

# Sentiment Based Trading Algorithm

—

Can Twitter be mined for profit?

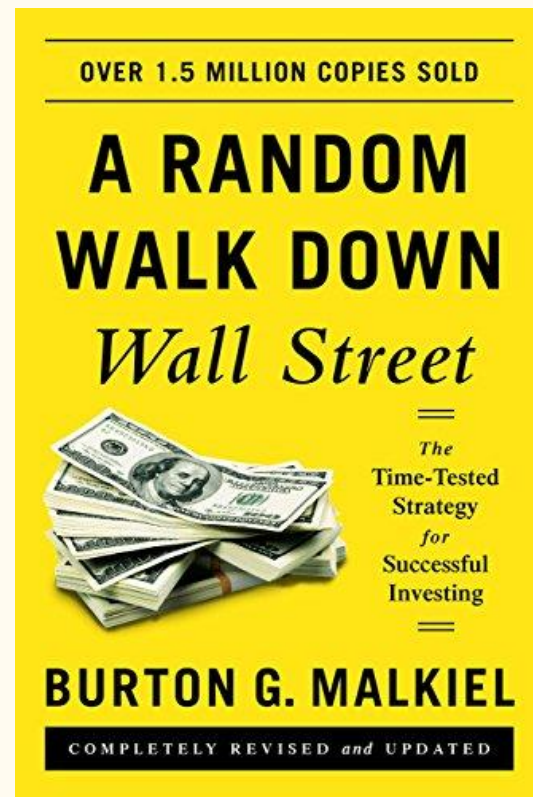
**Maybe.**

# VADER Sentiment Analysis

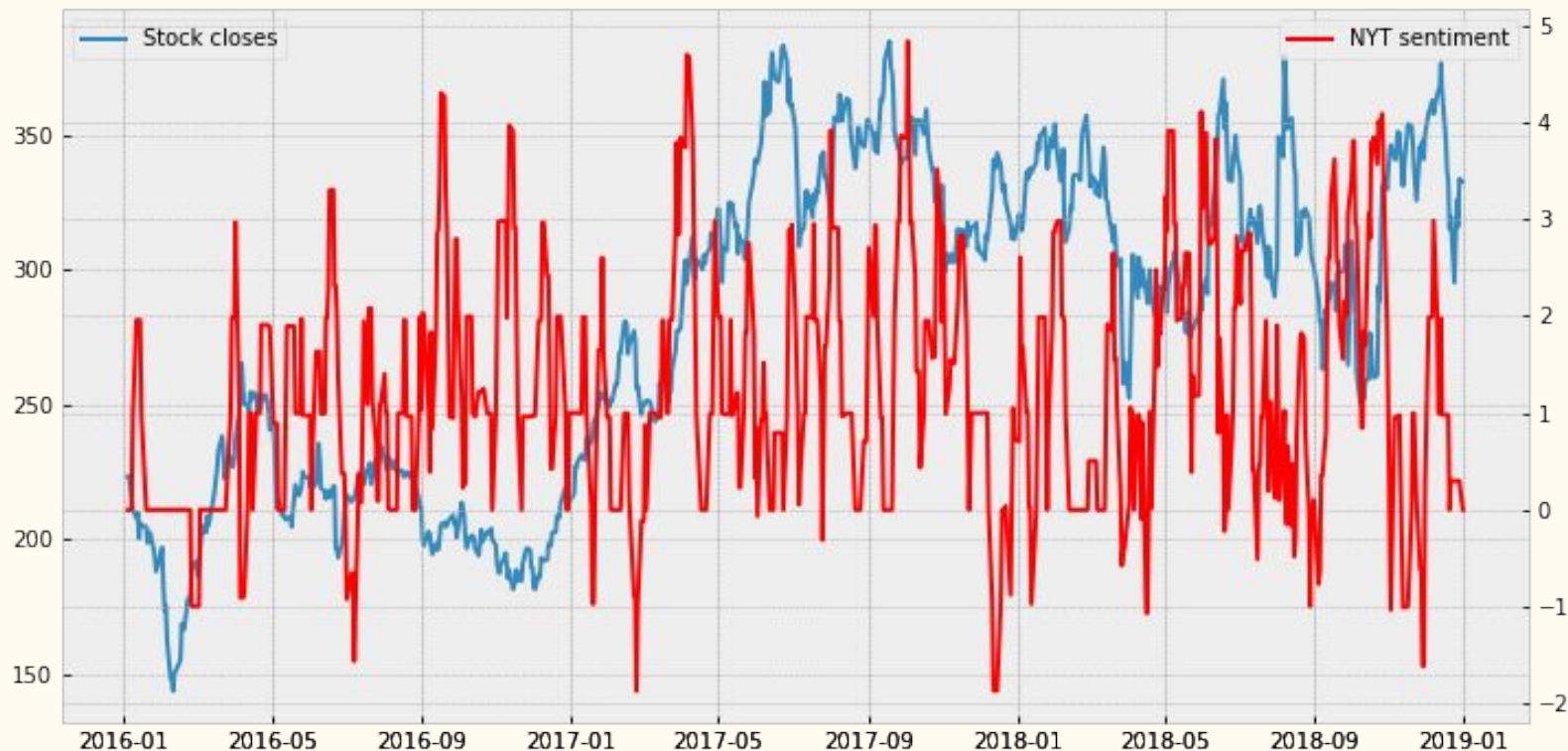
- Valence Aware Dictionary and sEntiment Reasoner
- ‘Lexicon’ and ‘Rule-based’
- Lexicon:
  - Built on a library of words/tokens that were rated by a group of human judges according to the sentiment expressed: from -4 (very negative) to 4 (very positive)
  - Includes slang/alt spellings as well as non-word tokens (the frowning emoticon :( has a rating of -2.2)
- Rule-based:
  - VADER has rules for handling sentences as whole structures - e.g. ‘not good’ is negative despite the word ‘good’
  - Pays attention to punctuation
- Pretty robust results even when applied to ‘messy’ text, such as from social media

