Using NLP to Analyze Call Center Data

Amit Arora
Georgetown University

Abstract

This project **analyzes customer feedback** data in a call center for a **telecommunications company** that provides internet access services for home users. We apply natural language processing (NLP) techniques to determine insights such as key named entities (**NER**), themes in customer feedback (**Topic Modeling**) and customer sentiment (**Sentiment Analysis**).

The analysis finds that video related services (Netflix) and Wi-Fi connectivity issues are key entities in customer feedback, based on trends in topic timeseries it appears that a specific theme of customers wanting unlimited data has been addressed and finally, the customer sentiment after a dramatic improvement in 2017 is now back at 2016 levels.

Methodology

A **CSV file for each call center** containing a column for customer feedback ("Please give us any additional feedback.") and a rating (1 – 10). **Total size ~245,000 rows**.

NER: done using the ***Spacy*** package with the "en" English language model. Results were post processed to account for spelling variants, for e.g. ATT and AT&T refer to the same entity.

Topic Modeling: NLTK regex tokenizer and stopwords corpus was used to tokenize and clean the text. ***Gensim*** for topic modeling using **LDA**. A set of **10 topics** were discovered after running LDA for 30 passes. The topics were then **manually given descriptive names**.

Sentiment Analysis: Lexicon based sentiment analysis was done using the ***sentimentr*** package. Augumented the **Jockers (2017) & Rinker's augmented Hu & Liu (2004) positive/negative word list** with frequently occurring bigrams and trigrams in the data. The **average weighted mixed sentiment method** was used to combine the sentiment of multiple sentences in the feedback to determine the overall sentiment.

Deep learning model Text is cleaned using ***NLTK*** and then converted to word vectors using the **en_vectors_web_lg** model (**GloVe**, common crawl, 300 dimensions) in ***Spacy***. Document vectors corresponding to each feedback are then used to train a model to classify feedback as positive (rating 6 to 10) or negative (rating 1 to 5). **Keras/Tensorflow** used to implement a **Multi Layer Perceptron** (MLP) model for classification.

Results

The analysis helped figure out several key insights, the more important ones are being listed here.

NER: **Figures 2, 3. Netflix, Facebook** frequently feature in customer feedback. Has been a recent spurt in **Wi-Fi** mentions which matches launch of a new product. Customers also mention **Verizon** and **AT&T** possibly as potential alternatives to their current service.

Topic Modeling: **Figures 4, 5.** Key themes such as "**need unlimited data, better price**", "**monthly billing**" identified. Visualized share of issues by product line. Key finding is that "need unlimited data" is not as much an issue for the latest product as it was for previous products and this **corroborates results of conscious business actions.**

Sentiment Analysis: **Figures 6-10. Deep Learning** model has a **73% accuracy** in predicting sentiment. The **lexicon based** sentiment analysis **captures the general trend** (Figure 2 and 3) but **misclassifies** a lot especially when the **sentiment changes** with each sentence in the **same feedback**. Appears that **Lexicon based sentiment "lags" the actual customer feedback rating.**

Discussion

This work represents implementation of a **broad range of NLP techniques** that can be applied to analyze customer feedback **for any industry**.

Several commercial products (**CLARABRIDGE, Qualtrix**) currently provide similar solutions, **this project** is an attempt to explore the same space and provide a **more customized solution**. A **differentiator** is creation of a **deep learning based sentiment analysis model** which can be used out of the box for determining sentiment for other types of customer feedback for the same product.

The results **accurately capture key findings** that are in line with business knowledge about the product, for e.g. appearance of **Netflix and Wi-Fi** as most important named entity in recent years, slow web browsing and **need for unlimited data** as key themes in customer feedback and finally the **sentiment analysis model** which can be reused.

An area of **future work** is the use of **Hierarchical Attention Network**[7] to improve the classification accuracy especially since we have evidence of sentiment changing during the length of a single feedback.

