Ronak Mehta
Instructor: Noah A. Smith
CSE 517: Natural Language Processing
Winter 2021
Assignment A4
Due: February 19, 2021

Problem 1

For this problem, I defined the score function as follows, letting $\mathbf{x}$ by the sentence, and letting $y, y' \in \mathscr{V} = \mathscr{L}$ (the label space in vocabulary are the same).

$$s(\mathbf{x}, i, y', y) = \begin{cases} \log p(y \mid y') & \text{if } x_i = \texttt{<mask>} \\ 0 & \text{else if } x_i = y \\ -\infty & \text{else} \end{cases}$$

This made it so that unmasked words would always be predicted as themselves by the vanilla Viterbi algorithm made to decode the labels $y_1, ..., y_\ell$, and the masked words would be predicted my maximum likelihood. THe outputed sentences are in `output.txt`.