust classification algorithms to optimize predictive accuracy in complex datasets.

The model of the approach which is proposed in this work is shown in Fig. 1.

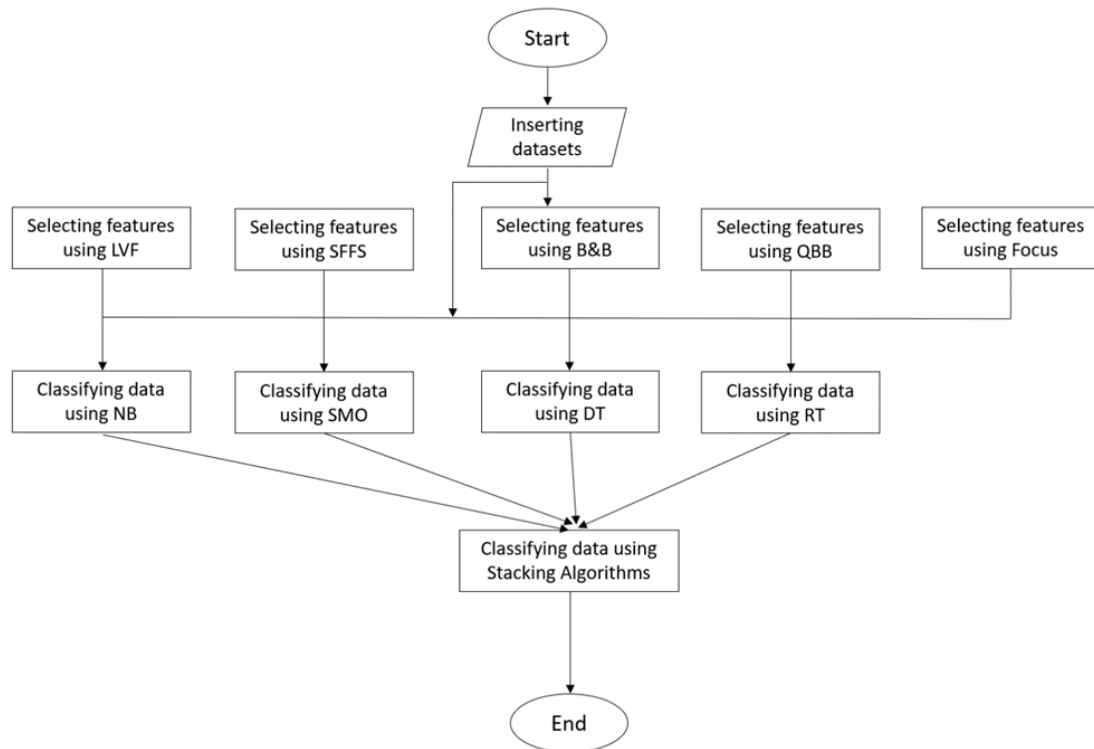


Figure 1: The structure of the proposed algorithm.

The following are the methods and materials used in this research work:

### 3.1 Data Collection

Patients' medical records from seven distinct examinations were compiled for this study, encompassing a variety of diagnostic procedures and their corresponding values. The datasets were sourced from one of the world's largest scientific dataset repositories, housed at the University of California, Irvine (UCI). Specifically, three datasets focused on gastrointestinal disorders, two on eye disorders, one on brain disorders, and one on skin disorders. Collectively, the databases included health data from 2,604 patients. To evaluate the proposed method, 70% of each dataset was allocated for training, while the remaining 30% was used for testing, ensuring a robust and reliable assessment of model performance.

The datasets selected for this study were chosen due to their representation of complex, real-world medical scenarios, including gastrointestinal, brain, and skin disorders. These datasets are frequently used as benchmarks in the field of medical data classification, allowing us to assess the effectiveness of our proposed approach across a variety of diagnosis tasks. Additionally, the diversity of these datasets ensures that the results are applicable to a wider range of classification problems.

### 3.2 Feature Selection

Various feature selection algorithms were employed in this study to reduce and eliminate irrelevant features from the datasets, ensuring that only the most significant attributes contributing to the results were retained. Specifically, the Las Vegas Filter (LVF), Sequential Floating Forward Search (SFFS), Branch and Bound (B&B), Quick Branch and Bound (QBB), and Focus algorithms were selected due to their proven effectiveness and impressive performance in optimizing feature selection processes.

