



University of
Pittsburgh

What do you think of vaping?

SANYA BATHLA TANEJA

Intelligent Systems Program

Introduction

Vape devices or e-cigarettes popular as a means to quit smoking.

- Addiction among non-smokers
- Effect on adolescents and minors
- E-cigarette or vaping product use-associated lung injury (EVALI)



Let us vape on airplanes you COWARDS @FAANews

4:35 AM · Sep 13, 2018 · [TweetDeck](#)



[censored]

[Follow](#)

why do people at school pee in the juul room

35 Retweets 126 Likes



my sister is addicted to her juul and she's trying to bribe me

Should I start vaping ?

9:37 PM · Sep 7, 2018 · [Twitter for iPhone](#)

72 Retweets 587 Likes



[censored]

[Follow](#)

i feel like high school bathrooms are just constant juul parties now

1 Retweet 8 Likes



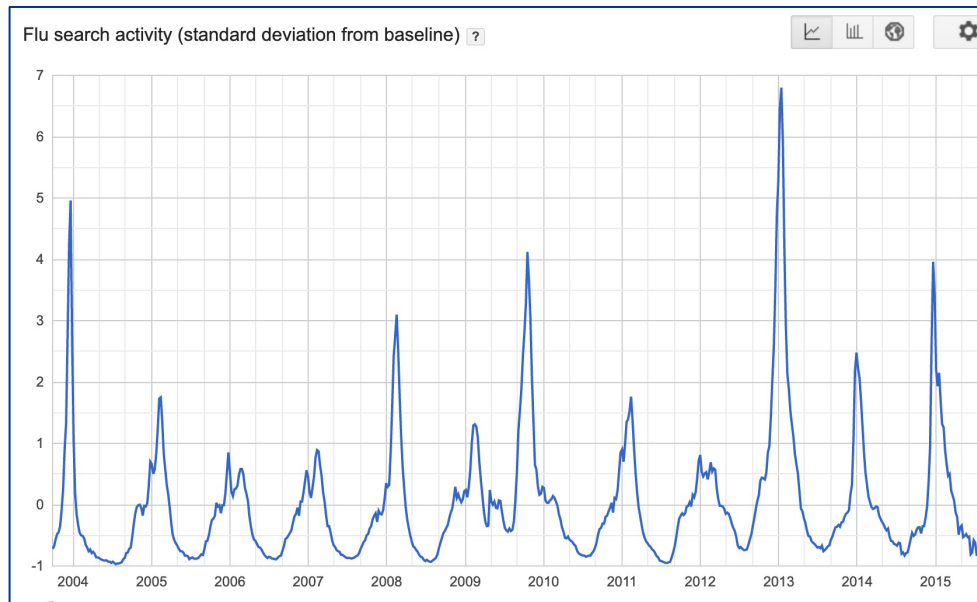
cool cucumber juul is far superior to every other flavor juul

5:56 PM · Sep 1, 2018 · [Twitter for iPhone](#)

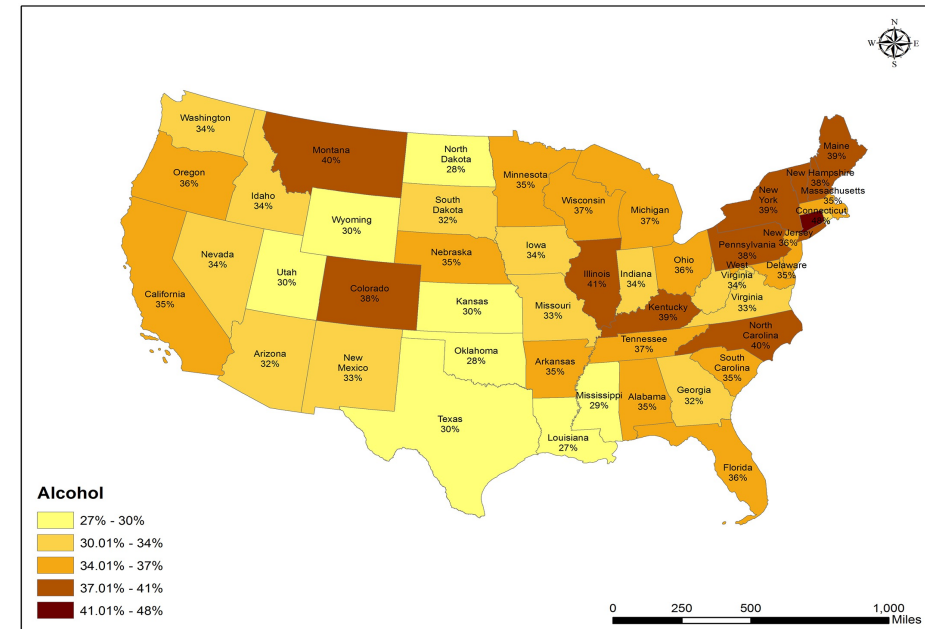
Motivation

Infoveillance - Social media as a public health surveillance tool

- ✓ Automatic assessment of current and future streaming data
- ✓ Analyze trends in 'real-time'
- ✓ Predict outbreaks



Google Flu Trends



National Substance Abuse Patterns

Machine Learning Methods for Twitter Stance Detection

- Assess stance on Twitter regarding vaping and significant policy events such as e-cigarette bans
- Analyze marketing and procurement of vape devices
- Opinions on EVALI and reported symptoms (2019)

Replying to @

Vaping should be banned completely, period. It's up to individuals to deal with their addictions, why should I put myself in harms way, to make it easier for vapers to stop killing them selves? If I can stop smoking without vaping anyone can.

1:56 PM · Aug 27, 2018 · [Twitter for iPhone](#)

Your Premium Source For Innovative [#Vape](#) Technology.

[#O2nano](#) VS [#O2flip](#)

We are running a great sale that goes through TUESDAY at midnight. It's 20% off everything on the site.

