Exploring and Learning Suicidal Ideation Connotations in Social Media with DL

Ramit Sawhney, Prachi Manchanda, Puneet Mathur, Rajiv Shah and Raj Singh

Research Goal

MOTIVATION: Sucide is currently the third largest leading cause of death of young people in the age 15-24 years, people who otherwise might have made an impact on the world. According to WHO, in the last 45 years, suicide rates have been up by 60% which calls for a framework that can help detect such people o that help can be offered.

PROBLEM STATEMENT: To come up with an automated framework that leverages Deep Learning techniques to classify a given tweet as Suicidal or Non-Suicidal and to do so with as good accuracy as possible so that timely help can be offered.

System Architecture
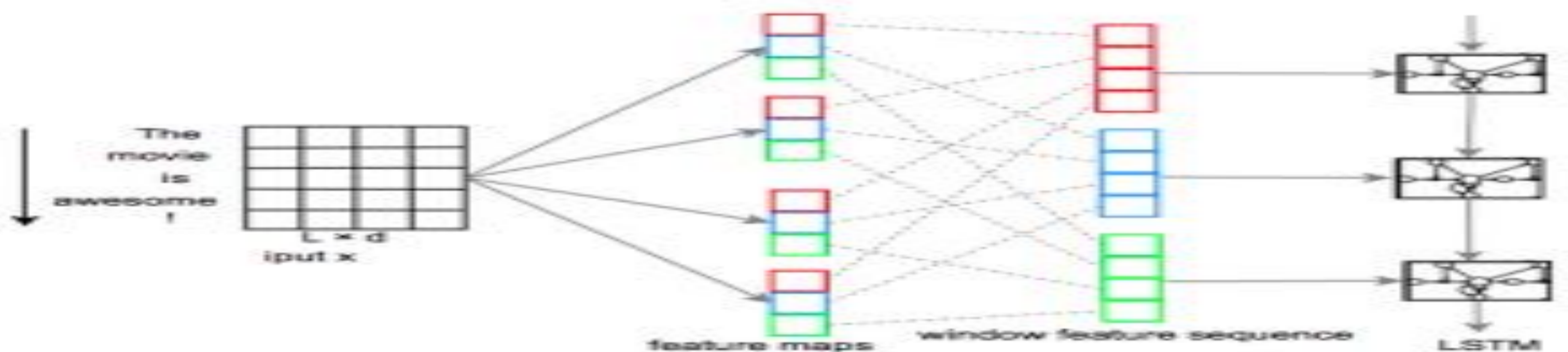
Stage 1
- Pre-process the tweets and split data into train, test & validation.
- An embedding layer is made using top N frequent words in tweet.

Stage 2
- For the C-LSTM model, a CNN and max-pooling layer are added after the embedding layer and the ouput is fed to the LSTM to claasify.

Stage 3
- Calculated the results of various models(Baselines as well as DL models) and compared it with state-of -the-art results. .

System Features

- The framework is able to achieve state of the art results using a C-LSTM framework.
- It combines efficient spatial encoding and feature extraction of CNN's with the efficient text classification by usiing long-range dependencies of LSTM.

- The framework can not only be used to detect suicidal tweets but also **depressional ideation in tweets** after re-training the models.
- Thus, it is of help for many organisations, police etc.
- Such a framework can also be build for other languages that aren't so morphologically rich.

Contributions

- Creation of an annotated dataset of tweets scraped from various suicide web-forums as well as micro-blogging sites such as Reddit and Tumblr. This dataset would encourage more research in this area.
- Basline results using various Machine Learning moddels to show that Deep Learning can be leveraged for this framework to detect suicidal ideation in tweets.
- Although the accuracy saf the model is 81.2%, the recall(what matters the most in these cases) is 87.2% which is very good compared to the previous results.
- We have shown why pre-processing is a crucial step when working with NN's and other DL frameworks.