Sentiment Analysis of Violent Radicalization on Social Media

Anna Singley[1]

[1]Central Catholic High School

Abstract

Identifying radicalized individuals in a timely manner grows increasingly important against the backdrop of the January 6th Capitol attacks and growing political polarity. By examining radicalization as a multifaceted psychological process, several key variables appear as indicative of radicalization. Identifying these variables with any real-time capabilities is difficult without automation. Natural language processing is a tool that can be used to identify psychological variables associated with radicalization, allowing for an automated means of collecting quantitative data from a qualitative source. Applications of nonlinear dynamics and time-series analysis can then be applied to this quantitative data to reveal underlying patterns and trends in the process of political radicalization.

Problem Background

The process of radicalization is a complex one with widely diverging theory regarding the necessary psychological traits to enable political violence. However, adapting commonly accepted theory allows for a relatively universal model of radicalization while creating tools that later may be used to quantify the efficacy of various radicalization models. Sentiment analysis allows for the quantification of qualitative data in an automated fashion – a trait that could prove invaluable to future research.

Psychological Variable	Definition
Resentment	Strong negative opinion directed towards an ideology, individual, or organization
Narcissism	Dangerously heightened sense self-importance that leads to a disregard of others
Capacity for Violence	Ability to enact violence or cause destruction
Social Alienation	Feelings of disconnection from social groups leading to a disenchantment with societal rules
Hopelessness	Feelings of abandonment and despair that allow for individuals to behave in a highly irrational manner
Fear	Concern or worry that enables drastic behavior

Support for the inclusion of narcissism, social alienation, and resentment can be sourced from the theory of Significance Quest, whereas backing for the inclusion of social alienation, hopelessness, and fear can be found from Focal Goal commitment theory. Support for capacity for violence can be found again from the theory of Significance Quest.

Future work may include a model of radicalization based on necessary and sufficient conditions to establish a greater understanding of the mechanics and dynamics of radicalization.

Data Collection

Sentiment analysis requires a large corpus of labeled training data, similar in linguistic structure to the data the model will be used upon. Acquiring such a corpus with accurate labels is a challenging aspect standing in the way of further application of sentiment analysis models. (cite Mohammed). There is additional difficulty associated with obtaining data for finely categorized and complex emotions or sentiments.

The collection of tweets by hashtag, wherein the used hashtag indicates the emotion or sentiment present in the document, allows for the creation of a relatively accurate, finely defined dataset of training tweets. The table below shows some of the data collected with this method.

Tweet ID	Sentiment	Tweet Content
3.75E+04	strong negative	Trump in conceding betrayed you. They are not like the common honest Christian folk
3.06E+04	hopelessness	When contextualized amid the two most unparalleled propaganda machines of the modern world there's no hope of any commentary being made in total good faith no one can hope to level a fair criticism without endlessly bouncing back and forth
8.50E+03	social alienation	Another relative gone. One more person I'll never get to see again and we're forced to grieve alone on phone. I hate being stuck alone #COVID19 #Lonely #lockdown

Model Selection and Mathematical Justification

Support Vector Machines (SVMs) are a rudimentary machine learning model that employs support vectors to create a canonical hyperplane between classes of data (Eq. 1). Each sentiment is classified against an opposite in a one vs one pipeline to allow for statements with multiple sentiments, such as a tweet that shows both resentment and narcissism.

Due to the highly nonlinear nature of the arbitrary plane on which data is plotted, a radial basis function kernel (Eq. 2) is used so that a linear support vector machine remains functional and applicable to the data. Platt scaling was used to obtain a continuous value from discrete classifications, allowing for comparisons in the strength of sentiments.

$$\operatorname{sgn}\left(\sum_{i=1}^{l} y_i \alpha_i K(x_i, x) + b\right) \qquad (1)$$

