

## Resources

<https://share.coursera.org/wiki/index.php/ML-004:Main>

```
[y fs] = wavread('file_location')  
mfcc
```

```
theta1 = [5 X 13]  
theta2 = [24 X 5]  
X = [m X 13] ← input  
Y = [m X 24] ← output
```

```
e = epsolon %rand_value  
theta1 = rand(5,13) * (2 * e) - e  
theta2 = rand(24,5) * (2 * e) - e  
delta2 = zeros(5,1);  
delta3 = zeros(24,1);  
Delta2 = zeros(5,1);  
Delta3 = zeros(24,1);
```

## Algorithm

1. Initialize weights(theta1, theta2) randomly. Initialize delta2, delta3 and Delta2, Delta3 to zero.
2. Extract training data(13 mfcc coefficients for each audio sample along with a 24-unit vector with corresponding language)
3. For each training example
  - a. Apply feedforward equations and get the output vector
  - b. Calculate cost
  - c. Calculate delta2 and delta3
  - d. Add each to Delta2 and Delta3 respectively
4.  $\Delta 2/m \rightarrow \theta 2$  ;  $\Delta 3/m \rightarrow \theta 3$
5. Finished Training. Now, for every test sample
  - a. Extract 13 mfcc coeffs
  - b. Run through the network(feedforward) and get the output vector
  - c. Whichever neural has highest value is the result of classification.