

# Classification and Recognition of Speech Emotion Recognition using ML Algo

Rishu Kumar<sup>1</sup>, Sandeep Singh<sup>2</sup>, Tejas Goje<sup>3</sup> Prof. V.P. Niwane<sup>4</sup>

E&TC Engineering Smt Kashibai Navale College of Engineering Vadgaon BK, Pune, India<sup>1,2,3,4</sup>

Guide, E&TC Engineering Dept, Smt. Kashibai Navale College of Engineering Vadgaon BK, Pune, India<sup>5</sup>

**Abstract:** *If we want to achieve an emotional-related response from some algorithm or other intelligent machines, the initial step is to fetch precise emotion recognition. This project deals with the implementation of the deep learning model of Convolutional Neural Networks (CNN). The architecture which is primarily based on an image processing CNN was developed in Python using Keras API which is based on the TensorFlow platform. The basic methods that lay the foundation for the classification of emotion recognition based on certain voice parameters are briefly described. As per obtained results, the model tries to obtain the average precision of 79.33% for five emotions namely (happy, fear, sad, neutral, and anger), which is comparable with performances reported in scientific literature.*

## I. INTRODUCTION

In today's digital landscape, speech has become a primary mode of communication between humans and computers which has been possible by several technological upgrades. Speech recognition techniques along with signal processing made exponential progress in Speech-to-Text (STT) technology which is used nowadays in most mobile phones. Speech Recognition is the fastest growing research domain in which attempts are being made to recognize and decipher speech signals. This in turn leads to Speech Emotion Recognition (SER) growing research topic in which several advancements can lead to progress in numerous fields such as automatic translation systems, machine to human interface, used in synthesizing speech from text and vice versa and so on. Contrary to that this project's focus is to survey and review various speech extraction features, emotional speech databases, classifier algorithms and so on. Problems present in various topics have been addressed.

Speech Recognition is the terminology that deals with various algorithms and computing processes to recognize the speech from the speech signals. Several technological strides in the field of AI and signal processing, and recognition of emotion made it easier and possible.

## II. MOTIVATION

Speech emotions basically refer to extracting the emotional factors of the speaker from their speech. Its approach is to study the speech of signals to detect suitable emotions based on their characteristics like tone, pitch, etc. To extract these nuances and test the speech signal, a decent number of algorithms have been designed with upgradation in the latest technologies.

CNN is one of the most used deep learning models that have resulted in massive success in research areas like 15 object recognition, face detection, and natural language pre-processing. Usually, CNN has 3 fundamental building blocks, convolutional layer, pooling layer, and fully connected layer. Thus, we depict these building blocks with some rudimentary concepts like softmax unit, rectified linear unit, and dropout.

## III. SYSTEM REQUIREMENT

For methodical use, every system needs configured hardware components along with respective drivers and dependency software systems. These prerequisites and requirements are popularly termed System requirements, these proposed requirements act as a regulative structure for software. In many cases, two different types of requirements are mentioned which are: Recommended and Minimal. Because of advancement in technology and industry norms these technical specifications continuous changes and increase over a period of time. We can also define system requirements as platform

specifications which must be met in order to run the given software smoothly without any technical obstacles and difficulties.

