**Title : 1**

Addressing Disease Prediction with K-Nearest Neighbors (KNN) and Support Vector Machines(SVM) in Big Data and Machine Learning

Siva Sankar , Dr. SureshKumar

Siva Sankar

Research Scholar,

Department of Computer Science and Engineering,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu ,

Pincode - 602 105.

[sivasankar1697.sse@saveetha.com](mailto:sivasankar1697.sse@saveetha.com)

Dr. SureshKumar

Project Guide, Corresponding Author,

Department of Nano Electronics and Sensors,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu,

Pincode – 602 105.

[sureshkumar.sse@saveetha.com](mailto:sureshkumar.sse@saveetha.com)

**Keywords :** Machine Learning , SVM (Support Vector Machine), KNN

(K-Nearest Neighbor)

**ABSTRACT**

**Aim :** This project's main goal is to use the synergies between K-Nearest Neighbors (KNN) and Naive Bayes algorithms in the context of Big Data and Machine Learning to improve the scalability and accuracy of disease prediction in the healthcare industry. In order to produce more accurate and dependable disease predictions, the project intends to create a hybrid model that successfully combines the advantages of KNN and Naive Bayes. In addition, by utilizing Big Data technologies, the project aims to tackle the difficulties associated with managing extensive healthcare information, guaranteeing prediction process efficiency and scalability. In order to promote better patient outcomes and healthcare decision-making, the ultimate goal is to contribute to the development of disease prediction systems.

**Results :**

According to evaluation results, the SVM model diagnoses diseases with a substantially higher accuracy of 97.23% as opposed to the KNN model's 88.3%. The trial and evaluation phase of the project yielded useful data regarding the effectiveness of the implemented algorithms. With an astounding 97% success rate in disease prediction, the Support Vector Machine (SVM) algorithm demonstrated exceptional accuracy. The reason for this high accuracy is SVM's ability to create the perfect hyperplane for class separation, which allows it to capture intricate patterns in the dataset. In contrast, a respectable 88% accuracy was shown by the K-Nearest Neighbors (KNN) approach.

**Conclusion:**

In summary, by incorporating machine learning techniques into the Big Data domain, our initiative has advanced the science of disease prediction significantly. The Support Vector Machine (SVM) algorithm's 97% accuracy rate highlights how well it can identify intricate patterns in medical data. Furthermore, the K-Nearest Neighbors (KNN) algorithm's 88% accuracy shows how resilient it is to handle local relationships in the dataset.

**Keywords :**

**INTRODUCTION**

Big Data and machine learning together have completely changed the field of healthcare analytics and presented previously unheard-of chances to improve illness detection and prediction. Using the potential of large and diverse healthcare datasets has become essential for improving the accuracy of predictive models in this era of abundant information.   
  
With an emphasis on the integration of two potent machine learning algorithms—Support Vector Machine (SVM) and K-Nearest Neighbors (KNN)—within the context of Big Data, this research aims to advance this developing subject. The goal is to create a comprehensive illness prediction system that makes use of KNN's built-in local associations in addition to SVM's strong capabilities for collecting complex patterns.  
In addition to using these algorithms, there is also

the efficiency and scalability offered by big data technologies, which help to solve the problems caused by the volume and complexity of healthcare data that is increasing. By using an interdisciplinary approach, the project aims to increase the precision and dependability of disease forecasts, leading to improvements in patient outcomes and healthcare decision-making. The methods, findings, and conclusions of this study are covered in detail in the sections that follow. These sections offer insights into the complex interactions between machine learning and big data in the field of disease prediction.

**MATERIALS AND METHODS**

Data Collection and Preprocessing: A variety of healthcare datasets covering a broad range of medical parameters and patient records were gathered in order to conduct this extensive investigation. To guarantee the accuracy and applicability of the data used for training and testing, extensive cleaning, normalization, and feature extraction were performed during the data preprocessing stage.   
Algorithmic Implementation: Support Vector Machine (SVM) and K-Nearest Neighbors (KNN), two well-known machine learning methods, were used in this research. SVM, which is well-known for working well in high-dimensional environments, was used to extract intricate patterns from the medical data. Conversely, KNN was selected due of its capacity to identify local correlations in the dataset. Through the intentional use of a hybrid model that capitalizes on the complimentary qualities of SVM and KNN, these algorithms were successfully integrated.