#### 3.2.1 Las Vegas Filter (LVF)

The algorithm of Las Vegas Filter arbitrarily governs and steers their search, ensuring a correct arrangement and result independent over whether bad decisions are made; there might take more time if it occurs [10].

### 3.2.2 Sequential Floating Forward Search (SFFS)

It can be thought that the older SFFS is a kind of "Trees Search" with nodes representing feature sets and edges representing possible transition steps. It is important to note that the graph does not follow the strict structure of a tree, as each node can be accessed through multiple pathways, and a single child node can have connections from multiple parent nodes. The transitions between parents and children are maintained by the use of edges. There are a number of transitions that can be used in both directions. One of the transitions from parent to child is Sequential Forward Selection (SFS), after a comprehensive scan, the best child from each parent node's offspring is chosen. Sequential Backward Selection (SBS) is another method that transitions from a child node to its parent by eliminating a feature. If the resulting parent node demonstrates superiority over the most recently visited node, an update is performed at level  $k$ . In the traditional Sequential Floating Forward Search (SFFS) framework, both Sequential Forward Selection (SFS) and Sequential Backward Selection (SBS) conduct a thorough exploration of all potential nodes to ensure optimal feature selection. [11].

### 3.2.3 Branch and Bound (B&B)

In a manner analogous to dynamic programming, Branch and Bound is an algorithmically generated search approach for exploring the space of all possible solutions. In the case of minimization, all feasible solutions are routinely partitioned into ever-smaller subgroups, and the minimum cost of a solution is determined for each subset. Subset borders that surpass the cost of a known possible solution are eliminated from further partitioning after each iteration of the partitioning procedure. If the cost of any partition exceeds the bound, the partitioning procedure is redone until a workable solution is found. That was found by a group of researchers led by Morrison et al. [12].

### 3.2.4 Quick Branch and Bound (QBB)

The Quick Branch & Bound (QBB) technique is designed to reduce the number of criterion function computations required at internal nodes of the search tree. The Branch and Bound search algorithm relies on the criterion function to satisfy the monotonicity condition for effective implementation. The QBB approach can be summarized as follows: the algorithm aims to predict and estimate criterion values without performing explicit computations, leveraging knowledge of how the criterion values change when individual features are removed. This predictive capability enhances computational efficiency while maintaining accuracy in feature selection [13].

### 3.2.5 Focus Algorithm

The algorithm of classic FOCUS evaluates each singleton's feature sets first, then each set of two features, and so on. the evaluation stops after a consistent solution has been found [14].

## 3.3 Data Classification

The following are the classifiers utilized in this research:

### 3.3.1 Naive Bayes (NB)

Naive Bayes is invented in 1960 for the first time and belongs to the field of text retrieval [15]. In the subject of text classification, it is regarded as a fundamental technique. "The NB classifier is a group or family of probabilistic classifiers in machine learning that are built on using simple Bayesian theory with assumptions about strong and independent features" [16]. It is typical to analyze frequency of the words in the text to determine the category of the document as a solution for document issues. When learning the problem, The NB classifier is very scalable and requires a large number of linear parameters as features (variables). Automated medical diagnostic applications are one of the most prevalent applications.

---

### 3.3.2 Sequential Minimum Optimization (SMO)

In 1998, this method was introduced as a solution to the general SVM problem of Quadratic Programming (QP) problems that arise throughout the process of training of SVM. It splits a large QP problem into smaller QP problems that are then analytically solved in order to save time. Because SMO uses linear memory, it can handle very large training sets [17].

### 3.3.3 Decision Table (DT)

A Decision Table serves as a quick and efficient visual tool for determining the appropriate actions to take in response to specific situations. The output of a Decision Table is a set of actions, representing the processes to be executed. Each decision corresponds to or validates a variable, where the appropriate values are determined based on predefined conditions. Unlike the rigid four-quadrant structure, Decision Tables can vary significantly depending on the conditions and actions they represent. They utilize diverse representations to describe situations, ranging from simple true/false values to numerical data and, in some cases, A Decision Table serves as a quick and efficient visual tool for determining the appropriate actions to take in response to specific situations. The output of a Decision Table is a set of actions, representing the processes to be executed. Each decision corresponds to or validates a variable, where the appropriate values are determined based on predefined conditions. Unlike the rigid four-quadrant structure, Decision Tables can vary significantly depending on the conditions and actions they represent. They utilize diverse representations to describe situations, ranging from simple true/false values to numerical data and, in some cases, probabilistic models [18].