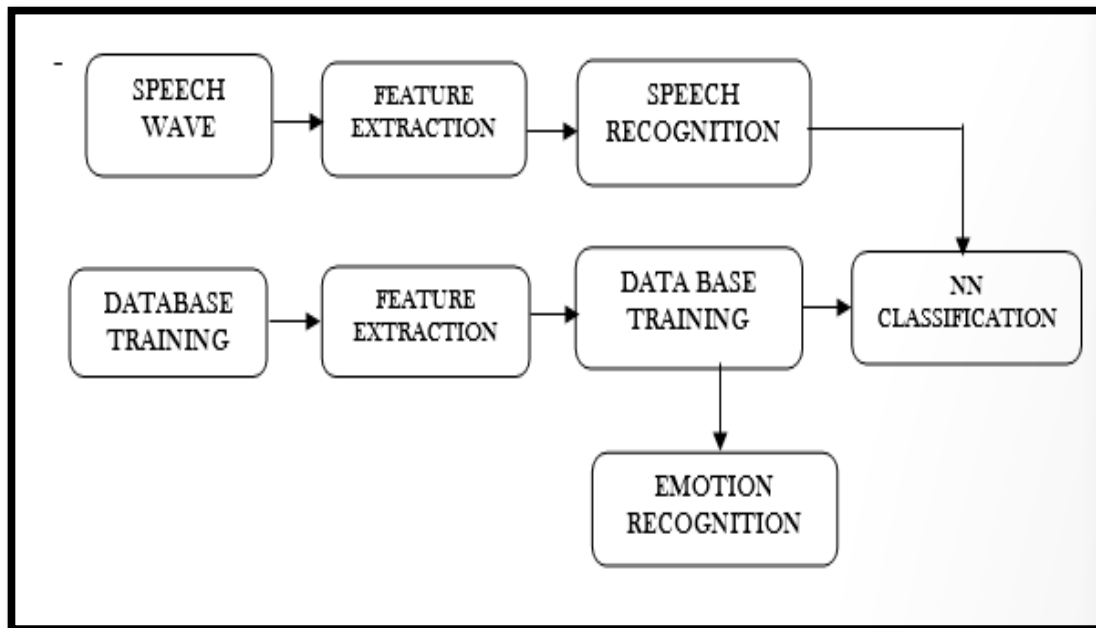
#### IV. SOFTWARE REQUIREMENTS

In accordance with the process of software engineering software requirements are specifications which are required to run programs on a system or dependencies which must be satisfied in order to execute software successfully. The summary of such requirements is as follows:

1. A prerequisite or potential that is preoccupied with software or its component to meet a standard specification or any other related criteria.
2. A prerequisite or a potential that is preoccupied with the system to complete the given task and finalize the outcome of intended objectives.
3. A well-documented depiction of a prerequisite or potential as in 1 or 2.

Software requirements can be assembled as follows:

Platform	Windows 8+ / Linux 16.04+
Drivers	Display updated drivers
Dependencies	Python 3.7, Django, Html CSS, MySQL
Libraries	NumPy, pandas, scikit-learn, Google, ML kit
ML Models	Multiple Linear Regression, k-nearest neighbors



**Figure 1:** Block Diagram of System

Although activity apprehension from accent may be an analogously new acreage of analysis, it's several abeyant applications. In human-computer or human-human alternation systems, activity acceptance systems ability accords users with bigger casework by getting adjustive to their emotions.

In virtual worlds, feeling recognition might facilitate simulating additional realistic avatar interaction. The body of labor on police work feeling in speech is sort of restricted. Currently, the researchers' area unit still debating what options influence the popularity of feeling in speech.

#### **IV. CONCLUSION AND FUTURE WORK**

Machine learning has made great progress so far, but in the field of speech signal processing, especially for building the SER system, there has not been much progress, SER is still a challenging problem. This paper proposes a new data enhancement method for SER and uses the proposed model that consists of the CNN, the LSTM, and the Attention Mechanism to classify speech emotions without using any traditional hand-crafted features.

#### **REFERENCES**

- [1] L. Kerkeni, Y. Serrestou, M. Mbarki, K. Raoof and M. A. Mahjoub, "A review on speech emotion recognition: Case of pedagogical interaction in classroom," 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Fez, 2017, pp. 1-7, DOI: 10.1109/ATSIP.2017.8075575.
- [2] Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3-4), 169–200. <https://doi.org/10.1080/02699939208411068>
- [3] Matilda's. Emotion recognition: A survey. *International Journal of Advanced Computer Research*. 2015;3(1):14-19.
- [4] Koolagudi SG, Rao KS. Emotion recognition from speech: A review. *International Journal of Speech Technology*. 2012;15(2):99-111 Ali H, Hariharan M, Yaacob S, Adom AH. Facial emotion recognition using empirical mode decomposition. *Expert Systems with Applications*. 2015.
- [5] Liu ZT, Wu M, Cao WH, Mao JW, Xu JP, Tan GZ. Speech emotion recognition is based on feature selection and extreme learning machine decision trees. *Neurocomputing*. 2018.