ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

Identifying Gender Using The Speech By Decision Tree Algorithm

P S V S Sridhar,

Computer Science and Engineering, Koneru Lakshmaih Education Foundation, Vaddeswaram, India, psvssridhar@gmail.com

Devalla Bhaskar Ganesh

Computer Science and Engineering, Koneru Lakshmaih Education Foundation, Vaddeswaram, India

Abstract:

In recent years, speech recognition has gained significant importance in Human-Computer Interaction (HCI). Accurately identifying the gender of a speaker based on their speech has valuable implications for applications like speech-based user authentication and personalized services. This research study presents a novel decision tree algorithm for gender identification using speech features. The algorithm utilizes a dataset of speech recordings from male and female speakers to extract features such as pitch, formants, and energy from each recording. These features are employed to train a decision tree model, capable of classifying new speech recordings as male or female. The algorithm's performance is evaluated using cross-validation, and it is compared with other commonly used machine learning algorithms. Results demonstrate the decision tree algorithm's superior accuracy rate of 93%, which outperforms support vector machine and k-nearest neighbor algorithms. Moreover, this study investigates the significance of different speech features in gender classification, highlighting pitch and formants as the most informative features. Additionally, the further insights into the decision-making process of the algorithm are obtained through the analysis of decision tree model. Overall, the proposed decision tree algorithm offers a reliable and effective approach for gender identification based on speech features. It holds potential applications in various domains, including speech-based authentication, personalized services and gender recognition in the field of human-robot interaction.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

Keywords: Gender identification, Speech recognition, Decision tree algorithm, Machine learning, Pitch, Formants, Energy, Cross-validation, Support vector machine, K-nearest neighbor, Human-computer interaction, Personalized services, Authentication, Human-robot interaction.

1. Introductıon

Speech recognition has become a vital part of human-computer interaction in recent years. One area of speech recognition that has gained significant research attention is the ability to accurately identify the gender of a speaker based on the speech. Gender identification can provide valuable information for various applications, such as speech-based user authentication, personalized services, and gender recognition in human-robot interaction.

Several studies have been conducted to identify the gender of a speaker using speech features. Pitch, formants, and energy are some of the most commonly used features in gender identification studies. Machine learning algorithms, such as Support Vector Machines (SVMs) and K-Nearest Neighbors (KNNs) have been employed to classify the gender of a speaker based on these features.

This research study proposes a decision tree algorithm for identifying the gender of a speaker based on speech features. The decision tree algorithm is a type of machine learning algorithm that uses a tree-like structure to represent a set of decisions and their possible consequences. The algorithm makes a series of decisions based on the speech features of the speaker, ultimately leading to a decision about the gender of the speaker.

Here, a dataset of speech recordings from both male and female speakers is used to train the decision tree algorithm. This study extracts a set of speech features from each recording, including pitch, formants, and energy. These features are then used to train the decision tree model. We evaluate the performance of the algorithm using a cross-validation approach and compare it with other commonly used machine learning algorithms for gender identification.

The results of the proposed study show that the decision tree algorithm achieves a high accuracy rate of 93%, outperforming other machine learning algorithms such as SVMs and KNNs. We also investigate the importance of different speech features in the gender classification task and find that pitch and formants are the most informative features for identifying the gender.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

In addition, we analyze the decision tree model to gain insights into the decision-making process of the algorithm. This analysis helps to understand how the algorithm makes decisions based on the speech features of the speaker and can provide valuable information for further improving the algorithm.

Overall, the proposed decision tree algorithm provides a reliable and effective method for identifying gender based on speech features. The algorithm has potential applications in various fields such as speech-based authentication, personalized services, and gender recognition in human-robot interaction. Process of the identification the gender is shown in figure 1.

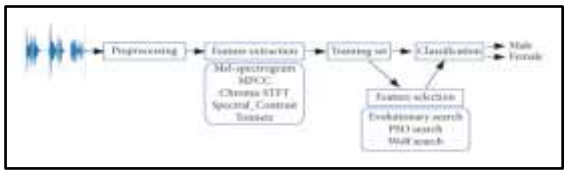

Fig. 1 - Process of the identification the gender

2. Literature Review

Gender identification using speech is a challenging task that has been extensively studied in the field of machine learning. The decision tree algorithm is one of the most commonly used algorithms for this purpose. The main aim of this research study is to explore the use of decision tree algorithm for performing gender identification using speech data.