Promo code is 'iworkhard'

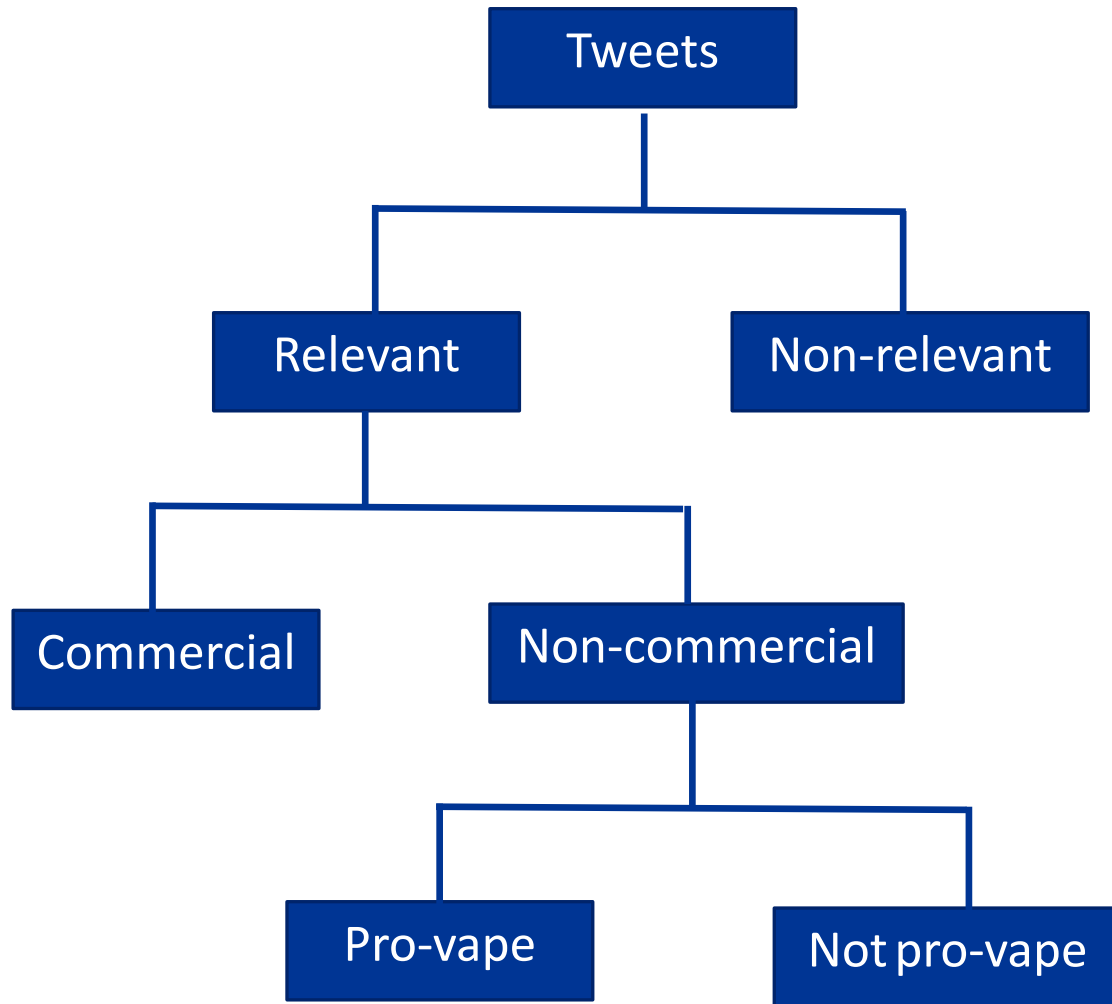
[O2vape.com](#)

[#vapepen](#) [#vapepens](#) [#vapemodel](#) [#vapefam](#)

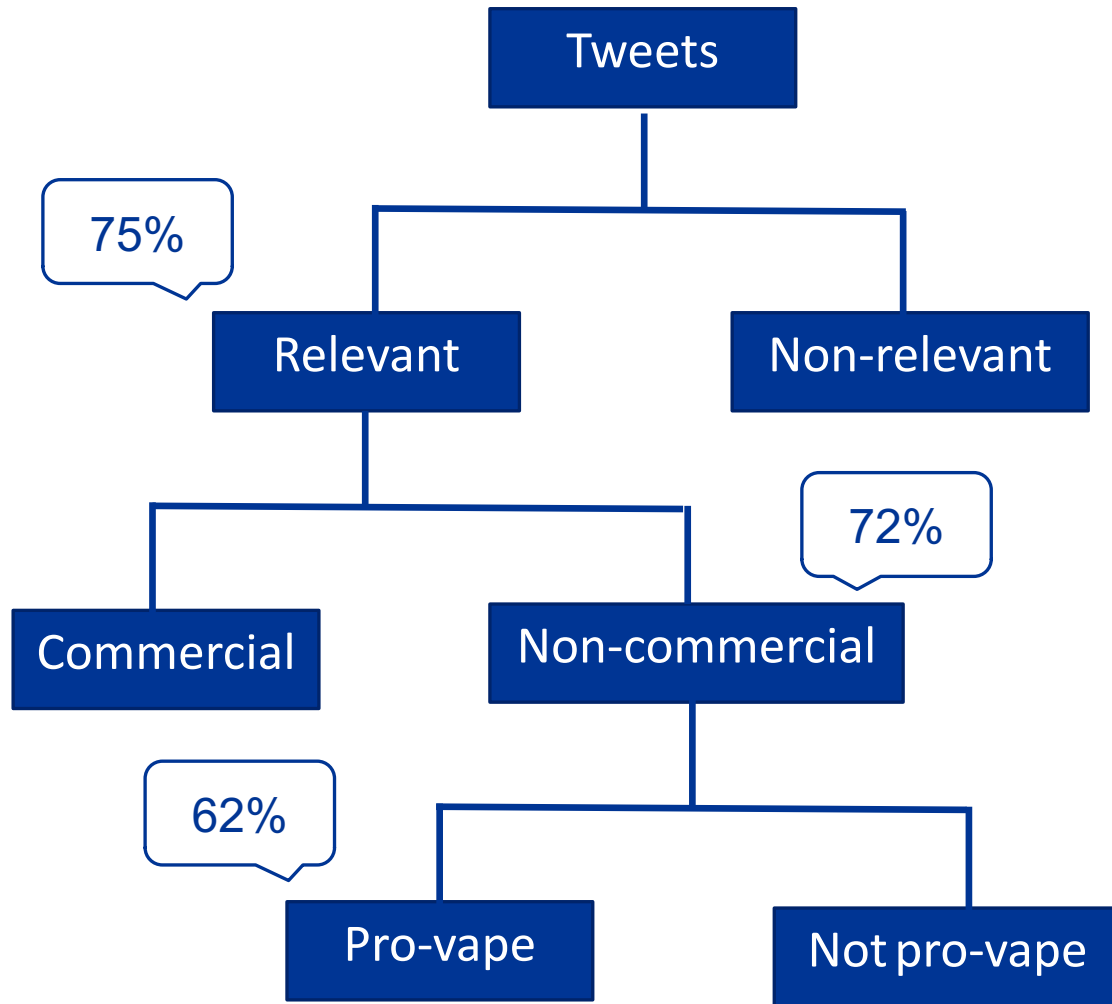
When I smoked, I coughed like I worked in a coal mine.
Now that I vape, I cough like I work at a chemical plant.

