

# Homework 2 Report

Jesus Garcia

March 2020

## 1 Introduction

For this assignment I decided to try out a programming language I've been meaning to learn. Julia is a high performing language designed for scientific computing. It's been around for a few years now and it was a good time learning the language while at the same time implementing a hidden markov model with viterbi decoding. The next section describes the process of downloading Julia.

## 2 Installing and Running Julia

I downloaded the long-term support release of Julia (v1.0.5). This can be found [here](#)(click on "here"). Instructions for finishing installation and being able to run it on your machine can be found [here](#)(click on "here"). Instructions for running my implementation can be found in the readme of my submission, "README.md".

## 3 Code structure

My entire implementation for the assignment exists in the file "hmm.jl". High level steps of my implementation: Fitting the model with the provided training set, this includes keeping track of all tags, word associations, tag-to-tag transitions, and tag counts. After fitting the model, a probability matrix (lattice) is constructed for each sentence. Each lattice is examined, returning a vector of tags that are predicted to be associated with the current sentence.

The first things in the script are two structs I've defined. The Lexicon struct is what's created from fitting the model. The fitted lexicon, along with a tagless corpora, is used to construct the probability matrix needed for Viterbi decoding. The entire process occurs in the function "main".

## 4 Evaluation: Reporting the Score

Training on "POS\_train.pos", tagging "POS\_dev.words":  
30733 out of 32853 tags correct  
accuracy: 93.547012

Training on "POS\_train.pos" and "POS\_dev.pos", tagging "POS\_dev.words":  
31583 out of 32853 tags correct  
accuracy: 96.134295