

Addressing Mental Health in the Digital Age

Apr 2019

Mental Health Trends:

300 Million

Global Depression Rate

16 Million

US Depression Rate

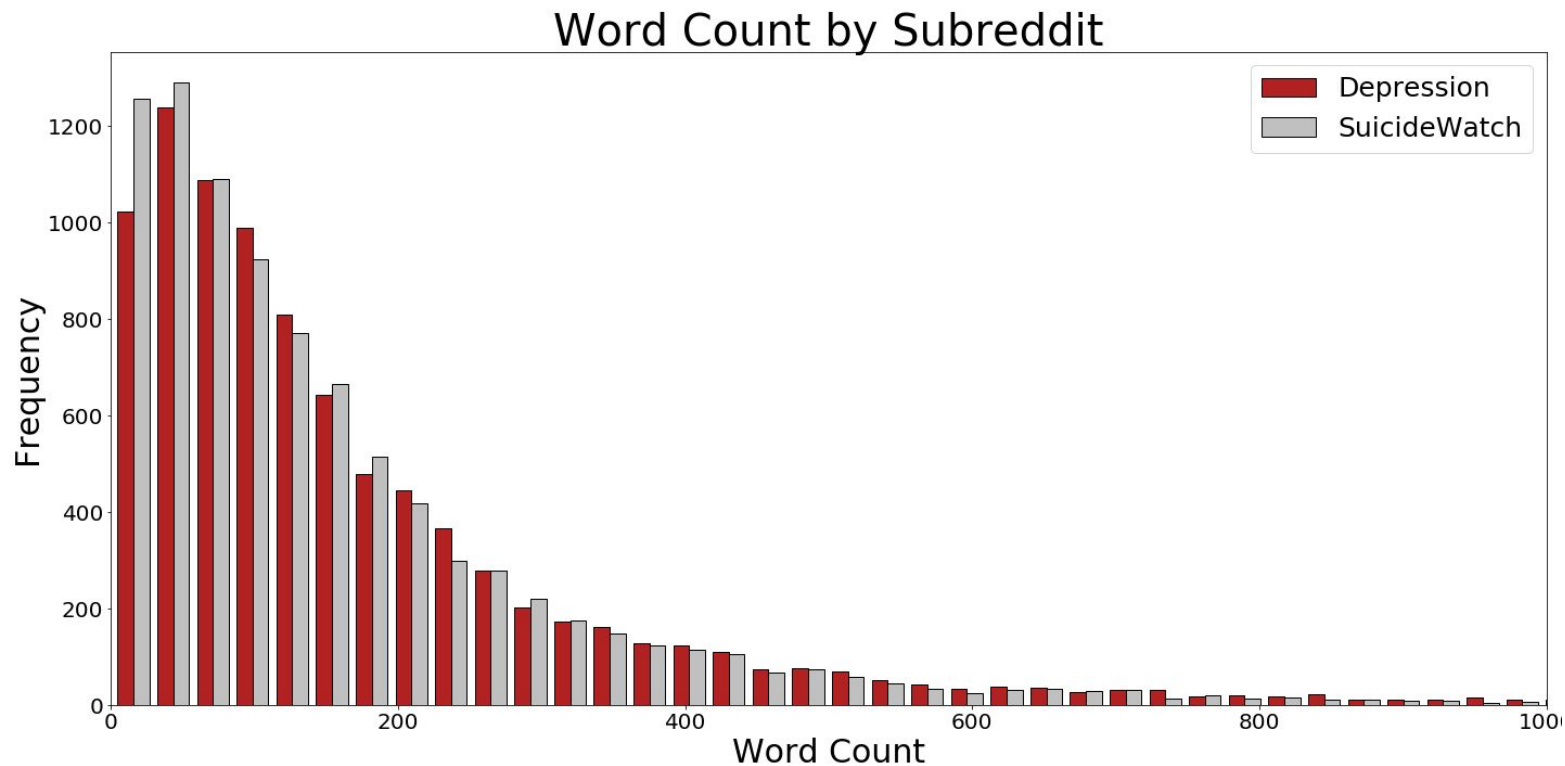
1 every 40 Sec.

