PREDICTION OF QUALITY OF RICE IN RICE MILL USING DECISION TREE COMPARED WITH SVM WITH IMPROVED ACCURACY

Chinnam Koti Pullarao[1], V.Nagaraju[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The main objective of this research article is to improve the accuracy rate in the Novel prediction of quality of rice in rice mills by using Decision Tree (D-Tree) in comparison with Support Vector Machine (SVM) Classifier.
Materials & Methods: The data set in this paper utilizes the publicly available Kaggle data set for Novel prediction of the quality of rice in rice mills. The sample size of Novel prediction of quality of rice in rice mill with improved accuracy rate was sample 80 (Group 1=40 and Group 2 =40), and calculation is performed utilizing G-power 0.8 with alpha and beta qualities are 0.05, 0.2 with a confidence interval at 95%. Novel Prediction of quality of rice in rice mill with improved accuracy rate is performed by Decision Tree (D-Tree) whereas some samples (N=10) and Support Vector Machine (SVM) were the number of samples (N=10).
Results: The Decision Tree (D-Tree) classifier has 92.7 higher accuracy rates when compared to the accuracy rate of Support Vector Machine (SVM) is 91.0. The study has a significance value of $p<0.05$, i.e., $p=0.035$ which infers they are statistically significant.
Conclusion: Decision Tree (D-Tree) provides better outcomes in accuracy rate when compared to Support Vector Machine (SVM) for Noval prediction of quality of rice in rice mill.

Keywords: Quality of Rice, Decision Tree (D-Tree), Support Vector Machine (SVM), Accuracy rate, Rice Mill, Novel Prediction

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105
[2*]Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105

1. Introduction

For around 85% of the Asia people, rice is viewed as the staple nourishment food(Bandumula 2018). Many countries accomplish self-capability in rice creation, better quality rice is the most extreme need for customers. It is incredibly monotonous for shoppers to investigate the excellent quality and rice evaluating on the lookout(Abdullah, Aziz, and Dos Mohamed 2000). Grain quality concludes valuing, and quality records vary based on end-use prerequisites. The suggested calculation has numerous applications on order and relapse-based issues. Ingrain takes care of units; quality is communicated based on actual attributes, such as size, shape, part hardness, dampness content, and visual properties like harmed, pervaded stained pieces, and unfamiliar materials. So the requirement for a programmed rice quality evaluating framework emerges, which can dispense with the deficiencies of the manual quality reviewing process. This study proposes a decision tree (D-Tree) algorithm to predict the rice grains(Lian et al. 2021). Experimental results show that the D-Tree model has exceptional execution with high order exactness on a testing subset(Nithya and Sundaram 2011).

Further research was completed to make programmed rice quality evaluating frameworks for the beyond five years. IEEE Xplore published 77 research papers, and Google Scholar found 87 articles. Chen et al.(Chen et al. 2012) suggested the Least Squares Support Vector Machines for the catagorization of head rice and broken rice. Mousavi Rad, Fayyazi, Prajapati and Patel, Gujjar and Siddappa(Fayyazi et al. 2013; Gujjar and Siddappa 2013; Mousavirad, Tab, and Mollazade 2012; Prajapati and Patel 2013; Kaur and Singh 2013) suggested neural networks and multi-class SVM for ordering and quality reviewing of Iranian and Indian rice assortments..(Kong et al. 2019) has suggested a novel strategy to recognize the thickness of rice utilizing the photogrammetry idea. The rice edges are formed by considering the elements like length, width, and consistency..(Singh, Vidyarthi, and Tiwari 2020) fostered another procedure that consolidates image processing and machine learning (ML) ensembles to quantify the size and mass of many rice pieces simultaneously. Here, with the assistance of a recursive strategy, an image processing algorithm was created to distinguish every rice piece from a image and compute the size of the portions in light of the pixels involved.(Parveen, Alam, and Shakir 2017), develops an algorithm to mark the rice grains based on characteristics like length, thickness, shading, and pasty region.Our team has extensive knowledge and research experience that has translated into high quality publications(K. Mohan et al. 2022; Vivek et al. 2022; Sathish et al. 2022; Kotteeswaran et al. 2022; Yaashikaa, Keerthana Devi, and Senthil Kumar 2022; Yaashikaa, Senthil Kumar, and Karishma 2022; Saravanan et al. 2022; Jayabal et al. 2022; Krishnan et al. 2022; Jayakodi et al. 2022; H. Mohan et al. 2022)