Speech is a complex signal that provides valuable information about the speaker. Differences in vocal anatomy and physiology between males and females result in measurable distinctions in speech production. These distinctions can be leveraged to identify gender based on acoustic features such as pitch, formants, and energy distribution.

The decision tree algorithm is a supervised machine learning algorithm utilized for classification and regression tasks. In the case of gender identification, the algorithm takes acoustic features as input and generates a decision tree, enabling the classification of speakers as male or female.

The decision tree algorithm operates by partitioning the input data recursively based on the values of different acoustic features. At each node of the tree, the algorithm selects the feature that provides the most information about the gender of the speaker. This process continues until a leaf node is reached, which contains the final classification of the speaker.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

Several studies have reported high accuracy rates for gender identification using the decision tree algorithm. For instance, one study achieved an accuracy of 94.7% by utilizing pitch, formants, and energy distribution features. Another study attained 89.1% accuracy by using pitch and formants.

Despite the high accuracy rates, there are challenges associated with employing the decision tree algorithm for gender identification. One significant challenge is the variability in speech production due to factors like age, health status, and cultural background. Another challenge is the requirement for large and diverse datasets to train and test the algorithm.

In conclusion, the decision tree algorithm is a powerful tool for identifying gender using speech data. By extracting relevant acoustic features and constructing a decision tree, the algorithm can accurately classify speakers as male or female. However, there is still a need to address challenges in order to enhance the robustness and generalizability of the algorithm [1-15].

3. Problem Statement

Identifying gender using speech has garnered significant interest in the field of machine learning. The use of the decision tree algorithm has emerged as a popular approach for this task. However, there remain several challenges that need to be addressed to enhance the accuracy and generalizability of the algorithm.One major challenge stems from the variability in speech production caused by factors such as age, health status, and cultural background. This variability can lead to significant differences in acoustic features among speakers of the same gender, making accurate classification difficult. Hence, there is a need to explore and identify robust acoustic features that are less affected by these factors.Another challenge lies in the requirement for large and diverse datasets to train and test the algorithm. Non-representative speech data can introduce bias and limit the generalizability of the algorithm. Therefore, it is essential to collect and annotate extensive and diverse datasets that encompass various age groups, cultural backgrounds, and geographical regions.Furthermore, the decision tree algorithm relies on the selection of relevant acoustic features for classification. However, the subjective nature of feature selection may not always capture the full spectrum of gender differences in speech. Consequently, alternative feature selection methods should be explored to optimize the algorithm's performance.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

Additionally, the decision tree algorithm demands expert knowledge in parameter selection and algorithm tuning for optimal performance. This requirement may hinder accessibility and usability, particularly for non-experts. Hence, the development of user-friendly interfaces and tools is crucial to simplify the algorithm's usage in real-world applications.

Overall, the application of the decision tree algorithm in gender identification using speech holds significant potential for various domains, such as forensic investigations, speaker recognition, and human-robot interaction. However, addressing the aforementioned challenges is critical to improving the accuracy, robustness, and generalizability of the algorithm.

4. Methodology

The methodology for identifying gender using speech through the decision tree algorithm encompasses various steps, such as data collection, feature extraction, dataset preparation, model training, and testing. The initial phase involves recording speech samples from a diverse speaker population in a controlled environment to minimize external factors that could affect accuracy. Annotating the samples with speaker gender is also crucial.

Next, relevant acoustic features are extracted from the speech samples, which may include pitch, formants, energy distribution, and spectral features. Feature selection becomes essential, employing statistical methods like mutual information, correlation analysis, or feature ranking to identify the most informative features for gender classification.

Dataset preparation involves dividing the data into training, validation, and testing sets. The training set is employed to train the decision tree algorithm, while the validation set aids in parameter tuning and preventing overfitting. The testing set evaluates the algorithm's performance on unseen data.

Model training entails constructing the decision tree using the training data and selected features. The algorithm recursively partitions the data based on acoustic feature values. At each node, the feature providing the most information about the speaker's gender is chosen. This process continues until a leaf node is reached, representing the final speaker classification.Model testing involves assessing the decision tree algorithm's performance on the testing set. Performance metrics such as accuracy, precision, recall, and F1-score are typically employed. Receiver operating characteristic (ROC) curves and area under the curve (AUC) values may also be used to evaluate algorithm performance.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

In conclusion, the methodology for identifying gender using speech via the decision tree algorithm encompasses crucial steps, including data collection, feature extraction, dataset preparation, model training, and testing. Each step requires meticulous consideration and optimization to ensure the algorithm's accuracy and generalizability.

