

# What are consumers complaining about?



## Financial Products Complaints

---

Since 2009, large US financial institutions have paid **\$187 billion in fines** in 140+ cases, with additional 250+ cases pending

*The Power of Complaints: Unlocking the Value of Customer Dissatisfaction, Deloitte Research*

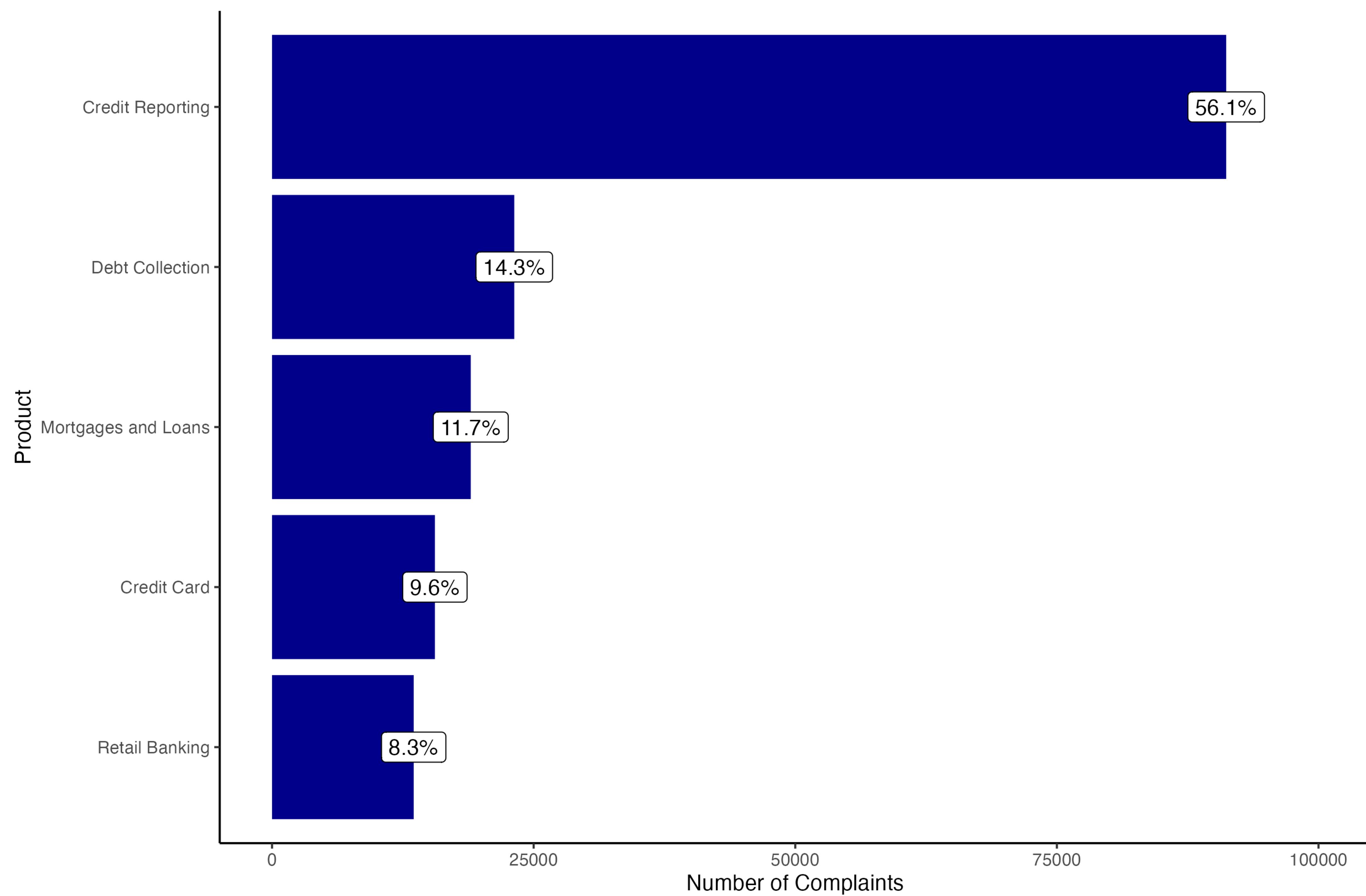
---

	<b>product</b>	<b>narrative</b>
0	credit_card	purchase order day shipping amount receive product week sent followup email exact verbiage paid two day
1	credit_card	forwarded message date tue subject please investigate comenity bank retailer card scam sent hello name
2	retail_banking	forwarded message cc sent friday pdt subject final legal payment well fargo well fargo clearly wrong need look
3	credit_reporting	payment history missing credit report specialized loan servicing sl made mistake put account forbearance
4	mortgages_and_loans	beginning mortgage held mb financial mb mortgage portfolio purchased bank first quarter mb financial customer
5	debt_collection	collection agency system inc mn phone sent letter regarding debt owe immediately sent system inc u postal

# Data

## Consumer Financial Protection Bureau (CFPB)

- CFPB is a US federal agency for mediating consumer complaints related to financial institutions or products
- Size: 99.46 MB
- Dimensions: 162, 420 rows; 2 columns
- Source: Kaggle (<https://www.kaggle.com/datasets/shashwatwork/consumer-complaints-dataset-fo-nlp>)
- March 2020 to March 2021



# Goals

- Develop an NLP model to predict product given a consumer complaint
  - Useful for businesses to “transfer” to the appropriate team
- Find what issues consumers are complaining about
  - Useful for developing new features

# Methodology

- TF-IDF to convert consumer complaints into a data matrix
  - 10,000 features; 162, 420 rows
  - Uni-gram and bi-gram word embeddings
- Predictive Model
  - Multinomial regression, Random forest and k-nearest neighbour to predict complaint category
- Topic Modelling with LDA
  - Five topics consumers are complaining about for each product category

# Retail Banking

make new cash  
opened asked fraud service still went  
able issue well get said customer never debit open  
contact balance year  
name payment number  
access credit one card told bank transaction made another  
department business sent back pnc time call information without  