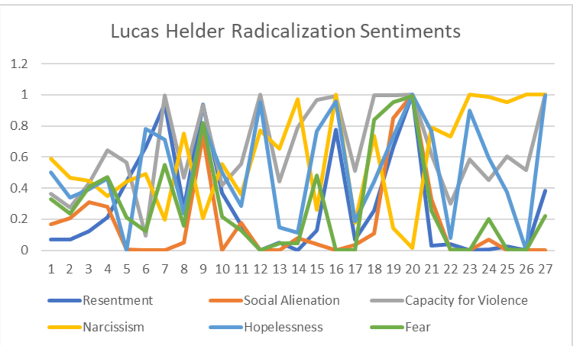
$$k(x_i, x_j) = \exp\left(\frac{\|(x_i, x_j)\|^2}{2l^2}\right) \qquad (2)$$

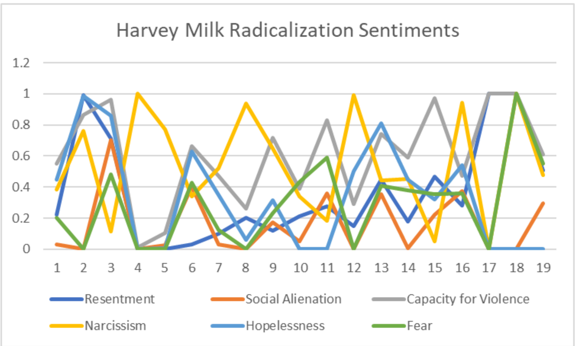
Abbreviated Results

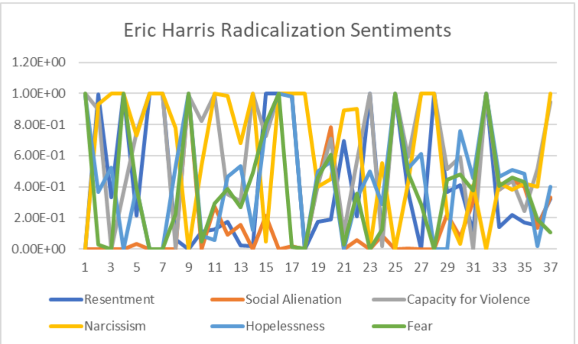
In order to evaluate the efficacy of machine learning models, it is necessary to test data that the model has not been exposed to. With this in mind, several terrorist manifestos were analyzed and compared with speeches, poems, and other writings from known pacifists. These pacifists held strong political or religious beliefs (a necessary trait violent radicalization) but lacked the will to bring about such violence.

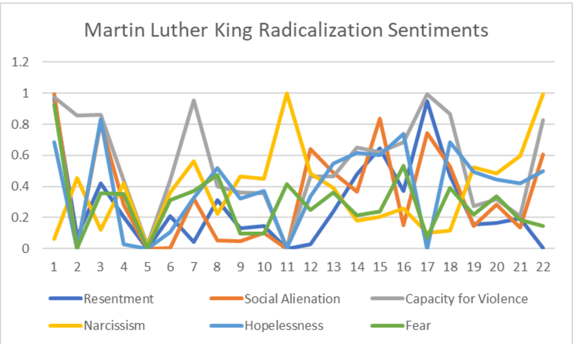
The following figures demonstrate the automated sentiment classification and quantification of the manifestos of several US domestic terrorists by paragraph compared with speeches from political figures with pacifistic beliefs.

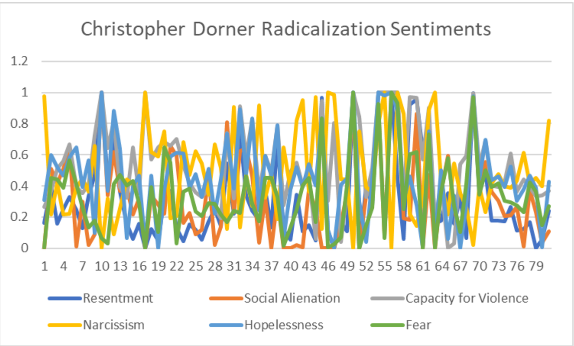
Note the differences in scale between manifestos – the results are not perfectly analogous but nevertheless applicable for future modeling purposes.