9:37 AM · Mar 5, 2019 · [Twitter for iPhone](#)

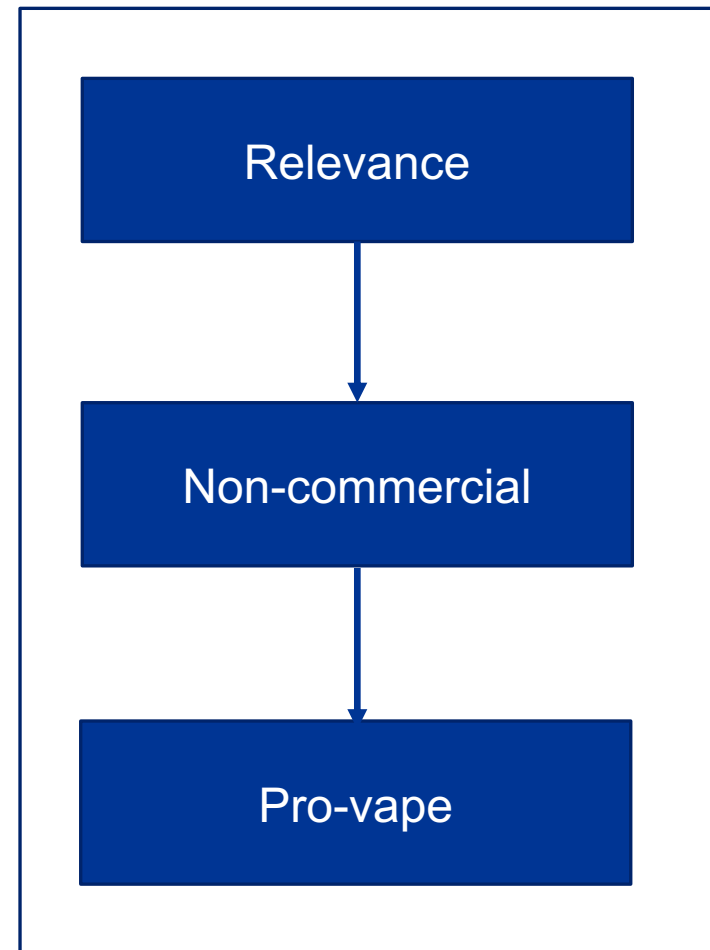
4000 tweets sampled and annotated
Keywords: vape* and juul*



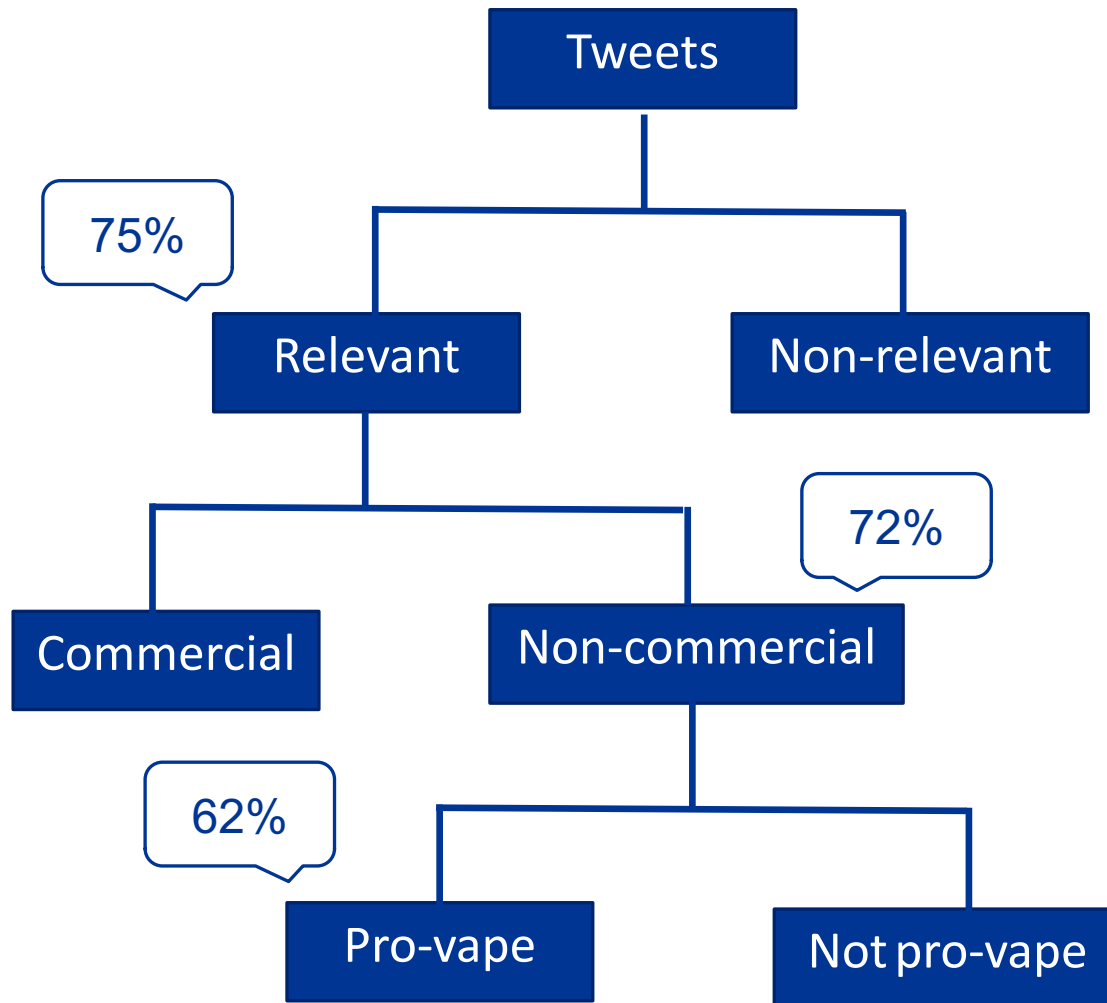
4000 tweets sampled and annotated
Keywords: vape* and juul*



Goal: 3-step classification of tweets



4000 tweets sampled and annotated
Keywords: vape* and juul*



Annotation

Pro-vape

Vader sentiment analysis

Neutral

Let us vape on airplanes you COWARDS @FAANews

4:35 AM · Sep 13, 2018 · [TweetDeck](#)

