Certain hexadecimal values are unreadable in UTF-8 file format and created decoding error for Unicode. To remove those unwanted sequence of non ASCII characters this code is written. The input to the code is the text file containing all the tweets and output will be further cleaned tweets to be given as input for training.