Suicide Rates

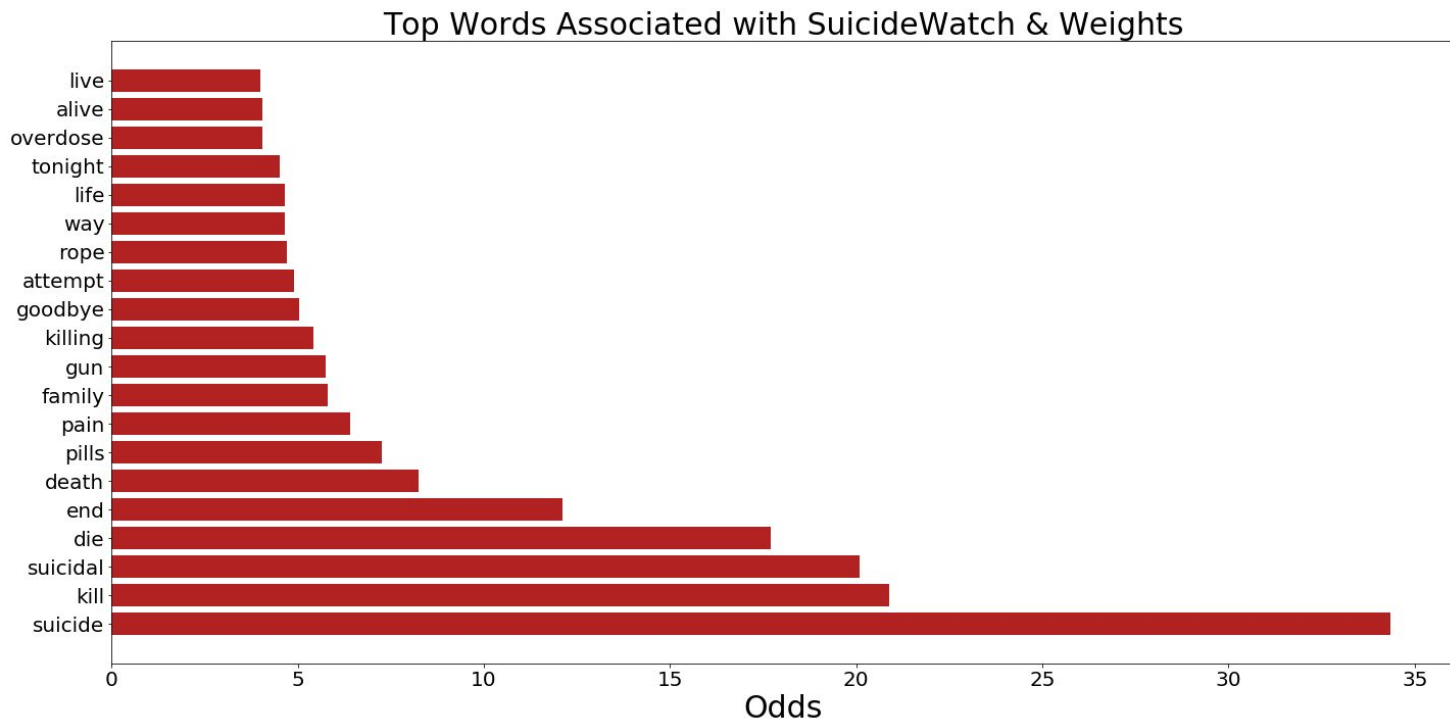
Problem Statement

- Can we identify the language/ language themes that are common for people dealing with depression?
- Can we identify the language/ language themes that are common for people that are contemplating self harm?
- Can we accurately distinguish between individuals struggling with depression vs those that intend to harm themselves?