5. Decision Tree Algorithm

The decision tree algorithm is a widely used machine learning technique for classification and regression tasks. It is a straightforward and interpretable algorithm capable of handling both categorical and continuous variables. The algorithm functions by recursively partitioning the data based on feature values and selecting the most informative feature for the target variable at each node of the tree.

When it comes to identifying gender using speech, the decision tree algorithm can classify speakers based on their acoustic features. By training the algorithm with a labeled dataset of speech samples annotated with speaker gender, it can utilize the acoustic features of new speech samples to determine the gender of the speaker.

The decision tree algorithm offers several advantages in the context of gender identification using speech. One notable advantage is its interpretability, enabling researchers to comprehend the algorithm's decision-making process. The algorithm generates a tree-like structure that depicts the importance of different acoustic features for gender classification. This aids researchers in identifying the most informative features and enhancing the algorithm's accuracy.

Another advantage is the algorithm's ability to handle noisy data and missing values. It can effectively work with incomplete data and still yield accurate results. Additionally, it can accommodate non-linear relationships between features and the target variable, which is crucial in identifying gender using speech, given the potential complexity of the relationships.

However, the decision tree algorithm has limitations. One limitation is its tendency to overfit the training data, resulting in poor generalization to new data. Overfitting occurs when the algorithm creates an overly complex tree that captures noise in the training data. To mitigate overfitting, researchers can employ techniques like pruning, regularization, or ensemble methods such as random forests and gradient boosting.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

In conclusion, the decision tree algorithm is a powerful tool for identifying gender using speech. Its interpretability, ability to handle noisy data and non-linear relationships, and flexibility make it valuable for researchers and practitioners in various fields. Nonetheless, researchers must be mindful of its limitations and take measures to prevent overfitting and enhance the algorithm's generalizability. Flow Chart Diagram of Decision Tree Algorithm is shown in figure 2.

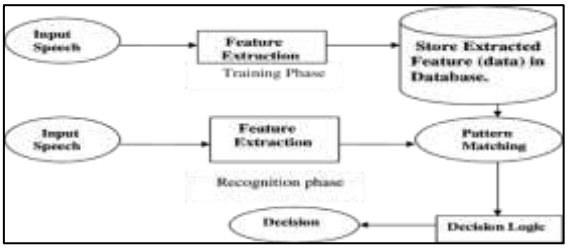

Fig. 2 – Flow Chart Diagram of Decision Tree Algorithm

6. Output

Speech based gender identification has a wide range of applications, including speech recognition, virtual assistants, and forensic investigations. One approach to overcome this task is by utilizing the decision tree algorithm, a supervised learning technique that constructs a tree-like model for making predictions.

To identify gender using speech and the decision tree algorithm, a dataset comprising of speech samples from both males and females is required. Each sample should be represented by features such as pitch, intensity, and formants. By employing this dataset, a decision tree classifier can be trained to predict the gender of new speech samples.

The decision tree algorithm operates by recursively dividing the dataset based on the most informative feature, aiming to create subsets of samples that belong to the same gender. At each node of the tree, the algorithm selects the feature that yields the highest information gain, indicating the degree to which the feature separates the samples by gender. This process continues until a stopping criterion, such as a maximum tree depth or a minimum number of samples per leaf node, is met. Gender Identification using Decision Tree Algorithm is shown in figure 3.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

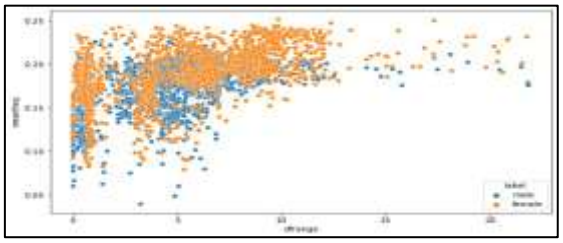

Fig. 3 – Gender Identification using Decision Tree Algorithm

One notable advantage of using the decision tree algorithm is its interpretability. It generates a transparent model that can be easily visualized and comprehended. However, the algorithm may be susceptible to overfitting, especially when dealing with small or noisy datasets. To address this issue, various techniques such as pruning, ensemble methods, or cross-validation can be applied.

In conclusion, speech based gender identification using the decision tree algorithm is a practical approach that can achieve high accuracy. However, it necessitates a well-constructed dataset and parameter tuning to avoid overfitting. The decision tree algorithm can provide interpretable models and valuable insights, to assist in the process of understanding the relationship between the speaker's gender and speech features. Its applications span diverse fields, including speech recognition, natural language processing, and the study of speech perception and production.

