NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
College of Engineering

AI For Healthcare

Analysis on Depressive Social Media Texts

Student: Neo Shun Xian Nicholas Supervisor: Associate Professor Erik Cambria

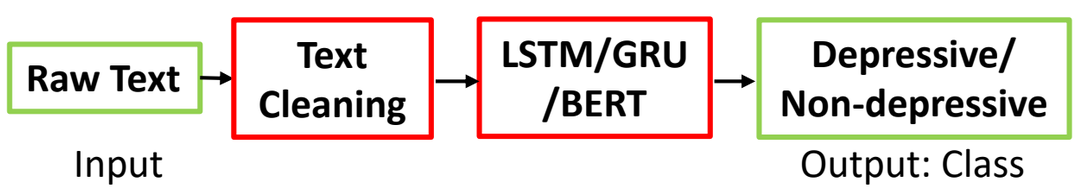
Abstract

Mental health has been an increasingly challenging issue to tackle in this era due to the stressful environment we are living in. One such example of mental health illness is depression. Depression is a mood disorder described as feelings of sadness, loss or anger that interfere with one's everyday activities. Depression has existed as a problem in this society for many years.

Project Objective

To leverage on the use of Deep Learning and Natural Language Processing (NLP) to do an in-depth analysis on depressive social media contents.

Emotion Classification



Classify if a text is depressive or non-depressive using both the sequential models (LSTM & GRU) and the transformer architecture (BERT).

Emotion Classification Results

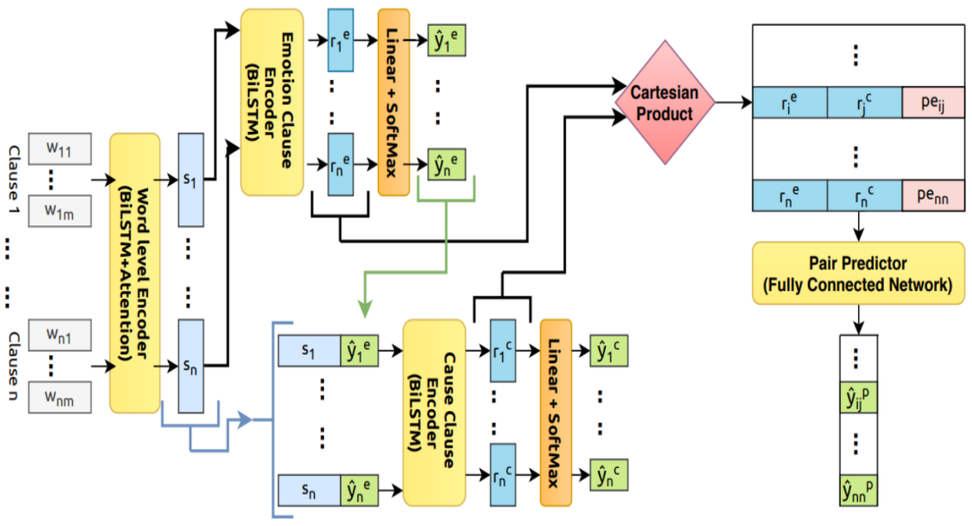
Model	Validation Accuracy	F1 Score	AUC
Bi LSTM	0.73539	0.7362	0.80849
Bi GRU	0.70698	0.7063	0.79508
BERT	0.7883	0.7874	0.8726

Twitter Dataset

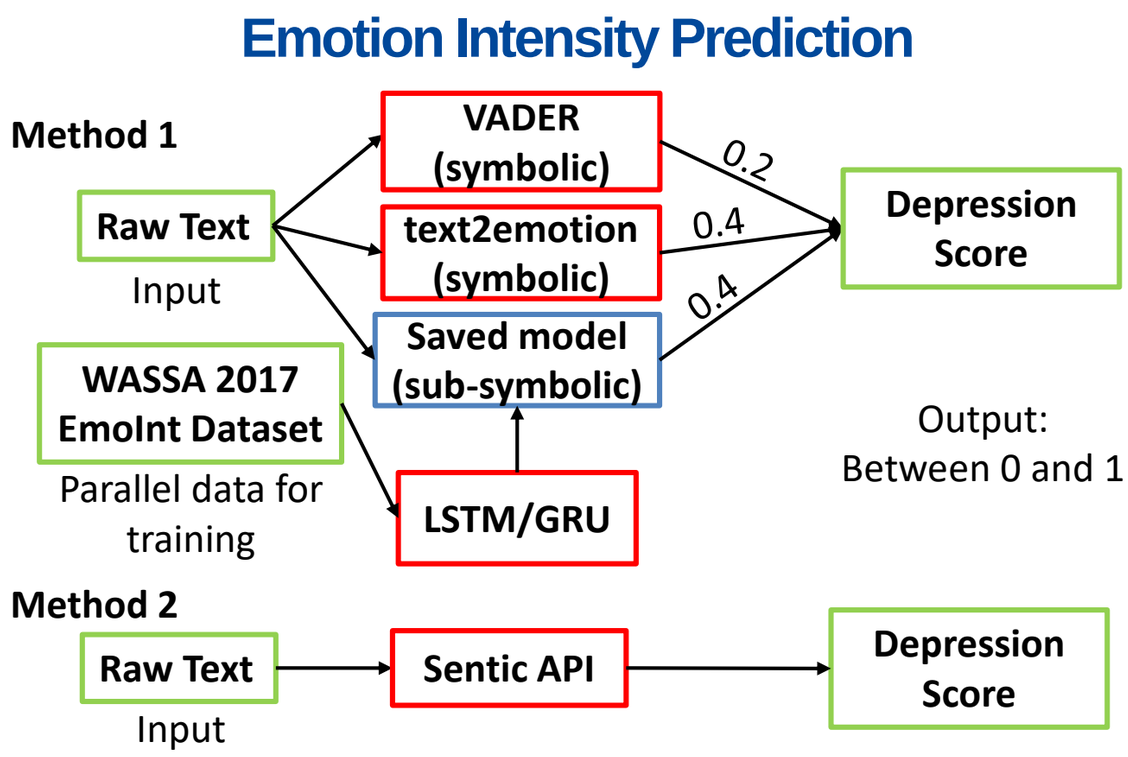
Model	Validation Accuracy	F1 Score	AUC
Bi LSTM	0.86642	0.8776	0.9408
Bi GRU	0.85723	0.8679	0.93939
BERT	0.8832	0.897	0.95

Reddit Dataset

E2E-Pext(E) Architecture



Emotion Intensity Prediction



Methods to estimate depression magnitude from text:

- **Method 1:** Combination of symbolic and sub-symbolic approaches
- **Method 2:** Sentic API on Depression Identification

Intensity Score for some Sample Data

Text	Symbolic + Sub-symbolic	Sentic API
So alone so tired so bored so ugly so depressed.	0.66375	0.866
Feeling bummed out rn. Family is disappointing and friends are too far away??	0.38461	0.5
Why is it that everyone gets what I want, I guess whatever	0.21962	0.33

Emotion-cause Pair Extraction

Focus on depressive emotion and the likely cause for it. Then, extract the emotion and the cause pair accurately.

Emotion-cause Pair Extraction F1 Scores

Variant	Dataset	Pair Extraction F1 Scores
E2E-PExt(E)	Twitter	0.5150
	Reddit	0.5488
E2E-PExt(C)	Twitter	0.5089
	Reddit	0.5333

E2E-Pext(C) Architecture