ANALYZING AND IMPROVING THE ACCURACY OF MUSIC GENRE CLASSIFICATION USING NOVEL SUPPORT VECTOR CLUSTERING ALGORITHM COMPARED WITH LOGISTIC REGRESSION CLASSIFIER

M. Pavan Venkata Naga Sai[1], S. Kalaiarasi[2*]

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

Abstract

Aim: To enhance the Music genre classification using Novel Support Vector clustering Algorithm and Logistic Regression with improved accuracy.

Materials and methods: Two groups such as novel Support Vector Machines Algorithm and Logistic Regression are applied. There will be a total of samples analyzed using this approach of 1000 music files. Among this sample dataset, 300 music files[70%] portion of the dataset was used as a training dataset. and 700 [30%] was taken as a testing dataset. Programming experiment was carried out for N=10 iterations for the Novel Support Vector clustering Algorithm and Logistic Regression algorithm respectively. Computation processes were executed and verified for exactness. Each group consists of a sample size of 10 Alpha value of the study's parameters is 2.020, beta value 1.359. The SPSS was used for predicting significance value of the dataset considering G-Power value as 80%.

Results and Discussions: Novel Support Vector clustering Algorithm shows a high accuracy and homogeneity for Music genre classification, and statistical significance difference is less than 0.001 ($p>0.05$).

Conclusion: The purpose of this essay is to adopt a novel method to the study of the subject Music genre classification. Comparison results show that efficiency Using the novel Support Vector Clustering technique is better than Logistic Regression Classifier for Music genre classification.

Keywords: Music Genre Classification, Novel Support Vector Clustering algorithm, Logistic Regression, Machine Learning, Acoustic features, Gtzan Database, Emotion.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical And Technical Sciences, Saveetha University, Chennai, Tamil Nadu. India. Pincode: 602105.

[2*]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

Computerized music handling is interested in music kind expectations. Advanced sign preparation procedures were used to eliminate music's acoustic components for the purpose of this evaluation and afterward music kind characterization and music proposals have been developed using machine learning methods. Moreover, Convolutional neural networks, Class classification and music suggestion and assessment, both of which are effective teaching strategies, were carried out using these techniques. Gtzan's database was used in the review, and the best results were achieved using the Gtzan information base with the Novel Support Vector Clustering calculation (Sousou et al. 2022). Since its widespread implementation, the Internet has had a significant impact on the music industry, including a wide range of progress. Instances of these turns of events incorporate the far reaching utilization of online music tuning in and deals stages, rights to copyrighted material, grouping of music kind, and music suggestions. In today's day and age, music broadcasting has had a significant impact on the industry, individuals can pay attention to music whenever and at anyplace and can come to a great many melodies through different music listening stages like Last.fm and Spotify. In this review, it is planned to make melodic proposals and music type arrangement utilizing acoustic components obtained by advanced sign preparation techniques from crude music disregarding the client's music profile or cooperative separation (Tate et al. 2013). Transformation of the original input set to higher dimensional feature space is done using the kernel function of Novel Support Vector Clustering algorithm, In the field of music signal processing, gtzan, which was initially introduced by G. Tzanetakis, is one of the most popular datasets. The advantage of Support Vector Clustering is that it requires less training data when compared to other models. The acoustic handling of the music data set improves the accuracy by a greater rate. The proposed algorithm helps in improving the accuracy of Music Genre Classification (Kawano et al. 2022).

There are about 34 articles in IEEE xplore and 20 in Scopus related to this study. In a study by G. Tzanetakis, In gtzan-based systems the probability was found using past frames and future frames which cover low-range in Blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, and rock are among the 100 genres represented in the gtzan. Support Vector Clustering is derived from the use of sequential data. Analyzing audio data using time-frequency is by far the most prevalent audio processing technique. (Nizamettin (2018)). The KNN (K-Nearest Neighbors) unfolds in time involved in its forward computation (Lim, Jang, and Lee 2016). KNN formulates as a dynamic classifier that analyzes each frame in turn to determine the class label. In a paper by (Hadjadji et al. 2019), it is said that Logistic Regression is suitable for learning time series data i.e long term temporal dependencies. Speech signal of the denoising autoencoder implemented with Logistic Regression is said to improve classification performance (Fayek, Lech, and Cavedon 2017). Previously there was wide-ranging research expertise in a wide range of academic fields. (Zhang et al. 2022).

