

**PROJECT REPORT**

**LOCATION PREDICTION FOR TWEETS**

Submitted to – Sanasam Ranbir Singh

Mentored by – LOITONGBAM GYANENDRO SINGH

IIT GUWAHATI

21-04-2021

Submitted by - Group MARD

Mukul Chaturvedi 204101035 Aditya Tandon 204101003

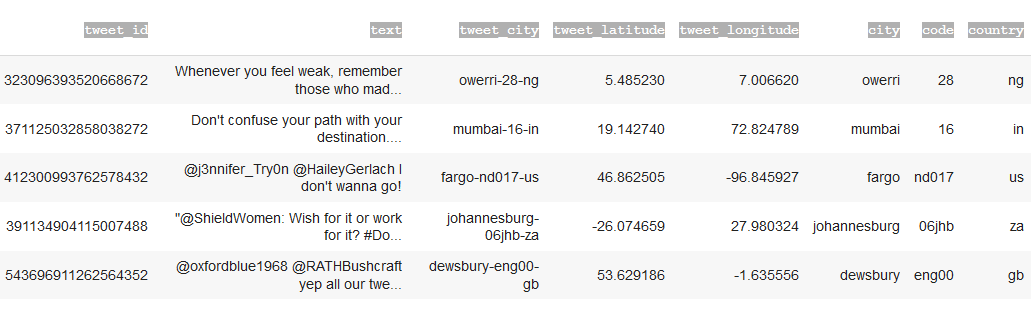
Ram Avtar Sahu 204101044 Dan Singh Pradan 204101020

**Problem statement -**

To predict the location (city and country) of a tweet using only the tweet text.

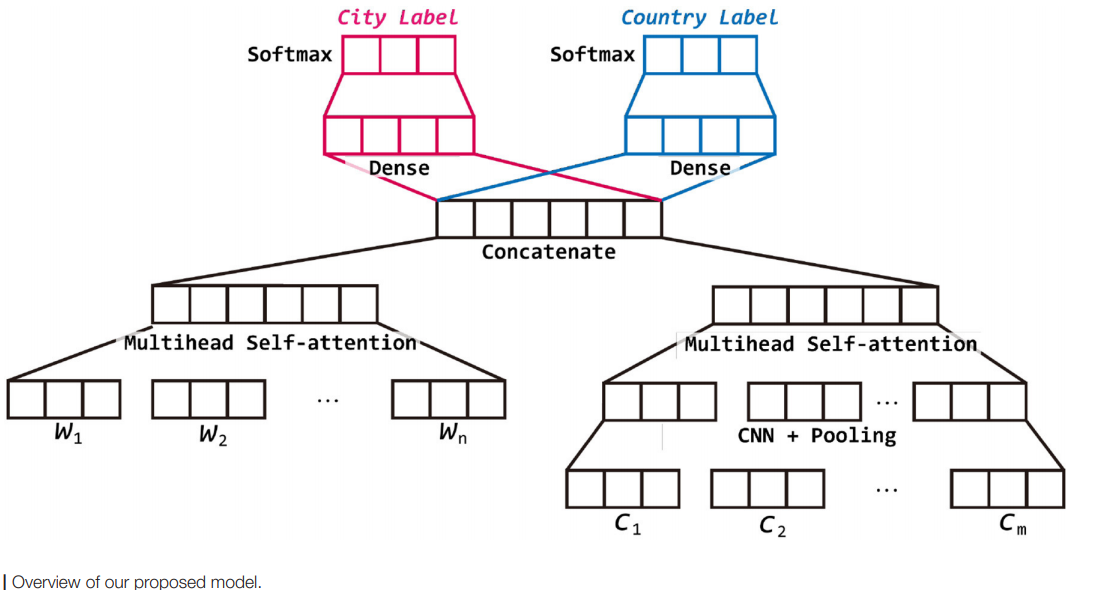
**DATA**

We have used a csv file with tweet text, country and city labels as fields. Number of data points were 5700 and we have split in 80:20 for training and testing purpose.



**Overview of our Model**

Given in the figure below is the overview of the model, which can be broken into two main parts of text representation and joint training. The text representation module consists of word representation and character representation. Both of the representation is encoded by multi-head self-attention layer but for character representation, we further use a CNN layer and pooling layer first to reduce the dimension and extract meaningful information. The word representation and character representation are then concatenated as a vector which represents the given tweet. In the second module, to utilize the relation between cities, we use the same concatenated vector but two different output layers to predict the country and city at the same time. However, the country classification is used only for training. In the testing phase, we use only the city part of the model for prediction.

