Complaint Assistant

Helping companies predict complaint escalation

Hao Lin

Dalton Chiscolm Lawsuit: Man Sues Bank Of America For "1,784 BILLION, TRILLION Dollars"

Problem

Complaint Escalation

- Huge money loss
- Lose customers
- Ruin reputation

After a positive customer experience:
- 69% would recommend the company to others
- 50% would use the business more frequently

After a negative experience:
- 58% will never use the company again
- 49% will tell friends not to use the business
- 34% would take revenge by posting a review online or sharing a poor experience on social media

Product

Developed a Web app

- Categorize the product type of the complaint
- Predict the chance of complaint escalation
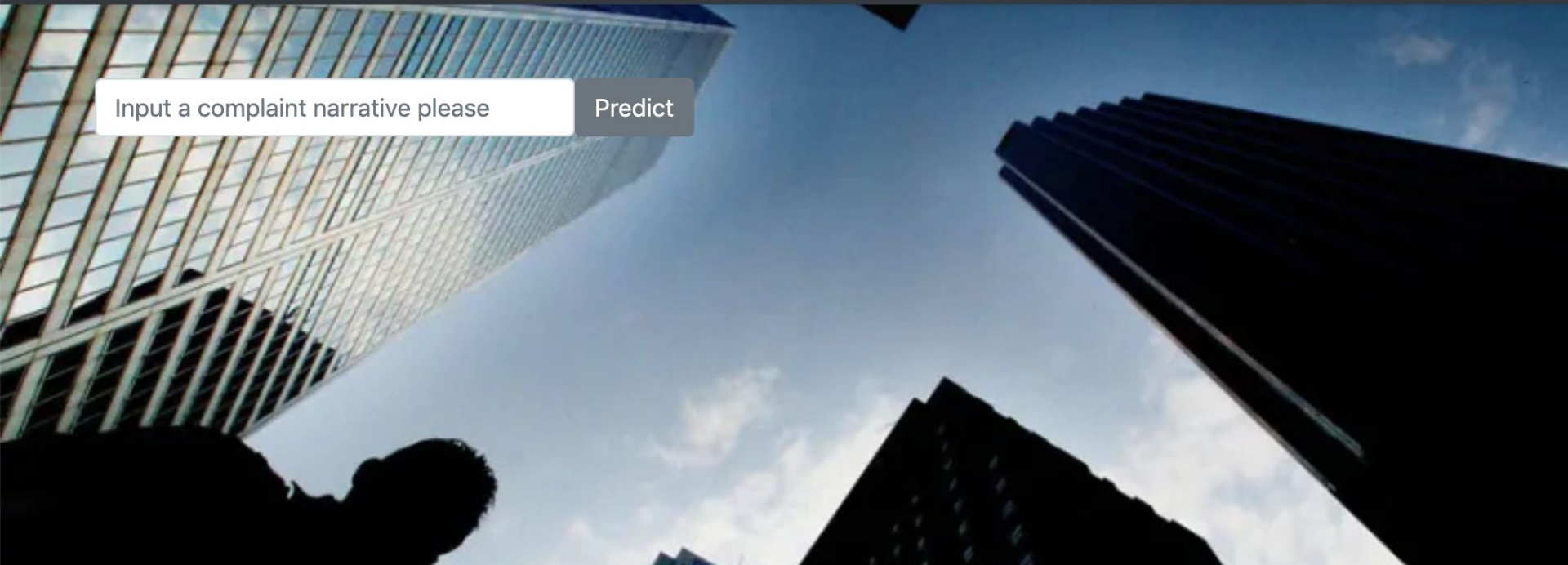- Suggest best response type to prevent escalation

Input a complaint narrative please
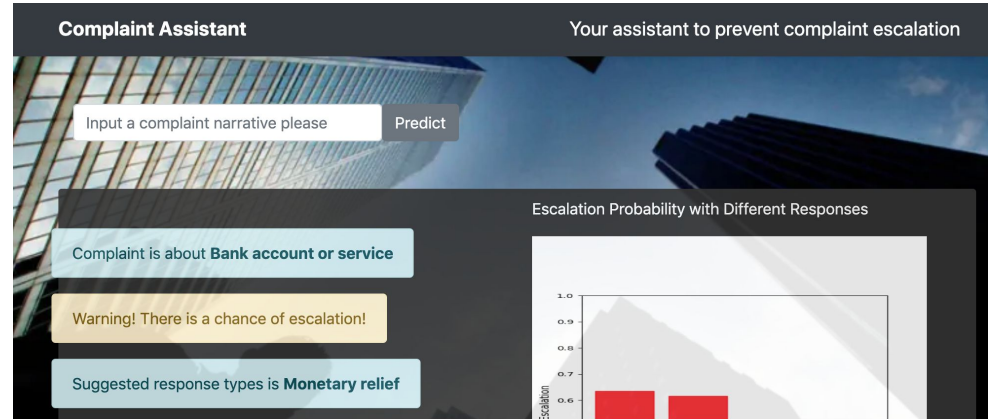
Predict

Complaint Assistant

Your assistant to prevent complaint escalation

I have a complaint regarding the overdra

Predict

I have a complaint regarding the overdraft fees that were billed to...