The main disadvantage with the conventional technique for estimating the rice aspects can be incredibly dreary, mainly whenever done at large scales. To conquer this drawback, this paper proposes a novel and robust strategy to precisely measure the nature of rice utilizing a decision tree (D-Tree) algorithm in comparison with support vector machine (SVM) algorithm. The performance analysis of the suggested rice quality examination gives preferred outcomes than the existing SVM method.

2. Materials and Methods

This work was carried out at Machine Learning Laboratory in Saveetha School of Engineering, SIMATS, Tamil Nadu, and India. In this study, the rice data set was collected from various rice mills. 13 different types of rice samples are taken. The Sample size was analyzed utilizing earlier works(Zareiforoush et al. 2016). Group 1 was a Decision Tree (D-Tree) algorithm and Group 2 was a Random Forest (RF) algorithm. In this work two groups are taken and 10 samples for each group, total samples considered are 20. The calculation is performed utilizing G-power 0.8 with alpha and beta qualities 0.05, 0.2 with a confidence interval at 95%.

Support vector machines algorithm

The sample preparation group 1 is the Support Vector Machine (SVM) Classifier is a supervised machine learning strategy used for classification, regression and outlier's detection. The steps involved in the implementation of the SVM algorithm are described as follows.

In data mining methodology, Support Vector Machine (SVM) Classifiers are supervised machine learning frameworks given the measurable learning hypothesis that investigate information and perceive designs in grouping and regression examination issues. Support Vector Machine (SVM) Classifier is a representation of the samples as focuses in space, planned to ensure that the standards of the specific classes are isolated by a reasonable limit that is as wide as be expected. In this methodology, the ideal limit, known as hyper-plane, of two sets in a vector space is gotten autonomously on the probabilistic conveyance of preparing vectors in the set. The hyperplane finds the limit beyond what many would consider possible from the two groups' closest vectors to the limit. The vectors arranged relative to the hyperplane are called support vectors. There might be no isolating hyperplane to recognize if the space isn't directly detachable. In such cases, a part capacity might be utilized to address the issue. The bit work assesses the connections inside the information and makes complex divisions(Cortes and Vapnik 1995). A support vector machine model portrays the examples as focuses in space, isolated by a reasonable limit. The ideal limit is known as hyper-plane. The vectors arranged close to the hyperplane are called support vectors.

Step 1: Separate the images based on the labels.

Step 2: For training the data 75% of the data is used for building the model.

Step 3: Remaning 25% of the data is used for testing the model.

Step 4: Remove all the outliers from the dataset, So the model does not overfit.

Step 5: Deep Support Vector Machine (SVM) Classifier relates to the data given.

Step 6: Predicts the nutrition analysis and calorie count significantly.

Decision tree (D-Tree) algorithm

The sample preparation group 2 is the novel decision tree (D-Tree) algorithm, and is the most powerful and popular tool for classification and prediction of rice quality. The experimental results show that the suggested D-Tree method has achieved better accuracy results.

A decision tree classifier is a proactive model having applications in various sections. It comprises a chart comprising nodes that address ascribes and seek clarification on some pressing issues. Then again, edges in a graph address the responses to the inquiries posed before. Finally, the leaves signify the actual outcome acquired in the wake of following away down the tree. This demonstrating tool can be executed in non-direct constructions to get the ideal result. Models are arranged from the root node to the leaf node. A classification tree or a decision tree illustrates a multistage choice process. Rather than utilizing the total arrangement of elements together to settle on a choice, various subsets of features are being used at different levels of the tree. These algorithms have been utilized in decision trees-based classification techniques(L.Gupta et al. 2012; Soltani and Omid 2015).

Step 1: Separate the images based on the labels.