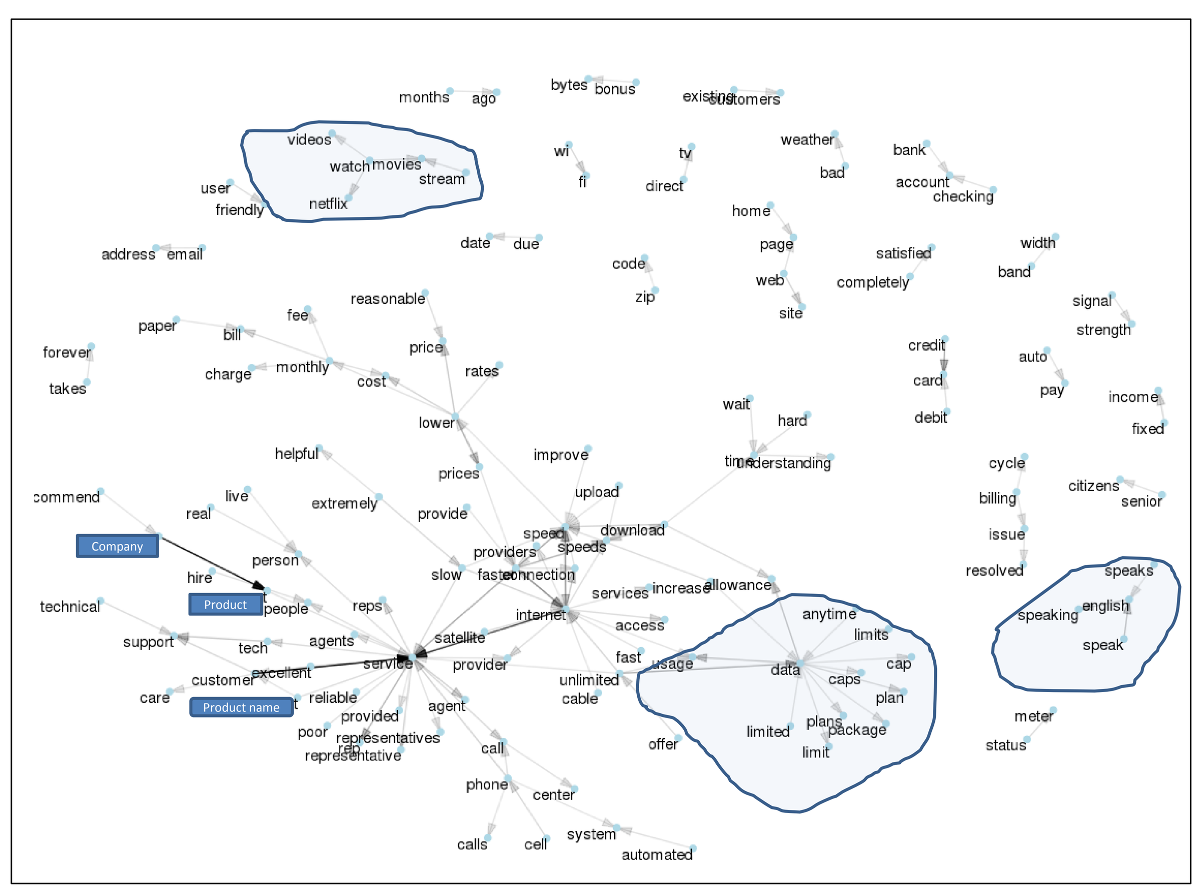

Figure 1. Important Bigrams {Company, Product}, {Customer Service}, {Internet Service} and clusters of bigrams highlighted