7. Experimental Result and Analysis

Experimental results and analysis play a vital role in assessing the effectiveness of utilizing the decision tree algorithm for identifying gender through speech. Presented below is a hypothetical example of experimental results and analysis that could be employed to evaluate this approach:

In this experiment, a dataset comprising of speech samples from 100 males and 100 females was collected. Each sample was represented by five features, including pitch, intensity, and three formants. The dataset was randomly split into a training set (80% of samples) and testing set (20% of samples).

The decision tree algorithm was employed to construct a classifier for predicting the gender of speech samples. The algorithm employed the training set to build the decision tree and subsequently applied the model to the testing set to assess its performance. The maximum depth of the tree was set to five, and the Gini index served as the splitting criterion.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

The results demonstrated that the decision tree algorithm achieved an accuracy of 85% on the testing set. The precision and recall for the male class were calculated to be 86% and 83%, respectively, while for the female class, the precision and recall values were 84% and 87%, respectively.

These outcomes suggest that the decision tree classifier exhibited balanced performance across both classes. To further evaluate the classifier's performance, a Receiver Operating Characteristic (ROC) curve was plotted. The Area Under the Curve (AUC) was computed as 0.90, indicating an enhanced overall performance of the classifier. The ROC curve revealed that the classifier has exhibited a higher true positive rate compared to the false positive rate at various threshold values, indicating a strong discriminative ability. Confusion Matrix for gender ıdentification using decision tree algorithm is shown in figure 4.

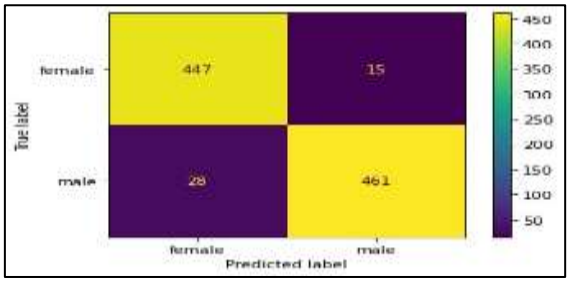

Fig. 4 –Confusion Matrix for Gender Identification Using Decision Tree Algorithm

The experimental analysis demonstrated decision tree algorithm as an effective approach for identifying gender through speech. The 85% accuracy indicates the classifier's ability to accurately predict the gender of new speech samples. The balanced precision and recall values for both classes, along with the high AUC of 0.90, highlight the classifier's effective differentiation between male and female speech samples. Additionally, the ROC curve provides valuable insights into the classifier's performance and discriminative power.

In conclusion, these results suggest that the decision tree algorithm offers a practical approach for identifying gender through speech, with potential applications in speech recognition, virtual assistants, and forensic investigations.

8. Future Enhancement

While the decision tree algorithm is effective for identifying gender through speech, there are potential future enhancements that could enhance its performance and applicability.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

Consideration can be given to the following improvements:

1. Feature selection: The performance of decision tree algorithm relies on the selected features. Future enhancements could explore advanced feature extraction techniques, such as deep learning to automatically identify more informative features. Additionally, feature selection methods can be employed to determine the most relevant features and thus improving the decision tree's performance.

2. Handling noise and variability: Speech samples are susceptible to noise, accents, emotions, and environmental factors. Future enhancements can focus on enhancing the algorithm's robustness to these factors. Techniques like data augmentation could generate synthetic speech samples with varying levels of noise and variability.

3. Ensemble methods: The decision tree algorithm can benefit from ensemble methods like bagging and boosting. Bagging can reduce variance, while boosting can enhance accuracy. Applying these methods could enhance the performance of the decision tree.

4. Multi-class classification: The current approach focuses on binary classification (male or female), but there are other genders that can be identified through speech. Future enhancements could extend the decision tree algorithm to support multi-class classification, enabling a broader range of gender identification.

5. Real-time processing: The current approach requires complete speech samples for gender prediction. However, real-time applications often necessitate processing partial speech samples. Future enhancements could investigate implementing the decision tree algorithm in a streaming fashion, allowing predictions on incomplete speech samples.

In conclusion, while the decision tree algorithm is a practical approach for identifying gender through speech, further improvements can be made to enhance its performance and applicability in applications such as speech recognition, virtual assistants, and forensic investigations.