Try it

balala

I have a complaint regarding the overdraft fees that were billed to my checking account. I have a complaint regarding the overdraft fees that were billed to mychecking account. I was charged XXXX overcharge fees for XXXX withdrawals in which I had funds in the account. I contact your office and spoke with a representativewho credited me with XXXX of the fees back. However, the XXXX fee was never credited. I just do n't understand how I can billed for an overdraft fee when the fundswere in my accounts. I contacted the office of the president for Flagstar Bank and my compliant was pushed aside. Flagstar has now filed a writ of garnishmentwith my employer.
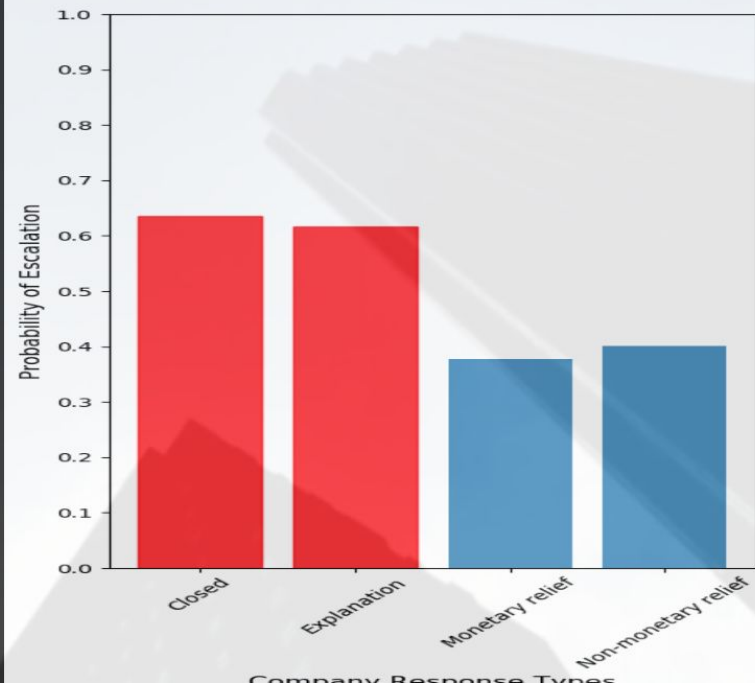
Data : 16k Historical Consumer Complaints

