

Using Natural Language Processing to predict suicidal ideation on Reddit

AUTHOR: Eugene Baraka

OUTLINE



BACKGROUND
AND CURRENT
RESEARCH



METHODS



RESULTS



DISCUSSION



CONCLUSION
AND LIMITATIONS

BACKGROUND AND CURRENT RESEARCH



BACKGROUND

Suicide is one of the leading causes of death globally¹

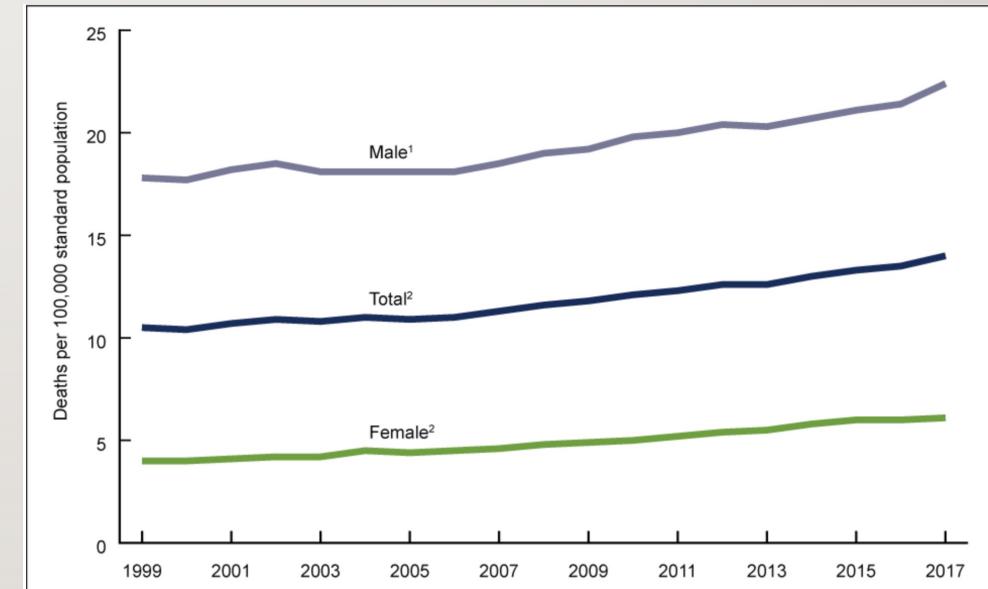
- An estimated 800,000 deaths annually → **One death every 40 seconds**

Among the youth, suicide is the second most common cause of death in the US²

The age-adjusted suicide rates have significantly increased over the past decade, worldwide²

Research suggests that self-harming youth are more active on social networks than the non-self-harming youth²

Age-adjusted suicide rates by sex in the US, 1999-2017³



1. Significant increasing trend from 2006 through 2017
2. Increasing trend from 1999 through 2017

¹WHO 2014

²Memon et al. 2018

³NCH 2018

CAN MACHINE LEARNING HELP?

- Active area of research in machine learning and medicine
 - Over 100 papers published on *PubMed* in the past decade
- Multiple studies^{4,5,6} explored the use of ML for early detection of suicidal thoughts and the progress has been impressive
- Overarching hypothesis: early suicide detection saves lives
- Primary aim: use state-of-the-art ML techniques to classify suicidal and non-suicidal thoughts

⁴Jashinsky et al. 2014

⁵Ueda et al. 2014

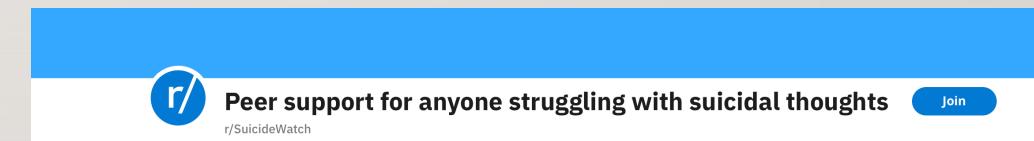
⁶Morese et al. 2022



METHODS

ABOUT THE DATA

- Collected from Reddit platform, which leverages group chats known as sub-reddits
- Each subreddit allocated to a specific topic and individuals interested in the topic join the group to participate in discussions
- The data collected using PushShift API from
 - "SuicideWatch" made of over 390k users (October 2022) → *labelled as suicidal*
 - "Teenagers" made of contains 2.9 million users (October 2022) → *labelled as non-suicidal*
- Data publicly available on Kaggle⁷



