

# **HACETTEPE UNIVERSITY**

## **COMPUTER SCIENCE AND ENGINEERING**

### **CMP670 – STATİSTİCAL NATURAL LANGUAGE PROCESSİNG**

#### **ASSIGNMENT-3**

**Mustafa KAYA**

**N18147271**

A language model can predict the probability of the next word in the sequence, based on the words already observed in the sequence. Neural network models are a preferred method for developing statistical language models because they can use a distributed representation where different words with similar meanings have similar representation and because they can use a large context of recently observed words when making predictions.

### **Language Model Design**

In this assignment, i developed a model of the text that i could then use to generate new 5 sentences of text. . In order to implement neural network in Python environment, Dynet framework used. For dataset, Triumph's speeches were used.

The language model will be statistical and will predict the probability of each word given an input sequence of text. The predicted word will be fed in as input to in turn generate the next word.

A key design decision is how long the input sequences should be. They need to be long enough to allow the model to learn the context for the words to predict. This input length will also define the length of seed text used to generate new sequences when we use the model.

### **Results**

In this section, the example sentences generated by the neural network language at the word level are listed below. Generated sentences have no meaning.

1 .sentence-> mined cheaply wealth ohio windchill moves expanded 37 promote fell sick comfortable abolishing al applause

2 .sentence-> unsalvageable replacing biden steel twenty-five file aggressively producer tax kid regimes recipients applying amt exact

3 .sentence-> nervous colleges definition never grave passed council listen 2000. criticism monkey signings interest accomplishments opposed

4 .sentence-> quitting strongest consultants mentioning foolishness endorses 4 romneycare  
rent up behind groups nastily discussions tremendously

5 .sentence-> convert proposal 2013 population commission isis d competitive wall himself  
insisted old deceit saturdays estate