Anti-vape

Neutral

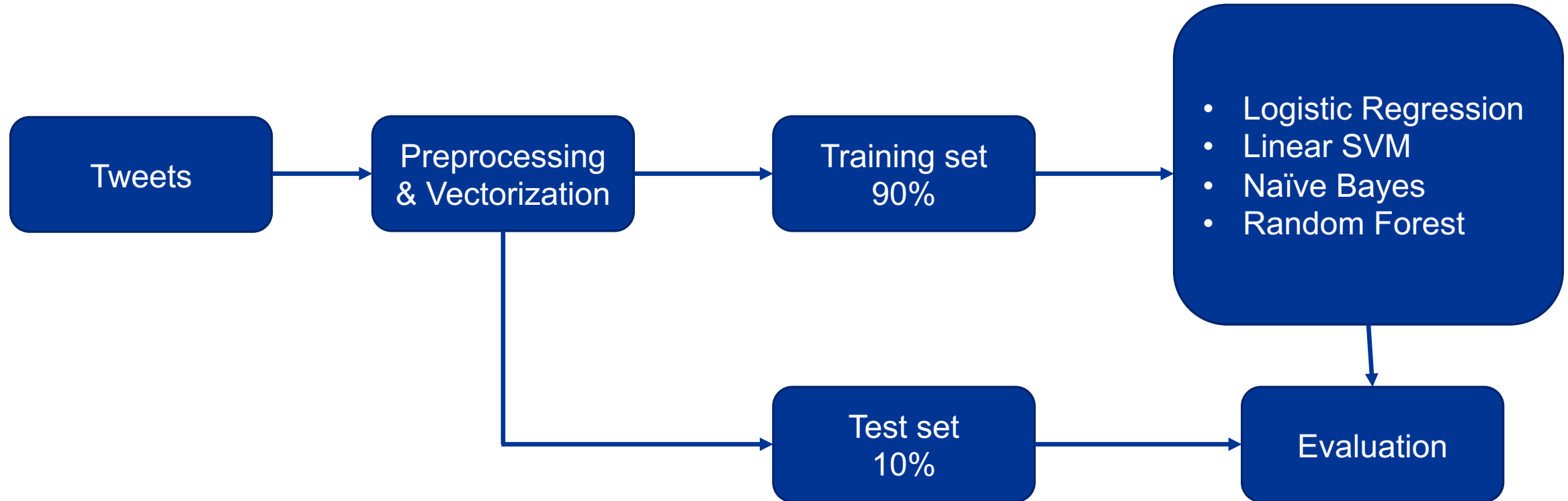
Replying to @

Vaping should be banned completely, period. It's up to individuals to deal with their addictions, why should I put myself in harms way, to make it easier for vapers to stop killing them selves? If I can stop smoking without vaping anyone can.

1:56 PM · Aug 27, 2018 · [Twitter for iPhone](#)

Binary Classification Experiments I - Relevance

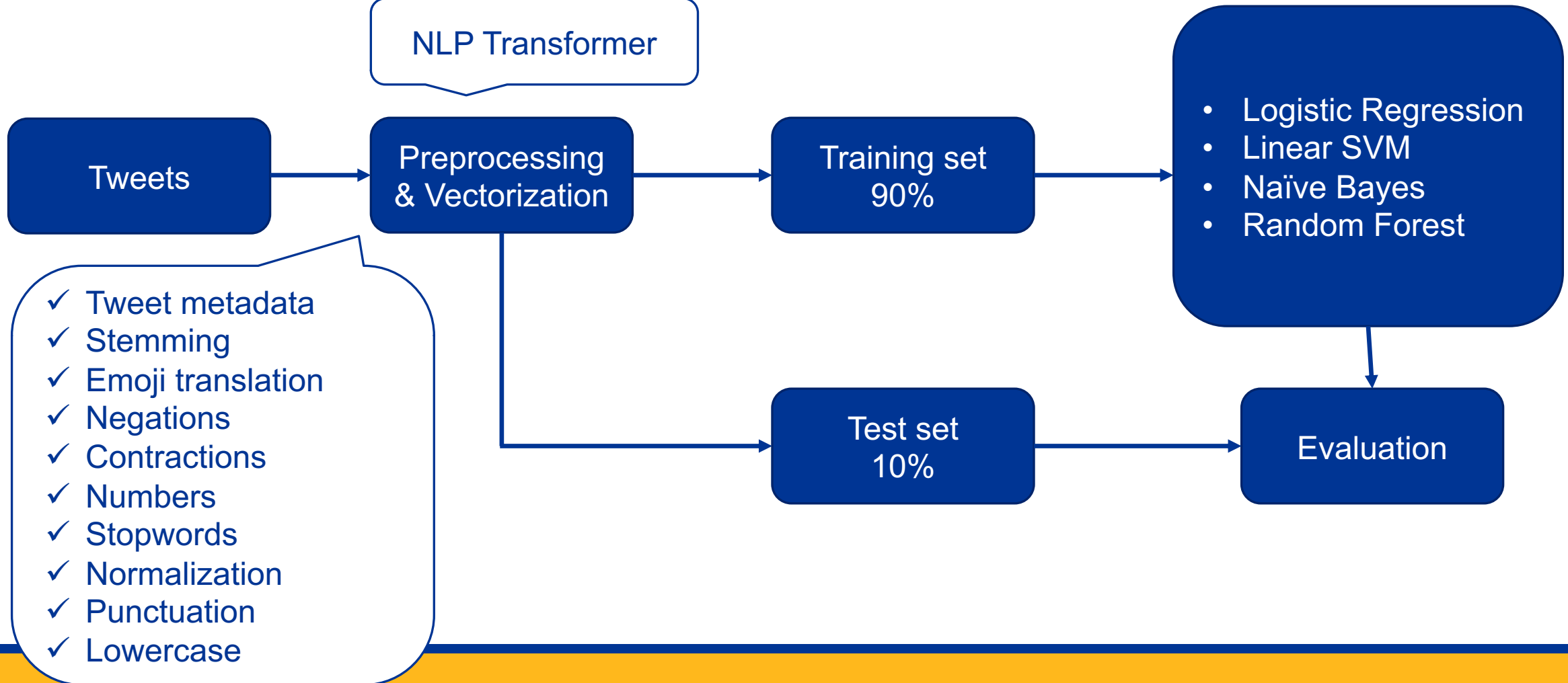
Grid Search Parameter Tuning – 5 fold Cross Validation



