HW1: N-gram Language models

Rahul R Huilgol (rrh2226)

February 7, 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	6.22
	test	24.05	16.51
WSJ	train	88.47	54.69
	test	294.52	199.81
Brown	train	110.47	61.27
	test	509.69	311.54

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	6.22	7.34
	test	24.05	16.51	13.47
WSJ	train	88.47	54.69	46.80
	test	294.52	199.81	140.51
Brown	train	110.47	61.27	60.76
	test	509.69	311.54	277.38

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

3 Conclusion