Our institution is keen on working on latest research trends and has extensive knowledge and research experience which resulted in quality publications (Rinesh et al. 2022; Sundararaman et al. 2022; Mohanavel et al. 2022; Ram et al. 2022; Dinesh Kumar et al. 2022; Vijayalakshmi et al. 2022; Sudhan et al. 2022; Kumar et al. 2022; Sathish et al. 2022; Mahesh et al. 2022; Yaashikaa et al. 2022).Some datasets are aimed at theoretical research rather than processing it as per the real life Emotion. Therefore defining the boundaries between the Music Genre. Most of the existing standard feature extraction processes are for short-term analysis, So researchers have made their own feature set. Finally a paper is proposed assuming all the limitations in Machine Learning. This paper solely focuses on enhancing the Music models to increase the accuracy of Music Genre classification (Liu et al. 2022).

2. MATERIALS AND METHODS

This work is done at Saveetha School of Engineering. The review comprises two example bunches i.e Support Vector grouping Algorithm and Logistic Regression. Each gathering comprises 10 examples with pre-test force of 0.18. The example size was gathered from (G. Tzanetaki, gtzan dataset) by keeping the edge at 0.05, G force of 80%, certainty stretch at 95%, and enrolment proportion as 1. The collection of data used to create the grouping is taken from the gtzan Database of Music classification, an open-source information vault for music signal preparation. It contains insights regarding the 100 tunes included in the gtzan, which includes blues, traditional, country, disco, hip-hop, jazz, metal, pop, reggae, and rock genres (Goienetxea et al. 2018).

The dataset contains 3 columns. In the Satz (Genre) column, it represents the Feature Dataset in wav files format. In the erkannt (sound wave) column, it represents the selection of the sentences that take place with at least 80% correctly recognized

emotion. whereas in natuerlich (frequency) column, it represents at least 60% of the assessments as natural. The dataset contains 816 instances. The independent variables in this study are audio wav-files and the dependent variables are recognized Music Genre Classification. It was necessary to divide the data set into training and testing sections using a predetermined test size of 0.2 (Hao 2019).
For training of the Support Vector Clustering, the sample size for the experiment is about 20% leftover data from the complete dataset 80% is used for the training set. The Support Vector Clustering training set determines a hyperplane to partition the training data into two groups, whereas the Logistic Regression model uses backpropagation for training. The whole dataset is fitted for training the Novel Support Vector Clustering algorithm and Logistic Regression model. Both models are evaluated using a small sample size of 10 using Python 2.7 (Ajoodha 2014; Wilkes, Vatolkin, and Müller 2021)**.**

Support Vector Clustering

Support Vector Clustering are no assumptions made regarding the quantity or form of the data clusters while using this nonparametric clustering approach. High-dimensional data frequently requires preprocessing, such as utilizing principal component analysis, which we've found to be most effective with low-dimensional data. libsvm is used for the implementation. With increasing sample size, the fit time quadruples, making it infeasible to fit more than 10,000 samples. Consider using Linear Support Vector Clustering or SGDclassifier instead, potentially after a Nystroem transformer for huge datasets. When it comes to multiclass support, we use a one-vs-one approach.
Presently, Prepared and tested sets of data are separated from one other, utilizing a test size of 0.2. Scaling the sets utilizing the standard scaler Fitting the preparation set with Novel Support Vector Clustering algorithm and fitting the test set to the model to finish last expectations. The accuracy of the precision of the forecast's Emotion is found using equation 1.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{——--------- (1)}$$

where,
TP= no of true positive model is used to classify
FP= no of false positive model is used to classify
TN= no of true negative model is used to classify
FN= no of false negative model is used to classify