Step 2: For training the data 75% of the data is used for building the model.

Step 3: Remaning 25% of the data is used for testing the model.

Step 4: Remove all the outliers from the dataset, so the model does not over fit.

Step 5: Deep decision tree (D-Tree) relates to the data given.

Step 6: Predicts the nutrition analysis and calorie count significantly.

Statistical Analysis

The output is obtained using Python software(Downey, n.d.). To train these datasets, required a monitor with a resolution of 1024×768 pixels (10th gen, i7, 12GB RAM, 500 GB HDD) and Python software with essential library functions and tool functions. For statistical implementation, the software tool used here is IBM SPSS (Verma 2012). The independent sample t-test was performed to find the mean, standard deviation, and standard error mean statistical significance. Then a comparison of the two groups with the SPSS software will give the accurate values for the two different, which will be utilized with the graph to calculate the significant value with maximum accuracy value (92.7%), mean value (92%), and standard deviation value (0.22843). Dependent variables are accuracy, and independent variables are image size.

3. Results

Figure 1 shows the simple bar graph for SVM Classifier accuracy rate is compared with D-Tree Classifier. The D-Tree Classifier is higher in terms of accuracy rate 92.7 when compared with SVM Classifier 91.0.

Table 1 shows the Evaluation Metrics of Comparison of SVM and D-Tree Classifier. The accuracy rate of SVM is 91.0, and D-Tree has 92.7.

Table 2 shows the statistical calculation such as Mean, standard deviation, and standard error Mean for SVM and D-Tree. The accuracy rate parameter was used in the t-test. The mean accuracy rate of SVM is 91.0, and D-Tree is 92.7. The Standard Deviation of Support Vector Machine (SVM) is 1.92394, and D-Tree is 0.22843. The Standard Error Mean of SVM is 0.93494, and D-Tree is 0.12948.

Table 3 displays the statistical calculations for independent samples test between SVM and D-Tree. The significance level for the accuracy rate is 0.035.

4. Discussion

Based on the above results, Support Vector Machine (SVM) Classifier accuracy rate is compared with Decision Tree (D-Tree) Classifier. The Decision Tree (D-Tree) Classifier is higher in terms of accuracy rate 92.7 when compared with Support Vector Machine (SVM) Classifier 91.0**.**

Several researchers have published papers in this area.(Devi, Neelamegam, and Sudha 2017) suggested a machine learning strategy considering the physical features to evaluate and sort rice grains. In the work, the attributes isolated were considered for reviewing and quality investigation of rice grains carried out in MatLab with the precision of 95%. Khunkhett and Remsungnen(Khunkhett and Remsungnen 2014) applied digital image examination for non-damaging identification of unadulterated reproducing rice seed. The presence of rice, for example, its shape and shading, is expected to be a massive element in rural propagation and quality testing. The proper order rates for the two phases are Good rice seeds 97.34% and unadulterated reproducing rice seeds 84.23%.Silva and Sonnadara(Silva and Sonnadara 2013) fostered a framework comprising of machine vision and neural networks for nine different rice seeds categorization. Various algorithms obtained fifteen morphological, six tone, and twenty textural highlights from each rice seed shading image test. The general exactness of order 93.21% was achieved with a consolidated list of features. Kaur and Singh(Kaur and Singh 2013) suggested a multi-class SVM for arrangement and rice reviewing. Multi-class SVM inspected shape, whiteness, and broken parts' level for grouping, precisely above 87.34%. El-Telbany et al.(El-Telbany, Warda, and El-Borahy 2006) explored Egyptian rice sicknesses utilizing the C4.5 decision trees algorithm.

The limitations of the suggested framework are not equipped for getting grain organization, inside bug pervasion, and organoleptic properties, factors that are primary deciders of food grain quality. In the future, new frameworks with higher exactness can be created with the assistance of genuinely progressing innovations of machine learning, deep learning and image processing utilizing a blend of various consolidated feature sets with a more extensive dataset of particular rice assortments.

