

# App Review Topic Modeling

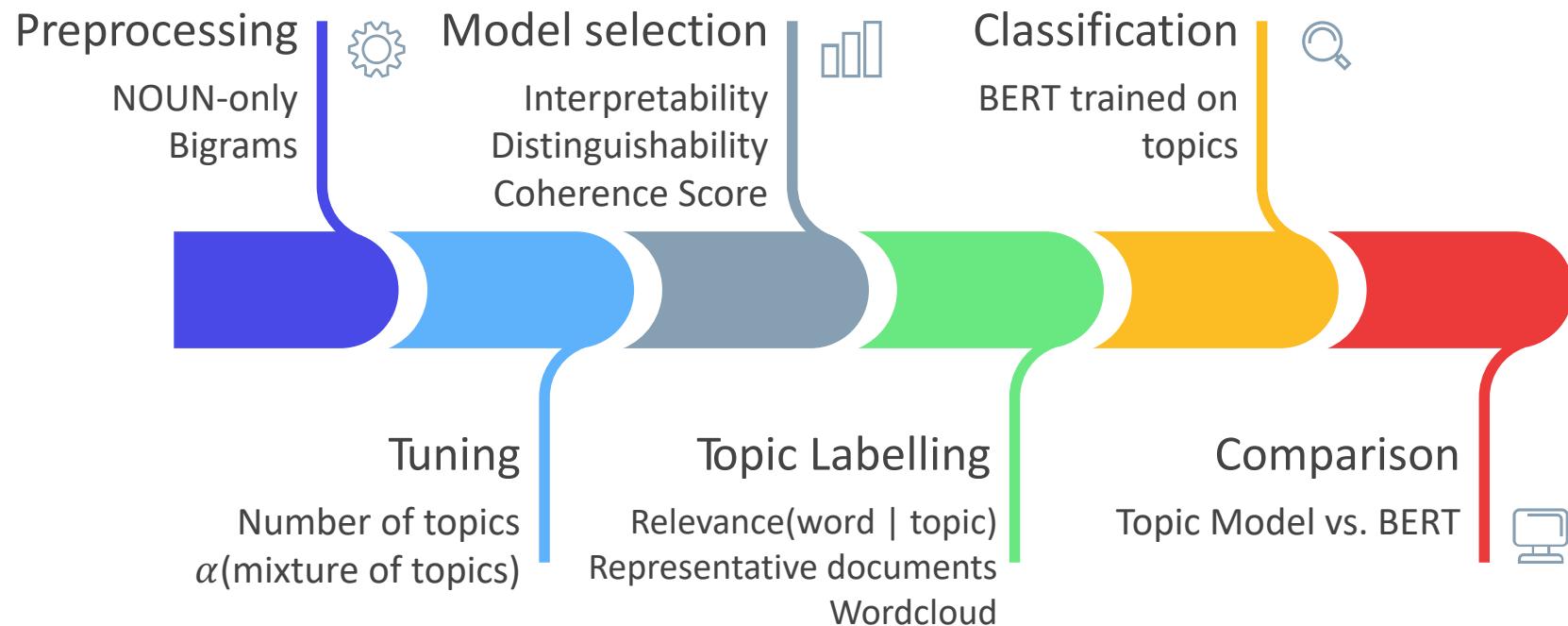
Jung-a Kim



# Proposal

- Background
  - Developers are paid 70% of the list price of the app or in-app purchase.
- The goal of this project
  - Fine-tune quality assurance
  - Debug critical errors in a prompt manner
- Dataset
  - "Netflix" which has one of the most reviews between 2010-11-04 and 2015-08-31.
  - 12,566 reviews for topic modeling and 6,283 reviews for classification