Logistic Regression

In the context of Supervised Learning, one of the most well-known machine learning computations is Logistic Regression. It's used to predict the unaffected ward variable by combining a number of independent variables. Using Logistic Regression, an unmitigated ward variable may be forecasted. As a consequence, the final value should be a straightforward or discrete one. Usually, it's either Yes or No, 0 or 1, True or False, and so on; nevertheless. When supplying particular values, it delivers probabilistic attributes that fall anywhere between 0 and 1 in value.
Logistic Regression is similar to Linear Regression but for the way it is applied. Regression concerns are addressed with Direct Regression, While Logistic relapse is used to deal with the characterisation problems. Logistic regression does not include the fitting of a relapse line, but rather we fit an "S" formed strategic capacity, where 0 or 1 are the two most important attributes to be predicted.
Logistic Regression is an important Machine Learning Algorithm since it provides probabilities and classifies new information based on endless and discrete datasets. Logistic Regression relapse utilizes the idea of prescient demonstrating as relapse; in this manner, it is called strategic relapse, however is utilized to characterize tests; Therefore, it falls under the characterization algorithm.The subordinate variable should be absolute in nature.The autonomous variable should not have multicollinearity.

Statistical Analysis

Statistical Package for the Social Sciences Version 26 software was used to do statistical analysis.. An independent sample T-test was performed to verify the results. It was also possible to determine standard deviation and mean errors using the SPSS software programme. The independent variables are audio wav-files and the dependent variables are recognized music genre classification from the dataset.

3. Results

The results of the statistical comparison of two groups demonstrate Novel Support Vector Clustering algorithm Table 8 has a greater mean accuracy than Logistic Regression table 9 and The standard error mean of Novel Support Vector Clustering is slightly lower than the standard error mean Fig.1 accuracy with standard deviation error is displayed for both methods in a bar chart The Support Vector Clustering algorithm had an accuracy rating of 63.0% as shown in Table 2 and Logistic Regression has scored 61.3% as shown in Table 3. The accuracies are recorded by testing the algorithms with 10 different sample sizes and the average accuracy is calculated for each algorithm.
From the results of this study, the Novel Support Vector Clustering algorithm is proved to be having

better accuracy than the Logistic Regression. Novel Support Vector Clustering algorithm has an accuracy of 63.0% whereas Logistic Regression has an accuracy of 61.3%. In Table 4, The statistical examination of the two groups demonstrates that Support Vector Clustering Table 8 has more mean accuracy than Logistic Regression Table 9 and the standard error mean including standard deviation mean is slightly less than Support Vector Clustering.

The limitation in this model is that the accuracy of the Novel Support Vector Clustering algorithm Inconsistent data and difficulties gathering the appropriate datasets for analysis might have an impact. Most of the data is simulated from nature which is far from reality. The availability of more cross-language related datasets of music, effective data preprocessing techniques, and the combination of Novel Support Vector Clustering algorithm Using additional methods for machine learning such as KNN and Random Forest may provide more accurate outcomes in the future.

4. Discussion

In this study, Music Genre Classification using the Novel Support Vector Clustering Algorithm has significantly higher accuracy of 63% in comparison to Logistic Regression Classifier with 46% Improved Accuracy.

The similar findings of the paper had an accuracy of 63% with the Novel Support Vector Clustering Algorithm which was used to detect the music (Ismanto, Kusuma, and Anggraini 2021). The proposed work of reported Novel Support Vector Clustering Algorithm has 63% accuracy which is used to predict the accuracy of Music Genre Classification and performance of music waves (Chaudary et al. 2021). The work proposed shows the Novel Support Vector Clustering Algorithm has a better accuracy of 63% (Barbosa et al. 2021). Logistic Regression is a parameter to measure Music Frequency which is used in both traditional and modern methods as per their research it opposes Mel-Frequency Cepstral Coefficients has highest accuracy and Logistic Regression will get least accuracy compared to other machine learning techniques which ranges between 63% when compared to other machine learning algorithms will get more accuracy than this (Brownlee 2018). By using Support Vector Clustering for forecasting Music Classification it will have key issues to pretend in this paper shows Logistic Regression has the least accuracy of 61%. Increasing the dataset's value only tends to get desired accuracy. Support Vector Clustering performs better with a combination of other machine learning algorithms (Hozano et al. 2017).