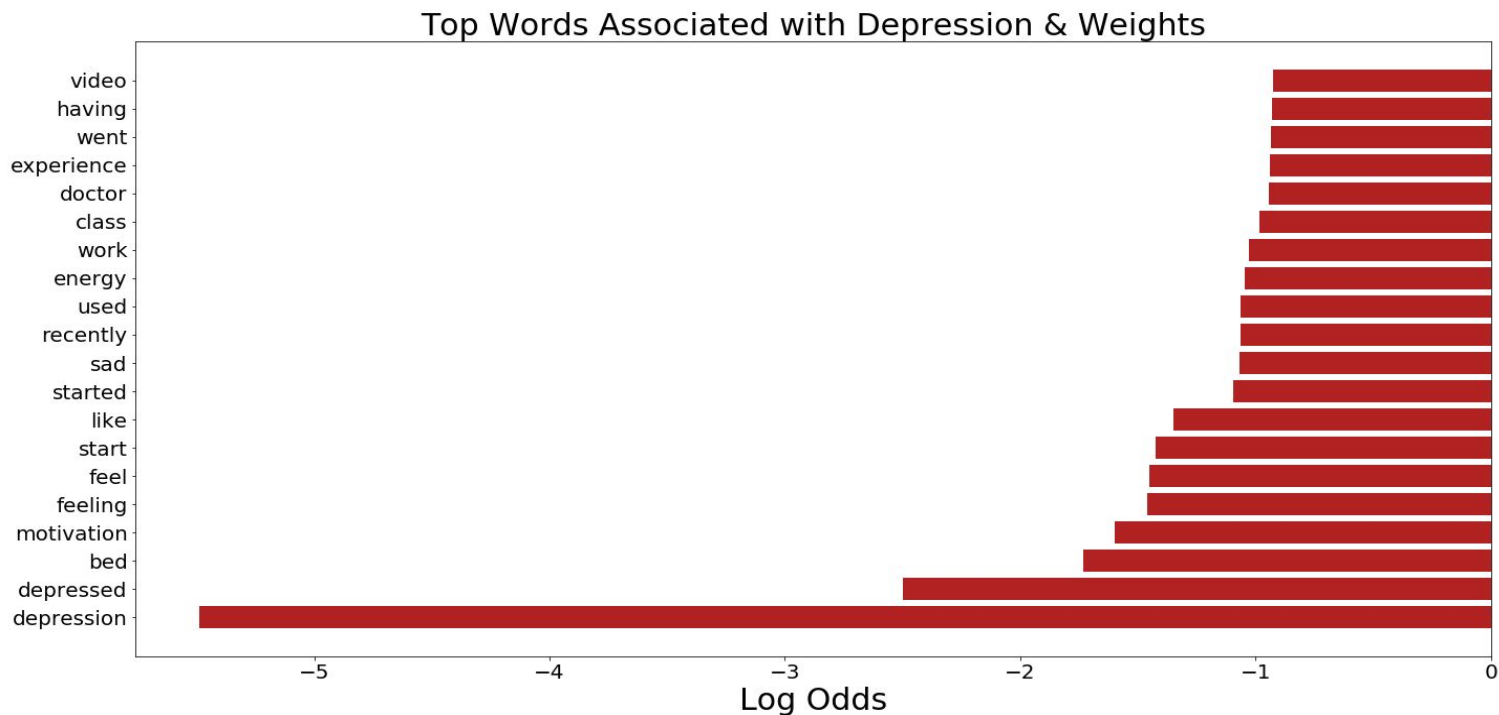
Word Count Distribution



Logistic Regression Overview

