Subreddit classifier

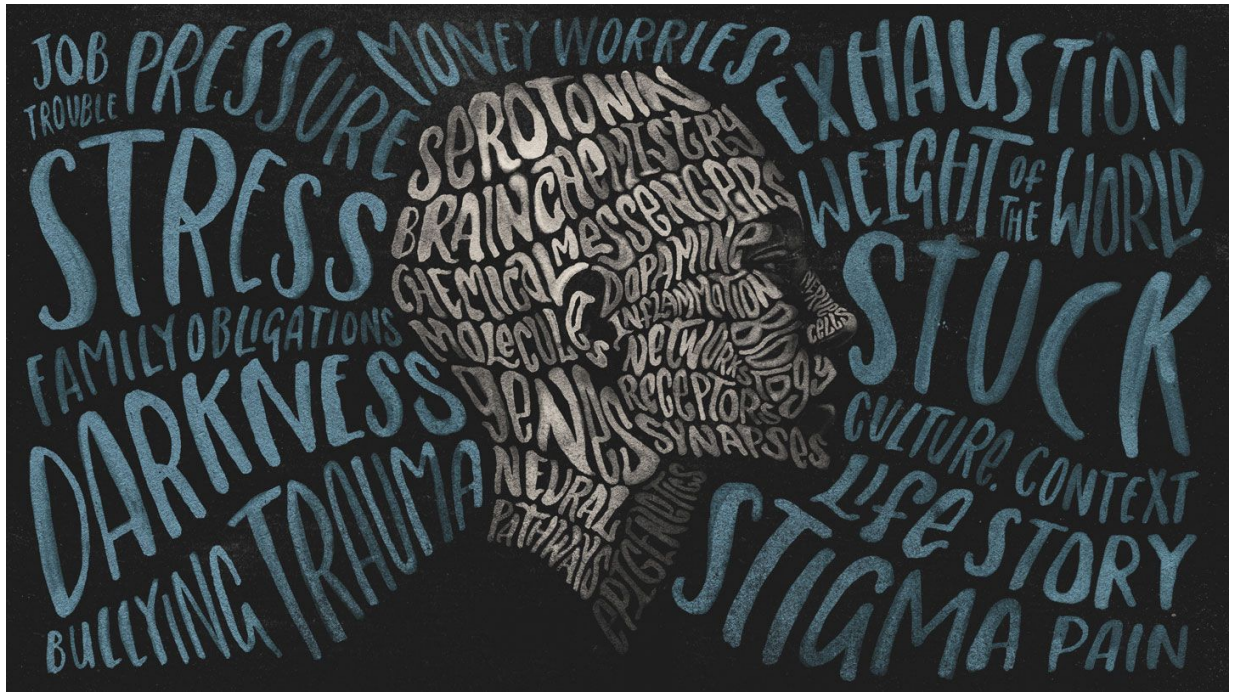
By
Benjamin, David, Swee Jin, Nicole
10th Aug 2023

Contents

- Background / Problem Statement
- Methodology
- Initial data insights
- Modelling
- Results
- Future Work