The limitation of this research is that smaller data sets are unable to provide meaningful findings. It is not possible to train this model using all of the specified feature variable settings. The projected project's eventual scope will include music genre classification. based on classification model using class labels for lesser time complexity (Bellinger, n.d.). Our future work will focus on faster and more accurate music detection through the use of advanced machine learning techniques.

5. Conclusion

Based on the experimental results, the Novel Support Vector Clustering has been proved to recognize speech emotion more significantly than Logistic Regression. It can be used in detecting Music Genre Classification systems According to the findings of the recommendations, the music suggestion is particularly effective for specific genres of music like classical, the performance of certain species is better than others falls in Machine Learning. Deep learning architectures and high-level feature extraction will be major areas of future study.

DECLARATIONS

Conflicts Of Interest

No conflicts of interest in this manuscript.

Author Contributions

Author MPVNS was involved in data collection, data analysis, data extraction, manuscript writing. Author SK was involved in conceptualization, data validation, and critical 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 this work successfully.

Funding

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

1. HKM Info Systems Pvt. Ltd.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

6. References

Ajoodha, Ritesh. 2014. *Automatic Music Genre Classification*.

Barbosa, Aaron, Elijah Pelofske, Georg Hahn, and Hristo N. Djidjev. 2021. "Using Machine Learning for Quantum Annealing Accuracy Prediction." *Algorithms*. https://doi.org/10.3390/a14060187.

Bellinger, David. n.d. "Learning from Science Lectures : Students Remember More and Make Better Inferences When They Complete Skeletal Outlines Compared to Other Guided Notes." https://doi.org/10.18297/etd/2513.

Brownlee, Jason. 2018. *Statistical Methods for Machine Learning: Discover How to Transform Data into Knowledge with Python*. Machine Learning Mastery.

Chaudary, Eamin, Sumair Aziz, Muhammad Umar Khan, and Paul Gretschmann. 2021. "Music Genre Classification Using Support Vector Machine and Empirical Mode Decomposition." *2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)*. https://doi.org/10.1109/majicc53071.2021.9526251.

Dinesh Kumar, M., V. Godvin Sharmila, Gopalakrishnan Kumar, Jeong-Hoon Park, Siham Yousuf Al-Qaradawi, and J. Rajesh Banu. 2022. "Surfactant Induced Microwave Disintegration for Enhanced Biohydrogen Production from Macroalgae Biomass: Thermodynamics and Energetics." *Bioresource Technology* 350 (April): 126904.

Fayek, Haytham M., Margaret Lech, and Lawrence Cavedon. 2017. "Evaluating Deep Learning Architectures for Speech Emotion Recognition." *Neural Networks: The Official Journal of the International Neural Network Society* 92 (August): 60–68.

Goienetxea, Izaro, José María Martínez-Otzeta, Basilio Sierra, and Iñigo Mendialdua. 2018. "Towards the Use of Similarity Distances to Music Genre Classification: A Comparative Study." *PloS One* 13 (2): e0191417.

Hadjadji, Imene, Leila Falek, Lyes Demri, and Hocine Teffahi. 2019. "Emotion Recognition in Arabic Speech." *2019 International Conference on Advanced Electrical Engineering (ICAEE)*. https://doi.org/10.1109/icaee47123.2019.9014809.

Hao, Yifei. 2019. *Music Genre Classification and Transfer Based on MusicXML with High-Level Features and Chord Vectors*.