Consumer Financial Protection Bureau

Consumer Tools   Practitioner Resources   Data & Research   Policy & Cor

Consumer Complaint Database ^{BETA}

Consumer complaints are added to this public database after the company has responded to the complaint, confirming a commercial relationship with the consumer, or after they've had the complaint for 15 calendar days, whichever comes first. We don't verify all the facts alleged in complaints, but we do give companies the opportunity to publicly respond to complaints by selecting responses from a pre-populated list. Company-level information should be considered in the context of company size and/or market share.
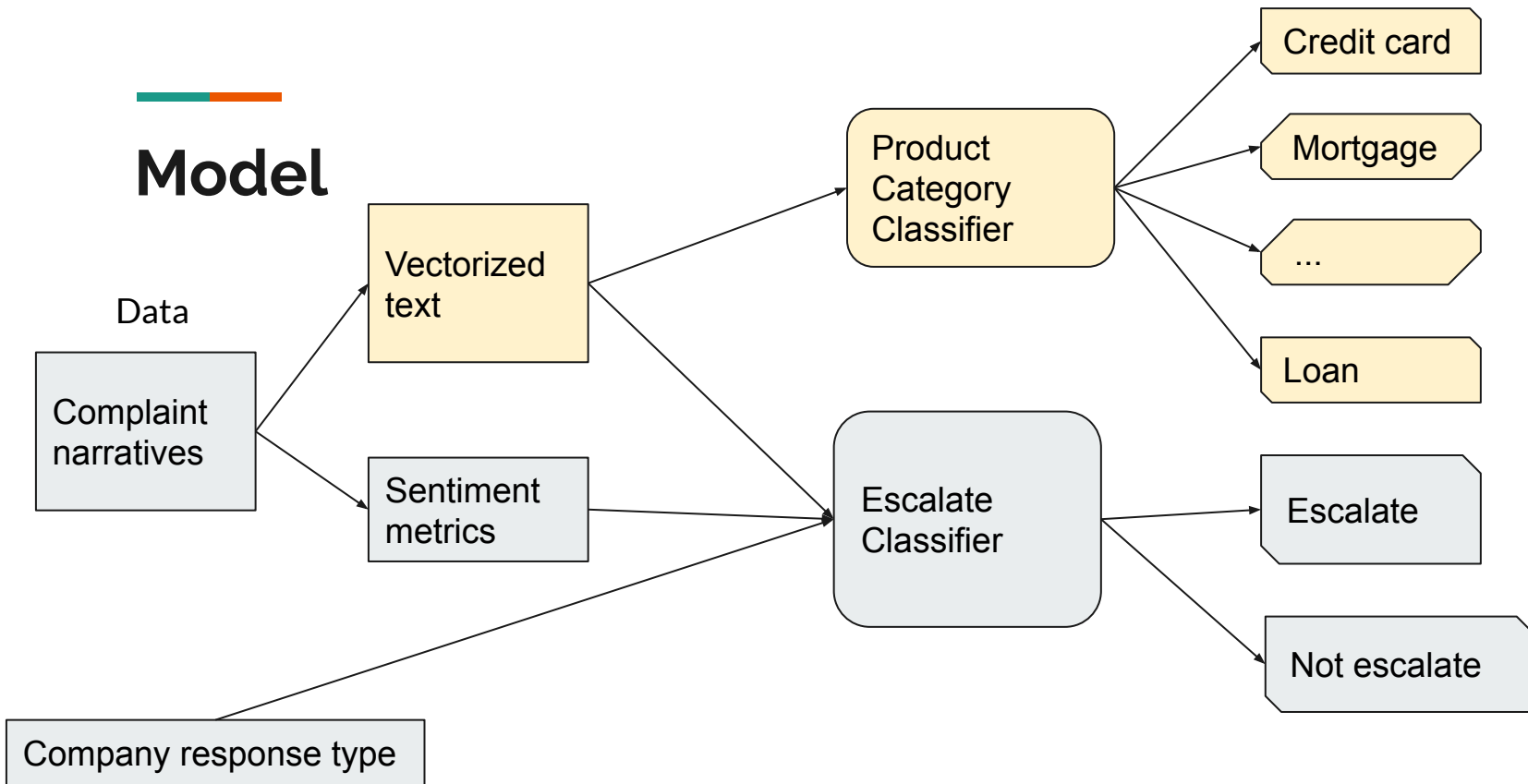
complaints about financial products

- Labeled (Dispute Yes / No)
- Narratives
- 13 product categories
- Company responses

