|  |  |
| --- | --- |
| **NGUYEN, Réal** | **COMP 472 SEC F** |
| **ID#275662673** | **DUE 04/12/2018** |

**MINI-PROJECT 3 REPORT**

**1. INTRODUCTION**

**2. BASIC SETUP**

- Explain why you chose Spanish as OT here and the corpora used

- Explain what constitutes as a correctly and incorrectly tagged sentence

- Explain how the language models were built

- Explain how the sentences are read for each model

**2.1 Results**

- Compare sentences w/ correct language model and actual language model side-by-side

- Compare unigram vs. bigram models when incorrectly tagged

- Use tables w/ languages and total sum probability

- Take sentences of interest, list them, and say why you want to talk about these

- Sentences incorrectly tagged in unigrams

- Sentences incorrectly tagged in bigrams

- Sentences incorrectly tagged in both

- Sentences using loanwords/uncommon letters

- Compare probability table with bigram models

- For each incorrect prediction, analyse and hypothesize why the model failed, and how it could be improved

**3. EXPERIMENTAL SETUP**

- Explain what each experiment is

- Follow same guidelines as previous section

**3.1 Results**

- Follow same guidelines as previous section

**4. CONCLUSION**