### 3.3.4 Random Tree (RT)

Similar to the Decision Tree method, the Random Tree algorithm only allows for a single random subset of attributes to be used at each fork in the tree. Data represented as a Tree has the benefit of being easy to understand and interpret. The algorithm's primary goal is to build a classification model for label value prediction using a large number of characteristics as inputs. As you navigate the tree's inner nodes, you'll see representations of each characteristic you supplied. Exactly as many edges connect each inner node as there are possible values for the associated input property. The values of the input attributes are shown as a route from the root to each leaf node, and this path represents the label value [19].

### 3.3.5 Stacking Algorithm

Stacking is a machine learning meta-algorithm designed to train a new model by leveraging the predictions of multiple weak learners. The outputs generated by these weak learners are used as inputs for subsequent layers, where they are combined to produce a new set of predictions. These aggregated predictions can either be further processed in additional layers or used directly to generate a final outcome, effectively enhancing the model's predictive accuracy by capitalizing on the strengths of individual weak learners [20].

## 4. Results And Discussion

The proposed method was trained and evaluated using a ten-fold cross-validation procedure on seven distinct datasets, utilizing five machine learning classifiers—Naïve Bayes (NB), Sequential Minimal Optimization (SMO), Decision Table (DT), Random Tree (RT), and Stacking. As a result, enhanced classification performance was achieved by integrating feature selection algorithms with machine learning classifiers. The primary objective of this study was to develop an optimized classification method by combining the most effective feature selection technique with the most accurate classifier. Initially, the classification algorithms were applied to the datasets without feature selection methods. The success rates of the various classifiers in this baseline evaluation are shown in Table 1. Following

this, the datasets were reclassified using the same classifiers, but this time incorporating feature selection methods. Table 2 presents the accuracy rates of the five classifiers after applying feature selection techniques. A direct comparison between Tables 1 and 2 reveals a substantial improvement in classification accuracy across all classifiers when feature selection techniques are utilized. Specifically, the accuracy rates improved as follows: the accuracy rate for the Naïve Bayes (NB) algorithm improved from 0.843 to 0.966 (+12.3%), for the Sequential Minimal Optimization (SMO) algorithm from 0.837 to 0.964 (+12.7%), for the Decision Table (DT) algorithm from 0.806 to 0.962 (+15.6%), for the Random Tree (RT) algorithm from 0.783 to 0.941 (+15.8%), and for the Stacking algorithm from 0.877 to 0.986 (+11.5%), demonstrating the effectiveness of these feature selection techniques in enhancing classification performance. These results clearly demonstrate the effectiveness of feature selection techniques in improving classification performance.

The findings of this study align with and surpass previous research on classification accuracy enhancement using feature selection. For instance, Tu et al. [21] compared three feature selection techniques and concluded that Bagging with Naïve Bayes achieved 94.0% accuracy. Similarly, Kim et al. [22] introduced an enhanced version of Naïve Bayes with feature selection, which resulted in an accuracy of 96.3%. In comparison, the proposed method outperformed both approaches, achieving an accuracy of 99.2% with the combination of Quick Branch and Bound (QBB) and Stacking. This demonstrates a 3.0% to 5.2% increase in classification performance over prior studies, highlighting the effectiveness of integrating advanced feature selection techniques with ensemble-based classification methods.

Table 1: The results of accuracy from classification algorithms before employing any feature selection algorithms.

Algorithm	Avg. Accuracy
NB algorithm	0.843
SMO algorithm	0.837
DT algorithm	0.806
RT algorithm	0.783
Stacking algorithm	0.877

Table 2: The results of accuracy from classification algorithms after employing feature selection algorithms.