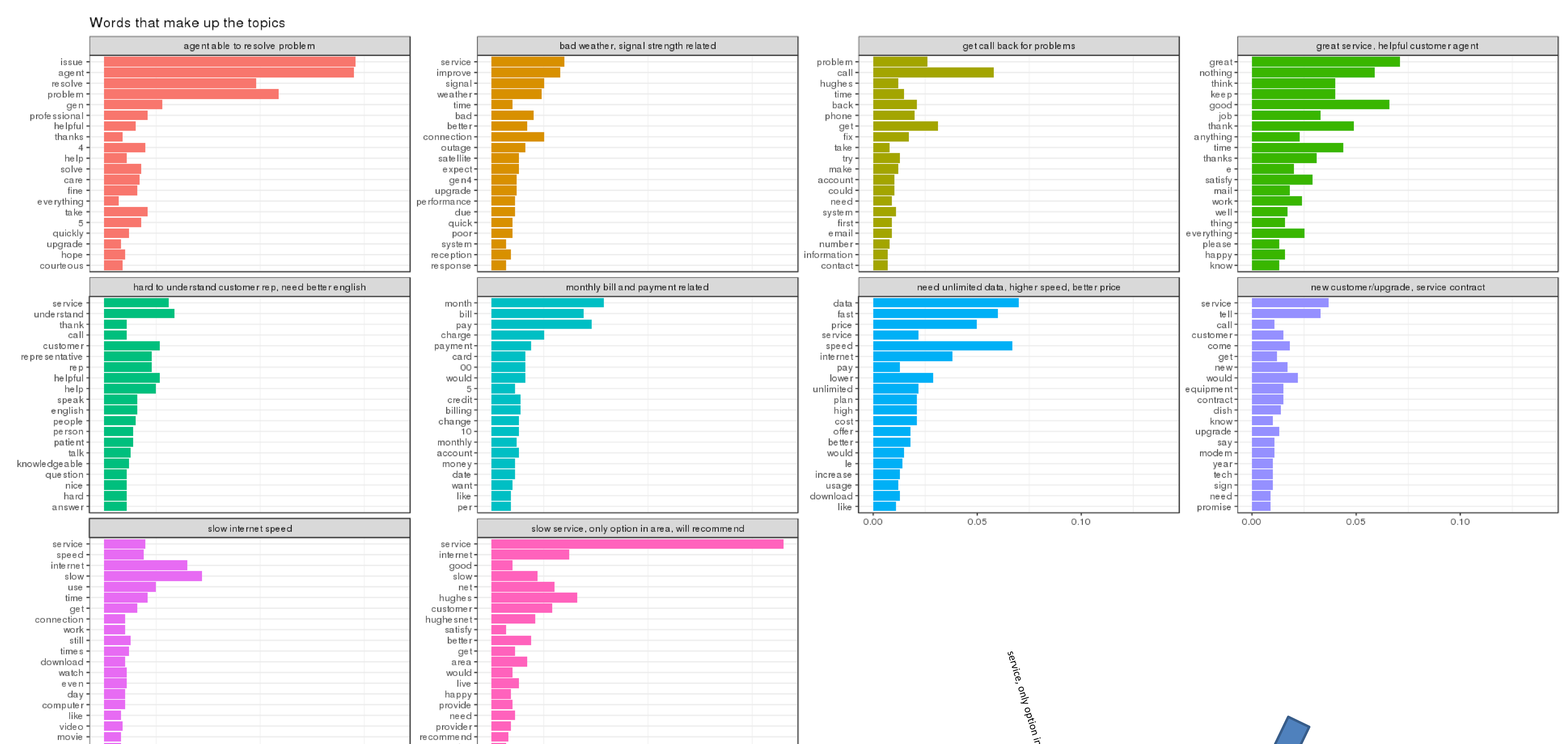

Figure 4. Topics discovered via LDA

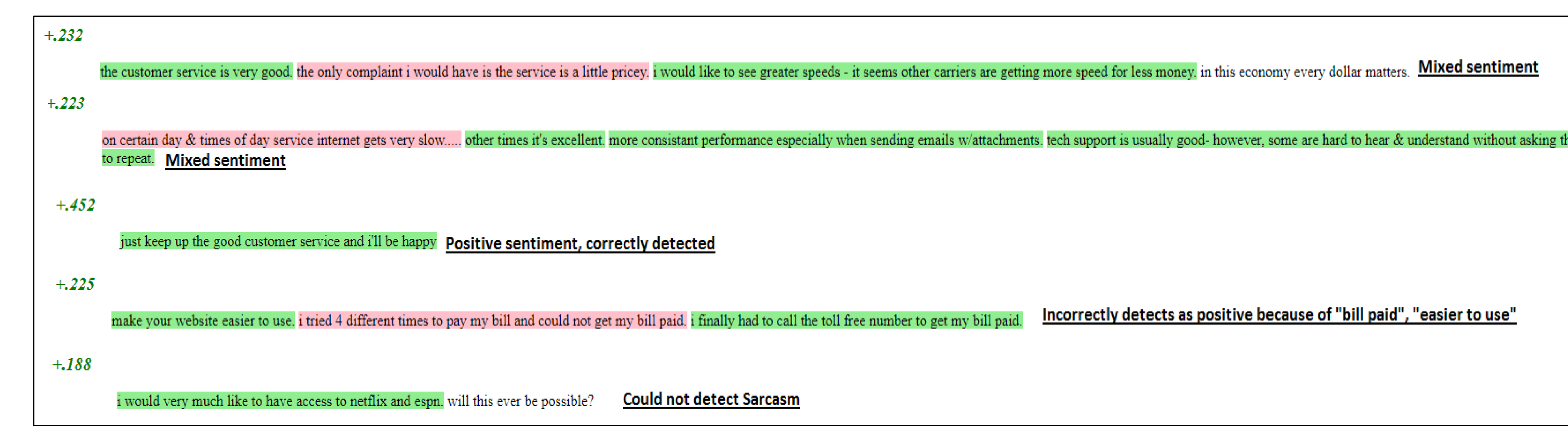