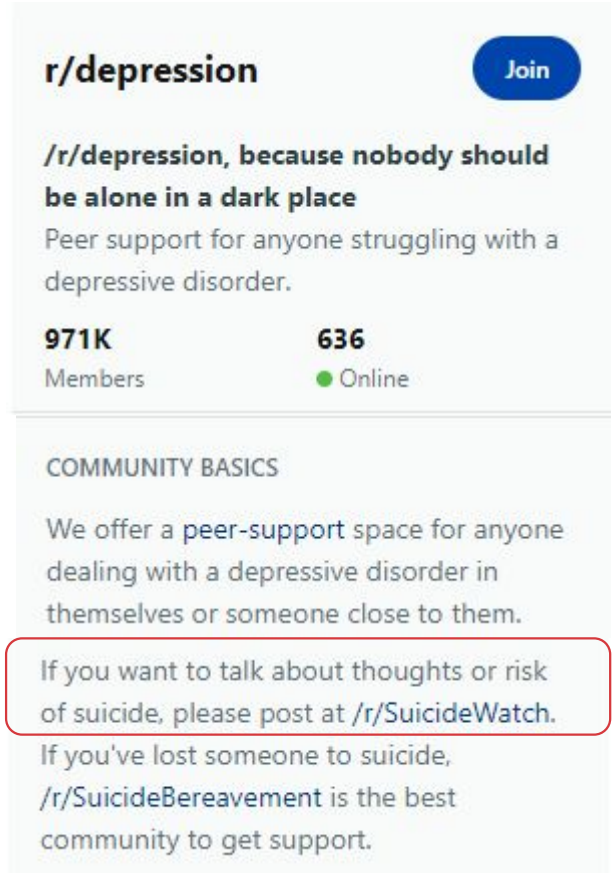
Background

What is Depression?



