Shared Task – 3
Report

Group Members

Srinidhi Bhat 14CS30035
Kinsuk Das 14CS10025
Vishnu Praneeth 14CS30031
Vikash Minz 14CS10052
Basava Raj 14CS10023

Instruction to run the code
The paths for the fasttext en dictionary needs to be set.
It was downloaded from here
https://s3-us-west-1.amazonaws.com/fasttext-vectors/wiki.en.vec

Also, check the path for wikitionary.csv file, which is the dataset.

Details
Train test split is 4:1

Each vector given by fasttext is of dimension 300. We take the 2 input words, form a concated vector which forms our input vector for the neural net.
The output vector is simply the fasttext vector of the portmanteau of the words.

Architecture of the Net
Input layer dimension : 600
Output layer dimension : 300
4 hidden layers, each of dimension 450

Algorithm to predict portmanteau given 2 input words
We fee the concatenation of fasttext vectors of the 2 input words to the neural net. Let the obtained output vector be Y_pred.
Now we form different possible word combinations considering the 2 input words and obtain its fasttext vector and take the cosine distance of it along Y_pred.

The word having the greatest cosine component is returned as the blend of the given words.

Cases where the fasttext vector doesn’t exist for the input words have been ignored.