Figure 6. Sentiment Analysis by Lexicon

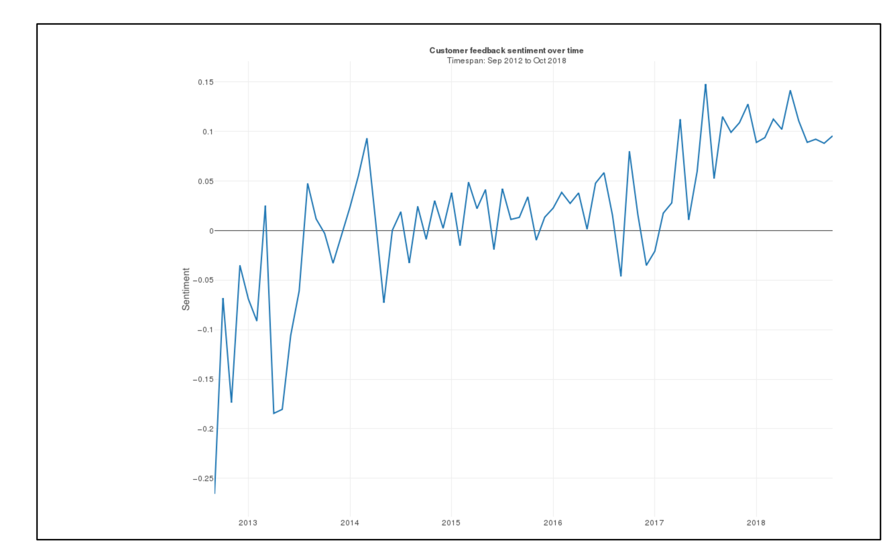

Figure 7. Lexicon based sentiment over time

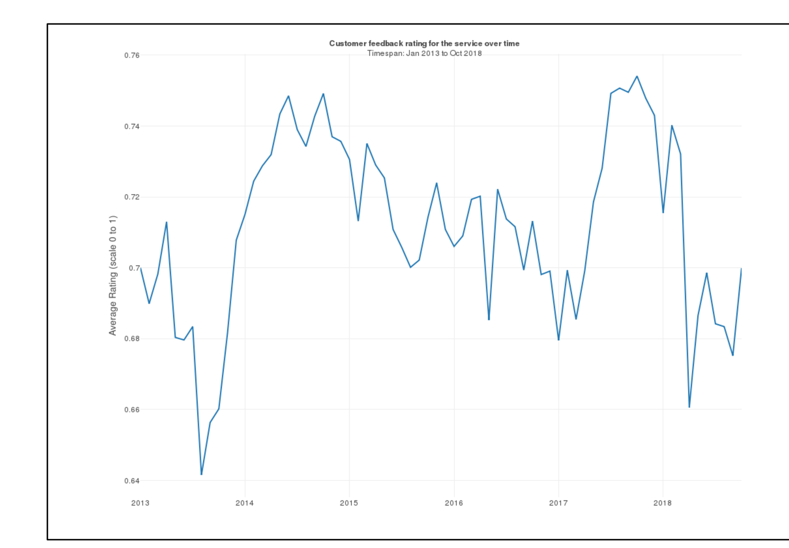

