

01.112 Machine Learning

Fall 2018

Design Project

Rayson Lim Jun Kai (1002026)

Angelia Lau Kah Mun (1002417)

Kimberlyn Loh (1002221)

# Table of Contents

[1 Estimating only Emission Parameters of Hidden Markov Model 2](#_Toc531076603)

[1.1 Approach 2](#_Toc531076604)

[1.2 Results 3](#_Toc531076605)

[2 Estimating Hidden Markov Model 4](#_Toc531076606)

[2.1 Approach 4](#_Toc531076607)

[2.2 Results 5](#_Toc531076608)

[3 Implementation of Second-Order Hidden Markov Model 6](#_Toc531076609)

[3.1 Approach 6](#_Toc531076610)

[3.2 Results 6](#_Toc531076611)

[4 Design Challenge 7](#_Toc531076612)

# Estimating only Emission Parameters of Hidden Markov Model

## Approach

The estimation of emission parameters of a Hidden Markov Model can be calculated by using Maximum Likelihood Estimate (MLE), whereby

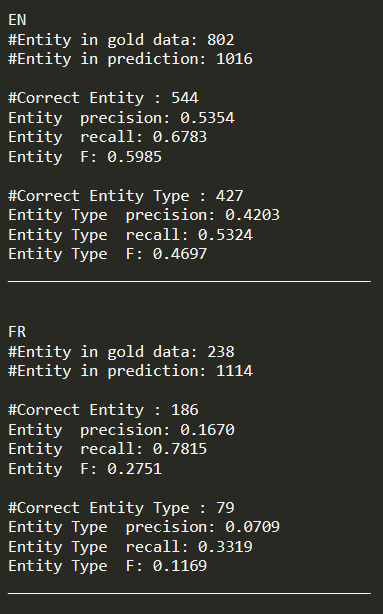
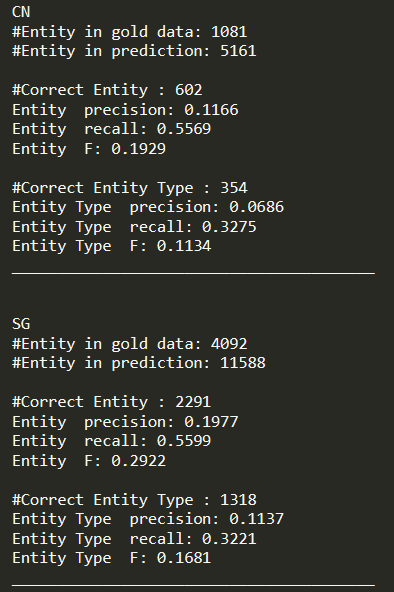
where is the number of times the observation was emitted from the tag and is the number of times the tag appears in the training data.

To account for rare words, a special word token of is used to replace rare word token in the training set that appear times, also known as smoothing. Thus, the smooth estimation of emission parameters can be written as

Given the training file, the function will calculate , utilising the formula above and return a dictionary in the form of . With the dictionary from the function, the tag generation for the test set can be performed, whereby for each word, the tag with the maximum probability is taken. In addition, to account for words only found in the test set / not found in the tag dictionary and for word token , the tag of is given.

## Results

The results of the above implementation of just estimating the emission is as shown below.

# Estimating Hidden Markov Model

## Approach

As observed in the previous section, utilising only emission parameters leads to low accuracy in the test set as transition parameters of the Hidden Markov Model, which dictates tag generation, is ignored. Hence, the transition parameters can be estimated, using MLE, to be as followed.

where is the number of times the two successive tag appears and is the number of times the tag appears in the training data. In addition, Hidden Markov Model can be defined as follows.

Thus, with the estimate of transition and emission parameters, the tags can be generated. However, as brute force enumeration is computationally expensive with a time complexity of , dynamic programming like Viterbi Algorithm with a time complexity of is implemented for decoding instead. In addition, the space complexity of the Viterbi Algorithm is .

The Viterbi Algorithm is as followed:

1. Base Case:

u

START 🡪

1. Recursive Case:

u 🡪

v

1. Final Case:

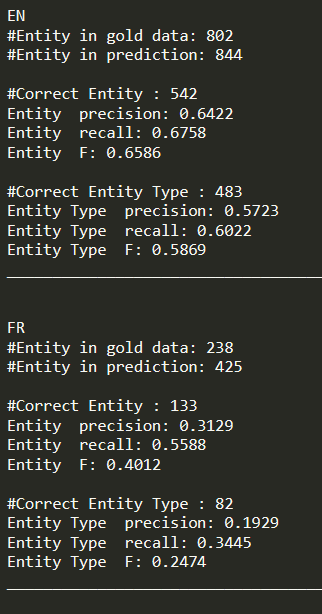
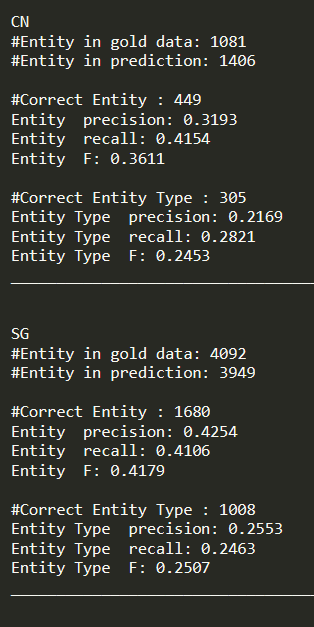
u

🡪 STOP

1. Backtracking to get the optimal sequence of tag

## Results

The results of the above implementation of Hidden Markov Model is as shown below.

# Implementation of Second-Order Hidden Markov Model

## Approach

## Results

# Design Challenge