# Workflow



## Tools

**gensim**

**spaCy**

**NLTK**  
Natural Language Toolkit  
Reference Guide



**K** Keras

gensim, spacy, NLTK for preprocessing

gensim, pyLDAvis for LDA, NMF, LSA

ktrain, bert\_embedding, huggingface/transformers for BERT

PIL, wordcloud, sklearn, umap, gensim for visualization

Jupyter lab for modeling

PyCharm for Flask

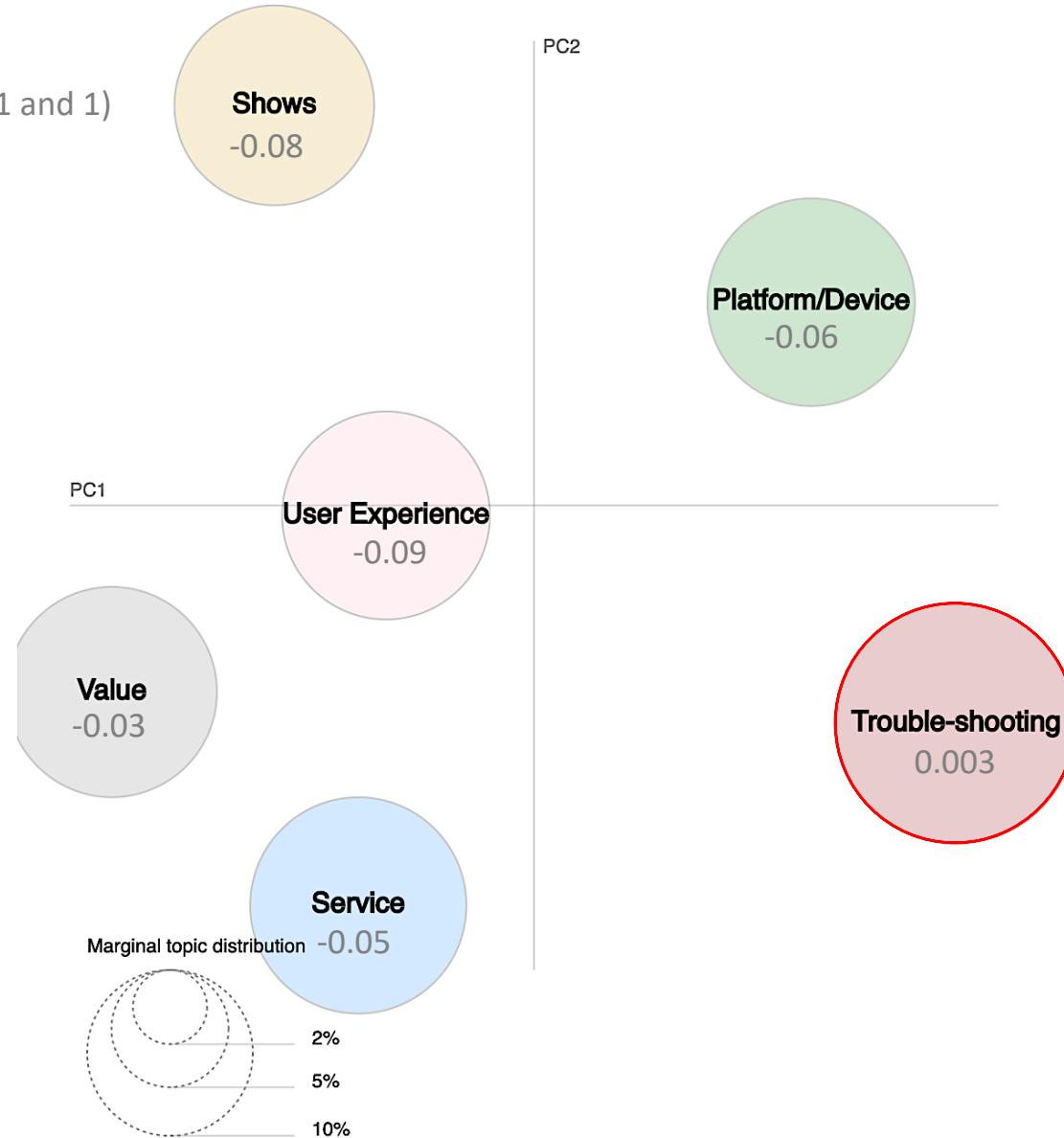


scikit  
**learn**

# LDA Mallet

(Coherence score(NPMI) between -1 and 1)

Intertopic Distance Map (via multidimensional scaling)