Figure 8. Customer feedback rating over time

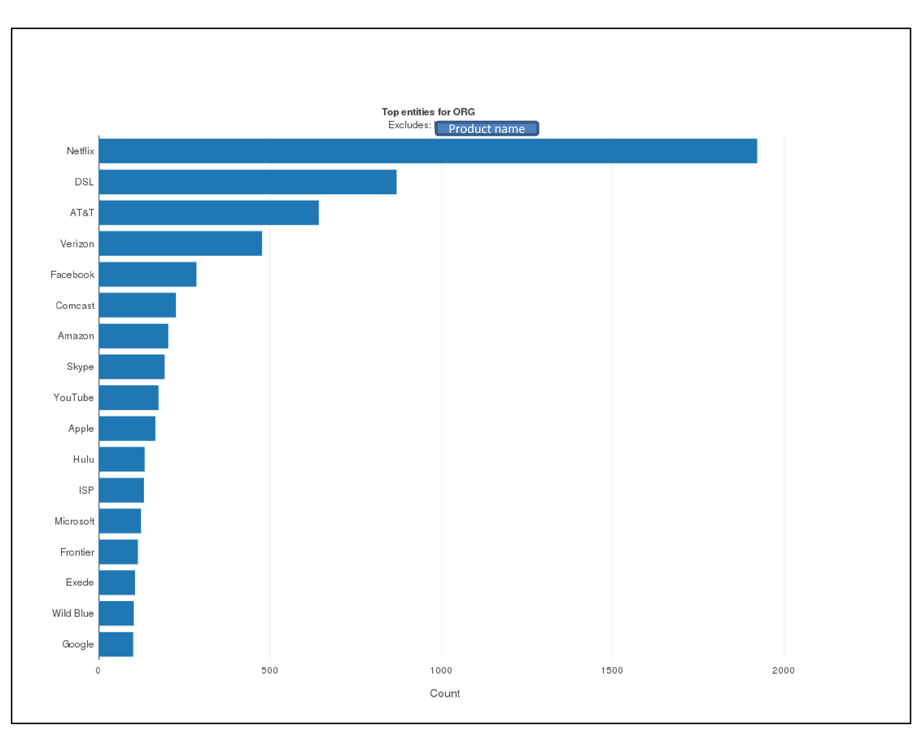

Figure 2. Named Entities (Organizations)