Hozano, Mário, Nuno Antunes, Baldoino Fonseca, and Evandro Costa. 2017. "Evaluating the Accuracy of Machine Learning Algorithms on Detecting Code Smells for Different Developers." *Proceedings of the 19th International Conference on Enterprise Information Systems*. https://doi.org/10.5220/0006338804740482.

Ismanto, Brizky Ramadhani, Tubagus Maulana Kusuma, and Dina Anggraini. 2021. "Indonesian Music Classification on Folk and Dangdut Genre Based on Rolloff Spectral Feature Using Support Vector Machine (SVM) Algorithm." *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*. https://doi.org/10.22146/ijccs.54646.

Kawano, Daisuke, Hitoshi Mori, Kenta Tsutsui, Yoshifumi Ikeda, Mitsuki Yamaga, Akane Kawai, Atsushi Sato, et al. 2022. "The Target Ablation Index Values for Electrical Isolation of the Superior Vena Cava." *Journal of Interventional Cardiac Electrophysiology: An International Journal of Arrhythmias and Pacing*, February. https://doi.org/10.1007/s10840-021-01112-w.

Kumar, J. Aravind, J. Aravind Kumar, S. Sathish, T. Krithiga, T. R. Praveenkumar, S. Lokesh, D. Prabu, A. Annam Renita, P. Prakash, and M. Rajasimman. 2022. "A Comprehensive Review on Bio-Hydrogen Production from Brewery Industrial Wastewater and Its Treatment Methodologies." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123594.

Lim, Wootaek, Daeyoung Jang, and Taejin Lee. 2016. "Speech Emotion Recognition Using Convolutional and Recurrent Neural Networks." *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*. https://doi.org/10.1109/apsipa.2016.7820699.

Liu, Ya-Min, Yu-Juan Wei, Xiao-Qing Lu, Yu-Fei Wang, Pei Wang, and Xian-Hui Liang. 2022. "Catheter-Related Superior Vena Cava Obstruction: A Rare Cause of Chylothorax in Maintenance Hemodialysis Patients." *The Journal of Vascular Access*, January, 11297298211073425.

Mahesh, Narayanan, Srinivasan Balakumar, Uthaman Danya, Shanmugasundaram Shyamalagowri, Palanisamy Suresh Babu, Jeyaseelan Aravind, Murugesan Kamaraj, and Muthusamy Govarthanan. 2022. "A Review on Mitigation of Emerging Contaminants in an Aqueous Environment Using Microbial Bio-Machines as Sustainable Tools: Progress and Limitations." *Journal of Water Process Engineering*. https://doi.org/10.1016/j.jwpe.2022.102712.

Mohanavel, Vinayagam, K. Ravi Kumar, T. Sathish, Palanivel Velmurugan, Alagar Karthick, M. Ravichandran, Saleh Alfarraj, Hesham S. Almoallim, Shanmugam

Sureshkumar, and J. Isaac JoshuaRamesh Lalvani. 2022. "Investigation on Inorganic Salts K2TiF6 and KBF4 to Develop Nanoparticles Based TiB2 Reinforcement Aluminium Composites." *Bioinorganic Chemistry and Applications* 2022 (January): 8559402.

Ram, G. Dinesh, G. Dinesh Ram, S. Praveen Kumar, T. Yuvaraj, Thanikanti Sudhakar Babu, and Karthik Balasubramanian. 2022. "Simulation and Investigation of MEMS Bilayer Solar Energy Harvester for Smart Wireless Sensor Applications." *Sustainable Energy Technologies and Assessments*. https://doi.org/10.1016/j.seta.2022.102102.

Rinesh, S., K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji. 2022. "Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms." *Journal of Healthcare Engineering* 2022 (February): 2761847.

Sathish, T., V. Mohanavel, M. Arunkumar, K. Rajan, Manzoore Elahi M. Soudagar, M. A. Mujtaba, Saleh H. Salmen, Sami Al Obaid, H. Fayaz, and S. Sivakumar. 2022. "Utilization of Azadirachta Indica Biodiesel, Ethanol and Diesel Blends for Diesel Engine Applications with Engine Emission Profile." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123798.