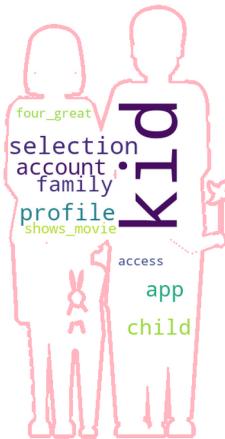


# LDA Topics

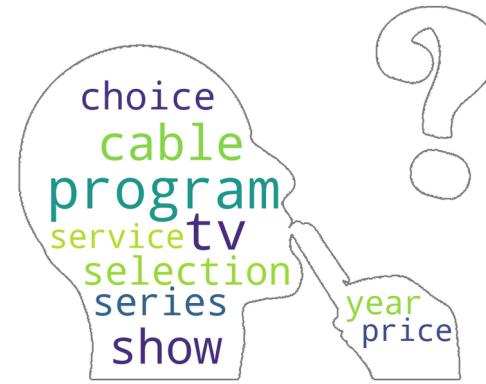
Platform/Device



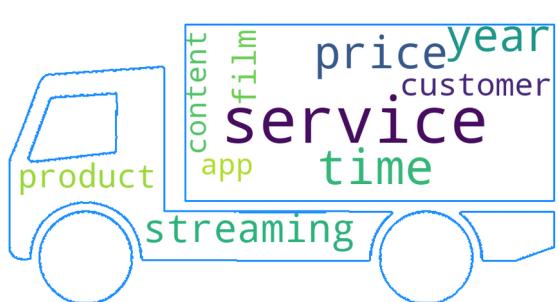
User Experience



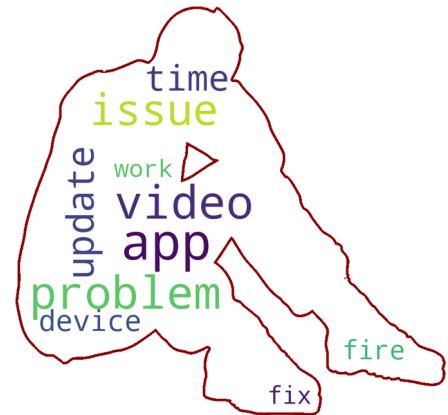
Value



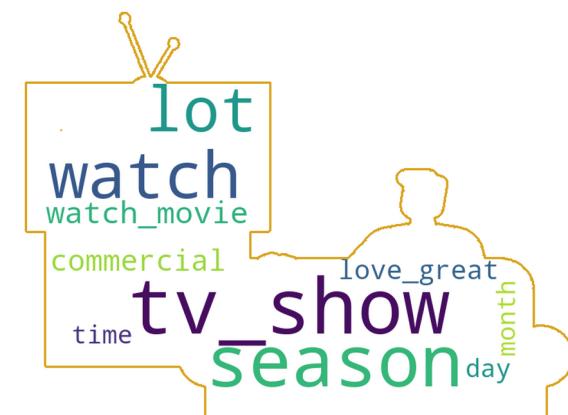
Service



Trouble-shooting(Video-related)



Shows



## Most representative reviews

### Platform/Device

...it runs great on our kindles, laptops, and I pods...

### User Experience

...my kids have access to R-rated, horror, etc...

### Value

...I can enjoy a movie without a considerable amount of commercials...

### Service

... Netflix has done a poor job getting new movies to their streaming service...

