Digit Recognition using Spectrogram in JDSP

NSF IUSE Workshop 2021

Uday Shankar Shanthamallu

Introduction

- Machine learning (ML) is EVERYWHERE !
- Emerging applications in speech and computer vision necessitate its introduction at the undergraduate level DSP classes
- Growing interest and demand in both academia and industry

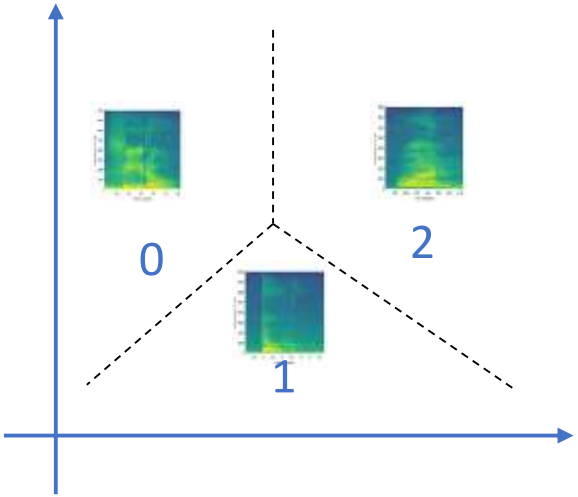
https://www.geeksforgeeks.org/machine-learning/

ML in Undergraduate DSP classes

- We introduce ML concepts in DSP class through learning modules

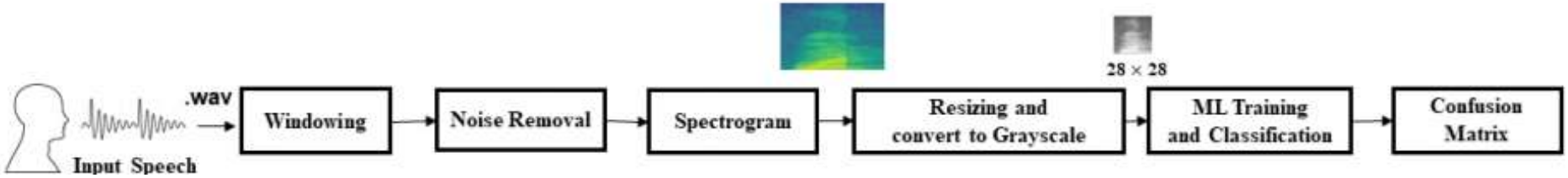
- The goal of the modules and exercise is to task students with performing spoken digit classification using ML algorithms and spectrograms

- Spectral representations which are covered in our DSP course are used to extract features



Overview of the Exercise

- Preprocessing and extraction of spectrogram from speech
- Feature extraction from resized spectrogram
- Training using ML algorithm such as Neural Networks
- Performance Evaluation: Accuracy and Confusion Matrix



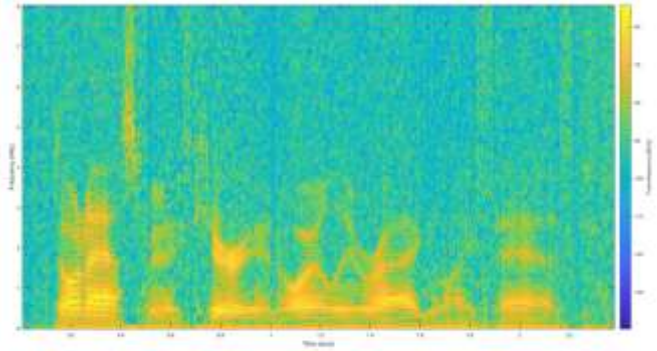
Learning Modules

- We provided students with lectures and notes on speech processing & spectral representations
- Online module developed for students covers the following topics:

- Special efforts were taken to teach students in class about different types of performance analysis such as accuracy measurements and confusion matrix

Speech Spectrogram and Feature Extraction

- We teach students the properties of speech (non-stationarity, formants, pitch, voice/unvoiced frames, AR models)
- Spectral characteristics are computed using the FFT on a frame-by-frame basis. Students are also exposed to Linear prediction
- Students learn about the and how the time-frequency plot for speech signal analysis.



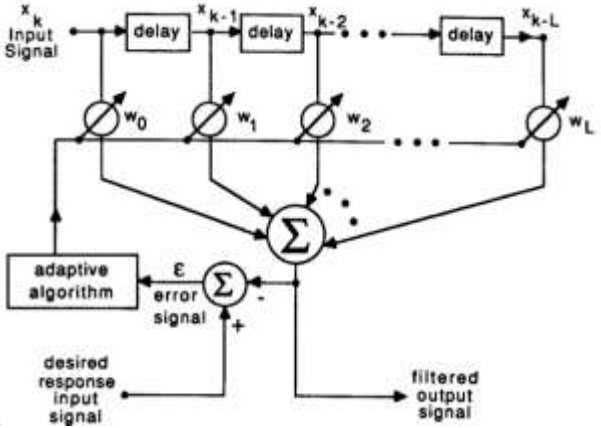
Creation of Training and Testing Dataset