Model
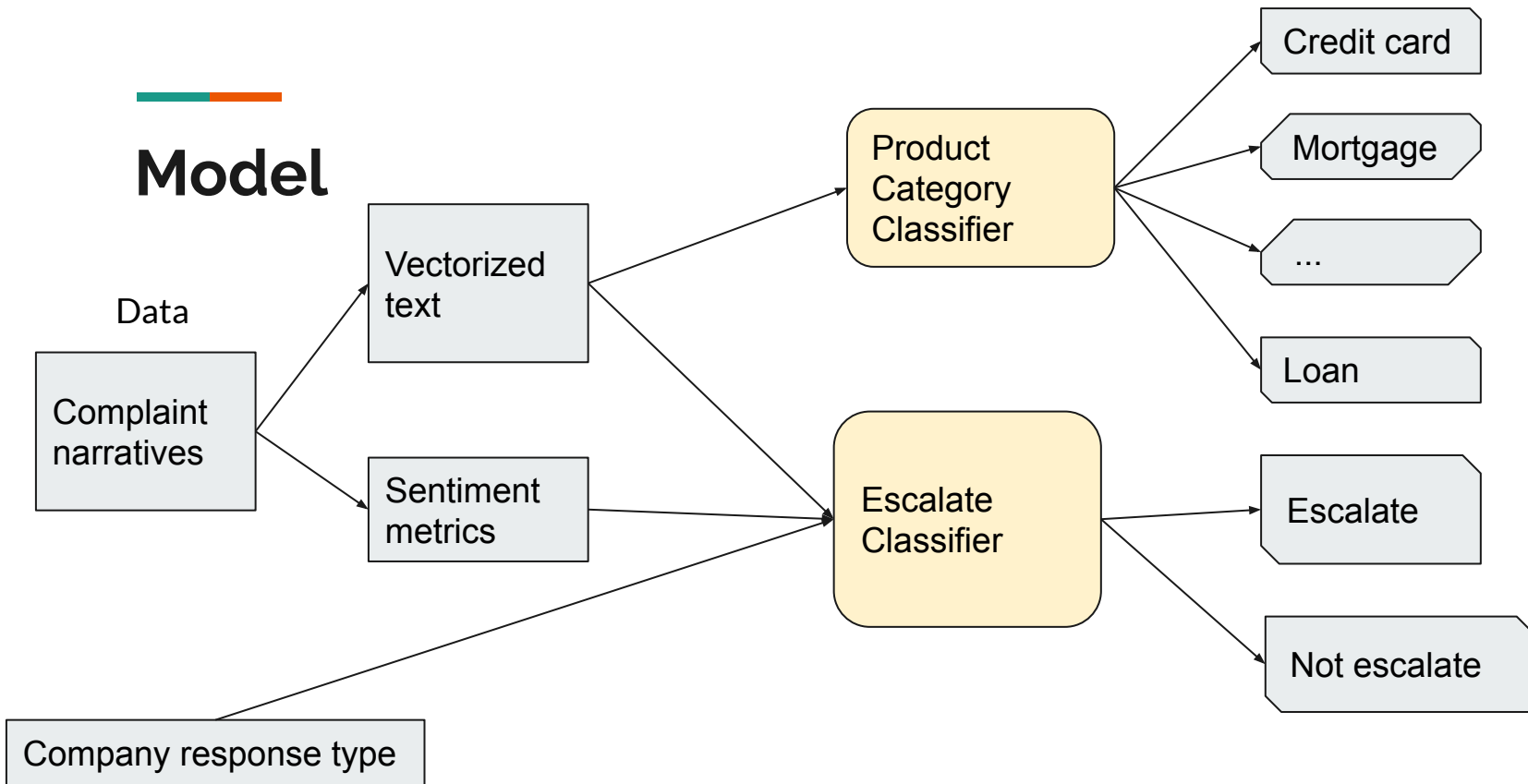
Model

Model

Tokenization
Lemmatization
Stopword
Tf-idf
vectorization

Data

Complaint narratives

Vectorized text

Sentiment metrics

Company response type

Product Category Classifier

Escalate Classifier

Credit card

Mortgage

...