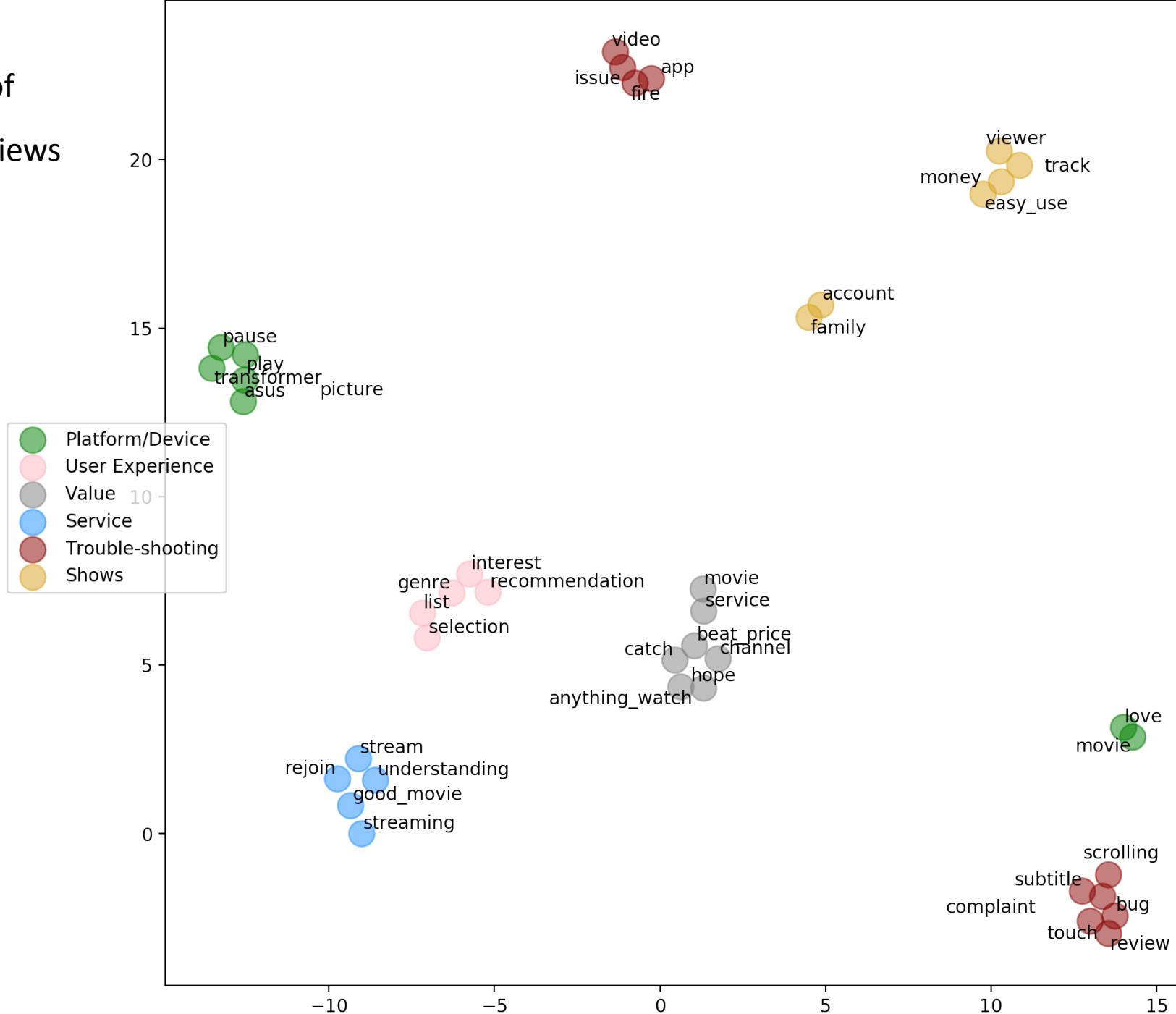
### Trouble-shooting

...this issue should be a high priority fix...

