

DRAFT

Unsupervised Text Classification With Chinese Social Media An Extension of King, Pan, and Roberts (2017)

Abstract

<150 words

I. Introduction

- Background on the 50c party. State of knowledge before KPR (2017) and since.
- Other quantitative approaches to studying the 50c party and Chinese internet “opinion guidance.”
- Summarize King, Pan, and Roberts (2017), specifically data source and results.
- Try to improve figures if I can replicate the original ReadMe results (but probably don’t have the human-coded posts).
- Describe impact/significance of the paper.

II. Extension

- Motivation/rationale for why the reassessment is warranted.
- Categories are informed by existing knowledge but still somewhat arbitrary...
- Define research question.
- How unsupervised text classification works (general approach) vis-à-vis ReadMe.
- Why I picked certain methodologies.

IDEAS * Automate + boost ReadMe (run it many times with different classifications?) *
Test my results with the out-of-sample batch. (knownWeibos_zg)

III. Results and Discussion

- What the findings were.
- Implications of findings.
- Directions for future work.

Acknowledgements

References