LOCATION PREDICTION FOR TWEETS

GROUP NAME - MAVERICK

AIM

- To develop a deep learning based solution to predict geographic information for tweets.

Motivation

The current approaches bear two major limitations, including-

- hard to model the long term information.
- hard to explain to the end users what the model learns.

Methods Used

- Text representation
- Multitask Learning

Applications

- Marketing Recommendation Systems
- Event Detection Systems

Text Representation

- Multi-head self Attention Mechanism
- Subword Feature

Why Multi-head self attention ?

- RNN-based model usually suffers from the extremely long training time
- CNN works better in modeling the local information
- Multi-head self-attention model utilizes only the attention mechanism, yet it enjoys the advantages of both RNN and CNN.

Subword Feature

- The subword feature is shown to be very useful for tasks built on social media since people tend to use lots of informal language on social media
- The idea of adding a subword feature is to infer the meaning of the low frequency word.

Multi-task learning

- Multitask Learning is a method to train a learning model with different targets.

Figure 1 explaination

Test representation consist of two part: Word representation & character representation

Both of these are encoded by a multihead self-attention model but for character representation we first use CNN and pooling to reduce dimension.

Then concatenate both word and character representation into a vector.

After that use that vector along with city and country label for training.

How to Remember position relation between words?

For this we use position encoding. The idea is to injecting the position information into the word vector.

Create the position embedding matrix and add this to the corresponding word vector.

Position encoding is applied in the beginning of the model.

Thanks!