Background

Background

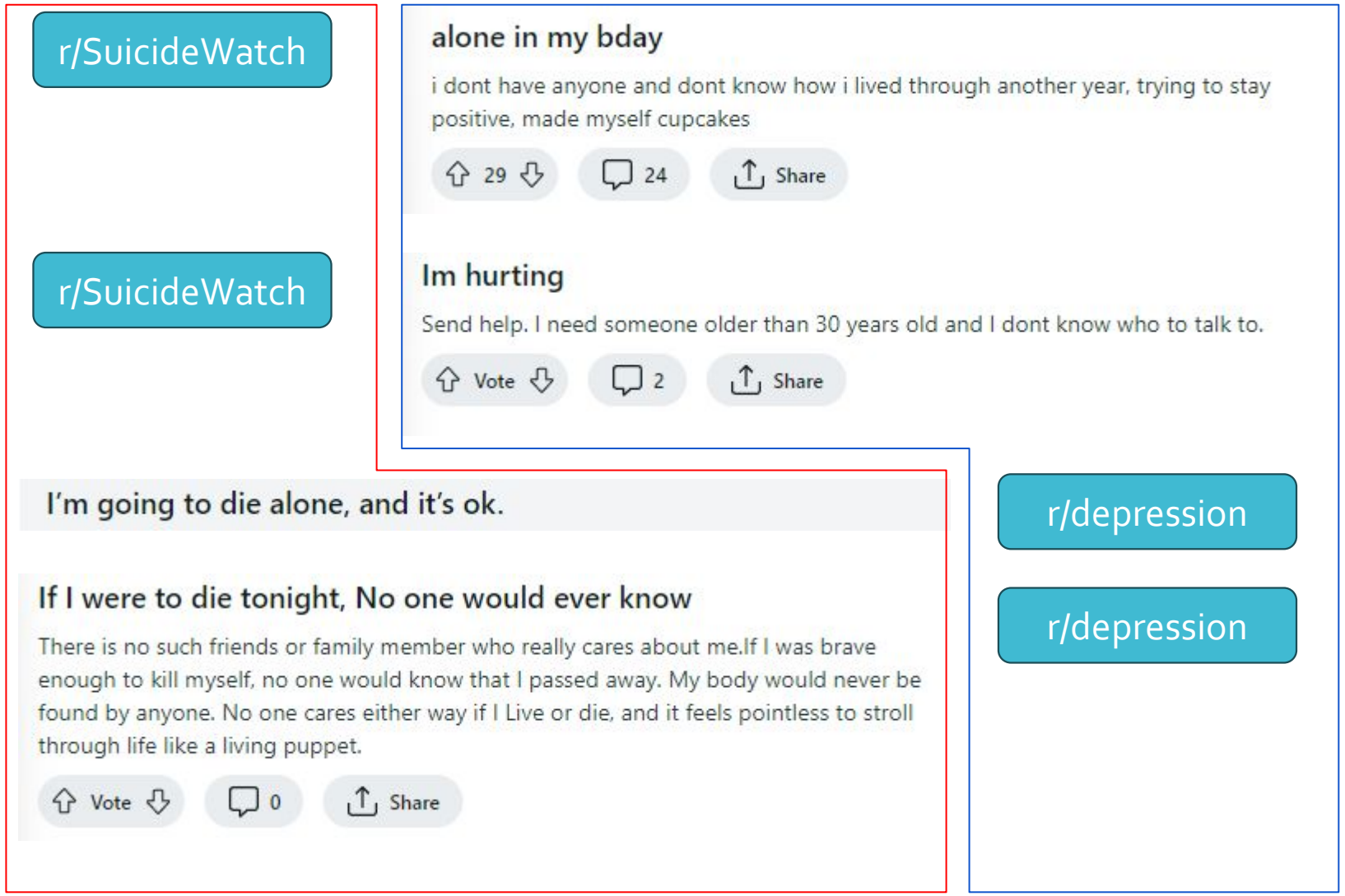

r/depression
Join
/r/depression, because nobody should be alone in a dark place
Peer support for anyone struggling with a depressive disorder.
971K
Members
636
Online
COMMUNITY BASICS
We offer a peer-support space for anyone dealing with a depressive disorder in themselves or someone close to them.
If you want to talk about thoughts or risk of suicide, please post at /r/SuicideWatch.
If you've lost someone to suicide, /r/SuicideBereavement is the best community to get support.


r/SuicideWatch
Join
Peer support for anyone struggling with suicidal thoughts
Peer support for anyone struggling with suicidal thoughts.
432K
Members
564
Online

Which subreddit?

r/SuicideWatch

r/SuicideWatch



r/depression

r/depression

Problem Statement

- Create a **classifier which identifies posts with suicidal ideations** to feed into an autosuggest mechanism on Reddit.
- Model as a "first pass" mechanism, with second pass by subreddit moderators
- Primary stakeholder: Reddit
  - Corporate Social Responsibility
  - Building and Supporting Communities
- Secondary stakeholders: Moderators and Members

Methodology

- **PRAW API**
  - ~5,500 posts
    - API limitations: **1000 post/day**
    - Scrapped over 1 week and glued together
    - Subreddits relatively active
  - 55% from r/SuicideWatch, 45% from r/depression
- **Word input**
  - Title + first 20 words + last 20 words
    - Resource efficiency
  - CountVectorizer, TFIDFVectorizer
    - Vectorise with specific stop words added