5. Conclusion

The suggested model exhibits the SVM and D-Tree, in which the D-Tree has the highest values. The accuracy Rate of the D-Tree is 92.7 is higher than the SVM, which has an accuracy rate of 91.0 in the analysis of Novel prediction of the quality of rice in rice mill with improved accuracy rate.

Declarations

Conflicts of Interest

No conflict of interest in this manuscript

Authors Contributions

Author KPR was involved in data collection, data analysis & manuscript writing. Author VN was involved in conceptualization, data validation, and critical review of manuscripts.

Acknowledgments

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 successfully carrying out this work.

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.

1. VEE EEE Technologies.
2. Saveetha University
3. Saveetha Institute of Medical And Technical Sciences
4. Saveetha School of Engineering

6. References

Abdullah, Mohd Zaid, Sabina Abdul Aziz, and Abdul Manan Dos Mohamed. 2000. “Quality Inspection of Bakery Products Using a Color-Based Machine Vision System.” *Journal of Food Quality* 23 (1): 39–50.

Bandumula, Nirmala. 2018. “Rice Production in Asia: Key to Global Food Security.” *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences* 88 (4): 1323–28.

Chen, Xiai, Shuang Ke, Ling Wang, Hong Xu, and Wenquan Chen. 2012. “Classification of Rice Appearance Quality Based on LS-SVM Using Machine Vision.” In *Information Computing and Applications*, 104–9. Communications in Computer and Information Science. Springer, Berlin, Heidelberg.

Cortes, C., and V. Vapnik. 1995. “Support Vector Machine.” *Machine Learning*. https://mlab.cb.k.u-tokyo.ac.jp/~moris/lecture/cb-mining/4-svm.pdf.

Devi, T. Gayathri, P. Neelamegam, and S. Sudha. 2017. “Machine Vision Based Quality Analysis of Rice Grains.” In *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, 1052–55.

Downey, Allen B. n.d. “Python for Software Design.” *Python for Software Design*. https://doi.org/10.1017/cbo9780511813832.002.

El-Telbany, Mohammed E., Mahmoud Warda, and Mahmoud El-Borahy. 2006. “Mining the Classification Rules for Egyptian Rice Diseases.” *Int. Arab J. Inf. Technol.* 3 (4): 303–7.

Fayyazi, S., M. H. Abbaspour-Fard, A. Rohani, H. Sadrnia, and S. A. Monadjemi. 2013. “Identification of Three Iranian Rice Seed Varieties in Mixed Bulks Using Textural Features and Learning Vector Quantization Neural Network.” In *1st International E-Conference on Novel Food Processing, Mashhad, Iran*, 26–27.

Gujjar, Harish S., and M. Siddappa. 2013. “A

Method for Identification of Basmati Rice Grain of India and Its Quality Using Pattern Classification." *International Journal of Engineering Research and Applications* 3 (1): 268–73.

Jayabal, Ravikumar, Sekar Subramani, Damodharan Dillikannan, Yuvarajan Devarajan, Lakshmanan Thangavelu, Mukilarasan Nedunchezhiyan, Gopal Kaliyaperumal, and Melvin Victor De Poures. 2022. "Multi-Objective Optimization of Performance and Emission Characteristics of a CRDI Diesel Engine Fueled with Sapota Methyl Ester/diesel Blends." *Energy*. https://doi.org/10.1016/j.energy.2022.123709.

Jayakodi, Santhoshkumar, Rajeshkumar Shanmugam, Bader O. Almutairi, Mikhlid H. Almutairi, Shahid Mahboob, M. R. Kavipriya, Ramesh Gandusekar, Marcello Nicoletti, and Marimuthu Govindarajan. 2022. "Azadirachta Indica-Wrapped Copper Oxide Nanoparticles as a Novel Functional Material in Cardiomyocyte Cells: An Ecotoxicity Assessment on the Embryonic Development of Danio Rerio." *Environmental Research* 212 (Pt A): 113153.

Kaur, Harpreet, and Baljit Singh. 2013. "Classification and Grading Rice Using Multi-Class SVM." *International Journal of Scientific and Research Publications* 3 (4): 1–5.