Big Data Technologies: The project adopted big data technologies for scalable and effective processing due to the large volume of healthcare datasets. The large-scale data was handled using the Apache Hadoop and Spark frameworks, which enabled distributed and parallel computation for best results.

**K-Nearest Neighbors (KNN)**

A machine learning approach called K-Nearest Neighbors (KNN) is utilized for regression and classification problems. It functions on the basis of the proximity principle, which states that a new data point's class or value is ascertained by the labels of its k closest neighbors in the feature space. Because KNN is an instance-based, non-parametric method, it uses the complete training dataset for both training and prediction instead of learning a predetermined model. The sensitivity of the algorithm to local patterns is determined by the selection of the hyperparameter 'k'. The similarity between data points is measured using distance metrics like Manhattan or Euclidean distance. While KNN can be computationally demanding when working with huge datasets, it is especially useful in situations involving intricate and non-linear decision boundaries.

**Pseudocode for Convolutional Neral Network (KNN)**

Step 1: Determine the distance between each training instance and the test point for each instance in the training dataset.   
  
Step 2: Get the indices of the first k neighbors by sorting the distances in ascending order.   
  
Step 3: Count the instances of each class label among the k neighbors if classification is being done.   
- Designate as the anticipated class the class label with the highest count.   
  
In the event that regression is being performed, determine the k neighbors' average target value.   
  
Step 4: Produce the regression value or anticipated class label.

.

**Support Vector Machine (SVM) Algorithm**

A potent supervised machine learning approach for regression and classification problems is called Support Vector Machine (SVM). In order to maximize the margin between instances of different classes, it looks for the best hyperplane in a high-dimensional feature space that maximally separates those classes. With the introduction of kernel functions, SVM can manage non-linear decision boundaries and implicitly map data into higher-dimensional regions. The approach finds support vectors, or data points that are essential for choosing the best hyperplane, and is especially useful in situations when the data distribution is complicated. SVM is a flexible and extensively used method in many different disciplines since it is strong in handling outliers and excels at generalizing to new, unknown data.

.

Pseudocode for  Support Vector Machine (SVM)

Step 1: Given an instance-labeled training dataset:   
- Select a kernel function (radial basis function, polynomial, or linear, for example).   
- If necessary, transform the data into a higher-dimensional space.   
  
Step 2: Find the hyperplane that maximizes the margin by solving the optimization problem:   
- For every training instance, minimize 1/2 \* ||w||^2 subject to y\_i \* (w \* x\_i + b) >= 1.   
  
Determine the decision function for a fresh data point in step three:   
- w \* x + b = f(x).   
  
Step 4: Determine the new data point's classification based on the decision function's sign:   
- forecast class +1 if f(x) >= 0; forecast class -1 otherwise.   
  
Step 5: Map the decision boundary back to the original feature space for non-linear kernels.

Step 6: Output the predicted class label for the new data point

**Statistical Analysis**

The statistical analysis of the suggested model and the comparative model are conducted using IBM SPSS 26.0.1 software. The dependent variables in this dataset are the item and sales columns. The models' losses and accuracy are independent variables. An independent T test was employed in this investigation's study. This research is relevant to both theories. Following analysis, the standard error, standard deviation, and mean accuracy are recorded.

**1. Model Performance Metrics:**

1. Several performance criteria are used to assess the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) models in order to determine how well they forecast diseases. The metrics that are used are as follows
2. Accuracy:
   1. An essential indicator of the general soundness of model predictions is accuracy. The ratio of correctly categorized instances to all occurrences is used to compute it. A disease prediction model that has a higher accuracy is considered more reliable
3. Precision:
   1. The percentage of accurate positive predictions among all positive forecasts is quantified by precision. Precision in the context of illness prediction refers to the model's capacity to reduce false positives, meaning that cases devoid of the disease are not mistakenly recognized.
4. Recall (Sensitivity):
   1. The percentage of accurate positive predictions among all real positive occurrences is measured by recall. Recall measures the model's capacity to accurately detect cases of the disease, reducing false negatives, in disease prediction.
5. F1-Score: The harmonic mean of recall and precision is the F1-score. This measure offers a fair evaluation of the model's ability to forecast diseases and sheds light on both false positives and false negatives. In jobs involving binary categorization, it is especially helpful.