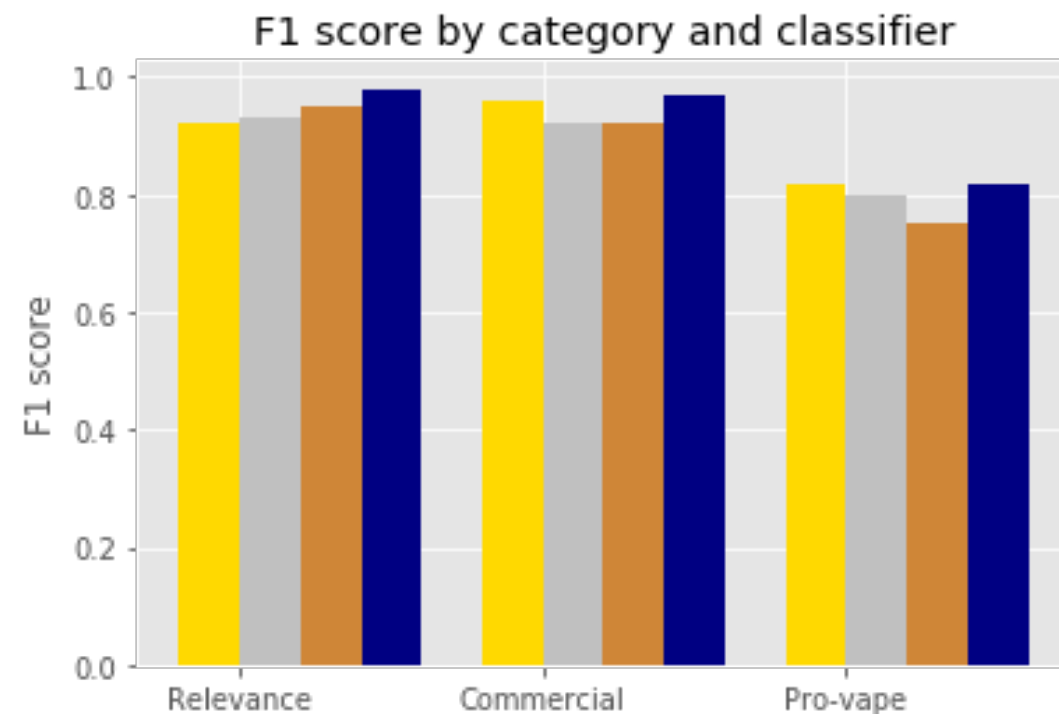
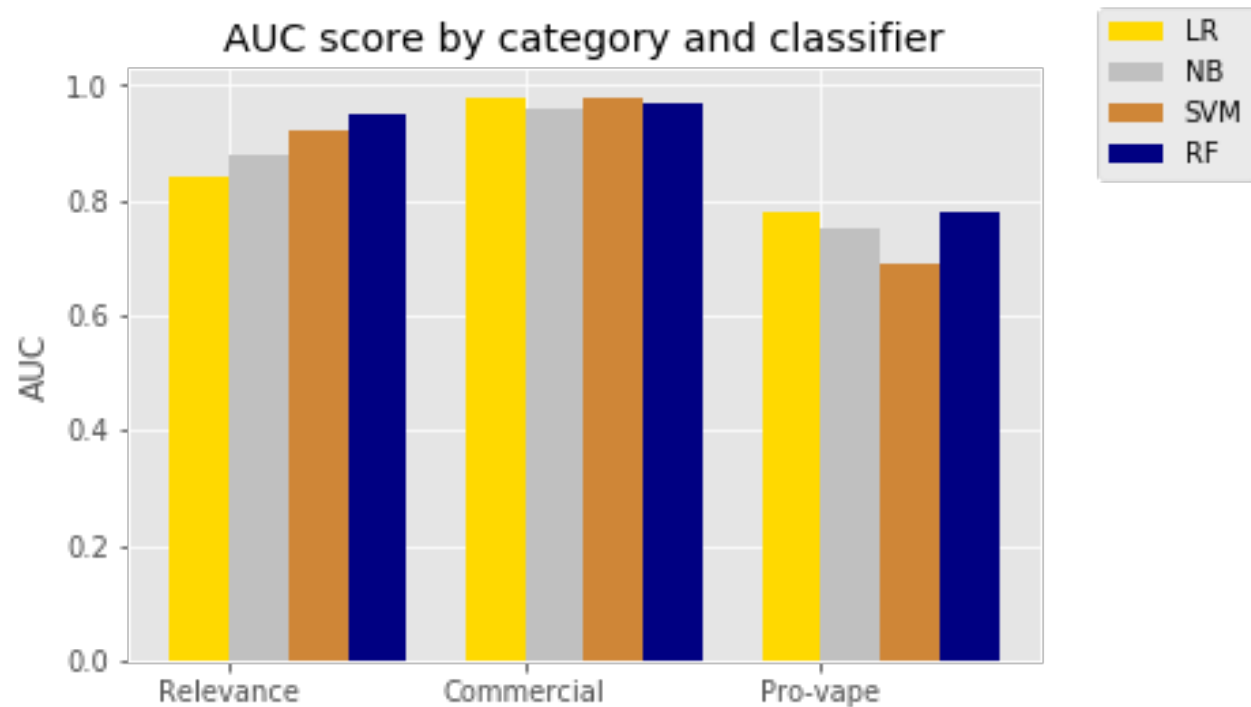
Binary Classification Experiments I - Relevance

