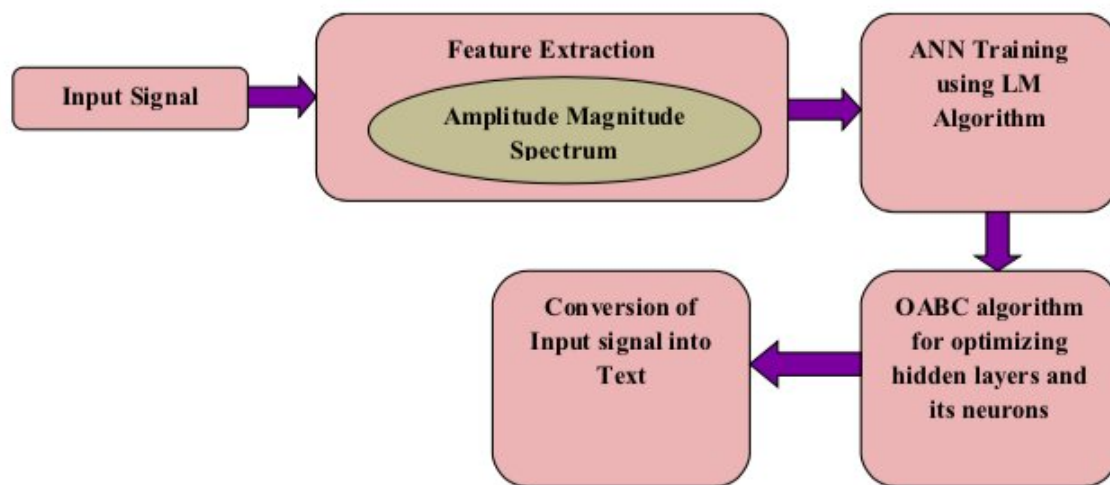


# Speech recognition using artificial neural networks and artificial bee colony optimization

October 7 2019, by Ingrid Fadelli



Block diagram of proposed model. Credit: Shukla & Jain.

Over the past decade or so, advances in machine learning have paved the way for the development of increasingly advanced speech recognition tools. By analyzing audio files of human speech, these tools can learn to identify words and phrases in different languages, converting them into a machine-readable format.

While several [machine learning](#)-based models have achieved promising results on speech [recognition](#) tasks, they do not always perform well in all languages. For instance, when a language has a vocabulary with many similar-sounding words, the performance of speech recognition systems can decline considerably.

Researchers at Mahatma Gandhi Mission's College of Engineering & Technology and Jaypee Institute of Information Technology, in India, have developed a [speech recognition system](#) to tackle this problem. This new system, presented in a paper published in Springer Link's *International Journal of Speech Technology*, combines an [artificial neural network](#) (ANN) with an optimization technique known as opposition artificial bee colony (OABC).

"In this work, the default structure of ANNs is redesigned using the Levenberg-Marquardt algorithm to retrieve an optimal prediction rate with accuracy," the researchers wrote in their paper. "The hidden layers and neurons of the hidden layers are further optimized using the opposition artificial bee colony optimization technique."

A unique characteristic of the system developed by the researchers is that it uses an OABC optimization algorithm to optimize the ANN's layers and artificial neurons. As the name would suggest, artificial bee colony (ABC) algorithms are designed to simulate the behavior of honey bees to tackle a variety of optimization problems.

"Generally, optimization algorithms randomly initialize the solutions in the matching domain," the researchers explained in their paper. "But this solution could lie in the opposite direction of the best solution, thereby increasing the computational overhead significantly. Hence this opposition-based initialization is termed as OABC."

The system developed by the researchers considers individual words

spoken by different people as an input speech signal. Subsequently, it extracts so-called [amplitude modulation](#) (AM) spectrogram features, which are essentially sound-specific characteristics.

The features extracted by the model are then used to train the ANN to recognize human speech. After it is trained on a large database of audio files, the ANN learns to predict isolated words in new samples of human speech.

The researchers tested their system on a series of human speech audio clips and compared it with more conventional speech recognition techniques. Their technique outperformed all the other methods, attaining remarkable accuracy scores.

"The sensitivity, specificity, and accuracy of the proposed method are 90.41 percent, 99.66 percent and 99.36 percent, respectively, which is better than all the existing methods," the researchers wrote in their paper.

In the future, the [speech](#) recognition system could be used to achieve more effective human-machine communication in a variety of settings. In addition, the approach they used to develop the system could inspire other teams to design similar models, which combine ANNs and OABC [optimization](#) techniques.

**More information:** Shilpi Shukla et al. A novel system for effective speech recognition based on artificial neural network and opposition artificial bee colony algorithm, *International Journal of Speech Technology* (2019). [DOI: 10.1007/s10772-019-09639-0](https://doi.org/10.1007/s10772-019-09639-0)

Citation: Speech recognition using artificial neural networks and artificial bee colony optimization (2019, October 7) retrieved 10 April 2024 from <https://techxplore.com/news/2019-10-speech-recognition-artificial-neural-networks.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.