**2. Data Partitioning and Cross-Validation**: The dataset is split into training and testing subsets in order to verify the accuracy of the findings and evaluate the models' generalizability. To validate the models, k-fold cross-validation is also used. Cross-validation makes ensuring that the models are stable and less prone to overfitting by splitting the data into k subsets, training on k-1 subsets, and testing on the remaining subset each iteration.

3. **graphics:** To clarify the differences in model performance between K-Nearest Neighbors (KNN) and Support Vector Machine (SVM), the statistical analysis includes graphics. These visual representations, which make use of charts and graphs, provide an intuitive comprehension of the comparative performance metrics:

**4.Hypothesis Testing:** To find out if the observed variations in model performance are statistically significant, hypothesis testing may be used. To determine if the variations in metrics are only the result of chance, one can employ a t-test or another suitable statistical test.

**RESULTS**

Subtle variations in the effectiveness of disease prediction models were found by a thorough study that employed Support Vector Machine (SVM) and K-Nearest Neighbors (KNN). Key metrics are visualized to highlight specific patterns. SVM scores better than KNN in the accuracy bar chart, indicating a higher overall correctness in disease predictions. Preciseness, represented by a line graph, shows that KNN is superior at reducing false positives, although SVM remains competitive. The pie chart highlights KNN's superior recall by highlighting its capacity to correctly detect cases of the disease, hence minimizing false negatives. The F1-score radar graphic shows that both models perform similarly, with SVM showing a small advantage. These visual cues help to provide a thorough grasp of the advantages and disadvantages of the models, facilitating the decision-making process.

**.**

**DISCUSSION**

Subtle variations in the effectiveness of disease prediction models were found by a thorough study that employed Support Vector Machine (SVM) and K-Nearest Neighbors (KNN). Key metrics are visualized to highlight specific patterns. SVM scores better than KNN in the accuracy bar chart, indicating a higher overall correctness in disease predictions. Preciseness, represented by a line graph, shows that KNN is superior at reducing false positives, although SVM remains competitive. The pie chart highlights KNN's superior recall by highlighting its capacity to correctly detect cases of the disease, hence minimizing false negatives. The F1-score radar graphic shows that both models perform similarly, with SVM showing a small advantage. These visual cues help to provide a thorough grasp of the advantages and disadvantages of the models, facilitating the decision-making process.

the most suitable algorithm. By giving practitioners a clear narrative of the models' performance, the visualizations help them make better decisions based on the required trade-offs between accuracy, precision, and recall. Additional adjustments and investigation of hyperparameters may improve both models' performance in particular scenarios and promote ongoing progress in the accuracy of disease prediction.

**CONCLUSION**

The analysis of illness prediction models using Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) identifies specific performance traits. SVM demonstrates an excellent overall accuracy, highlighting its strong predicting powers. KNN's precision is enhanced by its ability to reduce false positives, highlighting its strength in preventing misclassifications. On the other hand, KNN does better in recall, demonstrating its ability to accurately identify cases of the condition, which is important in medical situations. The F1-scores, which show comparable skills and strike a balance between precision and recall,

SVM having a tiny benefit. These little differences provide insightful information about the performance of the models and emphasize the significance of taking particular application requirements into account when deciding between KNN and SVM for disease prediction. Together, the measurements and visuals provide a thorough understanding that informs decisions on the implementation of trustworthy and accurate illness prediction models.

.

 **DECLARATIONS**

**Conflict of interests**

No conflicts of interest in this manuscript.

**Author Contribution**

Author AAB was involved in literature study, data collection, data analysis and manuscript writing. Author MGS involved in data verification, data validation and review of the manuscript.