DATA CLEANING AND PREPROCESSING

- Dataset made of 232,074 unique posts posted between December 16, 2008 to January 1, 2021
- Each post classified as either suicidal or non suicidal
 - No missing values and both classes were balanced (50%-50%)
- Contractions were expanded in each post (e.g., I'm → I am)
- Text cleaning involved removing:
 - Punctuations (e.g., quotations, question marks) and URLs
 - Stop words (e.g., I, to)
- Changed all text to lower case

DATA CLEANING AND PREPROCESSING (Cont'd...)

- More features were extracted from original text
 - Number of words, characters, and sentences
 - Average word and sentence lengths
- Sentiment analysis per post
- Compare the extracted features between the two classes during Exploratory Data Analysis

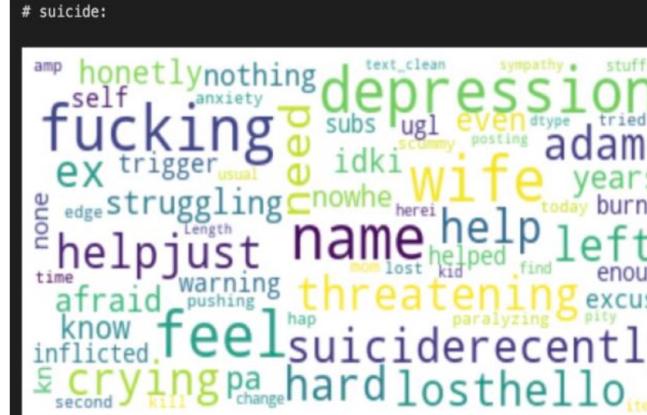
MODELLING

- Classification problem (either suicidal or not)
- Multiple classification models explored to choose the top performing models to fine tune
 - Logistic Regression
 - K-Nearest Neighbor (KNN)
 - Decision Trees
 - XGBoost
 - Random Forest
 - Support Vector Machines (SVM)
- Models compared to choose top 3 performing models
 - Precision, Recall, and F-1 score were the metrics used
- Hyperparameter optimization using GridSearch for the models to choose the best model
- Transfer learning using DistillBert was performed to compare its performance with the best model

RESULTS

COMPARISON BETWEEN CLASSES

The most common words by post category (left: suicidal, right: non-suicidal)

