

Hold Your Tongue!: Sentiment Analysis and Classification in the Legal Field



Sloan C. Stinnett
University of Oklahoma

Introduction

Introduction

Brief Literature Review

Data

Source

Contents

Methods

Sentiment Estimation

Classification

Findings

Preliminary Findings

Conclusion

Bibliography

Introduction

- ▶ Rhetoric has a deep historical connection with law.
- ▶ Psychologically we are predisposed to alter our opinions on factual matters based on subjective information including presentation.

- ▶ "Sentiment in Case Law" [2] by Danial Hoadley
 1. [5] In this paper the authors generated a general model for predicting supreme court decisions using only pretrial data. They manage to get a success rate of 70.2 percent well above the 60 percent range of human legal experts.
 2. [7] In this paper Sulea et al use classifier ensembles to predict: the area of law the case was in ,the judgement in the case, and the era that the case came from.
 3. [6] In this paper Liu and Chen Introduce the use of sentiment score as a feature in judgement classification

Data

- ▶ This data set was originally collected for "Echoes of Power: Language Effects and Power Differences in Social Interaction". [4]
- ▶ The outcome and vote data was originally collected from the Speath Supreme Court database. [3]
- ▶ Speech information came from the online releases of the oral arguments. [1]

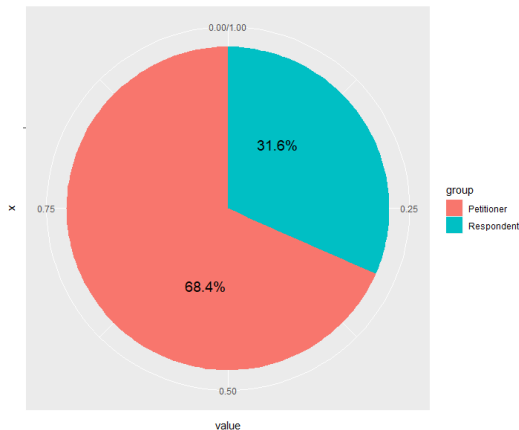
- The data set contained 51,000 utterances, for each utterance there was: a case ID, an utterance ID, the name of the speaker, whether they were a justice, if they were a justice how they ended up voting, whether the utterance followed the previous in a natural conversation, what side the speaker was associated with either petitioner or respondent, and who the case ultimately was found in favor of.

Methods

- ▶ For the purposes of sentiment analysis I used the nrc sentiment dictionary in the TidyText package for R. The dictionary files words into 10 different sentiments
- ▶ anger, disgust, anticipation, joy, fear, sadness, surprise, disgust, positive and negative.

Following the literature I will be testing two classification methods support vector machine and random forest. This classification task will involve two classes respondent wins and petitioner wins using the proportion of each sentiment in the speech, whether the speech was made by a representative of the respondent or petitioner and length of speech as features.

Findings



One interesting point is that Petitioners are far more likely to succeed in this sample than Respondents. This is consistent with the "always bet reverse" motto that some legal scholars have taken up.

Conclusion

- ▶ Practical applications in speech creation tools improving the sentiment profile of speech while they are being written.
- ▶ proof of concept for further research into the usefulness of sentiment analysis in the legal field and possibly other fields that are heavily reliant of persuasive speaking.

Bibliography

- [1] Oral arguments - supreme court of the united states.
- [2] Sentiment in case law.
- [3] The supreme court database.
- [4] Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. Echoes of power: language effects and power differences in social interaction. In *Proceedings of the 21st international conference on World Wide Web - WWW '12*, page 699. ACM Press.
- [5] Daniel Martin Katz, Michael J. Bommarito, and Josh Blackman. A general approach for predicting the behavior of the supreme court of the united states. 12(4):e0174698.
- [6] Yi-Hung Liu and Yen-Liang Chen. A two-phase sentiment analysis approach for judgement prediction. 44(5):594–607.
- [7] Octavia-Maria Sulea, Marcos Zampieri, Shervin Malmasi, Mihaela Vela, Liviu P. Dinu, and Josef van Genabith. Exploring the use of text classification in the legal domain.