Technical Report

Constructing a POS tagger

NLMI - Part A

Alex Khawalid (10634207)
Wessel Klijnsma (10172432)
Winand Renkema (10643478)

March 1, 2016

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quisque vulputate mi a libero luctus, ac maximus sem egestas. Pellentesque eget arcu hendrerit nisl pulvinar suscipit eu sit amet enim. In hac habitasse platea dictumst. Praesent volutpat massa massa, in eleifend dui mollis quis. Maecenas vulputate tortor non nibh bibendum, in aliquet ligula fringilla. Ut ac mauris nibh. Sed nisl diam, dictum quis condimentum vitae, euismod a felis. Nunc porttitor libero enim, eu mollis dolor venenatis ac. Pellentesque maximus risus vitae dolor bibendum ultrices. Aliquam erat volutpat. Pellentesque semper ipsum eget mauris suscipit vestibulum. Nam tincidunt ipsum et metus consectetur ullamcorper. Aenean fringilla ex turpis, a rutrum risus consectetur vel.

1 Introduction

2 Approach

2.1 Parsing training and test corpus

2.2 Constructing language and lexical model

2.2.1 Language model

2.2.2 Lexical model

2.3 Smoothing language and lexical model

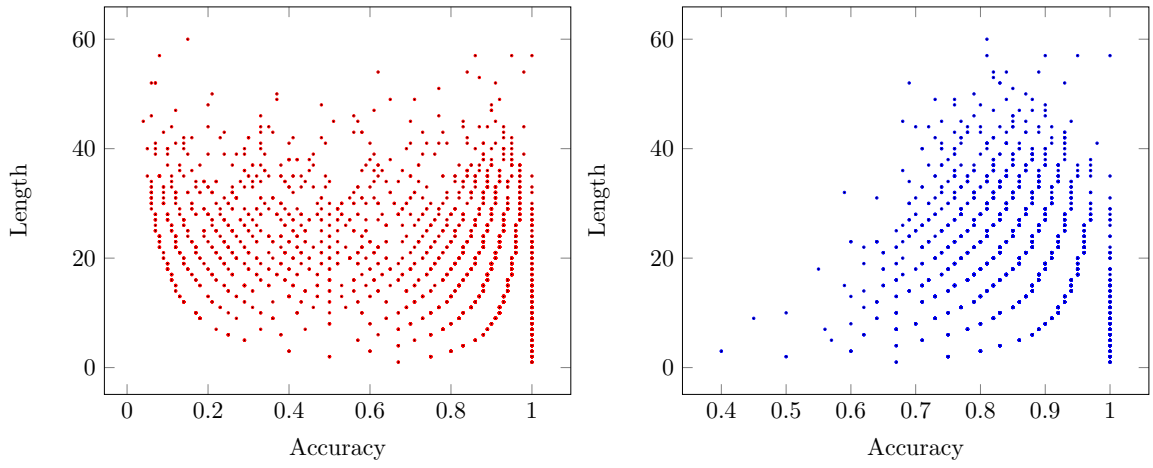
2.3.1 Language model

2.3.2 Lexical model

2.4 Computing most probable POS sequence

3 Results

3.1 Accuracy of POS tagger



In figure 1 one can see the dramatic result of smoothing.

4 Discussion

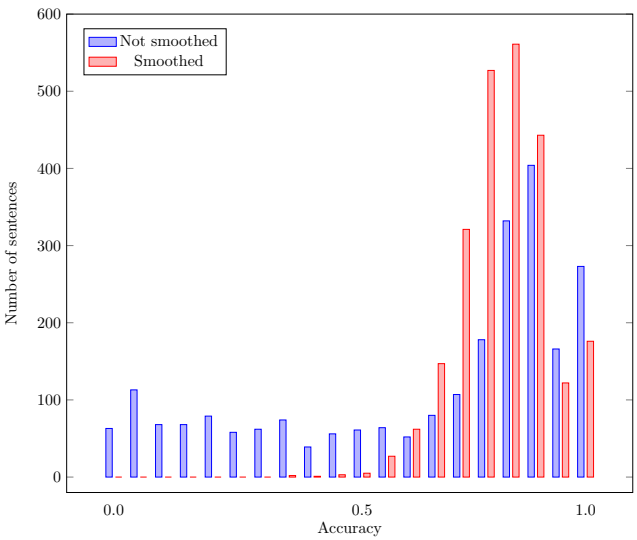


Figure 1: Accuracy without and with smoothing