LAB – 5 REPORT

Name: Ozan Berk Gultegin

Submission Date:

# Objective

The objective of this lab is to analyze and implement both linear and nonlinear models for speech signal prediction using artificial intelligence methods. A linear perceptron network is used for linear prediction, and a nonlinear feedforward artificial neural network (ANN) is used for nonlinear prediction. Each model is trained using past samples of a speech signal to predict its future values. The models are evaluated using visual plots and the mean squared error (MSE) metric

# Methods and Implementation

**1. Data Preparation**

* Speech signal was loaded from a .wav file and normalized to [-1, 1].
* A sliding window approach was used for each time step, the previous 20 samples were used to predict the next sample.
* The data was split into 70% training and 30% testing sets.

**2. Linear Prediction:**

* Implemented using PyTorch as a single-layer linear neural network.
* No activation function was used (pure linear transformation).
* Optimized using Adam and MSE loss over 400 epochs.

**3. Nonlinear Prediction:**

* Implemented as a two-layer network with a hidden layer of 128 neurons and ReLU activation.
* Trained with the same procedure as the linear model.

**4. Evaluation**

* Prediction quality was measured using Mean Squared Error (MSE) on the test set.
* Time-domain plots of predicted vs original signals and prediction error were visualized.

# Graphical Results and Evaluations

## II.I Signal – 1 - Voiced

A group of orange and blue lines

AI-generated content may be incorrect.

**Evaluation:**

Both the linear and nonlinear models demonstrate similar performance in predicting the voiced speech signal. The Mean Squared Error (MSE) values for the test set are very close, differing only in the order of 0.000005 to 0.000010. This small difference indicates that while the nonlinear model has a slightly lower MSE, the improvement over the linear model is minimal. Visually, the prediction plots show that both models capture the signal dynamics well, with overlapping predicted and original signals. The error plots further confirm that prediction errors remain low and comparable for both models, suggesting that linear prediction may be sufficient for this type of voiced speech signal.

## II.II Signal – 2 - Unvoiced

A line graph of a graph

AI-generated content may be incorrect.

**Evaluation:**

For the unvoiced speech signal, the linear perceptron model significantly outperforms the nonlinear ANN. The linear model achieves a much lower Mean Squared Error (MSE) of ~0.000013 compared to ~0.000155 for the nonlinear model. This large difference suggests that the nonlinear ANN struggles to effectively capture the characteristics of the unvoiced signal, whereas the simpler linear model provides more accurate predictions. Visual inspection of the prediction plots and error curves confirms this disparity, with the linear model’s predictions closely following the original signal, while the nonlinear model shows larger deviations and higher errors.

# Summary and Conclusion

In this lab, both linear and nonlinear neural network models were implemented for speech signal prediction using a sliding window of past samples. The linear model was realized as a single-layer perceptron without an activation function, while the nonlinear model was a two-layer feedforward ANN with a ReLU-activated hidden layer. Both models were trained and tested on normalized speech data, and their performance was evaluated using Mean Squared Error (MSE) and visual inspection of predicted signals.

The results indicate that for voiced speech, both models perform similarly, with nearly identical MSE values and overlapping prediction plots, showing that a simple linear model is sufficient for this type of signal. However, for unvoiced speech, the linear model significantly outperforms the nonlinear ANN, achieving a much lower MSE. The nonlinear model exhibits poorer generalization in this case, likely due to the higher variability and noise-like nature of unvoiced segments, which may not benefit from added model complexity.

Conclusion:  
Model complexity does not always lead to better performance. For speech prediction tasks, especially when the signal characteristics vary (e.g., voiced vs. unvoiced), simpler models like linear predictors can outperform more complex neural networks. The choice of model should consider the nature of the signal and the specific prediction task.