Arthur DE NEYER
Aymeric WIBO

Date: *May 6, 2023*

LINFO1104: Twit-Oz

1 Introduction

`Twit-Oz` is a simple tweet predictor based on a set of given data. Given a certain input, it will try to complete it with the most probable following words. In this report, we will explain all implemented extensions available in our production.

2 Implemented Extensions

2.1 Generalization of the n-gram formula **

A crucial aspect in this system is to have the best possible precision in its predictions. In order to do that, we allowed input of n-word(s). If there is no result, it will try to find one using n-1 Grams. We implemented it in such a way that it's fast on load up as well as when prediction is asked.

2.2 History **

A pratical feature in such a system is to keep a history. Every time the user enters a new input and ask for prediction, his input is stored in a seperated file. When loading the app, this file is read and all saved inputs are displayed.

2.3 Improve the existing GUI *

In order to make the user experience as best as possible, we made the app look fancier. It also looks like ChatGPT which is user-friendly and often familiar with users.

3 Other choices

3.1 Use of a dictionnary

Instead of using an Ordered Binary Tree, we used a dictionnary that has as keys the possible N-words combinaison (which corresponds to the input of the user) and as value another dictionnary: with as keys the next words and with value the corresponding probability. We used this method because dictionnaries are faster (due to it's C++ implementation in Mozart).

So this dictionnary works as follow: Depending on the input length, it will create a dictionnary where keys are