Loan

Escalate

Not escalate

Model

Data

Complaint narratives

Vectorized text

Sentiment metrics

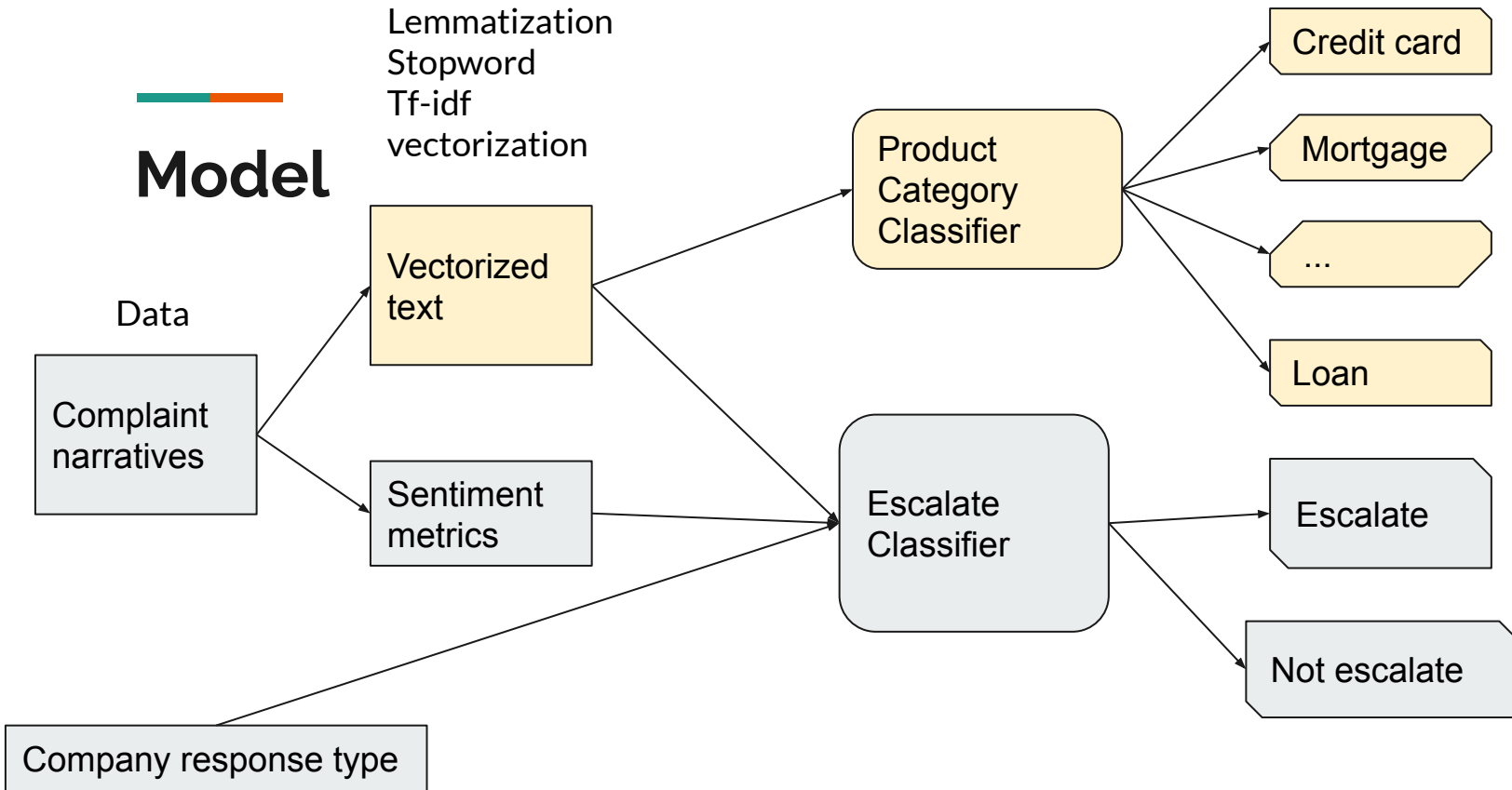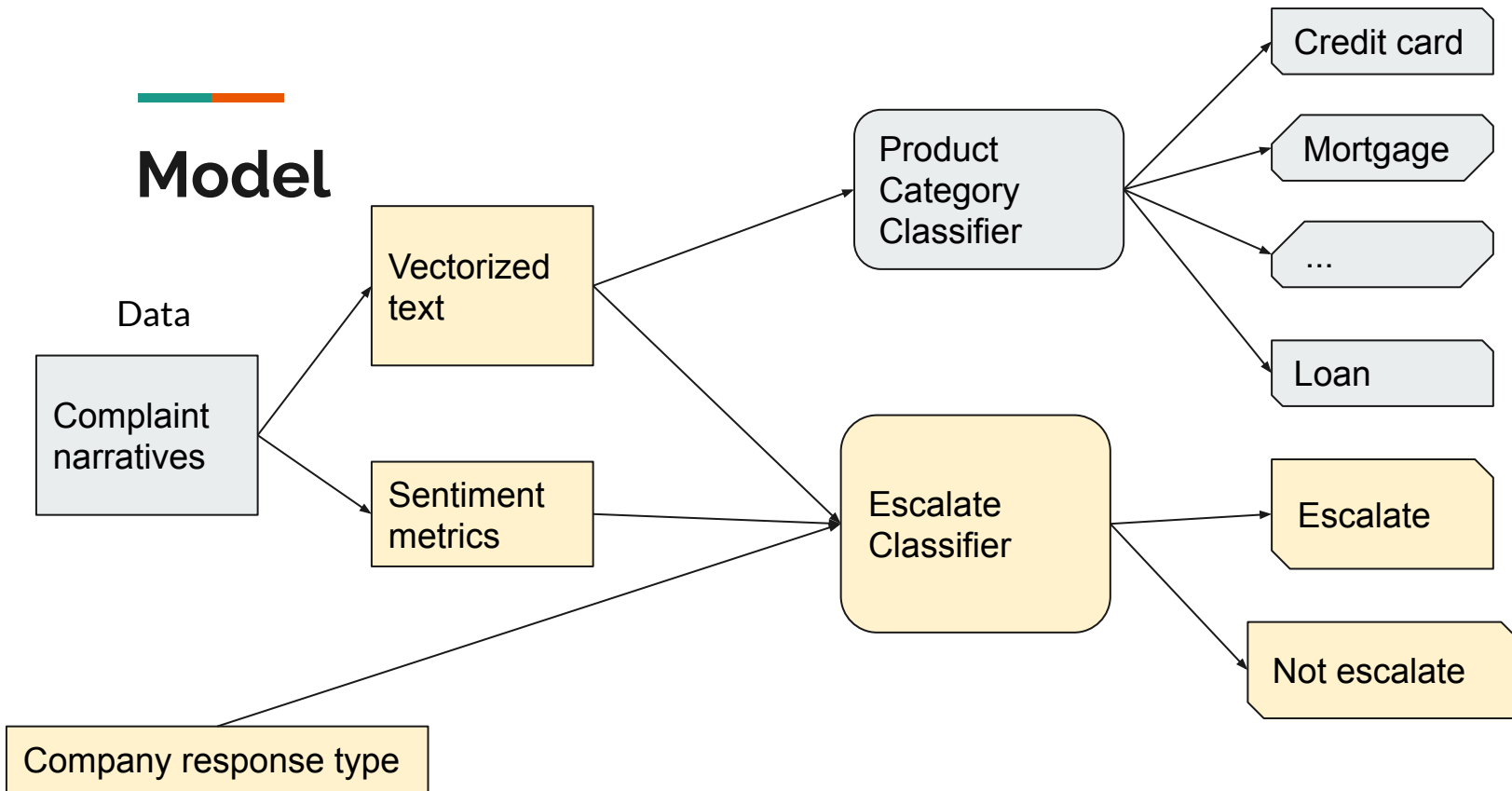Product Category Classifier
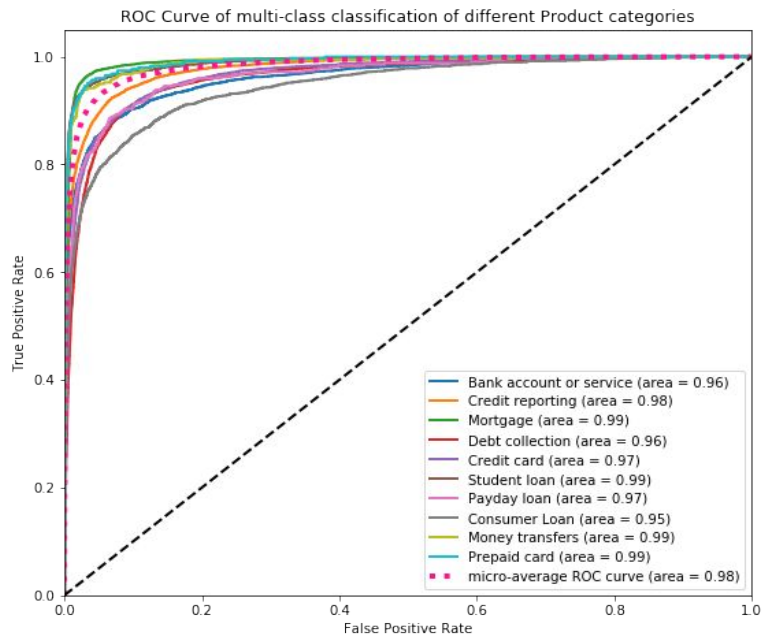
Credit card