Sousou, John M., Douglass M. Sherard, Jamie R. Edwards, and Elsio Negron-Rubio. 2022. "Successful Removal of a Thrombus in the Setting of SVC Syndrome Using the INARI FlowTriever Device." *Radiology Case Reports* 17 (3): 744–47.

Sudhan, M. B., M. Sinthuja, S. Pravinth Raja, J. Amutharaj, G. Charlyn Pushpa Latha, S. Sheeba Rachel, T. Anitha, T. Rajendran, and Yosef Asrat Waji. 2022. "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model." *Journal of Healthcare Engineering* 2022 (February): 1601354.

Sundararaman, Sathish, J. Aravind Kumar, Prabu Deivasigamani, and Yuvarajan Devarajan. 2022. "Emerging Pharma Residue Contaminants: Occurrence, Monitoring, Risk and Fate Assessment – A Challenge to Water Resource Management." *Science of The Total Environment*. https://doi.org/10.1016/j.scitotenv.2022.153897.

Tate, Jon, Angelo Bernasconi, Torben Jensen, Ian MacQuarrie, Ole Rasmussen, Matthew Robinson, Steven Tong, and I. B. M. Redbooks. 2013. *IBM SAN and SVC Stretched Cluster and VMware Solution Implementation*. IBM Redbooks.

Vijayalakshmi, V. J., Prakash Arumugam, A. Ananthi Christy, and R. Brindha. 2022. "Simultaneous Allocation of EV Charging Stations and Renewable Energy Sources: An Elite RERNN-m2MPA Approach." *International Journal of Energy Research*. https://doi.org/10.1002/er.7780.

Wilkes, Ben, Igor Vatolkin, and Heinrich Müller. 2021. "Statistical and Visual Analysis of Audio, Text, and Image Features for Multi-Modal Music Genre Recognition." *Entropy* 23 (11). https://doi.org/10.3390/e23111502.

Yaashikaa, P. R., P. Senthil Kumar, S. Jeevanantham, and R. Saravanan. 2022. "A Review on Bioremediation Approach for Heavy Metal Detoxification and Accumulation in Plants." *Environmental Pollution* 301 (May): 119035.

Zhang, Lin, Gonghao Ling, Yadong Gang, Zhaoxia Yang, Zhibing Lu, Xuedong Gan, Hongqin Liang, Yingting Zeng, and Xiaochun Zhang. 2022. "Classification and Quantification of Double Superior Vena Cava Evaluated by Computed Tomography Imaging." *Quantitative Imaging in Medicine and Surgery* 12 (2): 1405–14.

Tables and Figures

Table 1: Gtzan Dataset collection from Music Genre Classification

Genre	First 5 Songs	First 10 Songs
Blues	60%	48%
Classical	88%	90%
Country	40%	50%
Disco	48%	40%

Hip-hop	72%	60%
Jazz	52%	42%
Metal	84%	76%
Pop	60%	50%
Reggae	24%	26%
Rock	60%	50%

Table 2: Accuracy of Music Genre Classification using Support Vector Clustering for 10 Samples Out of 30(Accuracy=63.0%)

Test Size	Accuracy(%)
Test 1	63
Test 2	62
Test 3	61
Test 4	60
Test 5	59
Test 6	58
Test 7	57
Test 8	56
Test 9	55
Test 10	54

Table 3: Accuracy of Music Genre Classification using Logistic Regression model for 10 Samples Out of 30(Accuracy=63.0%)

Test Size	Accuracy(%)
Test 1	61
Test 2	60
Test 3	59
Test 4	58
Test 5	57
Test 6	56
Test 07	55
Test 8	54

Test 9	53
Test 10	52

Table 4: Group Statistics

	Groups	N	Mean	std.Deviation	std.Error
ACCURACY	SVC	10	58.52	3.050	.965
	LR	10	56.50	3.028	.957

Group Statistic analysis, representing Support Vector Clustering (mean accuracy 58.52%, standard deviation 3.050) and Logistic Regression (mean accuracy 56.50%, standard deviation 3.028).

