Su Zhang
Natural Language Programming Homework
Spelling Correction Additional Notes

- **<u>Key assumptions for my spelling correction model</u>**: In my model, I assume that there is only one error in the input string, and the potential correct words candidates have an edit distance of 1 from the error word. Additionally, I assume that deleting the first letter of word or inserting the first letter of word are extremely unlikely, which means that whenever the insertion and deletion errors happened before the position of first letter, we assumed very low probability to $P(x|w)$.
- **<u>Scenarios where my spelling corrector works well</u>**: I have run several test cases, such as "helo", "acress", "acommodate", "lovr", which work very well in my model, because these only have 1 error that are either "addition", "deletion", or "substitution" with edit distance of 1.
- **<u>Scenarios where my spelling corrector did not work well</u>**: My model would not work well for misspelled words where the edit distance is more than 1, such as "survelliance". My model would not work well for misspelled words that have more than two errors, such as "dilaogue". My model would not work well for spelling words with transposition error, such as "collaegue".
- **<u>Improvements that could be made:</u>** It would be better to add a recursion that could check multiple errors in the input strings. Moreover, to generate potential candidates with edit distance larger than 1 in order to correct misspelled word with larger edit distance. Moreover, to generate corrected transposition candidates in order to correct the transposition errors.