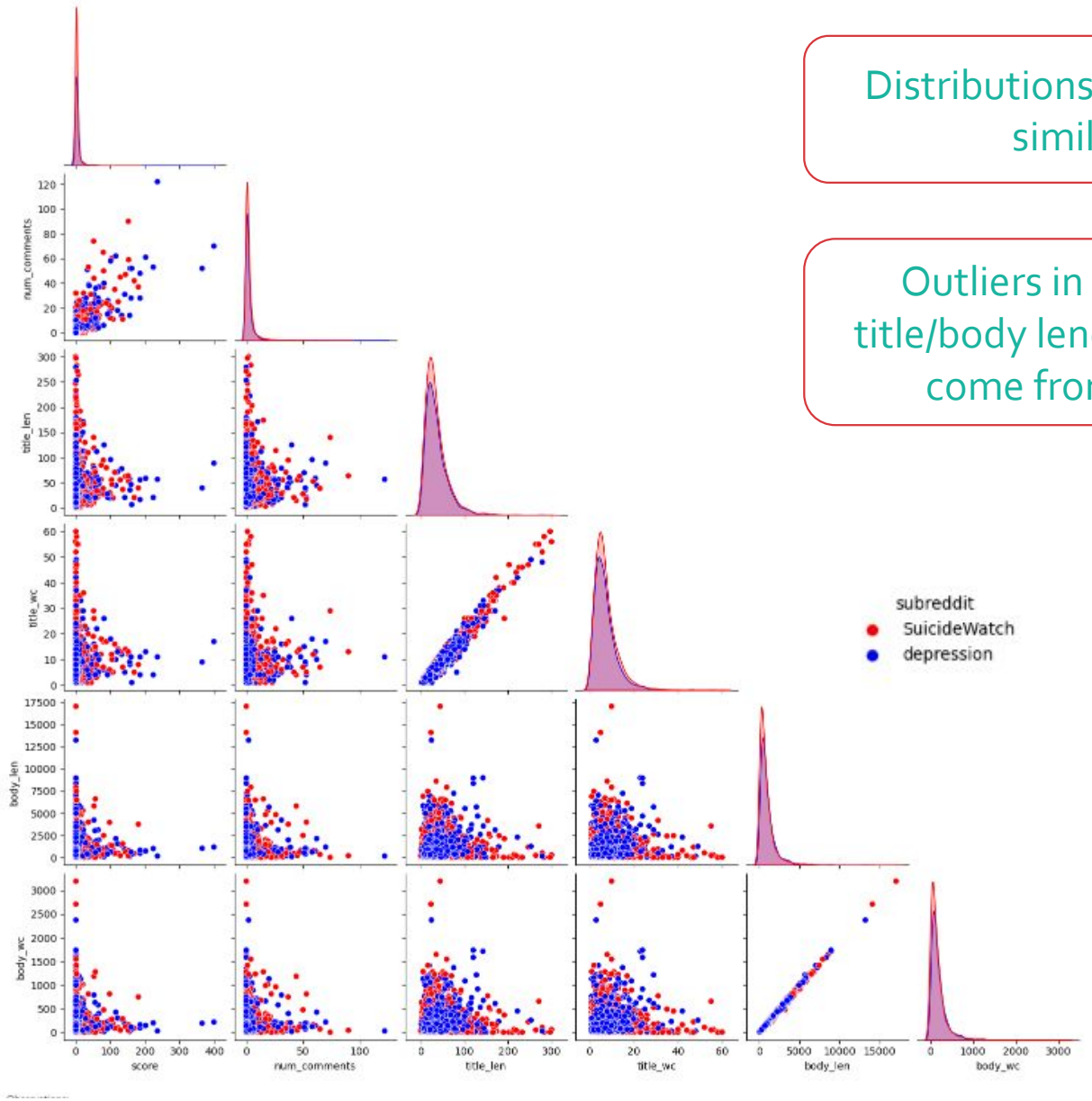
Methodology

- **Estimators:**
  - Naïve Bayes (Base)
  - Complement Naïve Bayes
  - Random Forest Classifier
  - ~~Boosting models~~
- **Evaluation:**
  - Recall
  - ~~F1 Score~~
  - Matthew's Correlation Coefficient

GridSearch parameters

- **Word vectorizers:** CVEC / TVEC
  - min_df: 1, 2
  - max_df: 0.9, 0.95, 0.98
  - max_feat: None, 5000, 7500
  - n_grams: (1,1), (1,2)
- **Stop words:** Custom stop words based on 'english' | None
- **Word processing:** Lemmatizer | Porter Stemmer