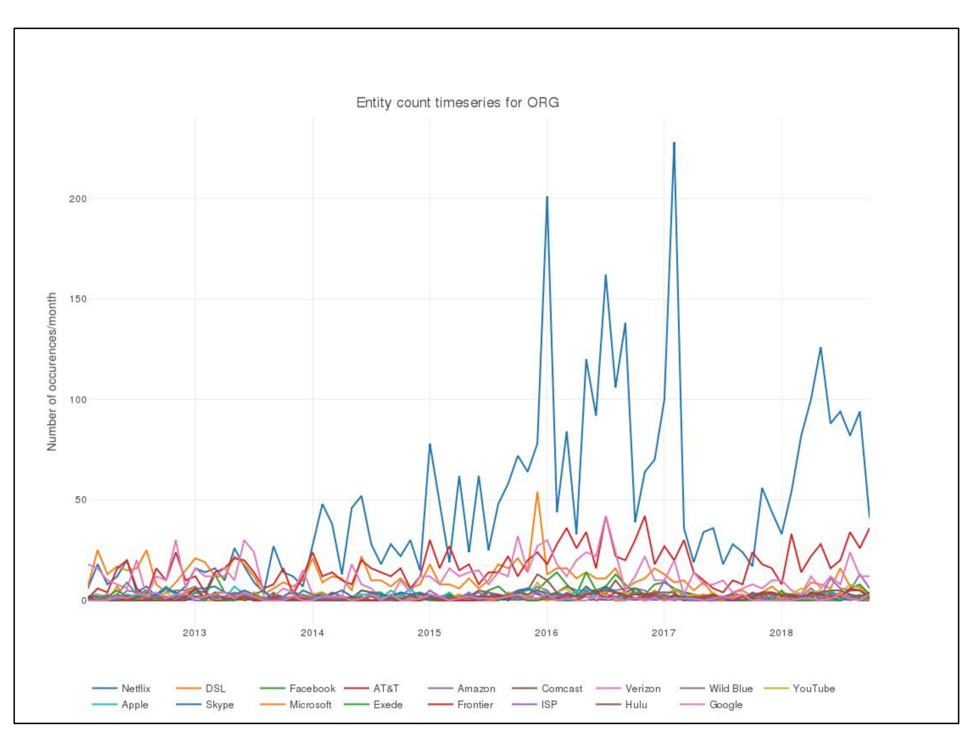

Figure 3. Named Entities (Organizations) over time

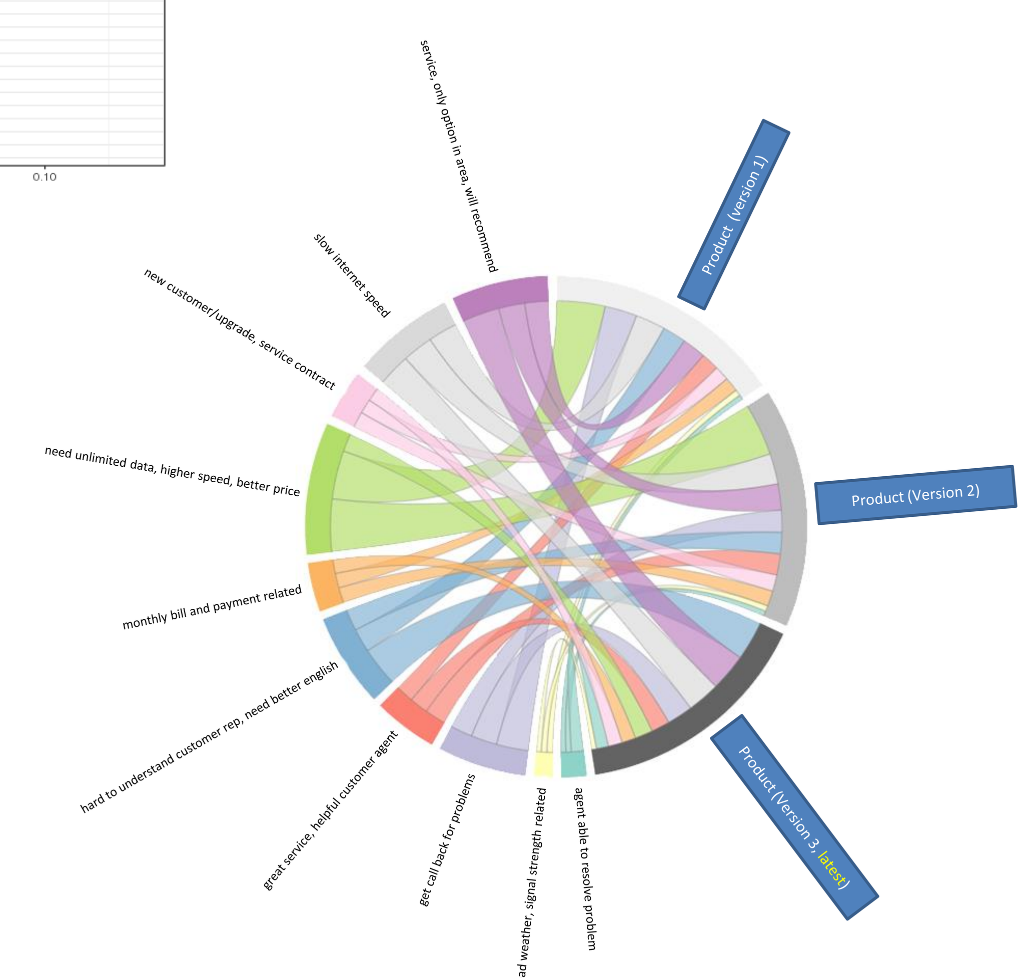

Figure 5. Topic distribution by product line (normalized to percentage). Width of a chord represents the percentage prevalence of a topic for a product line.