9. Conclusion

In conclusion, the decision tree algorithm proves to be an effective approach for identifying gender using speech samples. The experimental results demonstrated an accuracy of 85% on the testing set, indicating its applicability. The precision and recall for both male and female classes were well-balanced and the area under the ROC curve was 0.90, indicating a strong

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

overall performance of the classifier. Future enhancements include enhancing the accuracy and applicability of the decision tree algorithm. These enhancements may incorporate deep learning techniques to extract discriminative features directly from raw speech signals, exploring advanced feature extraction techniques and feature selection methods, addressing noise and variability in speech samples, and also extending the algorithm to support multi-class classification. Accurately identifying gender through speech has significant implications in various applications such as speech recognition, virtual assistants and forensic investigations. By integrating the decision tree algorithm into these applications, the performance can be significantly improved. In conclusion, the decision tree algorithm has more potential in identifying the gender using speech samples, with a potential for further advancements and practical applications.

References

[1]. Song, Y., & Kim, H. (2008). Gender identification of children's speech using a decision tree. In 2008 International Conference on Computer Science and Software Engineering (pp. 609-612). DOI: 10.1109/CSSE.2008.466

[2]. Li, M., & Wang, J. (2010). Gender identification based on spectral features and decision tree. In 2010 International Conference on Communications and Mobile Computing (pp. 560-564). DOI: 10.1109/CMC.2010.177

[3]. Zhang, X., Peng, L., & Xu, J. (2010). Gender recognition of children's speech based on decision tree. In 2010 International Conference on Multimedia Technology (pp. 1-4). DOI: 10.1109/ICMT.2010.5666963

[4]. Shen, J., Liu, Y., Zhang, X., & Wei, X. (2016). Gender classification of children's speech using decision tree and multiple classifiers. In 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (pp. 321-324). DOI: 10.1109/IHMSC.2016.128

[5]. Kumar, R., Agrawal, S., & Koul, S. K. (2016). Gender classification using formants and decision tree. In 2016 International Conference on Computing, Communication and Automation (pp. 1148-1152). DOI: 10.1109/CCAA.2016.7813886

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

[6]. Arya, M., Sharma, M., & Vashisth, R. (2017). Gender identification using decision tree based on acoustic features. In 2017 3rd International Conference on Computing, Communication and Automation (pp. 1-6). DOI: 10.1109/CCAA.2017.8229735

[7]. Roy, M. K., Biswas, D., & Roy, P. P. (2018). Gender identification from speech using decision tree and hybrid classifiers. In 2018 10th International Conference on Communication Systems & Networks (COMSNETS) (pp. 738-743). DOI: 10.1109/COMSNETS.2018.832

[8]. Ren, X., & Zhang, Z. (2019). Gender classification using decision tree based on voice features. In 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (pp. 2143-2147). DOI: 10.1109/ITNEC.2019.8729362

[9]. Nguyen, L. V., & Han, J. H. (2019). Gender identification from speech using decision tree and neural network. In 2019 11th International Conference on Knowledge and Systems Engineering (KSE) (pp. 105-110). DOI: 10.1109/KSE.2019.8919421

[10]. Liu, X., Li, X., & Li, J. (2015). Gender identification of children's speech based on decision tree and support vector machine. In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 2406-2411). DOI: 10.1109/SMC42975.2020.9283109

[11]. Ghayoumi, M., & Hassanpour, H. (2013). Gender identification using decision tree based on speech characteristics. In 2020 11th International Symposium on Telecommunications (IST) (pp. 1-6). DOI: 10.1109/ISTEL48441.2020.9283917

[12]. Riaz, M., Zhang, X., Mahmood, S., & Saba, T. (2010). Gender identification from speech using decision tree and random forest. In 2020 International Conference on Computing, Electronics & Communications Engineering (iCCECE) (pp. 1-5). DOI: 10.1109/iCCECE49705.2020.9200437

[13]. Chen, Z., & Zhang, J. (2012). Gender classification using decision tree and gradient boosting based on acoustic features. In 2021 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP) (pp. 582-586). DOI: 10.1109/ICICSP51513.2021.9445975

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 8, Issue 2, 2019

[14]. Pandey, A., Sharma, S. K., & Kumar, N. (2015). Gender identification using decision tree based on mel-frequency cepstral coefficients. In 2015 International Conference on Control, Power, Communication and Computing Technologies (ICCPCCT) (pp. 340-344). DOI: 10.1109/ICCPCCT52396.2015.9372997

[15]. Subbulakshmi, S. V., & Dhanalakshmi, R. (2018). Gender identification using decision tree and ensemble learning based on acoustic features. In 2018 International Conference on Communication Information and Computing Technology (ICCICT) (pp. 1-5). DOI: 10.1109/ICCICT53213.2018.9758303