Initial insights

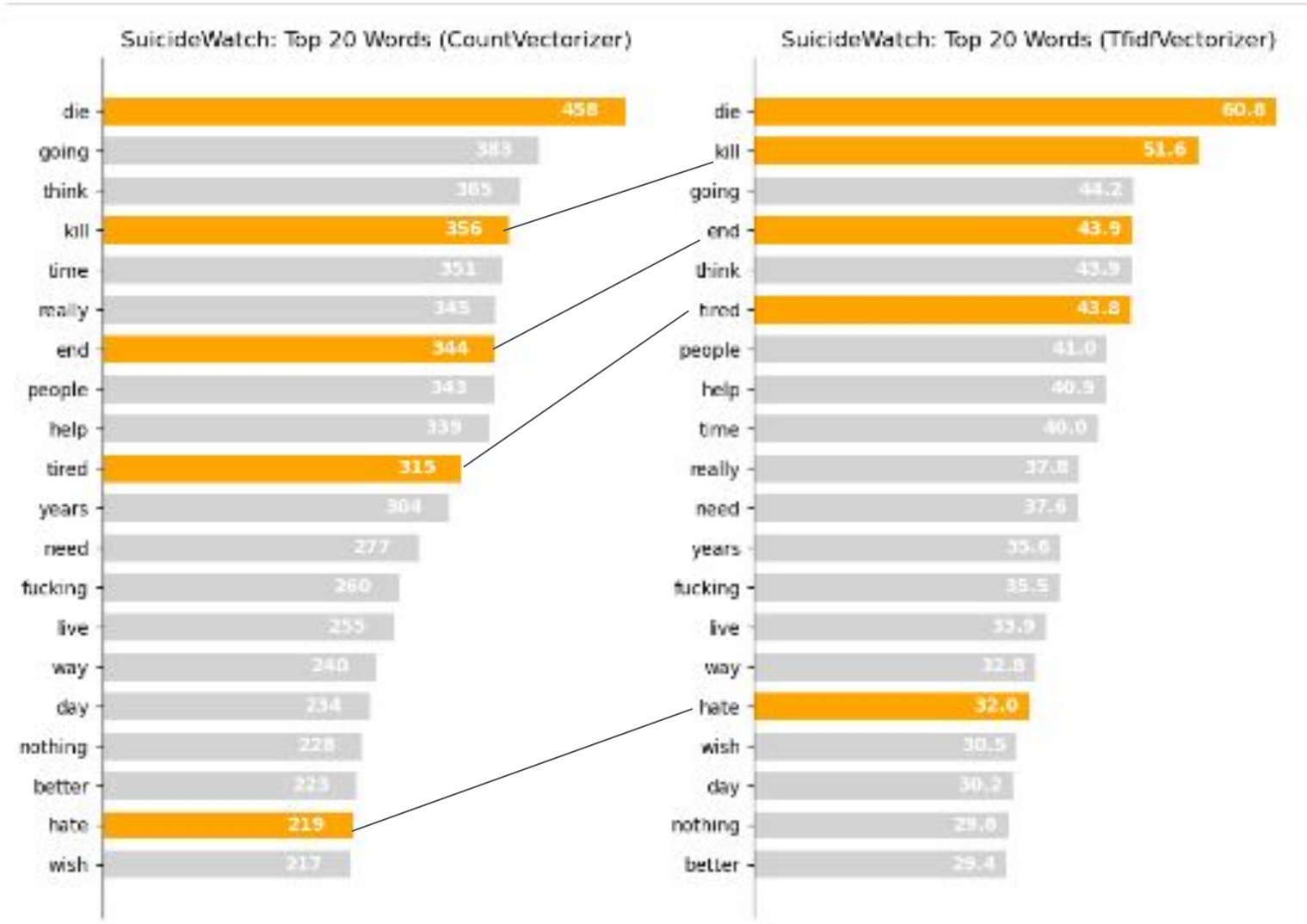


Distributions are pretty similar

Outliers in terms of title/body length tend to come from r/SW

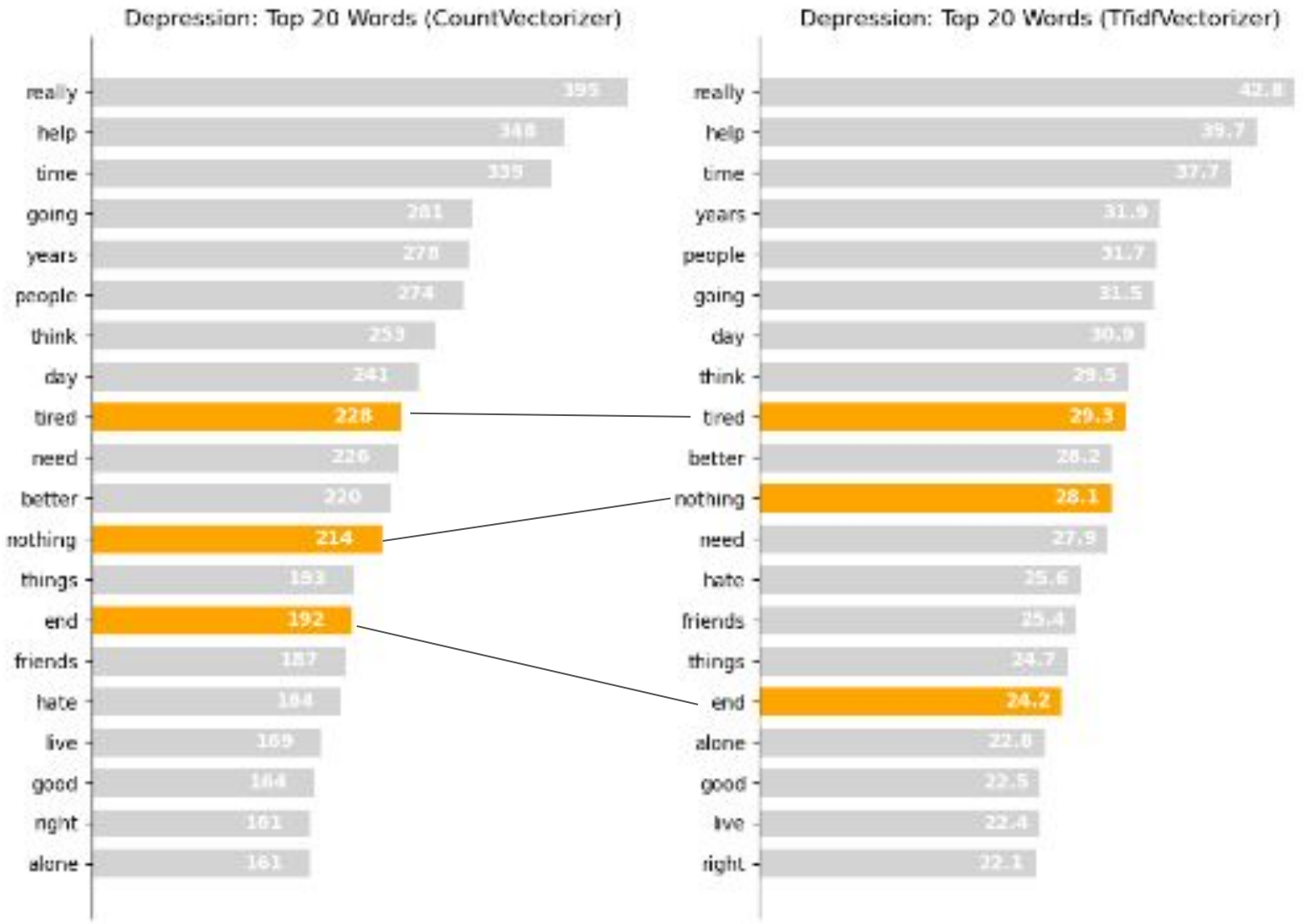
Top 20 words

- *CVEC vs TVEC*



Top 20 words

- *CVEC vs TVEC*

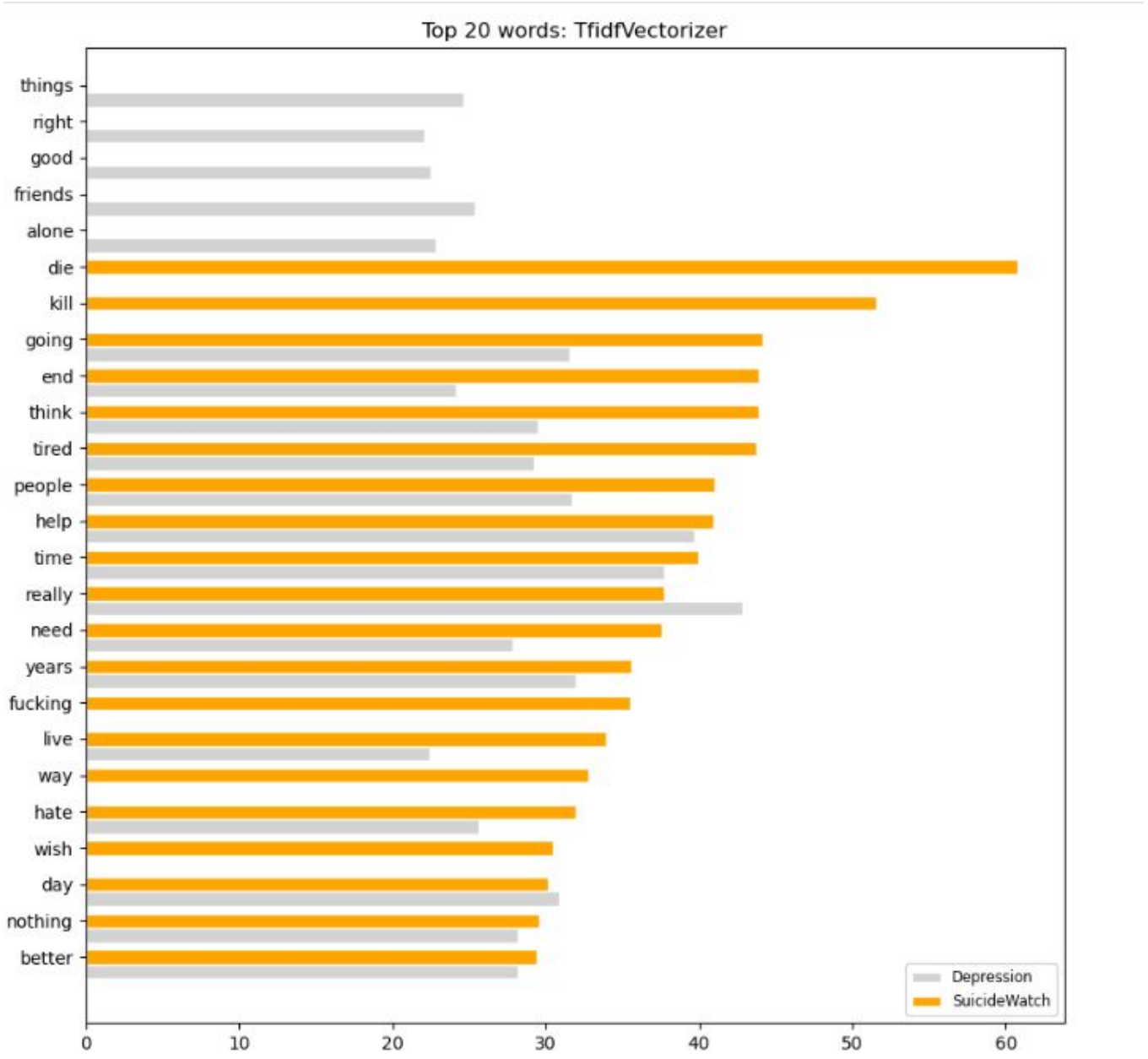


Top 20 words

SuicideWatch | depression

Top 20 words

SuicideWatch | depression



Model Results

Recall scores

Model	Train	Test
NB + CVEC	0.9907	0.7791
NB + TVEC	0.9996	0.9616
CNB + CVEC	0.9894	0.7712
CNB + TVEC	0.9978	0.9259
RF + TVEC	0.8243	0.7183

Matthew's correlation coefficient

Model	Train	Test
NB + CVEC	0.9428	0.3274
NB + TVEC	0.7438	0.2241
CNB + CVEC	0.9423	0.3323
CNB + TVEC	0.8307	0.2968
RF + TVEC	0.6657	0.2804

GridSearch:
Best parameters

- **Word vectorizers:** TVEC
  - min_df: 1
  - max_df: 0.95
  - max_feat: None
  - n_grams: (1,2)
- **Stop words:** None
- **Word processing:** Lemmatizer