Khunkhett, S., and T. Remsungnen. 2014. "Non-Destructive Identification of Pure Breeding Rice Seed Using Digital Image Analysis." In *The 4th Joint International Conference on Information and Communication Technology, Electronic and Electrical Engineering (JICTEE)*, 1–4.

Kong, Yuchen, Shenghui Fang, Xianting Wu, Yan Gong, Renshan Zhu, Jian Liu, and Yi Peng. 2019. "Novel and Automatic Rice Thickness Extraction Based on Photogrammetry Using Rice Edge Features." *Sensors* 19 (24). https://doi.org/10.3390/s19245561.

Kotteeswaran, C., Indrajit Patra, Regonda Nagaraju, D. Sungeetha, Bapayya Naidu Kommula, Yousef Methkal Abd Algani, S. Murugavalli, and B. Kiran Bala. 2022. "Autonomous Detection of Malevolent Nodes Using Secure Heterogeneous Cluster Protocol." *Computers and Electrical Engineering*. https://doi.org/10.1016/j.compeleceng.2022.107902.

Krishnan, Anbarasu, Duraisami Dhamodharan, Thanigaivel Sundaram, Vickram Sundaram, and Hun-Soo Byun. 2022. "Computational Discovery of Novel Human LMTK3 Inhibitors by High Throughput Virtual Screening Using NCI Database." *Korean Journal of Chemical Engineering*. https://doi.org/10.1007/s11814-022-1120-5.

L.Gupta, D., D. L. Gupta, A. K. Malviya, and Satyendra Singh. 2012. "Performance Analysis of Classification Tree Learning Algorithms." *International Journal of Computer Applications*. https://doi.org/10.5120/8762-2680.

Lian, Yi, Jin Chen, Zhuohuai Guan, and Jie Song. 2021. "Development of a Monitoring System for Grain Loss of Paddy Rice Based on a Decision Tree Algorithm." *International Journal of Agricultural and Biological Engineering* 14 (1): 224–29.

Mohan, Harshavardhan, Sethumathavan Vadivel, Se-Won Lee, Jeong-Muk Lim, Nanh Lovanh, Yool-Jin Park, Taeho Shin, Kamala-Kannan Seralathan, and Byung-Taek Oh. 2022. "Improved Visible-Light-Driven Photocatalytic Removal of Bisphenol A Using V2O5/WO3 Decorated over Zeolite: Degradation Mechanism and Toxicity." *Environmental Research*. https://doi.org/10.1016/j.envres.2022.113136.

Mohan, Kannan, Abirami Ramu Ganesan, P. N. Ezhilarasi, Kiran Kumar Kondamareddy, Durairaj Karthick Rajan, Palanivel Sathishkumar, Jayakumar Rajarajeswaran, and Lorenza Conterno. 2022. "Green and Eco-Friendly Approaches for the Extraction of Chitin and Chitosan: A Review." *Carbohydrate Polymers* 287 (July): 119349.

Mousavirad, S. J., F. Akhlaghian Tab, and K. Mollazade. 2012. "Design of an Expert System for Rice Kernel Identification Using Optimal Morphological Features and Back Propagation Neural Network." *Int J Appl Inf Syst* 3 (2): 33–37.

Nithya, A., and V. Sundaram. 2011. "Classification Rules for Indian Rice Diseases." *International Journal of Computer Science*. https://www.researchgate.net/profile/John-Paul-Yusiong/publication/265566527_Tetris_Agent_Optimization_Using_Harmony_Search_Algorithm/links/56bb948108ae47fa3956aa5f/Tetris-Agent-Optimization-Using-Harmony-Search-Algorithm.pdf#page=464.

Parveen, Zahida, Muhammad Anzar Alam, and Hina Shakir. 2017. "Assessment of Quality of Rice Grain Using Optical and Image Processing Technique." In *2017 International Conference on Communication, Computing and Digital Systems (C-CODE)*, 265–70.

Prajapati, Bhavesh B., and Sachin Patel. 2013.

"Algorithmic Approach to Quality Analysis of Indian Basmati Rice Using Digital Image Processing." *Int J Emerg Technol Adv Eng* 3 (3): 503–4.

