Homework Four: Natural Language Processing

Quin'darius Lyles-Woods

November 17, 2021

1 Lexical Translation Model using IBM Model 1

Given a corpus of two English sentences with their translation in Latin below.

English this course

Latin hoc utique

English Text processing

Latin Textus processus

Build your lexical translation model (IBM Model 1) and show each training (expectation maximization) step's computation. Assume that the model only trains for two iterations.

2 Translation Model using BLUE Score

Given candidate translation sentences and their references below.

Candidate One the cat the cat on the mat

Candidate Two the cat on mat mat

Reference One the cat is on the mat

Reference Two there is a cat on the mat

Evaluate your translation model performance using BLUE score where n-gram order, $N = 3$.

3   Passage Retrieval

Design a question answering system for finding definition of all the machine learning keywords from Wikipedia pages.

Example Query Text: Support Vector Machine

Your system should return the corresponding wiki page for this query. Please show overall diagram of your system and explain each individual component.

4   Entity Linking

What is entity linking? How the overall entity linking system works? How would you find matching between mention and entity in the linking system?

5   Machine Translation Issues

What is the machine translation issues? Please provide explanations and examples for each of the cases.