**Mini Project Report on**



**VOICE BASED GENDER IDENTIFICATION**



**Submitted in partial fulfilment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Neha Kumari** **University Roll No. 2016870**

***Under the Mentorship of***

**Mr. Ankit Gupta**

**(Assnt. Professor CSE)**



**Department of Computer Science and Engineering**

**Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand**

**January 2023**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Voice Based Gender Identification”** in partial fulfilment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Mr. Ankit Gupta**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Name: Neha Kumari University Roll no: 2016870

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction | 2-3 |
| Chapter 2 | Literature Survey | 4-5 |
| Chapter 3 | Methodology | 6-9 |
| Chapter 4 | Result and Discussion | 10-12 |
| Chapter 5 | Conclusion and Future Work | 13-15 |
|  | References | 16 |

**Chapter 1**

* **Problem Statement:** Project on Voice Based Gender Identification.
* **Objective:**

**-** Classifying a person’s gender based upon his/her acoustic

properties of voice.

- Machine learning classification techniques can be used on these type of problem statements.

- The learning used here should be supervised learning as we know the actual output and we can compare the predictions with it.

**Introduction**

Gender voice is considered one of the pivotal parts to be detected from a given voice, a task that involves certain complications.

In order to distinguish gender from a voice signal, a set of technique have been employed to determine relevant features to be utilized for building a model from a training set.

**This model is useful for determining the gender (i.e., male or female) from**

**a voice signals.**

**TECHNIQUES AND TECHNOLOGIES USED:**

* Data science libraries such as pandas, matplotlib, seaborn, numpy etc.
* Machine learning library sklearn used with many subordinate models such as model\_selection, classifiers, and metrics etc.
* Jupyter notebook used for implementation.
* Data set available on Kaggle.

**HOW THE DATASET LOOK LIKE?**

**Dataset Name** - Gender Recognition by Voice Dataset Kaggle

**Dataset Link** - <https://www.kaggle.com/datasets/primaryobjects/voicegender>

**Dataset Information:**

Gender Recognition by Voice and Speech Analysis

This database was created to identify a voice as male or female, based upon acoustic properties of the voice and speech. The dataset consists of

3,168 recorded voice samples, collected from male and female speakers.

**Features Explanation :**

**meanfreq**: mean frequency (in kHz)

**sd**: standard deviation of frequency

**median**: median frequency (in kHz)

**Q25**: first quantile (in kHz)

**Q75**: third quantile (in kHz)

**IQR**: interquantile range (in kHz)

**skew**: skewness

**kurt**: kurtosis

**sp.ent**: spectral entropy.

**sfm**: spectral flatness.

**mode**: mode frequency.

**centroid**: frequency centroid .

**meanfun**: average of fundamental frequency measured across acoustic signal.

**minfun**: minimum fundamental frequency measured across the acoustic signal.

**maxfun**: maximum fundamental frequency measured across the acoustic signal.

**meandom**: mean of dominant frequency measured across the acoustic signal.

**mindom**: minimum of dominant frequency measured across the acoustic signal.

**maxdom**: maximum of dominant frequency measured across the acoustic signal.

**dfrange**: range of dominant frequency measured across the acoustic signal.

**modindx**: modulation index.

**Label**: male or female.

The data was pre-processed,

it was clean and consisted of no null values/outliers. Machine

models applied were Decision Tree, Random Forest, SVM, and Naive Bayes. The random forest

has the best accuracy for this problem. (97.79 accuracy for test data)

**Chapter 2**

**Literature Survey:**

The voice of human speech is an effective communication method consisting of unique semantic linguistic and paralinguistic features such as gender, age, language, accent, and emotional state.

The sound waves consisting of human voice are unique among all creatures producing sound since every single wave carries a different frequency.

Identifying human gender based on voice has been a challenging task for voice and sound analysts who deploy numerous applications including

1. effective advertising and marketing strategies in customer relationship management (CRM) systems which depend on gender interoperability such as the user interface style as well as preferences of words and colours;
2. investigating criminal voice in crime scenarios; and
3. enhancing human-computer interaction (HCI) systems especially dialogue systems by customizing services that rely on gender voice and also improving the level of user satisfaction.

Because of the importance of identifying gender through voice recognition, the human voice should be converted from the analogue to the digital form to extract useful features and then to construct classification models.

The robustness and effectiveness of classifiers are determined by the quality of features that depend on a training set employing machine learning (ML) techniques.