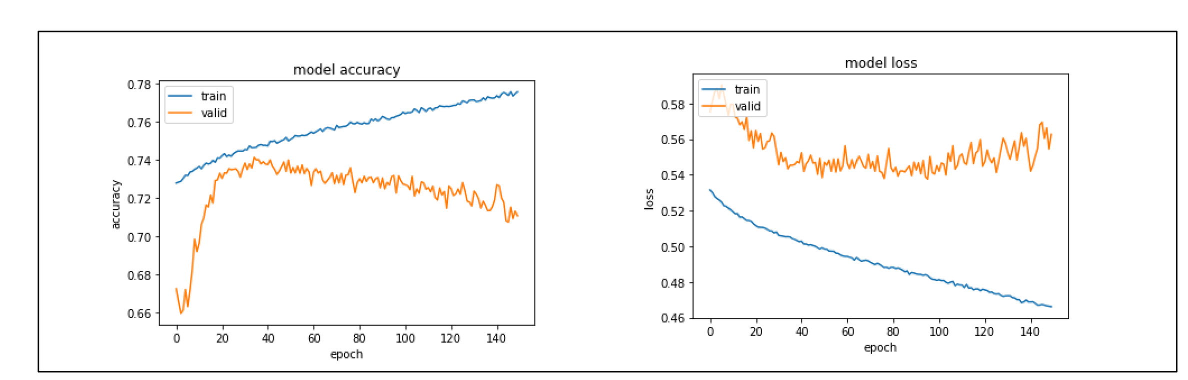

Figure 9. Multi Layer Perceptron (MLP) training/validation accuracy and loss. Validation set accuracy starts to drop after 35 epochs. **Model available after 35 epochs (accuracy 73%) is chosen for predictions.**

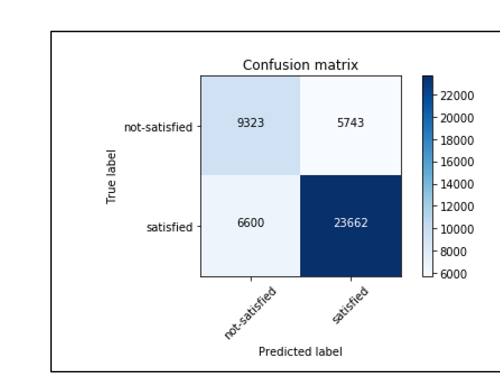

Figure 10. Confusion Matrix for predictions on holdout set

Contact Information

Amit Arora
Georgetown University
3700 O St NW, Washington, DC 20057
aa1603@georgetown.edu
301-525-6705

References

1. Quoc V. Le and Tomas Mikolov. 2014. Distributed representations of sentences and documents. CoRR, vol. abs/1405.4053.
2. Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space," CoRR, vol. abs/1301.3781.
3. Andrej Karpathy. 2015. The Unreasonable Effectiveness of Recurrent Neural Networks, http://karpathy.github.io/2015/05/21/rnn-effectiveness/.
4. David Nadeau, Satoshi Sekin. 2007. A survey of named entity recognition and classification.
5. Hamed Jelodar , Yongli Wang , Chi Yuan , Xia Feng. 2017. Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey.
6. Matthew L Jockers. 2013. Revealing Sentiment and Plot Arcs with the Syuzhet Package.
7. Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, Eduard Hovy. Hierarchical Attention Networks for Document Classification. https://www.cs.cmu.edu/~diyiy/docs/naacl16.pdf