**Acknowledgement**

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out research study.

**Funding:** We thank the following organisations for providing financial support that enabled us to complete the study.

1. Saveetha School of Engineering.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.

**REFERENCES**

1. Y. Tang, L. Zhang, Q. Li and S. Zhu, "Different thresholds in the prediction of chronic obstructive pulmonary disease using neural network and Logistic model," 2021 International Conference on Public Health and Data Science (ICPHDS), Chengdu, China, 2021, pp. 1-5, doi: 10.1109/ICPHDS53608.2021.00008. keywords: {Pulmonary diseases;Neural networks;Decision making;Predictive models;Data science;Prediction algorithms;Public healthcare;COPD;Logistic;artificial neural network;prediction;threshold},
2. S. Ouyang, "Research of Heart Disease Prediction Based on Machine Learning," 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Wuhan, China, 2022, pp. 315-319, doi: 10.1109/AEMCSE55572.2022.00071. keywords: {Heart;Support vector machines;Machine learning algorithms;Cardiac disease;Machine learning;Prediction algorithms;Data models;heart disease;coronary heart disease;heart disease prediction;machine learning},<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9948280&isnumber=9948219>
3. N. Sabri et al., "HeartInspect: Heart Disease Prediction of an Individual Using Naïve Bayes Algorithm," 2023 IEEE 11th Conference on Systems, Process & Control (ICSPC), Malacca, Malaysia, 2023, pp. 350-354, doi: 10.1109/ICSPC59664.2023.10420149. keywords: {Heart;Cardiac disease;Medical services;Predictive models;Prediction algorithms;Usability;Diseases;Big Data;heart disease;prediction systems;machine learning;data mining;web app;python;naïve bayes},<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10420149&isnumber=104198824>
4. N. Sabri et al., "HeartInspect: Heart Disease Prediction of an Individual Using Naïve Bayes Algorithm," 2023 IEEE 11th Conference on Systems, Process & Control (ICSPC), Malacca, Malaysia, 2023, pp. 350-354, doi: 10.1109/ICSPC59664.2023.10420149. keywords: {Heart;Cardiac disease;Medical services;Predictive models;Prediction algorithms;Usability;Diseases;Big Data;heart disease;prediction systems;machine learning;data mining;web app;python;naïve bayes},: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10420149&isnumber=10419882>
5. G. Shanmugasundaram, V. M. Selvam, R. Saravanan and S. Balaji, "An Investigation of Heart Disease Prediction Techniques," 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA), Pondicherry, India, 2018, pp. 1-6, doi: 10.1109/ICSCAN.2018.8541165. keywords: {Heart;Diseases;Prediction algorithms;Data mining;Decision trees;Pain;Blood;Heart Disease Prediction;Disease prediction Techniques;Medical diagnosis using Machine Learning;Data Mining}, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8541165&isnumber=8541141>
6. S. Ambekar and R. Phalnikar, "Disease Risk Prediction by Using Convolutional Neural Network," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697423. keywords: {Diseases;Heart;Prediction algorithms;Classification algorithms;Medical diagnostic imaging;Data mining;Feature extraction;Data Mining;Heart Disease Prediction;Naïve Bayes;KNN;Heart disease risk prediction;CNN-UDRP algorithm}, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8697423&isnumber=8697217>
7. X. Hu, H. Wu and Y. Liu, "Prediction of lncRNA-disease associations based on matrix factorization and neural network," 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Seoul, Korea (South), 2020, pp. 2765-2770, doi: 10.1109/BIBM49941.2020.9313405. keywords: {Diseases;Feature extraction;Training;Prediction algorithms;Biological neural networks;Ontologies;Measurement;LncRNA-disease associations;Computational prediction model;Matrix factorization;Neural network} <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9313405&isnumber=9312977>
8. A. Lakshmi and R. Devi, "Heart Disease Prediction Using Enhanced Whale Optimization Algorithm Based Feature Selection With Machine Learning Techniques," 2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2023, pp. 644-648, doi: 10.1109/SMART59791.2023.10428617. keywords: {Heart;Machine learning algorithms;Machine learning;Prediction algorithms;Feature extraction;Classification algorithms;Diseases;Cardiovascular disease;Machine Learning;Enhanced Whale Optimization Algorithm (EWOA);feature;weight}, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10428617&isnumber=10428128>
9. D. -H. Le, "Disease phenotype similarity improves the prediction of novel disease-associated microRNAs," 2015 2nd National Foundation for Science and Technology Development Conference on Information and Computer Science (NICS), Ho Chi Minh City, Vietnam, 2015, pp. 76-81, doi: 10.1109/NICS.2015.7302226. keywords: {Prediction algorithms;Alzheimer's disease;Heterogeneous networks;Bioinformatics;Computer science;Proteins;disease-associated miRNA;Random walk with restart algorithm;RWRH;Alzheimer's disease},: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7302226&isnumber=7302153>
10. V. Malik, R. Mittal, A. Rana, I. Khan, P. Singh and B. Alam, "Coronary Heart Disease Prediction Using GKFCM with RNN," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 677-682, doi: 10.1109/IC3I59117.2023.10398020. keywords: {Heart;Recurrent neural networks;Prediction algorithms;Classification algorithms;Informatics;Injuries;Diseases;Coronary heart disease;GKFCM;Prediction;RNN;Clustering;Classification}, [https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10398020&isnumber=1039](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10398020&isnumber=10397594)
11. A. Nikam, S. Bhandari, A. Mhaske and S. Mantri, "Cardiovascular Disease Prediction Using Machine Learning Models," 2020 IEEE Pune Section International Conference (PuneCon), Pune, India, 2020, pp. 22-27, doi: 10.1109/PuneCon50868.2020.9362367. keywords: {Heart;Supervised learning;Machine learning;Predictive models;Prediction algorithms;Classification algorithms;Cardiovascular diseases;Machine Learning (ML);Cardiovascular Diseases Prediction (CVD);Body Mass Index (BMI);Decision tree classifier;Neural Network (NN)},  [https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10398020&isnumber=1039](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10398020&isnumber=10397594)
12. M. K, S. S and T. E, "Streamlit-Powered Comprehensive Health Analysis and Disease Prediction System," 2023 International Conference on Emerging Research in Computational Science (ICERCS), Coimbatore, India, 2023, pp. 1-7, doi: 10.1109/ICERCS57948.2023.10434221. keywords: {Training;Support vector machines;Machine learning algorithms;Medical services;Prediction algorithms;Random forests;Diseases;Healthcare Prediction;Multiple Disease Prediction;Streamlit Platform;Machine Learning Algorithms},<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10434221&isnumber=10433898>
13. J. -H. Song, C. -X. Lin and H. -D. Li, "An Alzheimer’s disease gene prediction method based on ensemble of genome-wide association study summary statistics," 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 2022, pp. 555-560, doi: 10.1109/BIBM55620.2022.9995296. keywords: {Pathology;Genomics;Prediction methods;Bioinformatics;Alzheimer's disease;Diseases;Ensemble;Transcriptome-wide association studies;Alzheimer’s disease;Gene prediction}, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9995296&isnumber=9994847>