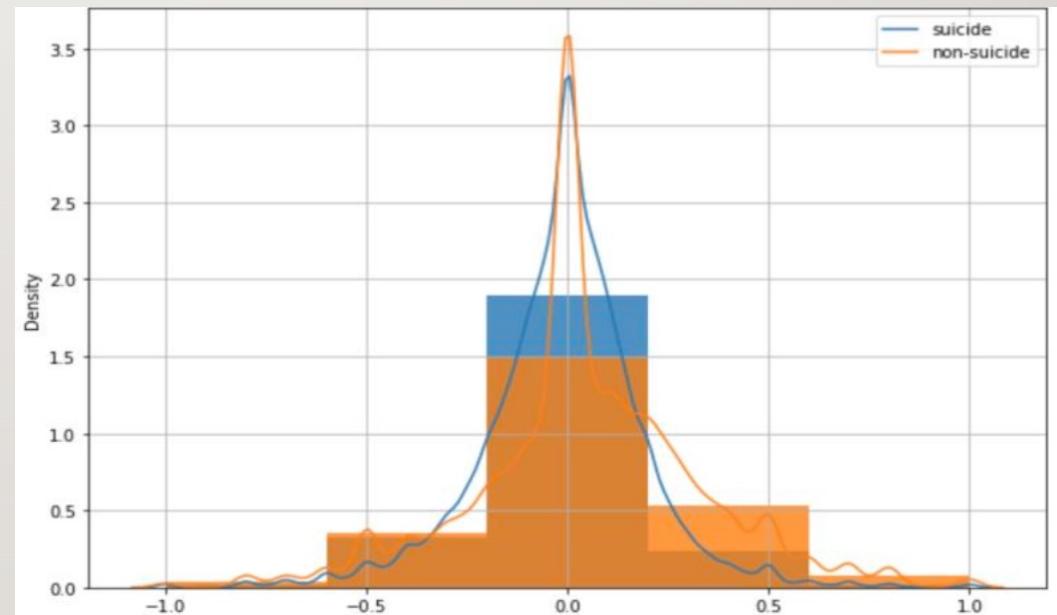


1. Suicidal posts generally contain more negative words than non-suicidal text
2. Suicidal posts involved discussions about struggling (e.g., breakups, suicide, depression, crying, being lost)
3. Non-suicidal posts included conversations around daily life (e.g., school, compliments, movies, family and friends, food)

COMPARISON BETWEEN CLASSES (Cont'd...)

- No difference in number of words, characters, or sentences used between the two groups
- No difference in the average post length
- Sentiment curves overlap suggesting no difference in sentiments between the two classes

Sentiment analysis by post class



MODEL RESULTS

- Logistic Regression, Random Forest, SVM had the highest F1 score (87%) followed by XGBoost (86%)
- Explored further through hyperparameter optimization to choose the best-performing model
- After tuning, *Logistic Regression* had the highest F1 score of **90%**

Results of preliminary non-tuned models

Model	class	Precision	Recall	F1 score
Logistic regression	0	0.86	0.88	0.87
	1	0.88	0.85	
KNN	0	0.88	0.75	0.82
	1	0.78	0.89	
Decision Trees	0	0.80	0.82	0.81
	1	0.82	0.79	
XGBoost	0	0.85	0.87	0.86
	1	0.87	0.85	
Random forest	0	0.86	0.88	0.87
	1	0.88	0.85	
SVM	0	0.86	0.88	0.87
	1	0.88	0.85	

DISTILLBERT

- Text cleaned using DistillBERT's cleaning and preprocessing modules
- Model trained on *5 total epochs* with *2 input layers*
 - Number of epochs and layers reduced due to small computation power
- F1 score of 92.3%

Results of DistillBERT classification model

F1 score 0.9232954545454547

Classification Report

	precision	recall	f1-score	support
non-suicidal	0.97	0.88	0.92	361
suicidal	0.88	0.97	0.92	336
accuracy			0.92	697
macro avg	0.92	0.92	0.92	697
weighted avg	0.93	0.92	0.92	697

DISCUSSION



DISCUSSION

- The best models were *Logistic Regression* (90% F1 score) and *DistillBERT model* (92.3% F1 score)
- Adding more epochs and input layers would significantly improve model performance
- However, results from Neural Networks lack interpretability. If interested in which words are more likely to be associated with suicidal thoughts, a logistic regression would be better suited
- It is possible to predict suicidal thoughts and the accuracy can be high with modern NLP techniques, as also found by previous studies^{4,5,6}
- However, multiple challenges remain

⁴Jashinsky et al. 2014

⁵Ueda et al. 2014

⁶Morese et al. 2022

LIMITATIONS

- Low data quality, associated with labelling costs
 - Lack of reliable publicly available datasets → privacy concerns, among others
 - This analysis assumed data from SuicideWatch are automatically suicidal and vice versa for Teenagers subreddit
- There is a thin line between anxiety, depression, and suicidal thoughts
- Need medical experts to label the data for more accurate predictions

Posted by u/lakrissockor 3 hours ago

I'm ugly.

I'm tired of being ugly , socially incompetent and unintelligent. I wish I was good at something. I'm just useless and a waste of space , don't understand why god created me. Im not suicidal but I just want to dissapear forever.

Post from SuicideWatch. Is this a suicidal thought, depression, anxiety?

An aerial photograph of a long bridge spanning across a large body of water. The bridge has multiple lanes of traffic, including several white cars and a few larger vehicles like trucks or buses. The water below is a vibrant turquoise color with visible ripples and reflections.

THANK YOU