### Shows

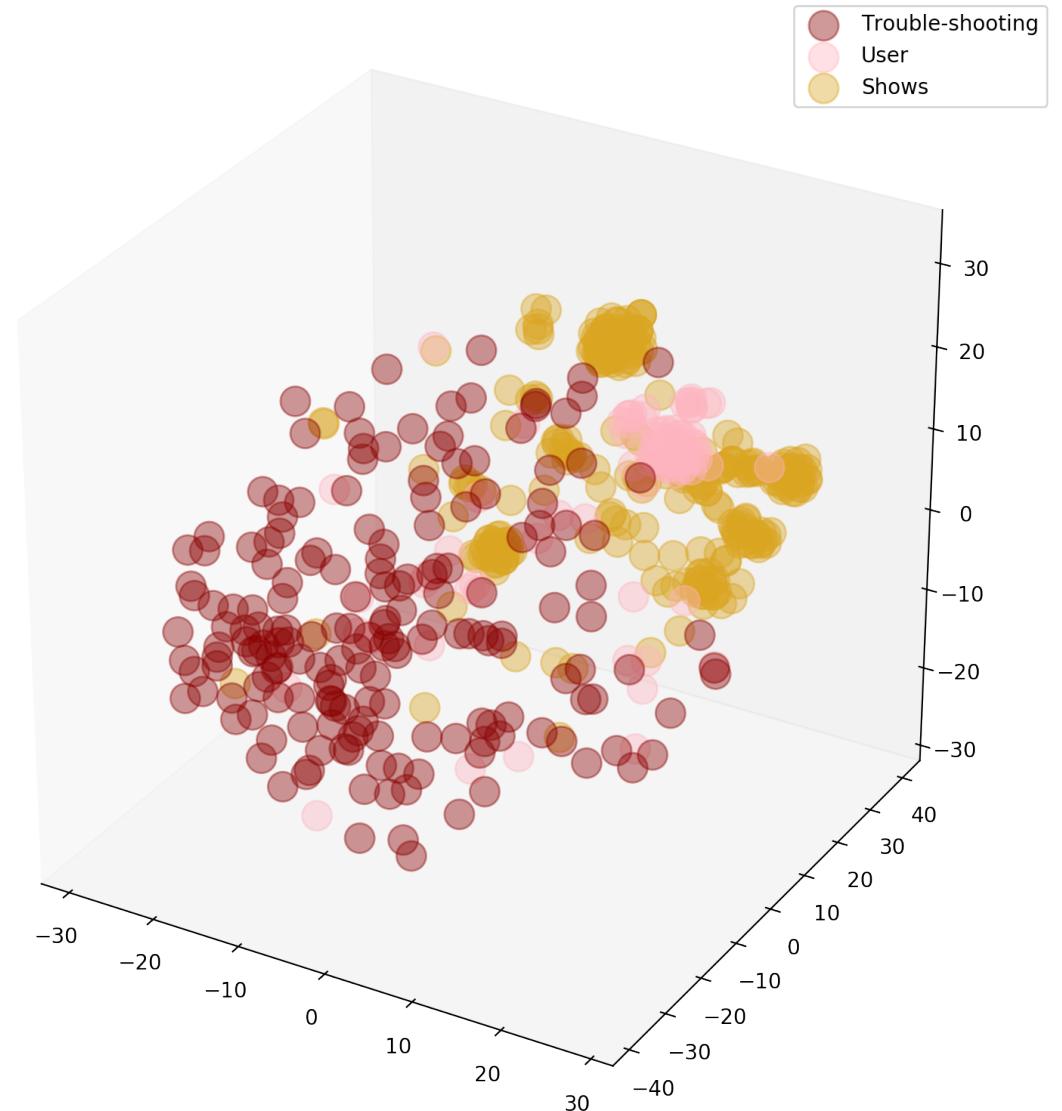
...who really doesn't love Netflix...