Table 5: Independent sample test

Independent sample test											
		Levene's Test for Equality of Variances		t-test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Accuracy	Equal variances assumed	.001	.976	1.486	18	.077	.155	2.020	1.359	-.835	4.875
	Equal variances not assumed			1.486	17.999	.077	.155	2.020	1.359	-.835	4.875

Independent Sample Tests results with confidence interval as 95% and level of significance as 0.05 (Support Vector Machine appears to perform significantly better than Logistic Regression model with the value of p=0.18).

Table 6: Independent Sample Effect Size

		Standardizer	Point Estimate	95% Confidence Interval	
				Lower	Upper
ACCURACY	Cohen's d	3.039	.665	-.247	1.559
	Hedges' correction	3.173	.637	-.236	1.493
	Glass's delta	3.028	.667	-.277	1.579

The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

Table 7: Pseudocode for Support Vector Clustering

// I : Input dataset records
1. Import the required packages.
2. Convert the audio files into numerical values after the extraction feature.
3. Assign the data to X_train, y_train, X_test and y_test variables.
4. Using train_test_split() function, pass the training and testing variables.
5. Give test_size and the random_state as parameters for splitting the data using SVM training.
6. Importing the Support Vector Clustering from sklearn library.
7. Using Support Vector Clustering, predict the output of the testing data.
8. Calculate the accuracy of the model.
OUTPUT **//Accuracy**

Table 8: PseudoCode for Logistic Regression

// I : Input dataset records
1. Import the required packages.
2. Convert the audio files and truncate the input sequences so that they are all the same length for modeling.
3. Assign the data to X_train, y_train, X_test and y_test variables.
4. Using train_test_split() function, pass the training and testing variables.
5. Give test_size and the random_state as parameters for splitting the data.
6. Data preprocessing step.

7. Fitting Logistic Regression to the Training set.
7. Evaluate the output using X_test and y_test function
8. Get the accuracy of the model.
OUTPUT **//Accuracy**

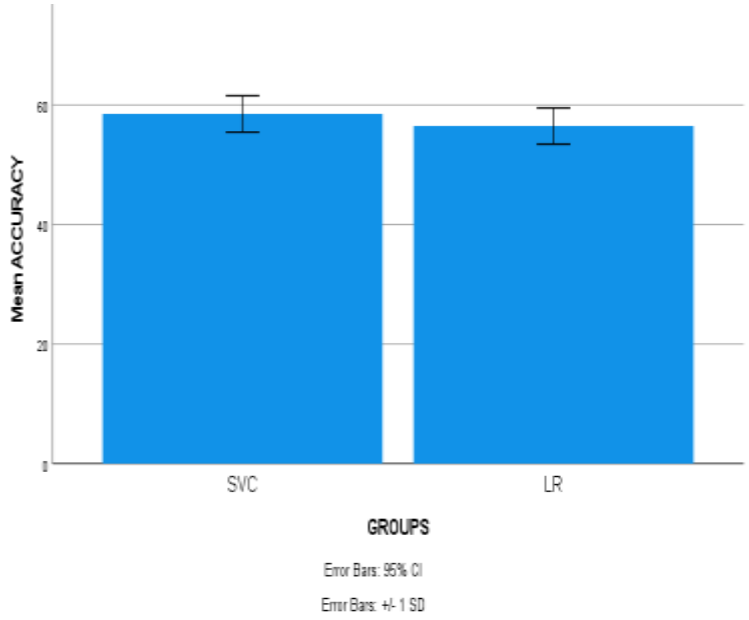

Fig. 1**.** Comparison of Support Vector Clustering and Logistic Regression in terms of accuracy. The mean accuracy of Support Vector Clustering is greater than Logistic Regression and the standard deviation is also slightly higher than Logistic Regression. X-axis: Support Vector Clustering vs Logistic Regression. Y-axis: Mean accuracy of detection $\pm$ 1 SD