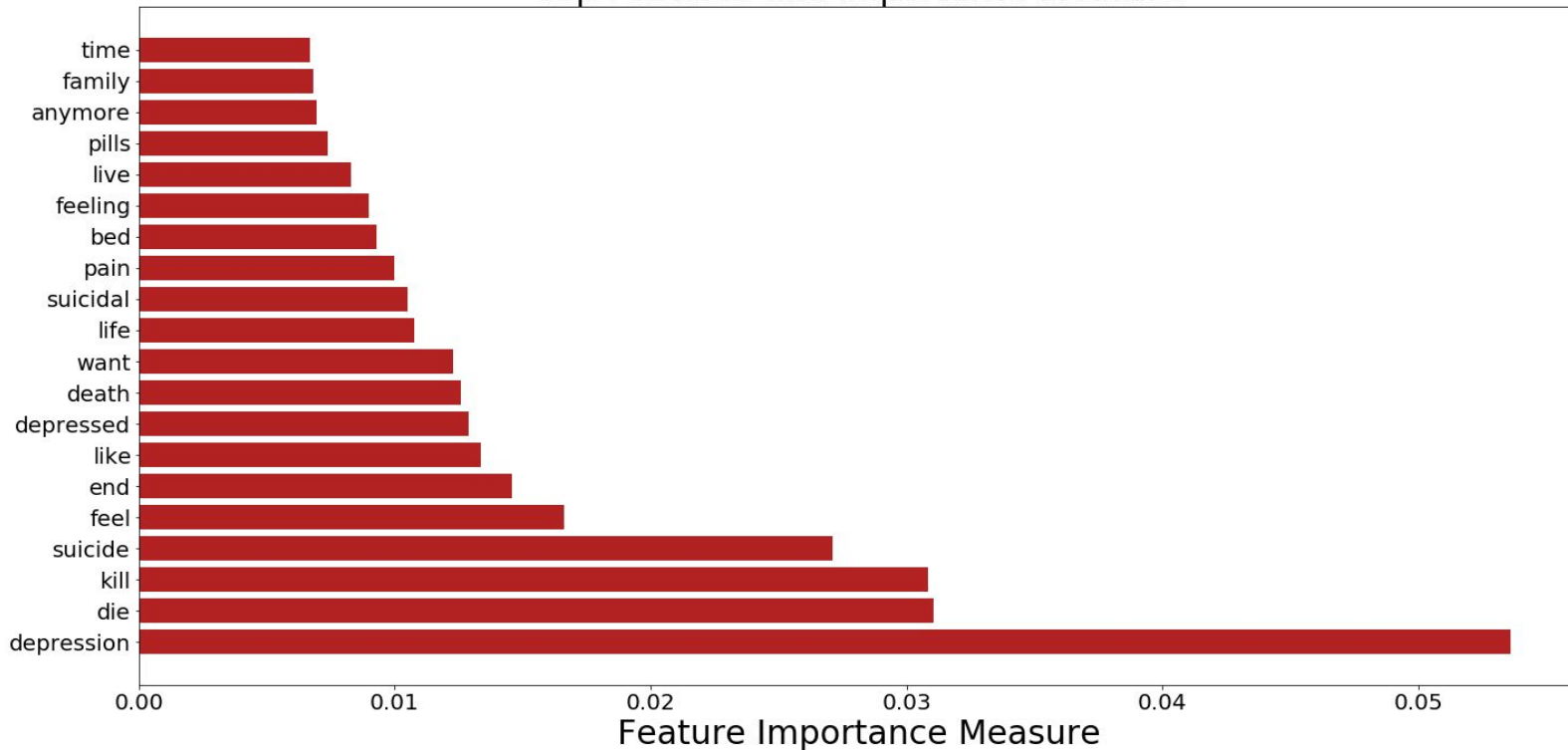


Logistic Regression Overview (cont.)