Reducing Overfit

- **Stop words:** Forcing through stop words + PorterStemmer

	Recall		Matt's Coeff	
	Train	Test	Train	Test
Stop words + Stem	0.9938	0.8585	0.9407	0.3041
Original	0.9978	0.9259	0.8307	0.2968

- **Logistic regression**

	Recall		Matt's Coeff	
	Train	Test	Train	Test
Log regression	0.8093	0.6362	0.6509	0.3284
Original	0.9978	0.9259	0.8307	0.2968

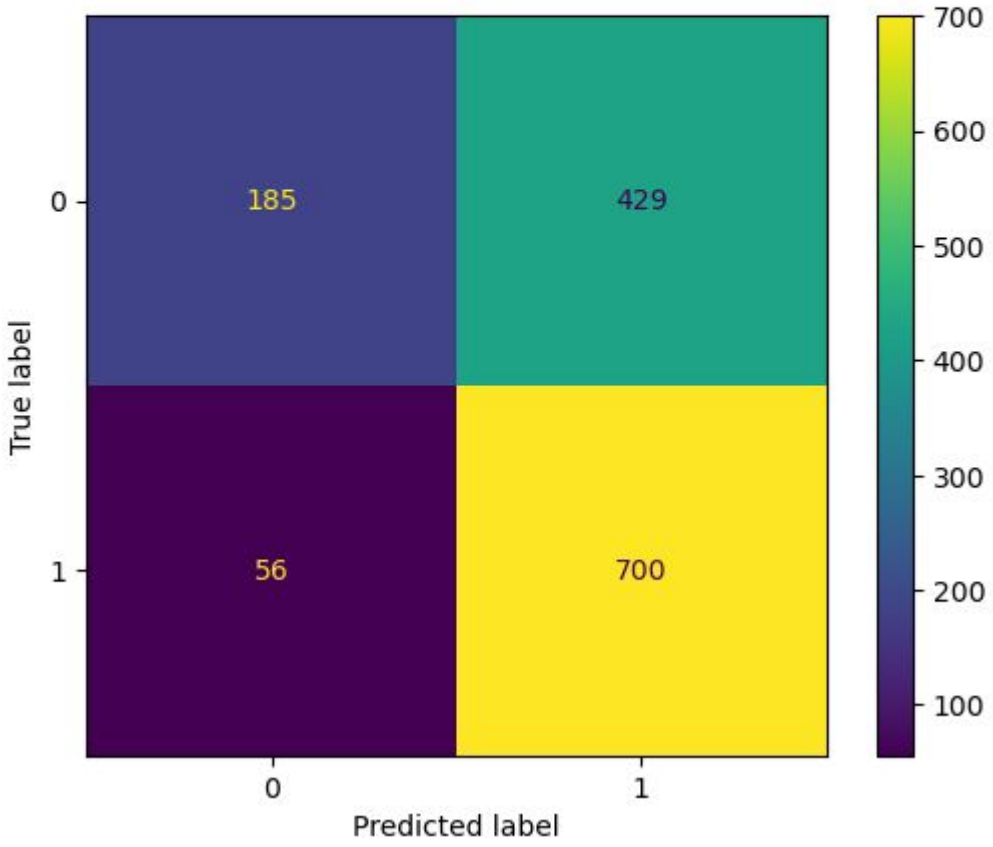
Original CNB-TVEC chosen

Conclusion

- CNB-TVEC

Baseline
SW: 55%
Dep: 45%

	precision	recall	f1-score	support
0	0.77	0.30	0.43	614
1	0.62	0.93	0.74	756
accuracy			0.65	1370
macro avg	0.69	0.61	0.59	1370
weighted avg	0.69	0.65	0.60	1370



1- r/SW
0 - r/dep

Further work

- Continued iterative assessment
  - With each iteration of model learning with better labelled data, results need to be re-reviewed
- Better labelled data would allow:
  - Boosting models
  - SVMs (Support Vector Models)
- Differentiating between suicidal ***intent*** and suicidal ***ideation***