Chapter 15
Classification of Indian Native English Accents

A. Aadhitya
Anna University, Chennai, India

K. N. Balasubramanian
Anna University, Chennai, India

J. Dhalia Sweetlin
https://orcid.org/0000-0002-3718-8006
Anna University, Chennai, India

ABSTRACT

The accent spoken by the people is generally influenced by their native mother tongue language. People located at various geographical locations speak by adding flavors to their native language. Various Indian native English accents are classified to bring out a classic difference between these accents. To bring a solution to this problem, a comparative classification model has been built to classify the accents of five distinct native Indian languages such as Tamil, Malayalam, Odia, Telugu, and Bangla from English accents. Firstly, the features of the five-second audio samples each from different accents are obtained and converted to images. The consolidated attributes are gathered. The VGG16 pre-trained model is fused with support vector model to classify accents accurately. Secondly, along with these features, mel frequency cepstral coefficient is added and trained. Then, the features obtained from VGG16 were reduced using principal component analysis. Highest accuracy obtained was 98.46%. Further analysis could be made to produce automated speech recognition for various aspects.

1. INTRODUCTION

Understanding accents has been a major issue in recent days, such that the Human-Machine interaction can be built to do the same. Accents are the speech patterns or pronunciations that are found in different languages. A person's pronunciation of words or usage of rhythm, intonation, and stress in speech are all

examples of what is generally referred to as their accent. The same kind of accents are identified among people of an identical national background or ethnic group (Tarun et al., 2022). Accent classification is a multidisciplinary field that involves the identification and categorization of accents based on various phonetic, linguistic, and sociolinguistic features. Accents can be classified along a dialect continuum, which is a range of dialects spoken across a large area or based on unique speech patterns. By listening to a speech of the person one could get to know more about their origin. However, humans can not accurately categorize the accents for the first time. Speech recognition (Mridha et al., 2022) is one of the important issues for many investigative agents because of the accented speech. There would be different kinds of pronunciations within the same language that cannot be recognized with ease. The way a person speaks, including their accent, is influenced by a complex interplay between various factors, including biological, environmental, and cultural factors. Certain methods can be done to modify or adjust the accent of a person such as speech therapy, mimicking native language speakers, immersion in language or dialect by living in a community, training in phonetics of the language and exposure to media.

The Speech Recognition System (Caballero et al., 2006) comes into play to detect the accent of the speaker when the accent spoken by the other is different from others. The accurate identification of the accent is very important and such a model that clearly distinguishes between each of the Indian accents is required. This system can be used by the telephone or call centres to forward the call to the respective support person based on the language and the accent spoken. English is the global language which is spoken by most people around the globe. On the other hand, many people find it difficult to understand the English spoken by the opposite person because of the accent that they are using. Such that to reduce the work of translation a model is developed to perform accent recognition of five different native languages.

The Python package librosa (Suman et al., 2022; McFee et al., 2015) is used to visualize the audio samples and the matplotlib package is used to convert the audio samples to wave plots. The wave plots have been constructed using the amplitude and the time feature of each frame respectively. Such that they could be transformed into features after passing through VGG16 layers. This is one of the pre-trained models such that this VGG16 model has been used to produce the features.

The features retrieved from the model were reduced further using the method of principal component analysis (Bodine & Hochbaum, 2022) to decrease the curse of dimensionality. The most widely used feature, Mel Frequency Cepstral Coefficient (MFCC) is a feature extractor which is used to provide the best representation of the frequency components in the form of waves. The machine learning algorithm Support vector Machine (SVM) is used to classify the features belonging to the different classes. This is best suited to regression and classification-type problems. The kernels present in SVM are used to work with large dimensional features. As the speech features might be very huge, such that this model can be very effective in this case.

2. RELATED WORK

Deshpande et al (2005) focused on developing a GMM classifier model to categorize between standard Indian accent and American accent. The formatted frequencies from the respective accents were used and achieved a higher accuracy. However, the GMM classifier model has disadvantages such as being prone to overfitting, no support for high-dimensional data, being sensitive towards initial model conditions, expensive to train. Upadhyay & Lui (2018) have developed a model to classify different non-native English accents spoken by people from six different countries. They have used MFCC to extract the

Related Content

A Layered Model for Building Ontology Translation Systems
Oscar Corchoand Asunción Gomez-Perez (2005). *International Journal on Semantic Web and Information Systems (pp. 22-48).*
www.irma-international.org/article/layered-model-building-ontology-translation/2807

Enabling Scalable Semantic Reasoning for Mobile Services
Luke Albert Steller, Shonali Krishnaswamyand Mohamed Methat Gaber (2009). *International Journal on Semantic Web and Information Systems (pp. 91-116).*
www.irma-international.org/article/enabling-scalable-semantic-reasoning-mobile/4115

Machine Learning-Based Distributed Denial of Services (DDoS) Attack Detection in Intelligent Information Systems
Wadee Alhalabi, Akshat Gaurav, Varsha Arya, Ikhlas Fuad Zamzamiand Rania Anwar Aboalela (2023). *International Journal on Semantic Web and Information Systems (pp. 1-17).*
www.irma-international.org/article/machine-learning-based-distributed-denial-of-services-ddos-attack-detection-in-intelligent-information-systems/327280

Knowledge Management
Salman Iqbal, Hayati Abdul Jalal, Paul Toulsonand David Tweed (2012). *Collaboration and the Semantic Web: Social Networks, Knowledge Networks, and Knowledge Resources (pp. 64-81).*
www.irma-international.org/chapter/knowledge-management/65687

Tightly Coupled Fuzzy Description Logic Programs under the Answer Set Semantics for the Semantic Web
Thomas Lukasiewiczand Umberto Straccia (2008). *International Journal on Semantic Web and Information Systems (pp. 68-89).*
www.irma-international.org/article/tightly-coupled-fuzzy-description-logic/2853