Human Gender and Emotion detection from live voice recordings using AI

Shubhamkumar Parekh, Adwait Avasekar, B Anil kumar, Vikas Bade, Prof. Prashant Karkalle
Smt. Kashibai Navle Sinhgad Institute of Technology and Science, Lonavala

Abstract: - Gender identification is considered to be one of the major problems in the field of signal processing. Formerly, this problem has been solved using various image classification techniques which typically includes information extraction from a set of images. However, gender classification using vocal features has recently been a topic of interest to a lot of researchers across the globe. A close scrutiny of some of the human vocal features reveals that classifying gender goes way beyond just the frequency and the pitch of a person. One of the most challenging problems faced in machine learning is feature selection or as is technically known as dimensionality reduction. A similar problem is faced while deciding gender-specific traits-which serve a significant purpose in classifying the gender of a person. This paper will inspect the efficiency and significance of machine learning algorithms to the voice-based gender identification problem. This voice based gender detection can be used for classifying user and displaying related products for online shopping.

Keywords: SDLC, SVM, KNN, Vocal Feature

1. Introduction

Machine learning has been a recent trend and a course of study. Institutions and business giants all over the world, from small-scale to large-scale, have started to shift gears and invest more in such techniques. It basically involves knowledge mining using various statistical learning approaches. Applying these methodologies in prediction and classification is seemingly interesting but can be gruesome in certain cases where the data is insufficient to elicit meaningful insights. Also, evaluating the impact of each feature on the final model is an involute task. It is quite easy for a human brain to differentiate between various voices. However, that similar task can be complex for a computer. Gender identification is one such interesting problem, the results of which can be found using standard machine learning techniques. Vogt and Andrè suggested that determining gender from a speech in turn help improve automatic emotion recognition from speech. This experimental study tries to determine the most performant models using our dataset.

1.1 Problem definition:

Gender identification is considered to be one of the major problems in the field of signal processing. Formerly, this problem has been solved using various image classification techniques which typically includes information extraction from a set of images but voice-based gender classification.

1.2 Objective:

Gender classification using vocal features has recently been a topic of interest to a lot of researchers across the globe. A close scrutiny of some of the human vocal features reveals that classifying gender goes way beyond just the frequency and the pitch of a person. One of the most challenging problems faced in machine learning is feature selection or as is technically known as dimensionality reduction. A similar problem is faced while deciding gender-specific traits-which serve a significant purpose in classifying the gender of a person.

1.3 Technical Approach:

In this project features data set is used with 110 features and 1 label (male and female) this dataset is pre-processed and trained and tested with machine learning algorithm (SVM and Neural Network) then model is saved to system. Using this saved model when user given voice input prediction is performed with stored model and male or female result is returned.

This result is used in online shopping website to help user to show only related products in website.

1.4 Application:

Male and female classification is useful in online shopping website for product recommendation.

Male and female classification can be used for call allocation inside call centres.

2. LITERATURE SURVEY:

Since there are only two categories of response values in the dataset, the problem is narrowed down to binary classification. Any general classification algorithm such as Logistic Regression, Support Vector Machine, Nearest Neighbours, and Discriminant Analysis etc. can be applied on the data. These techniques are by far, the most commonly used machine learning algorithms.

Kuynu Chen, in his work, shows that the discriminant analysis classifier gives the most interesting results in terms of test error rate and precision. However, even this model still suffers from a test error rate of greater than 10%. Albeit, Chen states that running backward selection on the feature set can minimize the test error rate. They extracted the audio features, 24 in total, from Yaafe their dataset consists of 12004 data points, amongst which 6286 are labeled ‘female’ and 5718 are labeled ‘male’.

Becker’s study on a dataset similar to ours, shows that in order to gain a deeper understanding of the model and to determine the exact properties that indicate a gender of a person, a classification and regression tree (CART) should be applied. His results indicate that the mode frequency (mode) serves as a root node for detecting the gender of a person. Traversing further down the tree evinces that minimum fundamental frequency, maximum dominant frequency, first quantile hertz, skewness, median frequency, additionally correspond to gender classification. The CART model results in an accuracy of 81% on the training set and 78% on the test set. He further takes CART, a step further and applies Random Forest to the data. This achieves a positive boost over CART with an accuracy of 100% on the training set and 87% on the test set.

Training machine learning models is not much of a challenge when working with Python’s libraries. However, a rather difficult task is to achieve a trade-off between bias and variance. Over-fitting is a common problem which occurs when the model function is too complicated to be able to predict accurate values. This implies that the variance of the model is too high and that it is trying to fit every point inside the resultant curve, eventually performing poorly on unseen data. One typical method to evaluate the performance measure of a model is to analyse the training as well as the test error rate. A good model not only has a good training accuracy but also a good test accuracy, and if both the accuracy scores are closer to each other, in terms of their values, then it is likely that the model is not over-fitting the data.