Grid Search Parameter Tuning – 5 fold Cross Validation

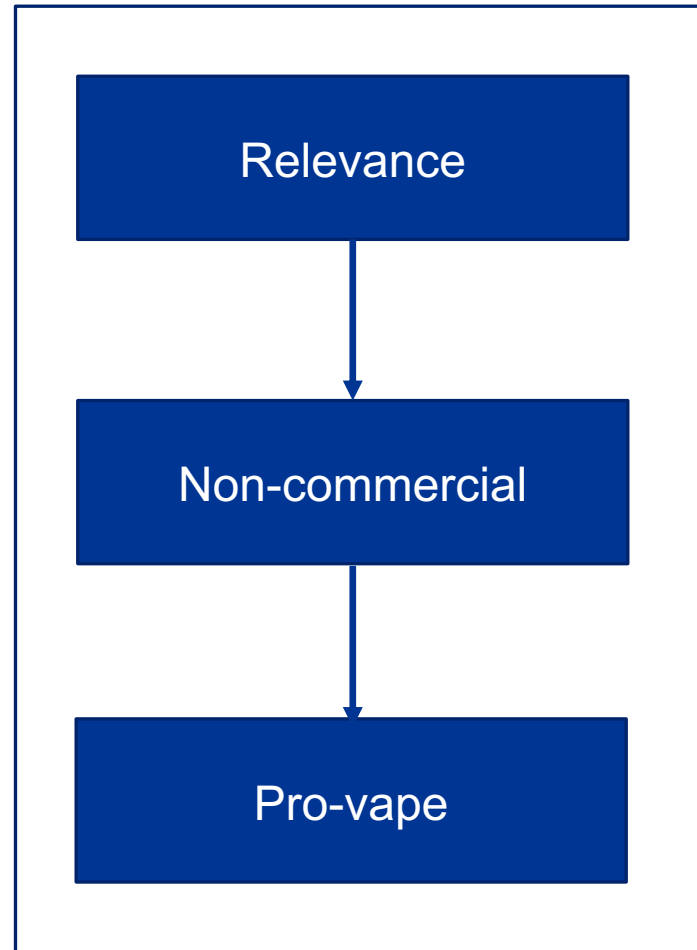
NLP Transformer



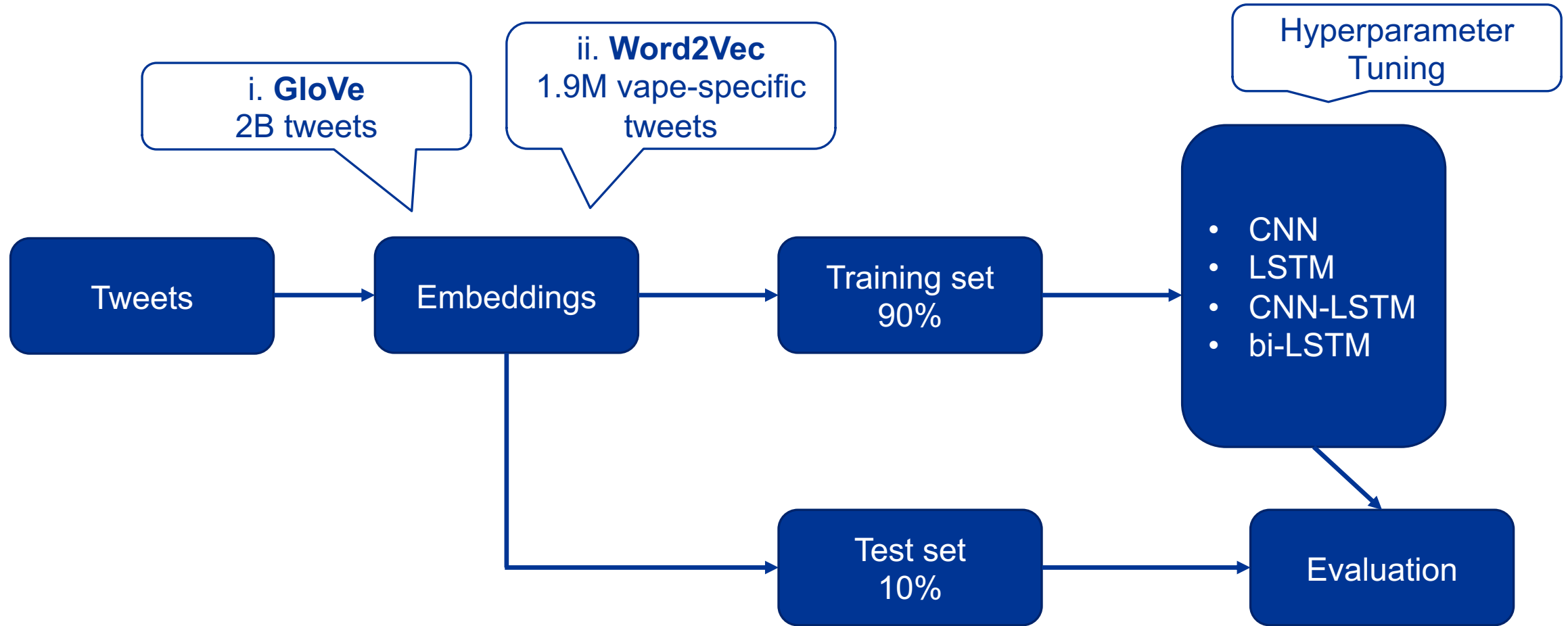
Results – Experiments I



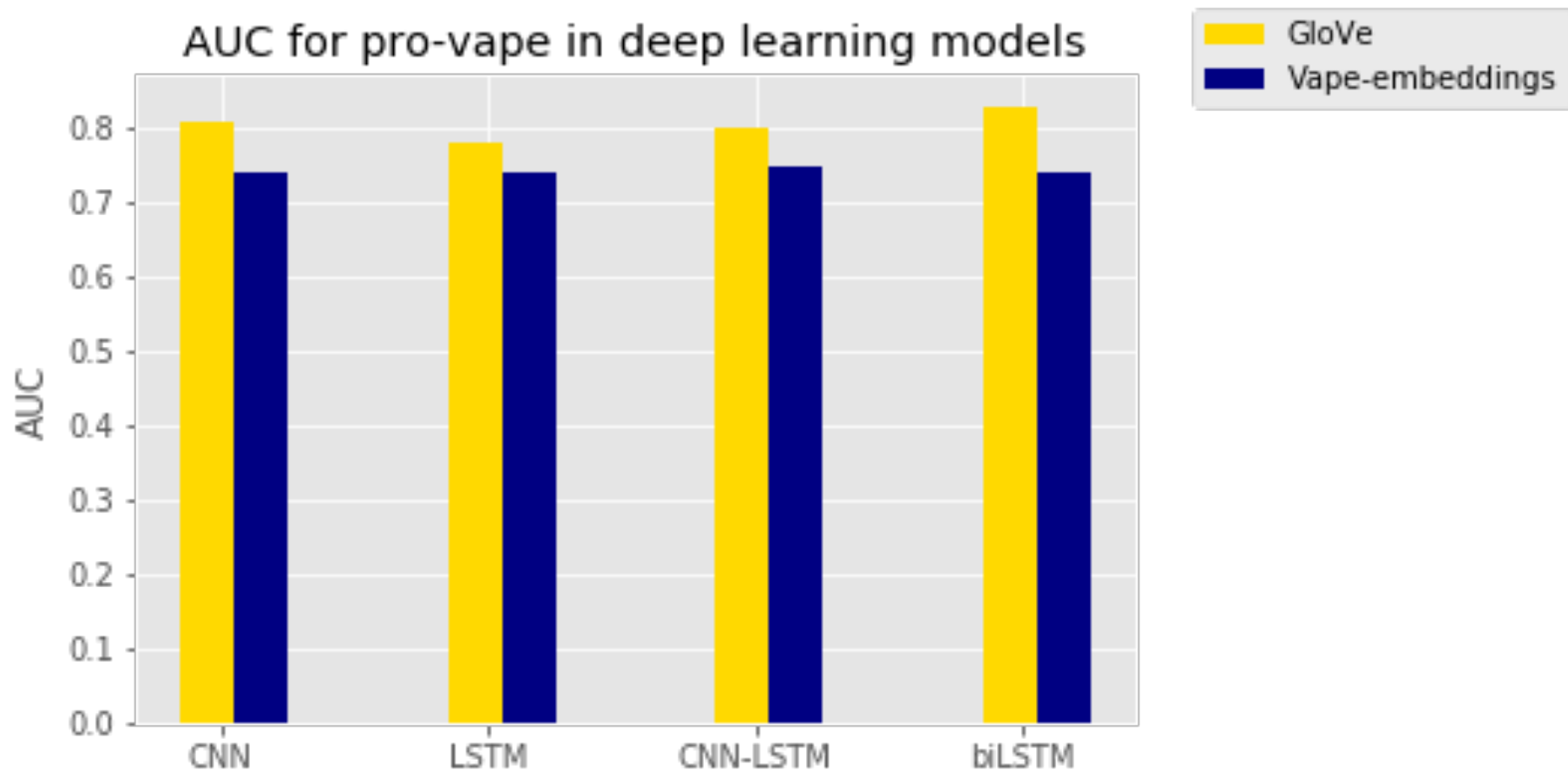
All experiments repeated for non-commercial and pro-vape tweets