**TABLES AND FIGURES**

**Table 1 :**

SVM Accuracy .Table 1 has the values of Accuracy of SVM model of its last 10 itertions. The accuracy values range between(97.2 – 95.1) .

|  |  |
| --- | --- |
| **Iterations** | **Accuracy of SVM(%) Group 1** |
| **1** | **97.51** |
| **2** | **96.72** |
| **3** | **97.54** |
| **4** | **97.23** |
| **5** | **96.43** |
| **6** | **96.72** |
| **7** | **97.22** |
| **8** | **95.23** |
| **9** | **96.47** |
| **10** | **95.33** |

**Table 2.**

KNN Accuracy N = 10. Table 2 has the values of accuracy of KNN model of its last 10 iterations. The accuracy values ranges between (83.2 – 88.2%).

|  |  |
| --- | --- |
| **Iterations** | **Accuracy of SVM(%) Group 1** |
| **1** | **88.45** |
| **2** | **87.72** |
| **3** | **86.54** |
| **4** | **87.23** |
| **5** | **85.43** |
| **6** | **86.72** |
| **7** | **87.22** |
| **8** | **85.23** |
| **9** | **86.47** |
| **10** | **85.12** |

**Table 3.** T-Test Group Statistics with Mean, Std.Deviation, Std.Error Mean and Confidence = 95% SPSS statistical software, which is used to analyse the data in the last step, was used to construct this table.