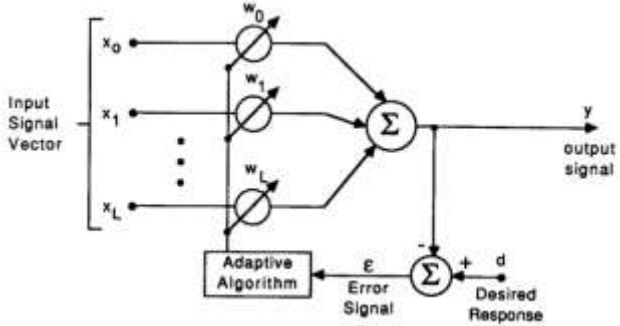
- Five spoken digits (0,1,2,3 and 4) are used in the exercise. Students evaluate first the spectrograms.
- To reduced training complexity and for simplicity, RGB converted to 28 x 28 grayscale image reshaped to 784 dimensional vector.
- 10000 such vectors created for the entire dataset (10000x784)

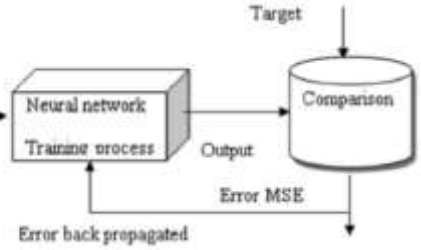
Perceptron and the Adaptive Filter

- At the heart of the neural network, there is a process similar to an adaptive filtering algorithm which are used to optimize the weights/filter coefficients: FIR adaptive filter

- The structure of a neuron and that of an LMS adaptive filter are very similar with a few differences.



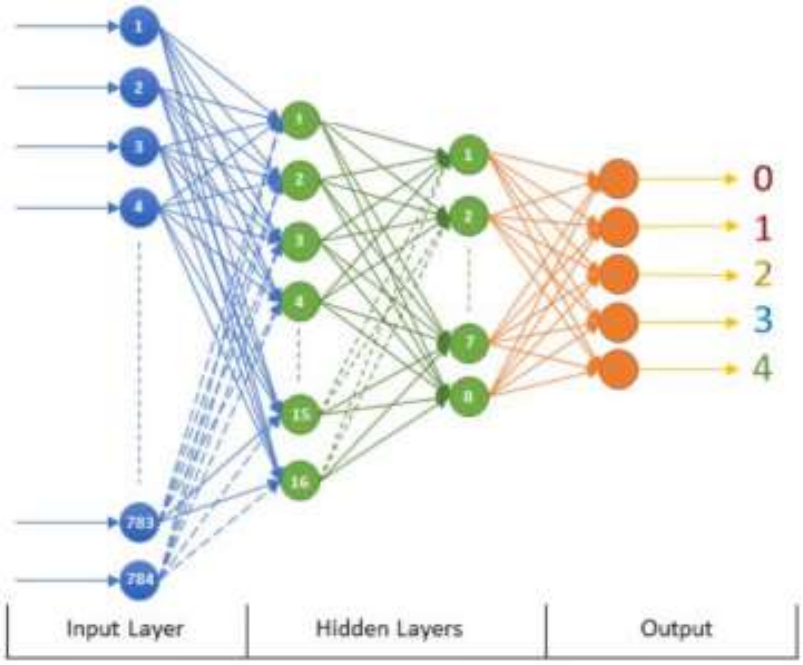




Widrow, Bernard, and Rodney Winter. "Neural Nets for Adaptive Filtering and Adaptive." (1988).

Training and Testing

- 3 layer Multi-Layer Perceptron
- Input: 10000 x 784 image matrix
- Output: Digit predicted (0,1,2,3 or 4).
- 70% - Training , 30% - Testing

