HW1: N-gram Language models

Rahul R Huilgol (rrh2226)

February 14, 2016

1 Introduction

A simple Bigram Model works by predicting the probability of a word given the context of the previous word.

		Word Perplexity	
Data	**Set**	*Forward*	*Backward*
ATIS	train	10.59	9.91
	test	24.05	24.79
WSJ	train	88.89	53.82
	test	275.12	180.39
Brown	train	113.36	61.43
	test	310.67	189.16

Table 1: Comparing perplexities of Backward and Forward Bigram Models

2 Bidirectional Bigram Model

		Word Perplexity		
Data	**Set**	*Forward*	*Backward*	*Bidirectional*
ATIS	train	10.59	9.91	7.34
	test	24.05	24.79	11.95
WSJ	train	88.89	53.82	46.41
	test	275.12	180.39	121.14
Brown	train	113.36	61.43	61.13
	test	310.67	189.16	160.22

Table 2: Comparing perplexities of Forward, Backward and Bidirectional Bigram Models

3 Conclusion