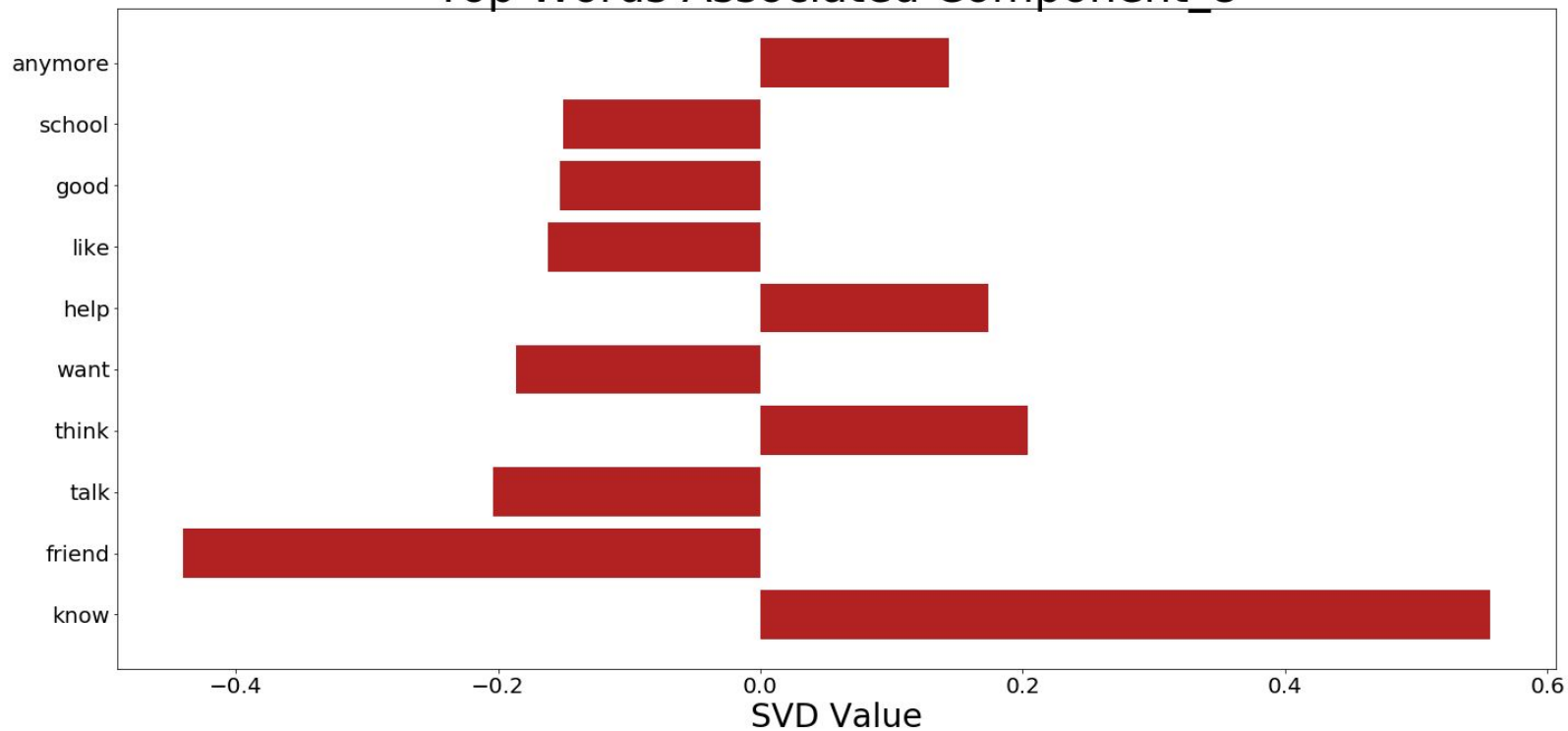
Random Forest Feature Importance

Top Features and Importance Measure



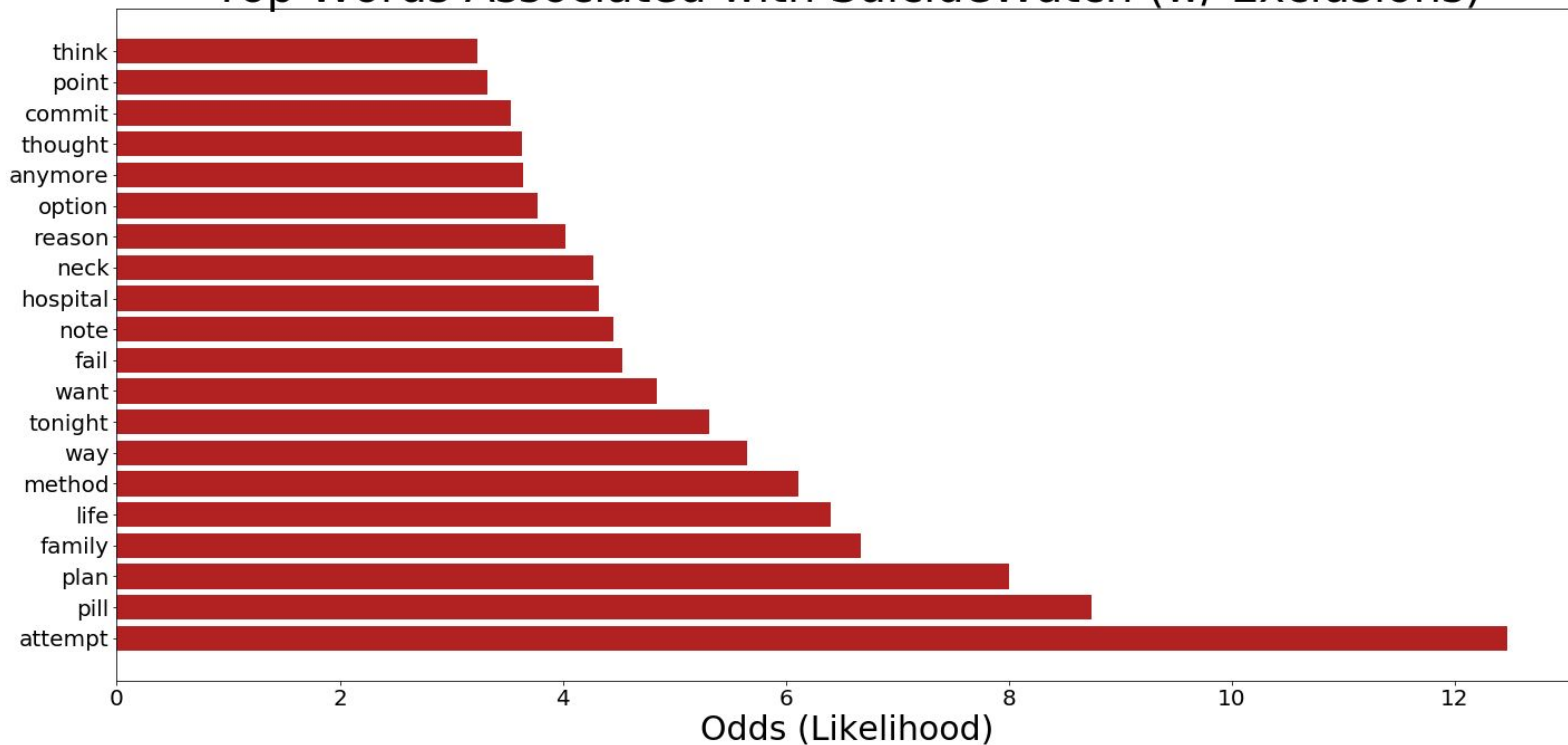
Topic Modeling - LSA | SVD

Top Words Associated Component_8