Saravanan, A., P. Senthil Kumar, B. Ramesh, and S. Srinivasan. 2022. "Removal of Toxic Heavy Metals Using Genetically Engineered Microbes: Molecular Tools, Risk Assessment and Management Strategies." *Chemosphere* 298 (July): 134341.

Sathish, T., R. Saravanan, V. Vijayan, and S. Dinesh Kumar. 2022. "Investigations on Influences of MWCNT Composite Membranes in Oil Refineries Waste Water Treatment with Taguchi Route." *Chemosphere* 298 (July): 134265.

Silva, Chathurika Sewwandi, and D. U. J. Sonnadara. 2013. "Classification of Rice Grains Using Neural Networks." http://192.248.16.117:8080/research/bitstream/70130/4861/1/ipsl13-02.pdf.

Singh, S. K., S. K. Vidyarthi, and R. Tiwari. 2020. "Machine Learnt Image Processing to Predict Weight and Size of Rice Kernels." *Journal of Food Engineering*. https://www.sciencedirect.com/science/article/pii/S0260877419304728.

Soltani, Mahmoud, and Mahmoud Omid. 2015. "Detection of Poultry Egg Freshness by Dielectric Spectroscopy and Machine Learning Techniques." *LWT - Food Science and Technology* 62 (2): 1034–42.

Verma, J. P. 2012. *Data Analysis in Management with SPSS Software*. Springer Science & Business Media.

Vivek, J., T. Maridurai, K. Anton Savio Lewise, R. Pandiyarajan, and K. Chandrasekaran. 2022. "Recast Layer Thickness and Residual Stress Analysis for EDD AA8011/h-BN/B4C Composites Using Cryogenically Treated SiC and CFRP Powder-Added Kerosene." *Arabian Journal for Science and Engineering*. https://doi.org/10.1007/s13369-022-06636-5.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Algal Biofuels: Technological Perspective on Cultivation, Fuel Extraction and Engineering Genetic Pathway for Enhancing Productivity." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123814.

Yaashikaa, P. R., P. Senthil Kumar, and S. Karishma. 2022. "Review on Biopolymers and Composites – Evolving Material as Adsorbents in Removal of Environmental Pollutants." *Environmental Research*. https://doi.org/10.1016/j.envres.2022.113114.

Zareiforoush, Hemad, Saeid Minaei, Mohammad Reza Alizadeh, and Ahmad Banakar. 2016. "Qualitative Classification of Milled Rice Grains Using Computer Vision and Metaheuristic Techniques." *Journal of Food Science and Technology* 53 (1): 118–31.

Tables and Figures

Table 1. Comparison of SVM and D-Tree Classifier for predicting the quality of rice in rice mills with improved accuracy rate. The accuracy rate of SVM is 91.0 and D-Tree is 92.7.

S.No.	Test Size	ACCURACY RATE	
		Support Vector Machine (SVM) Classifier	Decision Tree (D-Tree) Classifier
1	Test1	90.344	92.004
2	Test2	91.083	92.031
3	Test3	91.075	92.675
4	Test4	90.192	92.192

5	Test5	91.014	91.484
6	Test6	91.012	91.872
7	Test7	91.083	91.483
8	Test8	91.094	92.394
9	Test9	91.093	92.493
10	Test10	90.393	92.393
Average Test Results		91.0	92.7

Table 2. The statistical calculation such as Mean, standard deviation and standard error Mean for SVM and D-Tree. The accuracy rate parameter used in the t-test. The mean accuracy rate of SVM is 91.0 and D-Tree is 92.7. The Standard Deviation of SVM is 1.92394 and D-Tree is 0.22843. The Standard Error Mean of SVM is 0.93494 and D-Tree is 0.12948.

Group		**N**	**Mean**	**Standard Deviation**	**Standard Error Mean**
ACCURACY	DECISION TREE (D-TREE)	10	92.7	0.22843	0.12948
	SUPPORT VECTOR MACHINE (SVM)	10	91.0	1.92394	0.93494

Table 3. The statistical calculations for independent samples test between SVM and D-Tree. The significance level for signal to noise ratio is 0.035. Independent samples T-test is applied for comparison of SVM and D-Tree with the confidence interval as 95% and level of significance as 0.2323.