Algorithm	Feature Selection					Avg. Accuracy
	LVF	SFFS	B&B	QBB	Focus	
NB algorithm	0.964	0.953	0.978	0.962	0.974	<b>0.966</b>
SMO algorithm	0.972	0.956	0.968	0.964	0.962	<b>0.964</b>
DT algorithm	0.969	0.953	0.967	0.950	0.972	<b>0.962</b>
RT algorithm	0.939	0.953	0.951	0.930	0.932	<b>0.941</b>
Stacking algorithm	0.985	0.988	0.980	0.992	0.989	<b>0.986</b>

The performance improvements observed, particularly in the combination of QBB and Stacking, can be attributed to the complementary strengths of these algorithms. The QBB algorithm's ability to efficiently reduce irrelevant features reduces noise in the data, which enhances the accuracy of the Stacking classifier. We also observed that in datasets with high-dimensional features, algorithms like Decision Tree struggled due to overfitting, whereas SMO performed consistently better. The variation in performance across different datasets highlights the importance of choosing the right classifier and feature selection method based on data characteristics.

## 5. Conclusion and Future Works

Data classification represents a critical and nuanced area of research, reflecting its significance across various applications. While considerable progress has been made in this domain, further research and development remain imperative to address ongoing challenges and refine existing methodologies. This study introduces a novel approach that combines feature selection techniques with classification algorithms to enhance performance. The proposed method is rigorously evaluated using seven diverse datasets, demonstrating its effectiveness and versatility. The datasets utilized in this study encompassed a diverse range of disease-related data, including conditions affecting the lungs, heart, and kidneys. As detailed in Section 4, the integration of the Quick Branch and Bound (QBB) algorithm with the Stacking algorithm significantly enhanced outcomes and improved accuracy rates. In conclusion, this research demonstrates that employing QBB as a feature selection algorithm and Stacking as a classification algorithm effectively produces accurate results across various domains, highlighting the robustness and applicability of the proposed approach. During the development of the proposed approach, an additional perspective for future work emerged. It was observed that the proposed algorithm achieved an accuracy rate exceeding 99.2%, demonstrating its exceptional performance. Furthermore, this finding highlighted the absence of inherent limitations in designing algorithms capable of achieving even higher accuracy rates. Consequently, future research could explore the development and application of alternative algorithms with similarly high or improved accuracy rates to address diverse challenges in this domain. The 99.2% accuracy rate was achieved in medical datasets with fewer outliers and a well-balanced feature set, which may not be the case in all datasets. This result was most pronounced in datasets where the QBB algorithm successfully eliminated irrelevant features, thereby reducing noise and improving classification accuracy. Future work will explore whether this method maintains similar accuracy levels in more complex datasets with imbalanced data or noisy features.

## 6. Author's Contribution:

We confirm that all named authors have read and approved the manuscript. We also confirm that each author has the same contribution to the paper. We further confirm that all authors have approved the order of authors listed in the manuscript.

## 7. Conflict of Interest:

There is no conflict of interest for this paper.

## References

- [1] Khachidze, M. & Tsintsadze, M. & Archuadze, M. Natural language processing based instrument for classification of free text medical records. *BioMed research international*. 2016. <https://doi.org/10.1155/2016/8313454>
- [2] Bottou, L. & Curtis F.E. & Nocedal, J. Optimization methods for large-scale machine learning. *SIAM review*. 2018, 60(2), 223-311. <https://doi.org/10.1137/16M1080173>