Therefore, eliciting voice features plays a vital role in improving the efficiency of classifiers since the human voice is liable for nonuseful features.

Research on improving the efficiency of voice classifiers is copious, particularly studying the process of extracting efficient features from voice including identifying the linguistic content of speech signal components and disposing of nonuseful contents such as background noise.

**Chapter 3**

**METHODOLOGY**

**Classification Model**

* **Decision Tree Model**

It is one the classification techniques used in data mining and machine learning where decision trees are constructed via algorithmic approach that identifies ways to split a data set based on different conditions.

• Decision tree algorithm falls under the category of supervised learning. They can be used to solve both regression and classification problems.

• Decision tree uses the tree representation to solve the problem in which each leaf node corresponds to a class label and attributes are represented on the internal node of the tree.

• We can represent any Boolean function on discrete attributes using the decision tree.

* **Random Forest Model**

Random Forest is an ensemble technique capable of performing both regression and classification tasks.

It combines hundreds or thousands of decisions trees, trains each one on a slightly different set of the observations, splitting nodes in each tree considering a limited number of features.

The final predictions of the random forest are made by averaging the predictions of each individual tree.

The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees.

Random Forest has multiple decision trees as base learning models.

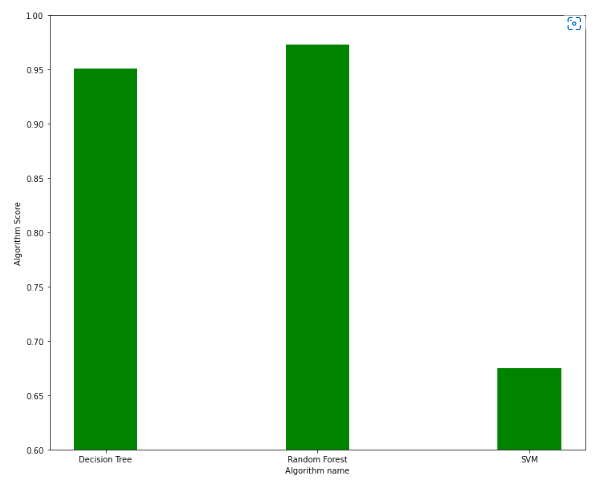
* **Support Vector Machine (SVM)**

“Support Vector Machine” (SVM) is a supervised [machine learning algorithm](https://courses.analyticsvidhya.com/courses/introduction-to-data-science-2?utm_source=blog&utm_medium=understandingsupportvectormachinearticle) that can be used for both classification or regression challenges. However, it is mostly used in classification problems.

In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate.

Then, we perform classification by finding the hyper-plane that differentiates the two classes very well.