# Sentiment as a financial metric

- Is it possible to glean systemically actionable insights from public data? Maybe...
- Looking primarily at tech/startup type companies where valuation doesn't seem tied specific financial metrics
- Non-conventional data sources:
  - Require compiling
  - Require processing/parsing
- Avoiding reverse causation



# Sentiment vs Stock close prices



# Model Building

- Hard to see, but sentiment does appear to be linked to price performance:
  - Average next day price performance: .04%
  - Next day price performance following high NYT sentiment: .2%
  - Next day price performance following low NYT sentiment: -.4%
- ‘Effect’ seems to be non-linear once viewing NYT and Twitter sentiment together
- Stock movement data inherently noisy
- Created classification model with a trading strategy in mind: above average and below average performance

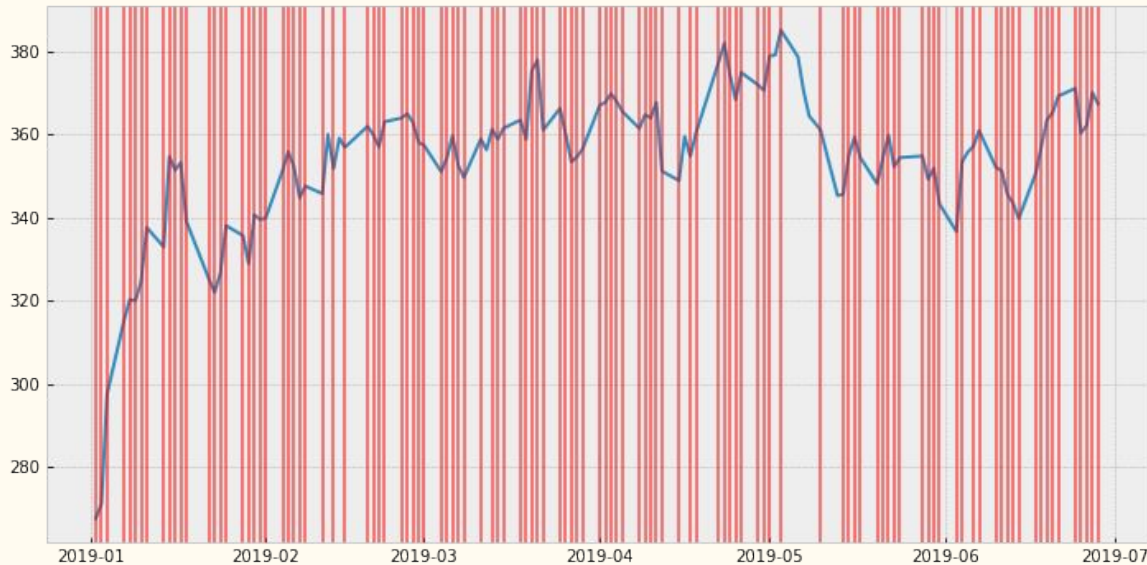
# Model performance: Tesla

- Trading Tesla based on my best model would net you \$39
- But Tesla stock performed horribly in the first half of the year!
- Buy and hold strategy over the same period earnings: \$-267



# Model performance: Netflix

- Unlike with Tesla, Netflix model much more active trader
- Model earnings: \$421
- Buy and hold strategy over the same period earnings: \$372
- Outearns simple strategy by managing to avoid a handful of drops
- Probably not statistically significant





# Model performance: Facebook

- Facebook model not successful
- Model earnings: \$82
- Buy and hold strategy over the same period earnings: \$381
- Difficulty cleaning text: many articles reference comments made on Facebook
- Facebook is more financially mature?



# Next steps

- Better text - cleaning practice
- Alternate sources of sentiment:
  - More distinct sources
  - Filling out day to day sentiment for companies with sparse news footprint
- Thresholds for intra-day positions
- Operationalizing and making bank