- 
- [3] Bell, D.A. & Wang, H. A. formalism for relevance and its application in feature subset selection. *Machine learning*. 2000, 41, 175-95. <https://doi.org/10.1023/A:1007612503587>
- [4] Chmiela, S & Tkatchenko, A. & Sauceda, H.E. & Poltavsky, I. Schütt, K.T. & Müller, K.R. Machine learning of accurate energy-conserving molecular force fields. *Science advances*. 2017, 3(5). <https://doi.org/10.1126/sciadv.1603015>
- [5] Thaseen, I.S. & Kumar, C.A. Intrusion detection model using fusion of chi-square feature selection and multi class SVM. *Journal of King Saud University-Computer and Information Sciences*. 2017, 29(4), 462-72. <https://doi.org/10.1016/j.jksuci.2015.12.004>
- [6] Allahyari, M. & Pouriyeh, S. & Assefi, M. & Safaei, S. & Trippe, E.D. & Gutierrez, J.B. & Kochut, K. A brief survey of text mining: Classification, clustering and extraction techniques. 2017. <https://doi.org/10.48550/arXiv.1707.02919>
- [7] Revathy, N. & Amalraj, D. Accurate cancer classification using expressions of very few genes. *International Journal of Computer Applications*. 2011, 14(4), 19-22. <https://doi.org/10.5120/1832-2452>
- [8] Nadkarni, P.M. & Ohno-Machado, L. & Chapman, W.W. Natural language processing: an introduction. *Journal of the American Medical Informatics Association*. 2011, 18(5), 544-51. <https://doi.org/10.1136/amiajnl-2011-000464>
- [9] Korde, V. & Mahender, C.N. Text classification and classifiers: A survey. *International Journal of Artificial Intelligence & Applications*. 2012, 3(2), 85. <https://doi.org/10.1136/amiajnl-2011-000464>
- [10] Wei, A. Optimal las vegas approximate near neighbors in  $\ell_p$ . In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms 2019*, 1794-1813. <https://doi.org/10.1137/1.9781611975482.108>
- [11] Lee, J. & Park, D. & Lee, C. Feature selection algorithm for intrusions detection system using sequential forward search and random forest classifier. *KSII Transactions on Internet and Information Systems (TIIS)*. 2017, 11(10), 5132-48. <http://dx.doi.org/10.3837/tiis.2017.10.024>
- [12] Morrison DR, Jacobson SH, Sauppe JJ, Sewell EC. Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning. *Discrete Optimization*. 2016 Feb 1;19:79-102. <https://doi.org/10.1016/j.disopt.2016.01.005>
- [13] Kabbaj, M.M. & El Afia, A. Towards learning integral strategy of branch and bound. In *2016 5th International Conference on Multimedia Computing and Systems (ICMCS) 2016*, 621-626. <https://doi.org/10.4018/979-8-3693-7117-6.ch006>
- [14] Elankavi, R. & Kalaiprasath, R. & Udayakumar, D.R. A fast clustering algorithm for high-dimensional data. *International Journal Of Civil Engineering And Technology (Ijci et)*. 2017, 8(5), 1220-7. [https://doi.org/10.34218/IJCIET\\_16\\_01\\_004](https://doi.org/10.34218/IJCIET_16_01_004)
- [15] Russell, S.J. & Norvig, P. *Artificial intelligence: a modern approach*. Pearson, 2016. <http://dx.doi.org/10.1016/j.artint.2011.01.005>
- [16] Vijayarani, S. & Dhayanand, S. Liver disease prediction using SVM and Naive Bayes algorithms. *International Journal of Science, Engineering and Technology Research (IJSETR)*. 2015, 4(4), 816-20. <https://doi.org/10.1016/j.procs.2020.03.226>
- [17] Reddy, K. & Shiva, K.L. & Abhilash, K. & Yoganandam, Y. Database Assisted Automatic Modulation Classification Using Sequential Minimal Optimization. 2018. <https://doi.org/10.48550/arXiv.1806.07566>
- [18] Zhang, X. & Miao, D. Three-layer granular structures and three-way informational measures of a decision table. *Information sciences*. 2017, 412, 67-86. <https://doi.org/10.1016/j.ins.2017.05.032>
- [19] Kevric, J. & Jukic, S. & Subasi, A. An effective combining classifier approach using tree algorithms for network intrusion detection. *Neural Computing and Applications*. 2017, 28(1), 1051-8. <https://doi.org/10.1007/s00521-016-2418-1>
- [20] Sikora, R. A modified stacking ensemble machine learning algorithm using genetic algorithms. In *Handbook of research on organizational transformations through big data analytics*. 2015, 43-53. <https://doi.org/10.58729/1941-6679.1061>
-

- 
- [21] Tu, M.C. & Shin, D. & Shin, D. A comparative study of medical data classification methods based on decision tree and bagging algorithms. In 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing 2009, 183-187. <https://doi.org/10.1109/DASC.2009.40>
- [22] Kim, S.B. & Han, K.S. & Rim, H.C. & Myaeng, S.H. Some effective techniques for naive bayes text classification. IEEE transactions on knowledge and data engineering. 2006, 18(11), 1457-66. <http://dx.doi.org/10.1109/TKDE.2006.180>
-