Group	**Levene's Test for Equality of Variances**		**t-test for Equality of Means**						
	F	**Sig.**	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**	**95% Confidence Interval (Lower)**	**95% Confidence Interval (Upper)**

Accuracy	Equal variances assumed	9.02	).035	18.39	18	.001	12.9823	0.87655	12.72342	15.83455
	Equal variances not assumed			12.33	14.67	.001	12.0233	0.12497	10.34444	13.01246

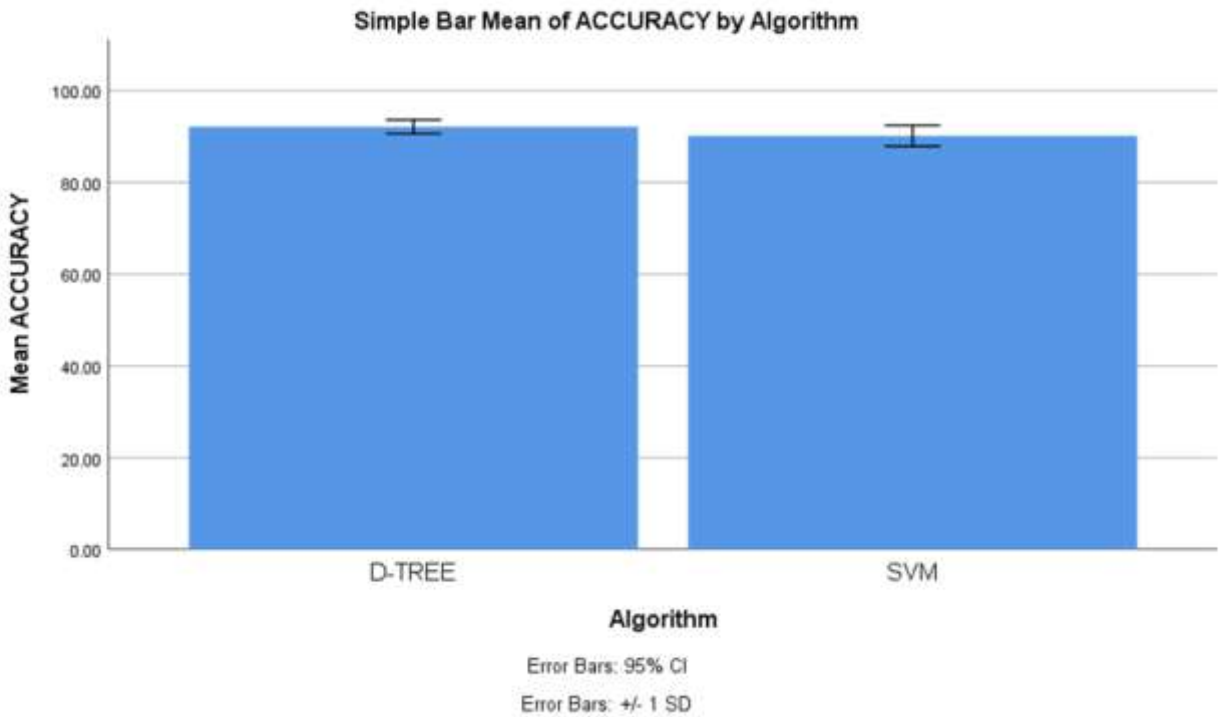

Fig. 1. Simple Bar graph for SVM Classifier accuracy rate is compared with D-Tree Classifier. The SVM Classifier is higher in terms of accuracy rate 91.0 when compared with D-Tree Classifier 92.7. Variable results range from 80 lower to 90 higher SVM Classifier where D-Tree Classifier standard deviation ranges from 90 lower to 100 higher. There is a significant difference between the SVM Classifier and D-Tree Classifier ($p<0.05$ Independent sample test). X-axis: D-Tree Classifier accuracy rate vs. SVM Classifier Y-axis: Mean accuracy rate for identifying keywords ± 1 SD with 95 % CI.