3. SOFTWARE MODEL OR ARCHITECTURE ANALYSIS:

Structured project management techniques (such as an SDLC) enhance management's control over projects by dividing complex tasks into manageable sections. A software life cycle model is either a descriptive or prescriptive characterization of how software is or should be developed. But none of the SDLC models discuss the key issues like Change management, Incident management and Release management processes within the SDLC process, but, it is addressed in the overall project management. In the proposed hypothetical model, the concept of user-developer interaction in the conventional SDLC model has been converted into a three dimensional model which comprises of the user, owner and the developer. In the proposed hypothetical model, the concept of user-developer interaction in the conventional SDLC model has been converted into a three dimensional model which comprises of the user, owner and the developer. The ―one size fits all‖ approach to applying SDLC methodologies is no longer appropriate. We have made an attempt to address the above mentioned defects by using a new hypothetical model for SDLC described elsewhere. The drawback of addressing these management processes under the overall project management is missing of key technical issues pertaining to software development process that is, these issues are talked in the project management at the surface level but not at the ground level.

3.1 Functional requirements

Outputs from computer systems are required primarily to communicate the results of processing to users. They are also used to provide a permanent copy of the results for later consultation. The various types of outputs in general are:

- External Outputs, whose destination is outside the organization,.
- Internal Outputs whose destination is within organization and they are the
- User's main interface with the computer.
- Operational outputs whose use is purely within the computer department.
- Interface outputs, which involve the user in communicating directly.
- Understanding user's preferences, expertise level and his business requirements through a friendly questionnaire.
- Input data can be in four different forms - Relational DB, text files, .xls and xml files. For testing and demo you can choose data from any domain. User-B can provide business data as input.

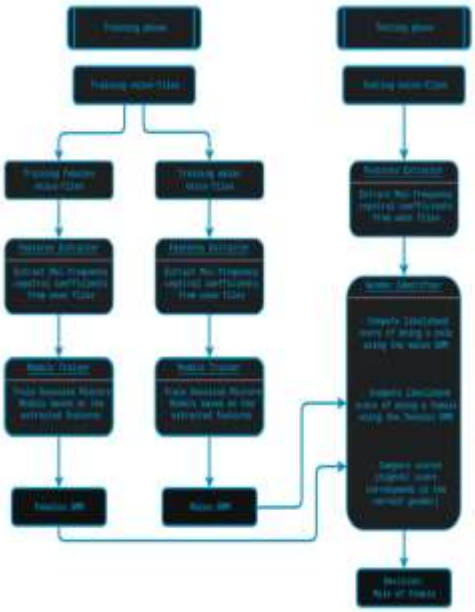
3.2 Non-Functional Requirements

1. Secure access of confidential data (user's details). SSL can be used.
2. 24 X 7 availability.
3. Better component design to get better performance at peak time

4. Flexible service based architecture will be highly desirable for future extension

4 .SYSTEM DESIGN:

4.1 SYSTEM ARCHITECTURE



At the top of the system style all the main knowledge structures, file formats, output formats, and also the major modules within the system and their specifications square measure set. System design is that the method or art of process the design, components, modules, interfaces, and knowledge for a system to satisfy such as needs. Users will read it because the application of systems theory to development.

The purpose of the design phase is to arrange an answer of the matter such as by the necessity document. This part is that the opening moves in moving the matter domain to the answer domain. The design phase satisfies the requirements of the system. The design of a system is probably the foremost crucial issue warm heartedness the standard of the software package. It's a serious impact on the later part, notably testing and maintenance.

The output of this part is that the style of the document. This document is analogous to a blueprint of answer and is employed later throughout implementation, testing and maintenance. The design activity is commonly divided into 2 separate phases System Design and Detailed Design.

System Design conjointly referred to as top-ranking style aims to spot the modules that ought to be within the system, the specifications of those modules, and the way them move with one another to supply the specified results.

5. CONCLUSION

In conclusion, most of the above models have performed well but in some cases, interpretability outweighs inference. In order to understand gender-specific characteristics, it is important to eliminate all the insignificant features from the model. By the end of the experimental study, it can be concluded that a great level of accuracy can be achieved by selecting some specific features. This will reduce the overall model training time, model complexity and also increase inference simplicity.

6. FUTUREWORK

One of the obstructions in classification using audio cues is that the audio samples are usually obtained from noisy environments. Such environmental or artificially generated noise detrimentally limits the accuracy of classification. Different and more efficient ways to eliminate noise can be found out and this becomes a course for future research

7. REFERENCES:

[1] Udry J. R., "The nature of gender," Demography, vol. 31, pp. 561-573, 1994.

[2] Demirkus M., Garg K., and Guler S., "Automated person categorization for video surveillance using soft biometrics," in Biometric Technology for Human Identification VII, 2010, p. 76670P.

[3] Jain A. K., Ross A., and Prabhakar S., "An introduction to biometric recognition," IEEE Transactions on circuits and systems for video technology, vol. 14, pp. 4-20, 2004.

[4] Paluchamy M., Suganya D., and Ellammal S., "Human gait based gender classification using various transformation techniques," IJRCCT, vol. 2, pp. 1315-1321, 2013.

[5] Gnanasivam P. and Muttan S., "Gender classification using ear biometrics," in Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012), 2013, pp. 137-148.