Mortgage

...

Loan

Escalate Classifier

Escalate

Not escalate

Company response type

ROC Curve of Product Category Classifier

ROC Curve of multi-class classification of different Product categories

True Positive Rate

False Positive Rate

Bank account or service (area = 0.96)
Credit reporting (area = 0.98)
Mortgage (area = 0.99)
Debt collection (area = 0.96)
Credit card (area = 0.97)
Student loan (area = 0.99)
Payday loan (area = 0.97)
Consumer Loan (area = 0.95)
Money transfers (area = 0.99)
Prepaid card (area = 0.99)
micro-average ROC curve (area = 0.98)

OneVsRest, Logistic Regression

AUC score 0.98

ROC Curve of Escalate Classifier

ROC curve for escalate classifier for lgrg
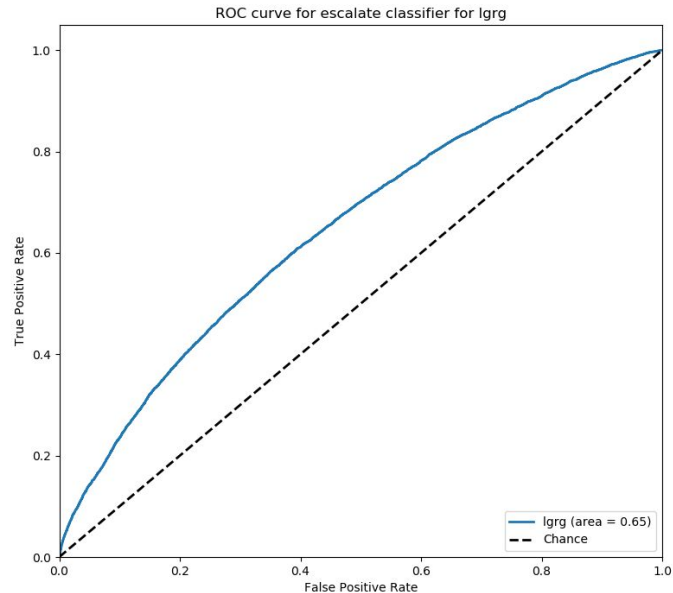
Logistic Regression(C=1)