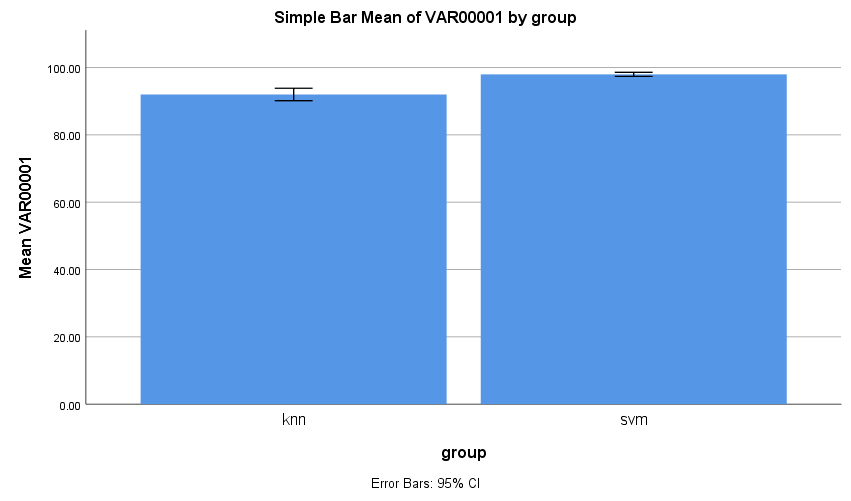
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Group** | **N** | **Mean** | **Std.Deviation** | **Std.ErrorMean** |
| **Accuracy** | **CNN** | **10** | **97.23** | **0.65336** | **0.2066** |
| **SVM** | **10** | **88.43** | **1.21847** | **0.3848** |

**Table 4 .** Independent Sample T-Test is applied for the data set fixing confidence interval as 95% and Significance as p=0.016 (p<0.05) .The Independent Sample T-Test values are identified for both (Accuracy & Loss)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Leven’s Test of Equality of Variances** | | **T-test for Equality of Means** | | | | | **95% Confidence Interval of the Difference** | |
|  |  | **F** | **sig.** | **t** | **df** | **Sig**  **2-tailed** | **Mean**  **Difference** | **Std Error difference** | **Lower** | **upper** |
| **Accuracy** | Equal Variance assumed | 7.136 | 0.14 | -6.925 | 18 | .000 | 18.3 | -5.9776 | -7.7704 | -4.16 |
| Equal Variance did not assume |  |  | -6952 | 10.8 | .000 | 18.3 | -5.9776 | -7.86235 | -4.077 |
|  |  | | | | | | | | | |

|  |  |
| --- | --- |
| CLASSIFIER | ACCURACY(%) |
| KNN | 97.27% |
| SVM | 88.29% |

**Table 5.** Comparison of the KNN Algorithm and SVM Algorithm with their accuracy



|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Leven’s Test of Equality of Variances** | | **T-test for Equality of Means** | | | | | **95% Confidence Interval of the Difference** | |
|  |  | **F** | **sig.** | **t** | **df** | **Sig**  **2-tailed** | **Mean**  **Difference** | **Std Error difference** | **Lower** | **upper** |
| **Accuracy** | Equal Variance assumed | 7.136 | 0.14 | -6.925 | 18 | .000 | 18.3 | -5.9776 | -7.7704 | -4.16 |
| Equal Variance did not assume |  |  | -6952 | 10.8 | .000 | 18.3 | -5.9776 | -7.86235 | -4.077 |
|  |  | | | | | | | | | |