

Homework 1: N-gram Language Models

Ankit Goyal
ankit@cs.utexas.edu
CS380L

February 14, 2015

1 Introduction

N gram Model assumed that the probability of observing the i^{th} word w_i in the context history of the preceding $i-1$ words can be approximated by the probability of observing it in the shortened context history of the preceding $n-1$ words (n^{th} order Markov property).

In this assignment, the forward bigram model was modified to get the backward bigram model. Backward model aims at parsing sentences from right to left as compared to left to right for forward model. The performance of both the models were tested on three different corpus based on Word Perplexity Measures. In the next part of the assignment the two models were combined to a bidirectional model which combined both forward and backward models with a pre-decided weight given to each model.

Atis	Forward	Backward	Bidirectional (F = 0.5, B = 0.5)	Bidirectional (F = 0.4, B = 0.6)	Bidirectional (F = 0.6, B = 0.4)
<i>Perplexity_{Train}</i>	9.04	9.012	NA	NA	NA
<i>WordPerplexity_{Train}</i>	10.59	11.63	7.23	7.32	7.25
<i>Perplexity_{Test}</i>	19.34	19.36	NA	NA	NA
<i>WordPerplexity_{Test}</i>	24.05	27.16	12.70	12.94	12.69
Brown	Forward	Backward	Bidirectional (F = 0.5, B = 0.5)	Bidirectional (F = 0.4, B = 0.6)	Bidirectional (F = 0.6, B = 0.4)
<i>Perplexity_{Train}</i>	93.52	93.51	NA	NA	NA
<i>WordPerplexity_{Train}</i>	113.36	110.78	61.47	61.85	62.22
<i>Perplexity_{Test}</i>	231.30	231.20	NA	NA	NA
<i>WordPerplexity_{Test}</i>	310.67	299.68	167.48	168.31	169.86
WSJ	Forward	Backward	Bidirectional (F = 0.5, B = 0.5)	Bidirectional (F = 0.4, B = 0.6)	Bidirectional (F = 0.6, B = 0.4)
<i>Perplexity_{Train}</i>	74.27	74.27	NA	NA	NA
<i>WordPerplexity_{Train}</i>	88.89	86.66	46.51	46.83	47.11
<i>Perplexity_{Test}</i>	219.71	219.52	NA	NA	NA
<i>WordPerplexity_{Test}</i>	275.12	266.35	126.113	127.07	127.81