Binary Classification Experiments II - Relevance

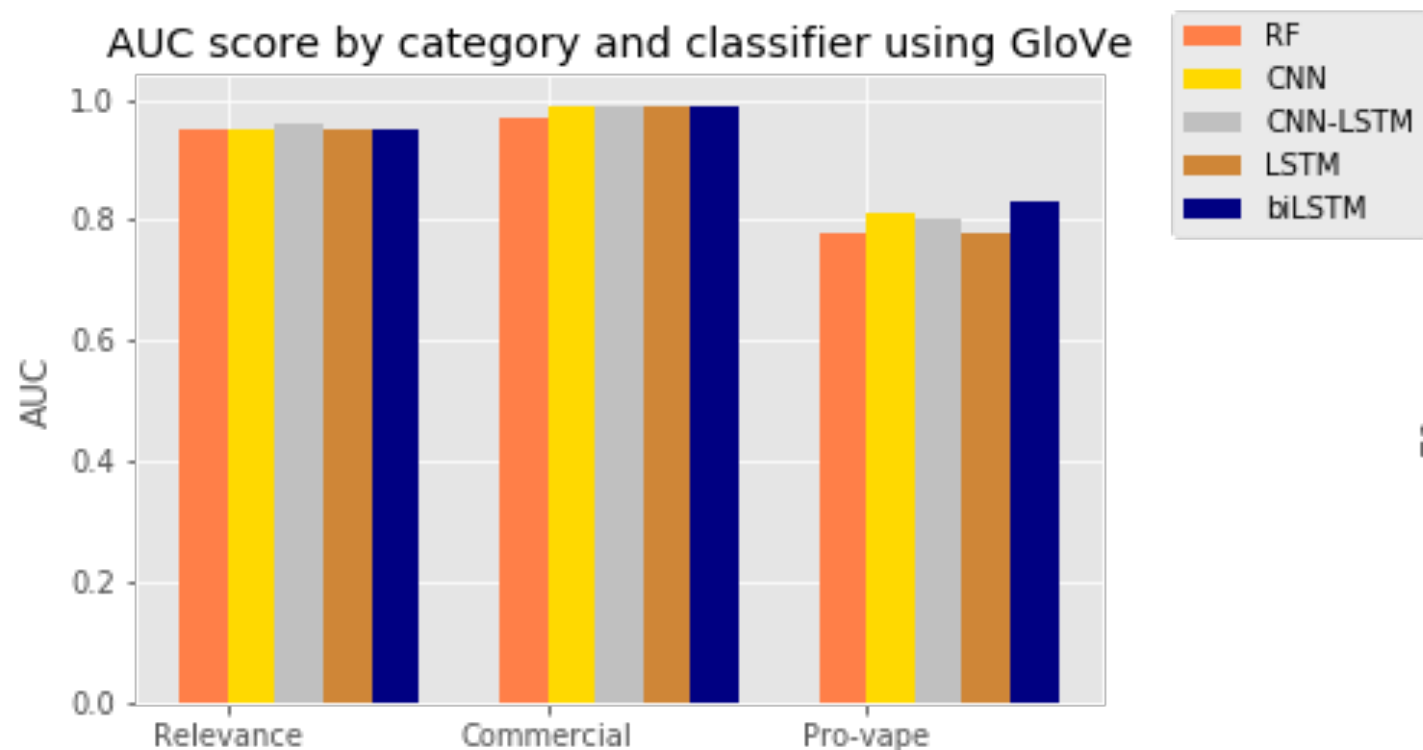


Results – Experiments II

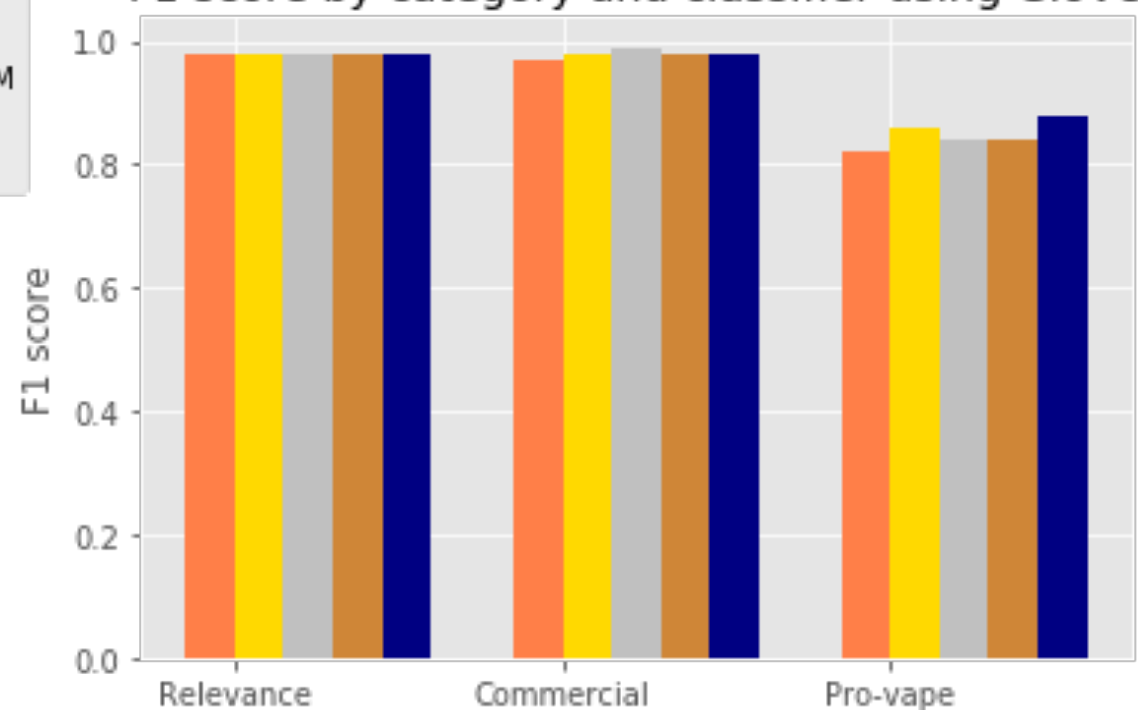


Results – Experiments II

AUC score by category and classifier using GloVe

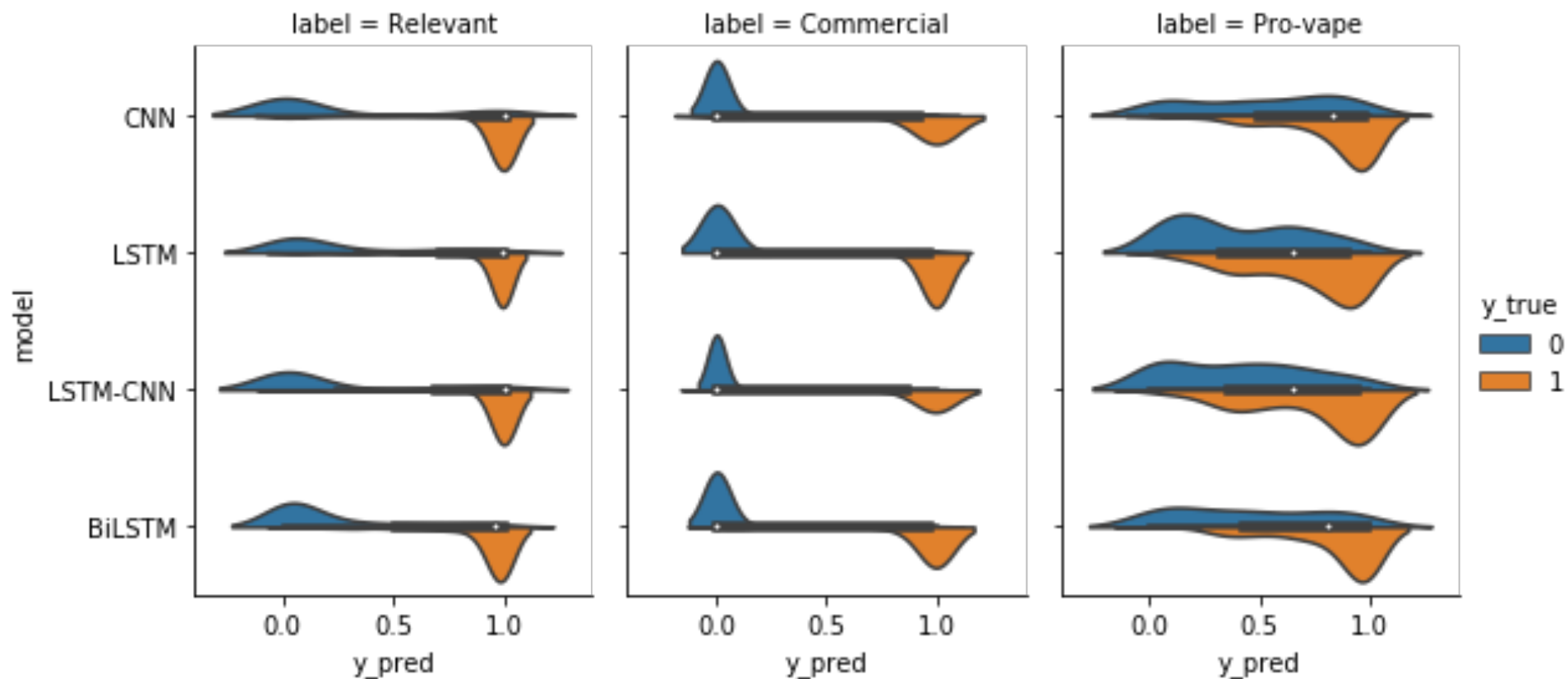


F1 score by category and classifier using GloVe



Results – Experiments II

Probability Density Curves



Conclusion

- Processing pipeline
 - Minimal preprocessing for all classifiers
 - Tweet metadata is important – hashtags, URLs, usernames, unicodes and emojis
- GloVe outperforms context-specific embeddings
- **Relevance and Commercial classification**
 - Random forest at par with deep learning classifiers using GloVe (statistically significant)
 - Commercial tweets use distinct vocabulary features (SHAP models in paper)
- **Pro-vape**
 - Subjective and challenging
 - AUC's from 0.78-0.83
- Best classifier is as good as the worst performing one in the 3 tasks
 - Classify with stricter thresholds using Youden's index
 - Choose parameters and classifier for 'real-time' classification

Future Work

- Time series analysis of pro-vape tweets
 - variability in stance surrounding events such as flavor bans, FDA inspections in 2018
 - extend to 2019
- Clinical cases of EVALI at UPMC and Twitter reported symptoms – preliminary findings
- Geolocation with national and state-level infoveillance
- Evaluate classifier with more human-annotated tweets and calibrate the models



Jason B. Colditz, MEd

Jaime Sidani, PhD, CHES

Kar-Hai Chu, PhD

Brian Primack, MD, PhD



Shyam Visweswaran, MD, PhD

Patrick O' Halloran, MS

Na-Rae Han, PhD (Department of Linguistics)

Joel Welling, PhD (Pittsburgh Supercomputing Center)