AUC for model is: 0.648

Validation

Validate using 1000 recently submitted complaints

- 240 are labeled as escalation in the data
- 118 escalation are detected
- 117 out of of 118 are suggested to adopt a response different from current response

Half of the escalation could be eliminated !!

Conclusion

Complaint Assistant

- Leveraged by NLP and machine learning
- Could be customized to specific company

Other Application

- Product category classifier -> Customer segmentation
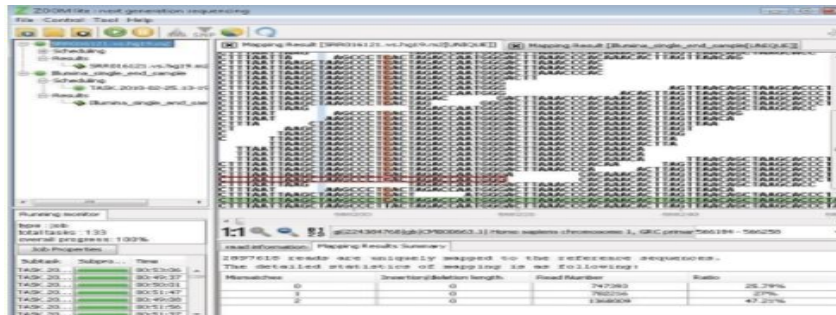- Imbalanced data -> Fraud detection

Complaint Assistant

Hao Lin

PhD in Computer Science

NLP processing of narrative

Tokenization

Stop word

- Combine two nltk stop word library and sklearn.feature_extraction
- Remove stop words with strong sentiment. Eg. shouldn't, never

Lemmatization

TF-IDF vectorization   (ngram=(1,3) max_feature_num = 50000

Sentiment Metrics for each narrative

Based on sentiment score generated by Vader for each sentence

- # of sentences
- # of words
- the most negative score
- Negative ratio
- sum of sentiment score
- average sentiment score
- # of ?
- # of !

Future Plans

More accurate sentiment analysis

- Train sentiment analysis model other than using Vader as base

Incorporate more semantic information

- Fine tune doc2vec
- Bi-LTSM

Generate dangerous word library

Customer complaint content and sentiment

'credit card number : XXXX set up automatic payment to " Pay amount due on due date \'\' for this Bank of America Visa Card (under XXXX XXXX XXXX). However, on XXXX, only minimum payment of {$25.00} was deducted. This resulted in I being charged interest on XXXX (and potentially future date), and potentially affect my credit score. I contacted BOA but it is a mess there. I was being directed here and there with no resolution. \nThe mistake is unexcusable. My family have all kind of credit cards and have never seen a mistake like this. I like BOA to:1. correct his promptly and refund any penalty interest charged. \n2. explain to CFPB why this happen and how many customers have been affected by this mistake, how much fee they have collected out of this mistake. \n3. explain to me whether this affect my credit score and correct promptly if it does.'
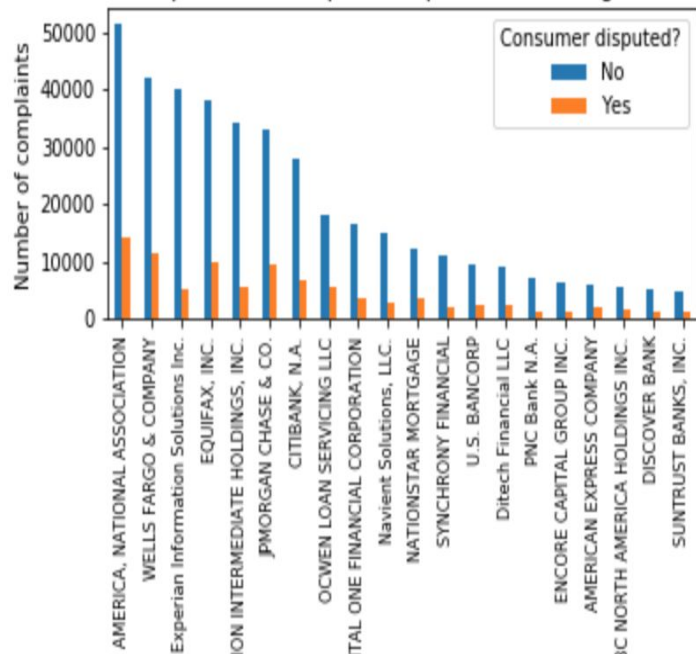
In [ ]:

The bar chart of complaints number in different Product categories
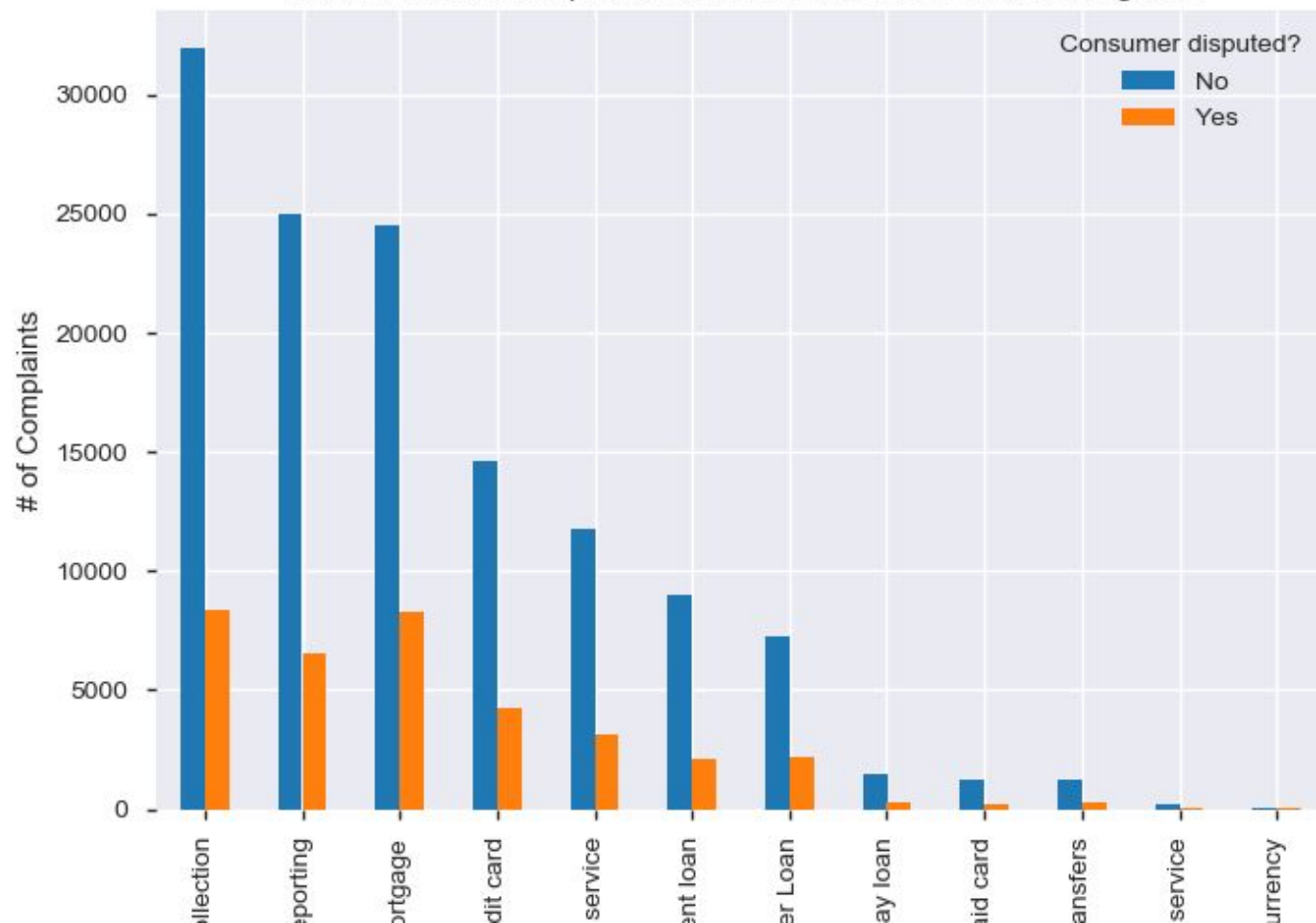
How the company response matters

Number of complaints from top 20 companies containing most complaints

Dispute Ratio of complaints of 20 companies with most number of Complaints