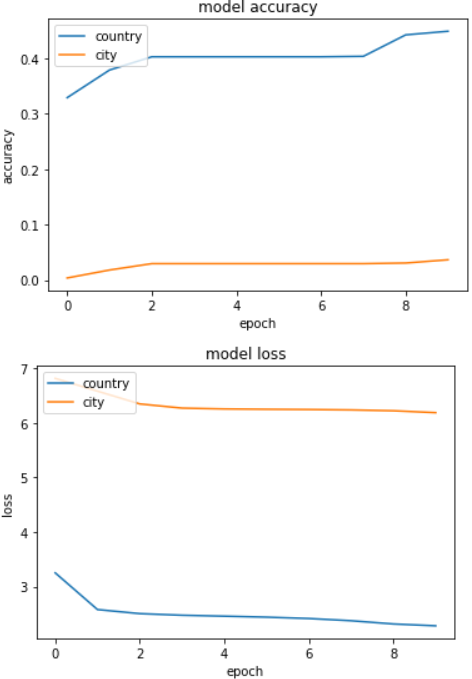


**RESULTS**

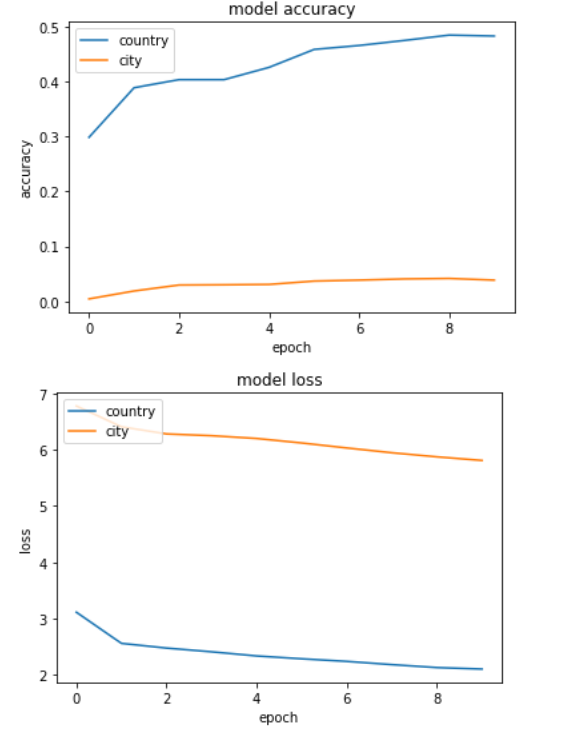
|  |  |  |
| --- | --- | --- |
|  | City Prediction | Country Prediction |
| Paper Model 1 | 4.0905 | 43.080 |
| Paper Model 2 | 4.003 | 45.691 |
| Paper Model 3 | 4.003 | 46.388 |

**Observations**

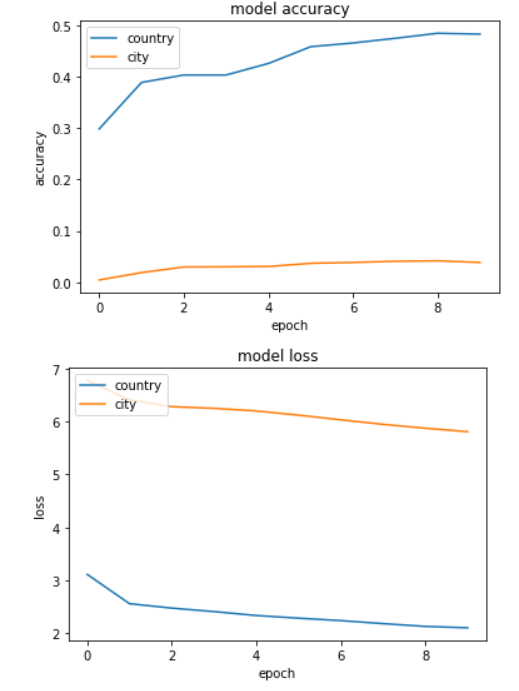
**Paper Model 1**



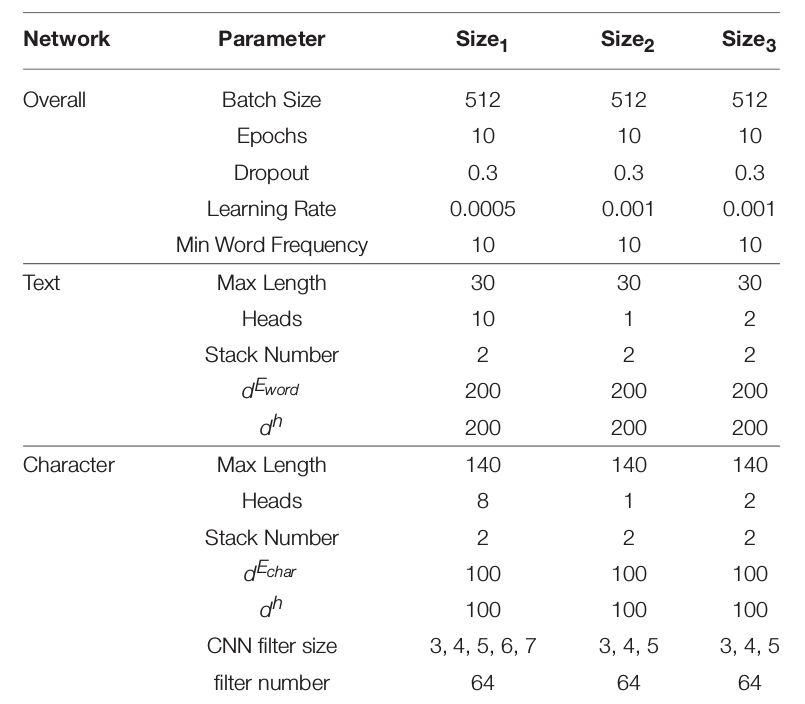
**Paper Model 2**



**Paper Model 3**



**Paper Models Summary**



CNN filter size is only three in our implementation.

**Additional work Results**

|  |  |  |
| --- | --- | --- |
|  | City Prediction | Country Prediction |
| Additional Model 1 | 4.003 | 40.731 |
| Additional Model 2 | 2.872 | 37.771 |
| Additional Model 3 | 3.655 | 45.343 |

**Additional Model 1 Summary**

The base model is the same with modification being of using stack of 3 multiheaded self-attention with dense layers instead of 2. Also, we used CNN and pooling layers in word side representation much like what was done in characters side in the research paper.

**Additional Model 2 Summary**

The base model is the same with modification being of using Bidirectional LSTM instead of stacks of multiheaded self-attention in both character and word representation.

**Additional Model 3 Summary**

The base model is the same with modification being of using stack of 3 multiheaded self-attention with dense layers instead of 2.