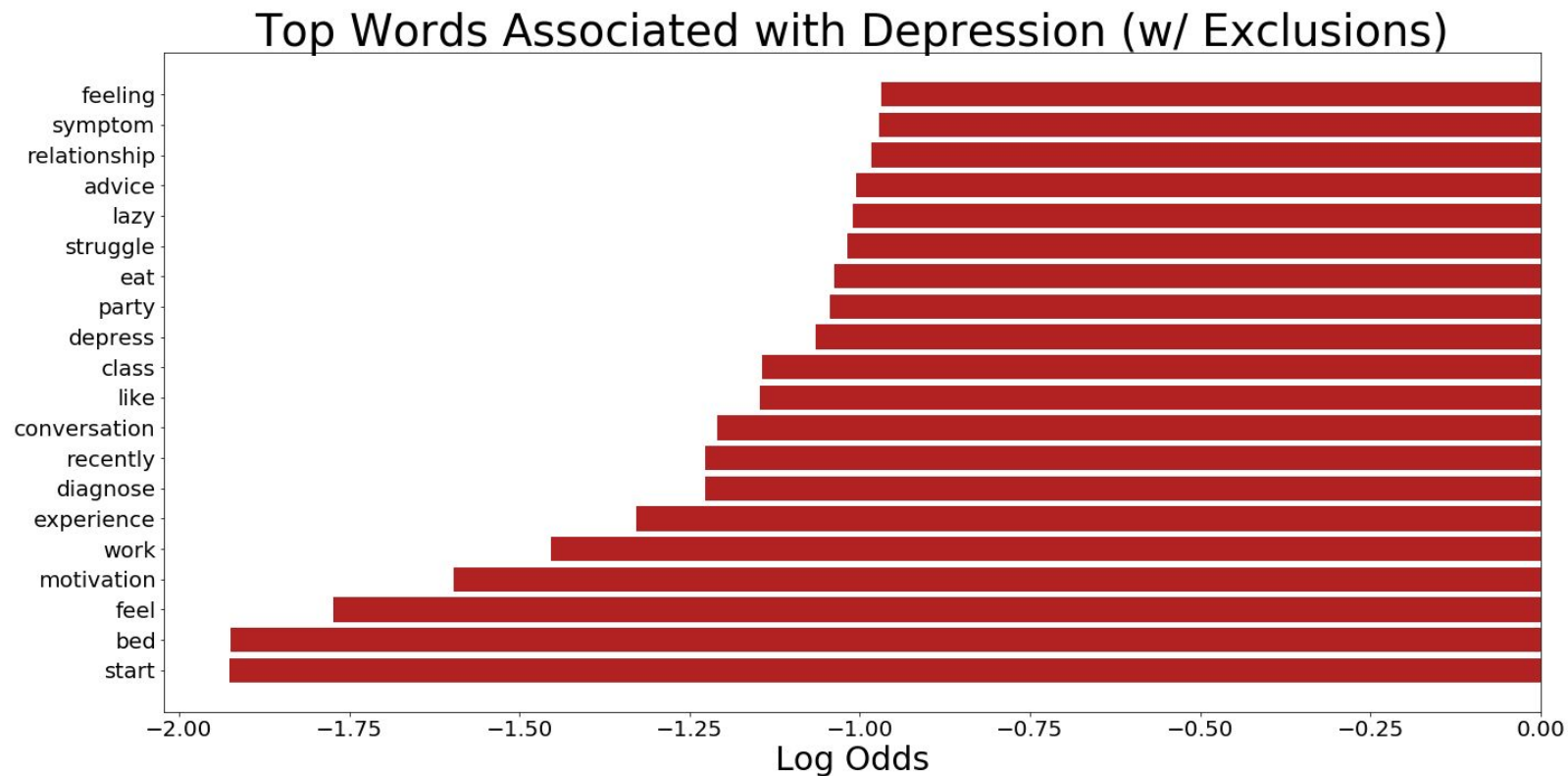


Logistic Regression with Exclusions

Top Words Associated with SuicideWatch (w/ Exclusions)



Logistic Regression with Exclusions



Recommendations & Next Steps

- ❑ Continue to test other models to improve accuracy (XGBoost, SVM)
- ❑ Use model to A/B test ad targeting on social sites like Facebook and Twitter
- ❑ Test using the most predictive features for rule-based ad or bot targeting
- ❑ Revisit topic modeling techniques