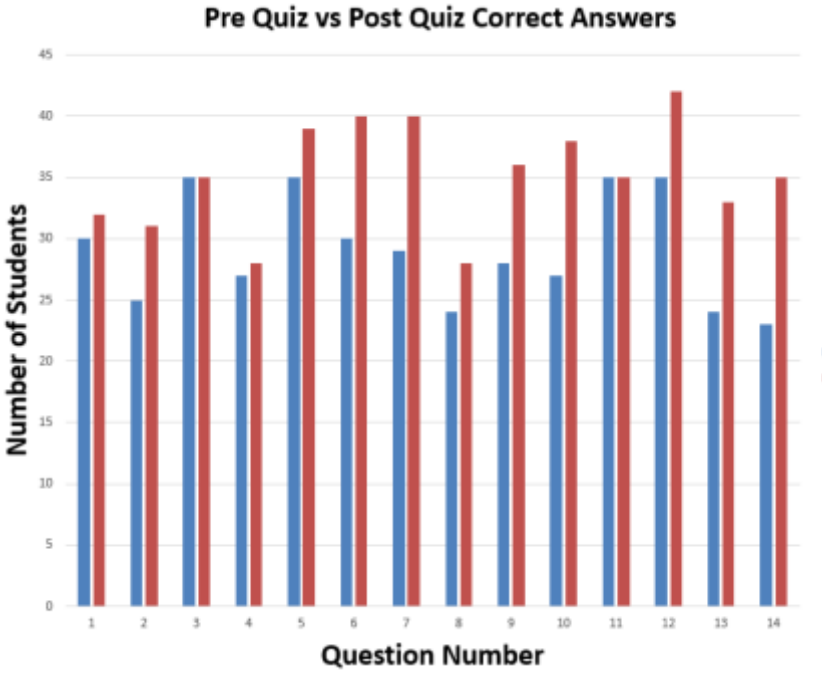


Algorithm Evaluation

- Confusion matrix for the speech data.
  - Students determine the per class accuracy scores and misclassifications
- Recognition Accuracy scores
  - Students determine the overall training and test accuracy
- Experiment with hidden layer size to understand overfitting and underfitting.

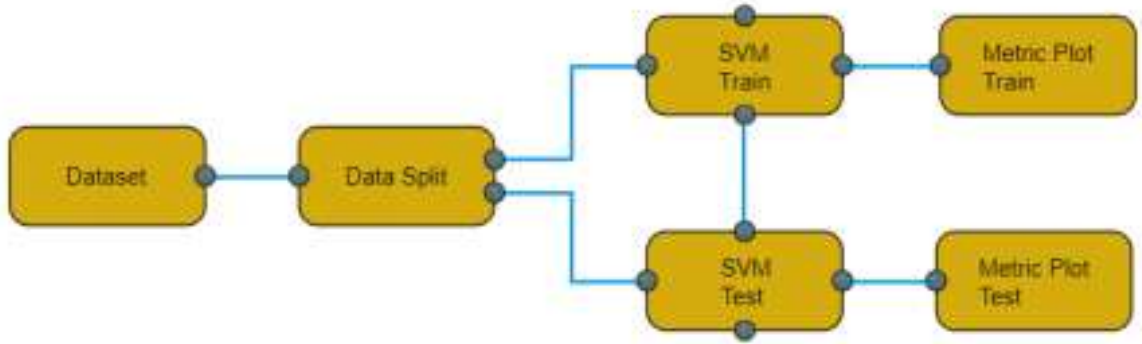
Assessment of Modules and Exercise

- Pre-quiz and Post-quiz evaluation
- The questions in the quiz are related to learning modules and exercise



- 45 students participated
- Students gained knowledge on speech properties
- Students scored better on the post quiz
- Average Performance Improvement in SP Questions (1-6) ~ 14 %
- Average Performance Improvement in ML Questions (7-14) ~ 28 %

Work in progress



- Implementation in HTML 5 Java-DSP
- Simple Interface, Easy to learn and visualize

Summary

- We presented a MATLAB based exercise to introduce Electrical Engineering students taking the DSP class to machine learning
- We illustrate the fundamentals of every block in the digit recognition pipeline
- Students understood pre-processing and feature extraction
- Students were able assess properties of speech from spectrograms
- Students got exposed to neural nets and deep learning
- We describe the teaching pedagogy
- We perform pre and post quiz evaluations on students and observe significant improvement.
- Implementation in JDSP (Future Work)

Acknowledgement

The work at Arizona State University is supported in part by the NSF- DUE award 1525716 and the SenSIP Center.

References

[1] A. Dixit, U. Shanthamallu, A. Spanias, S. Rao, S. Katoch, M. Banavar, G. Muniraju, J. Fan, P. Spanias, A. Strom, C. Pattichis, H. Song, "Multidisciplinary Modules on Sensors and Machine Learning," " Proc. 2018 ASEE Annual Conference, Salt Lake City, June 2018.

[2] A. Spanias, "An introductory signal processing course offered across the curriculum, “IEEE Signal Processing and Signal Processing.

[3] A. Spanias and J. Blain Christen, “A STEM REU Site on The Integrated Design of Sensor Devices and Signal Processing Algorithms',” Proc. IEEE ICASSP 2018, Calgary, April 2018.

[4] A. Dixit, U. S. Shanthamallu, A. Spanias, V. Berisha, and M. Banavar, "Online Machine Learning Experiments in HTML5", IEEE Frontiers In Education (FIE), San Jose, October 3-6, 2018.

[5] M. Stanley and Jong Ming Lee, Sensors for IoT Applications, ISBN 9781627054638, Synthesis Lectures, Morgan and Claypool Publishers, 113 Pages, March 2018.

[6] Diagnostic Quiz, Linear Prediction and Machine Learning, Website: http://jdsp.engineering.asu.edu/MLExercise/MLQuiz.pdf.

[7] U. S. Shanthamallu, A. Spanias, C. Tepedelenlioglu and M. Stanley, "A brief survey of machine learning methods and their sensor and IoT applications," 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA), Larnaca, August 2017.

Thank You