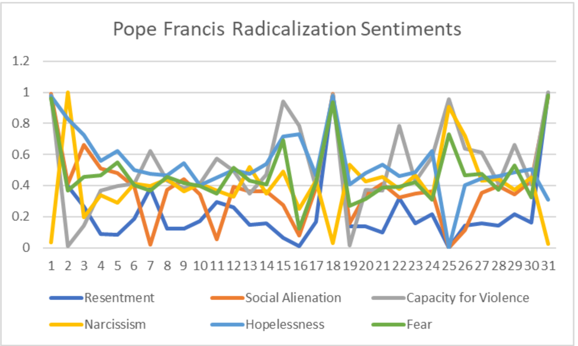












Individualized Progression Model

Examining the process of political radicalization as a dynamic system with multiple stable states gives way to new modeling potential. A model as simple as the one proposed by Noy-Meir (Eq. 3) illustrates the concept at hand. Instead of viewing simply radicalization as an attractor, envisioning both pacifism and violent radicalization as dueling attractors allows for a quantifiable and mathematical understanding of the process of radicalization.

$$\frac{dV}{dt} = G(V) - c(V)H \qquad (3)$$

Original, 1975 Noy-Meir Alternative Stable State equation for the rate at which vegetation is consumed by animal-life. This gives two stable states or attractors – created by two dueling forces. A more complex, but similarly structured model could prove applicable to violent radicalization once more data is acquired.

Using data from sentiment analysis of social media will allow for the justification of a more heterogenous and multifaceted model of radicalization.

Conclusions and Future Work

There are a wide variety of potential applications for sentiment analysis in the fields such as quantitative psychology or applied linguistics. Natural language processing techniques can be applied to problems ranging from detection of emotion, to automated topic analysis. The issue of political radicalization presents both as a grave problem, and one that can be predicted and mathematically modeled with automated sentiment analysis.

Future work is planned to create a larger corpus of manifestos and journal entries – including those that show progression from a stable psychological state into a potentially violent, unstable state. Critical slowing down is a nonlinear dynamics concept associated with the progression and actualization of suicide, in addition to ecological phenomena such as population extinctions, or deaths of ecosystems. The application of critical slowing down via metrics such as autoregressive conditional heteroskedasticity, autocorrelation, or even simple variance can serve as early warning signs of a phase transition before an individual is likely to cause harm.

The application of a quantitative model of radicalization is a novel development to the field that could contribute to a greater theoretical understanding of political violence.

References

Eubank, W., & Weinberg, L. (2001). Terrorism and democracy: Perpetrators and victims. *Terrorism and political violence*, *13*(1), 155-164.

Jasko, K., LaFree, G., & Kruglanski, A. (2017). Quest for significance and violent extremism: The case of domestic radicalization. *Political Psychology*, *38*(5), 815-831.

Zainuddin, N., & Selamat, A. (2014, September). Sentiment analysis using support vector machine. In *2014 international conference on computer, communications, and control technology (I4CT)* (pp. 333-337). IEEE.

Trappenberg, T. P. (2019). Machine learning with sklearn. In *Fundamentals of Machine Learning* (pp. 38-65). Oxford University Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, 2825-2830.

Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, *10*(3), 61-74.

Scheffer, M., & Jeppesen, E. (1998). Alternative stable states. In The structuring role of submerged macrophytes in lakes (pp. 397-406). Springer, New York, NY.

Noy-Meir, I. (1975). Stability of grazing systems: an application of predator-prey graphs. The Journal of Ecology, 459-481.

Nazarimehr, F., Jafari, S., Perc, M., & Sprott, J. C. (2020). Critical slowing down indicators. *EPL (Europhysics Letters)*, *132*(1), 18001.

van de Leemput, I. A., Wichers, M., Cramer, A. O., Borsboom, D., Tuerlinckx, F., Kuppens, P., ... & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, *111*(1), 87-92.

Dakos, V., Van Nes, E. H., d'Odorico, P., & Scheffer, M. (2012). Robustness of variance and autocorrelation as indicators of critical slowing down. Ecology, 93(2), 264-271.