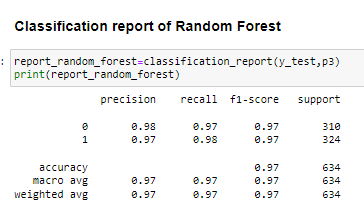
* **CLASSIFICATION SCORE BAR CHART–**

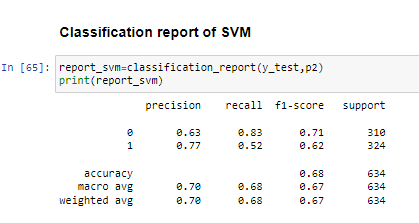


* **CLASSIFICATION REPORT OF ALL MODELS –**

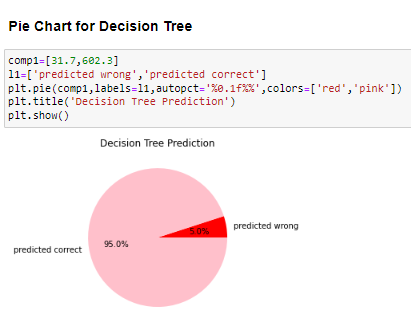
Table

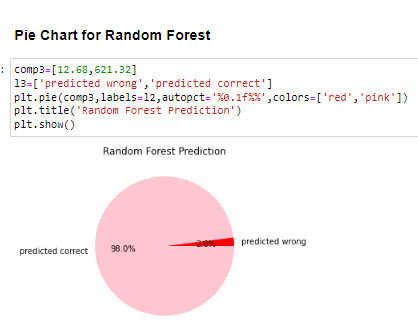
Description automatically generated

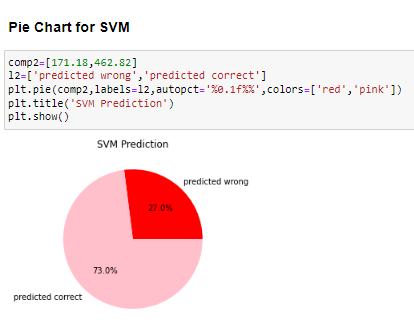




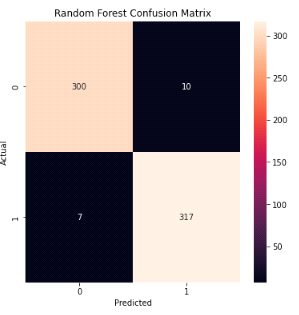
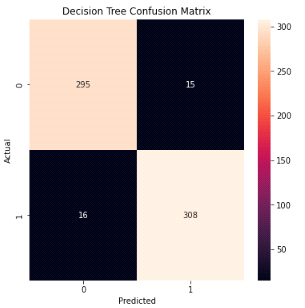
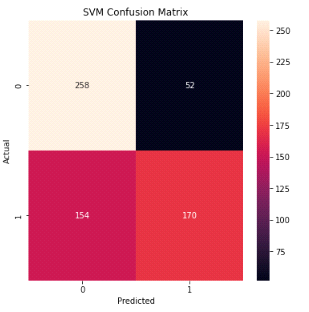
* **PIE CHART FOR MODEL PREDICTIONS–**



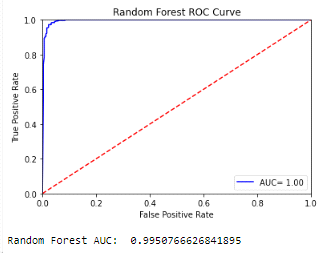


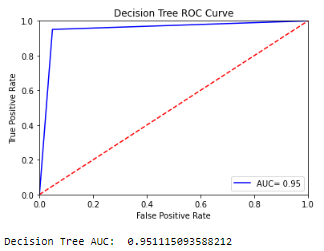


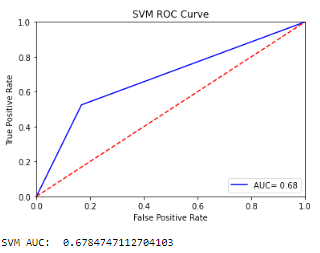
**CONFUSION MATRICS FOR ALL MODELS–**



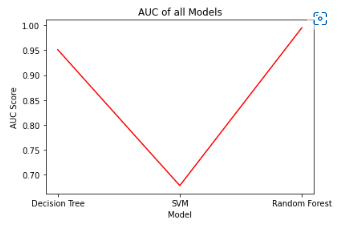
* **ROC CURVE FOR ALL MODELS–**







* **AUC LINE GRAPH FOR COMPARISION–**



**ACCURACY-**

* Accuracy for decision tree model is: - 95%
* Accuracy for Support vector machine model is: - 73%
* Accuracy for Random Forest model is: - 98%

**Chapter 4**

**Result and Discussion**

* In the project, it is clear that our model works well on recognizing whether a

person is male or female with a high accuracy of about 98% using Random

Forest model.

* Using other model such as Decision tree and Support vector machine we got

accuracy of 95% and 73% respectively.

* Based on precision, F1-Score, Recall Values, prediction pie chart, Accuracy

Score, confusion matrix, ROC curve and AUC line graph, we can say that

**Random Forest Classification is best suited** for this problem where we are

Predicting the person as male/female based upon they voice signals. **Support**

**Vector machine is least suitable** based on same analysis.

**Chapter 5**

**CONCLUSION**

* Random Forest model itself is a super accurate model, and hence has more potentials than the other model in the voice-based gender classification, and possibly in other classification applications.
* For text dependent type of classification, the SVM could be the best choice.
* For text independent type of classification, the Random Forest is one of the choices.

**FUTURE WORK**

* Investigate the reasons why such a super SVM model can’t perform well for the text independent gender classification.
* Explore the possible voice features which might improve the SVM text independent classification performance.
* It could be meaningful to compare SVM performance with other classification model.
* Examine SVM model for other voice based classification applications, such as age and spoken language.

**References**

[1] [**www.kaggle.com**](http://www.kaggle.com/)

[2] [**https://towardsdatascience.com**](https://towardsdatascience.com/)

[3][**https://stackoverflow.com**](https://stackoverflow.com/)