# LDA embeddings of representative reviews

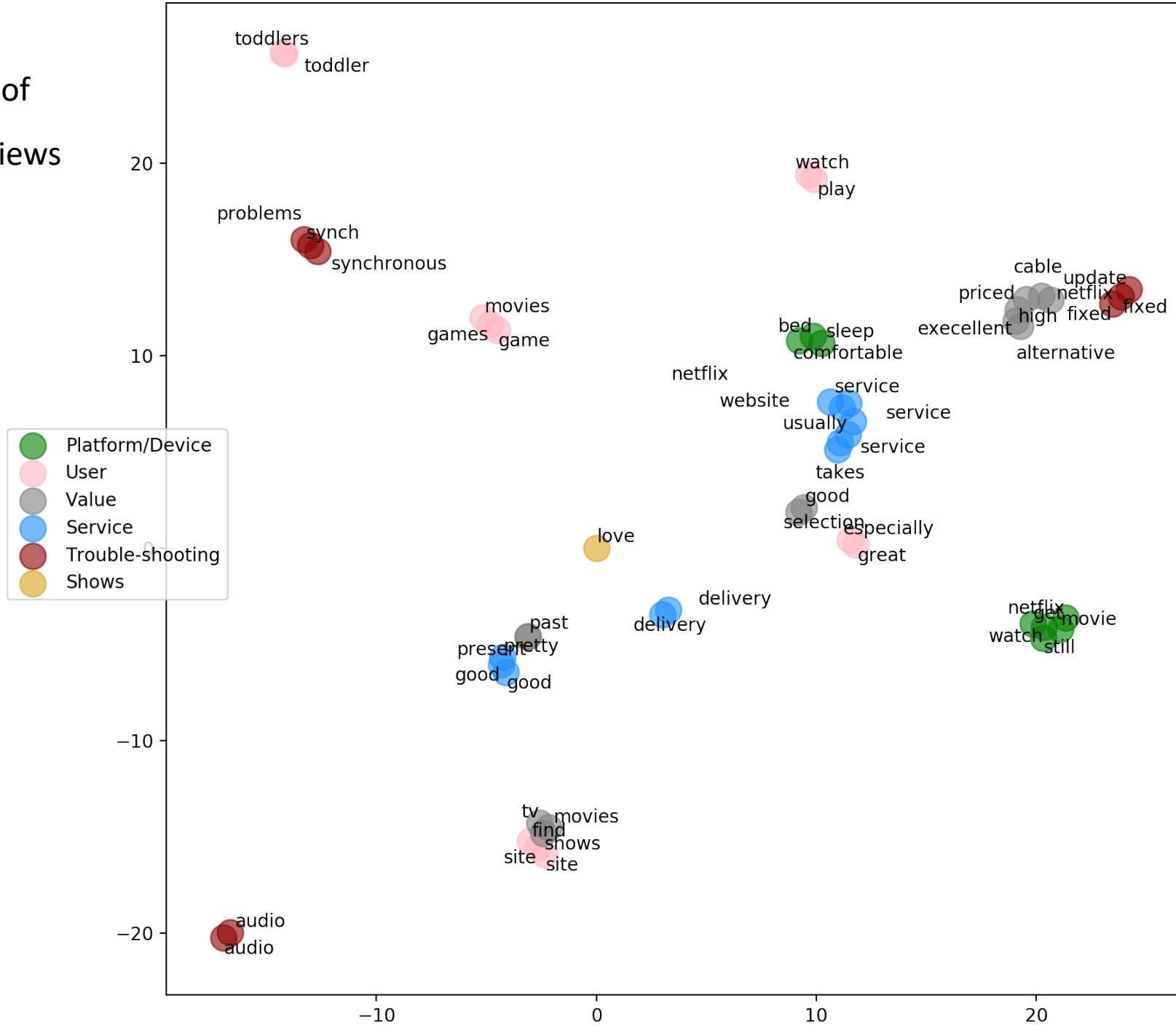


# Bidirectional Encoder Representations from Transformers

- Trained with LDA topics
- 5 epochs, Stochastic Gradient Descent
- 43% topic agreement with LDA
- Attentive to context



## BERT embeddings of representative reviews



Please write your review about this app.

We will refer you to the right service.

Your name

Your Email Address

Type your message here....

Submit

Designed by [Colorlib](#)

Open

## *Conclusion*



Topic models are good for reviews that have similar keywords.



BERT is sensitive to sentiments.

## *Future work*



Re-cluster with BERT's classifications



Collect more sarcastic reviews

# BERT Wordcloud on unseen data

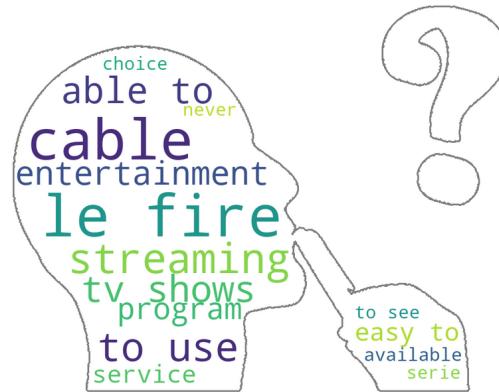
## Platform/Device



# User Experience



## Value



## Service



## Trouble-shooting(Video-related)



## Shows



## Examples from the app

### Bad feelings:

I hate this app! all it does is mess up! I'll be trying to watch a show when all the sudden my kindle screen goes blank! that is just one of the few reasons why I hate this app

### Mixed feelings:

convienent but not always reliable Love the service, having shows and movies available on my computer, TV even my smart phone, but Netflix doesnt always load, kicks you off in the middle of movies stating its not available right now choose another title until its avaialbel again. Its afordable and for \$8 a month i will put up with glitches and minor annoyances for the convienence.

### Sarcastic(good):

Warning. Extremely addictive! I love Netflix! You have unlimited access to the shows or movies you enjoy and you can watch them whenever you want.

### Sarcastic(bad):

This app is amazing. The "team" apparently believes giving people fewer options is better because it's all just too advanced for us sheep to understand.