see month checking fund chase need like week  
app help transfer due go u  
dispute even branch called got day check could since hold deposited  
take america claim received email paypal phone charge saving  
complaint closed deposit also amount person  
tried letter two contacted case receive

# Mortgages and Loans

another property stated want  
contacted checknumber contact  
phone made work asked forbearance  
insurance back pay get due paid still escrow know  
representative amount like received since process  
provided help able spoke account loan see complaint  
statement sent new u one credit email need  
balance tax told also time fee bank end please  
response week letter call even fargo  
lender could late mortgage would company year issue  
application rate interest home month day two called never money financial  
refinance customer well first make said information request  
requested take document date service report applied  
vehicle monthly going

# Credit Card

department transaction information request date capital first received contacted customer never receive issue make used said payment card bank called interest complaint citibank still get bill back account would call fee like pay fraud spoke new number dispute u told balanceservice phone another asked paid year chase full late tried amount statement help need charged please email contact merchant representative

# Credit Reporting

response  
complaint inaccurate received equifax open  
get need block balance told listed name debt still investigation  
address fraudulent bureau day also letter law sent made accurate  
original charge inquiry file payment information contacted  
service number within remove card since  
deleted theft reporting credit report collection following  
removed fcra closed score well never year statement  
copy act company consumer date request several  
loan late agency item fair dispute identity right  
amount month would one proof section call please result  
violation experian creditor reported due back  
immediately transunion document even  
may mine must disputed filed

# Baseline Models

Model	Training Accuracy	Test Accuracy	Notes
Multinomial NB	83.60%	83.30%	No overfitting; acceptable performance
Random Forest	79.90%	76.18%	Model biased to the most popular class
kNN (k = 5)	99.49%	64.32	High overfitting

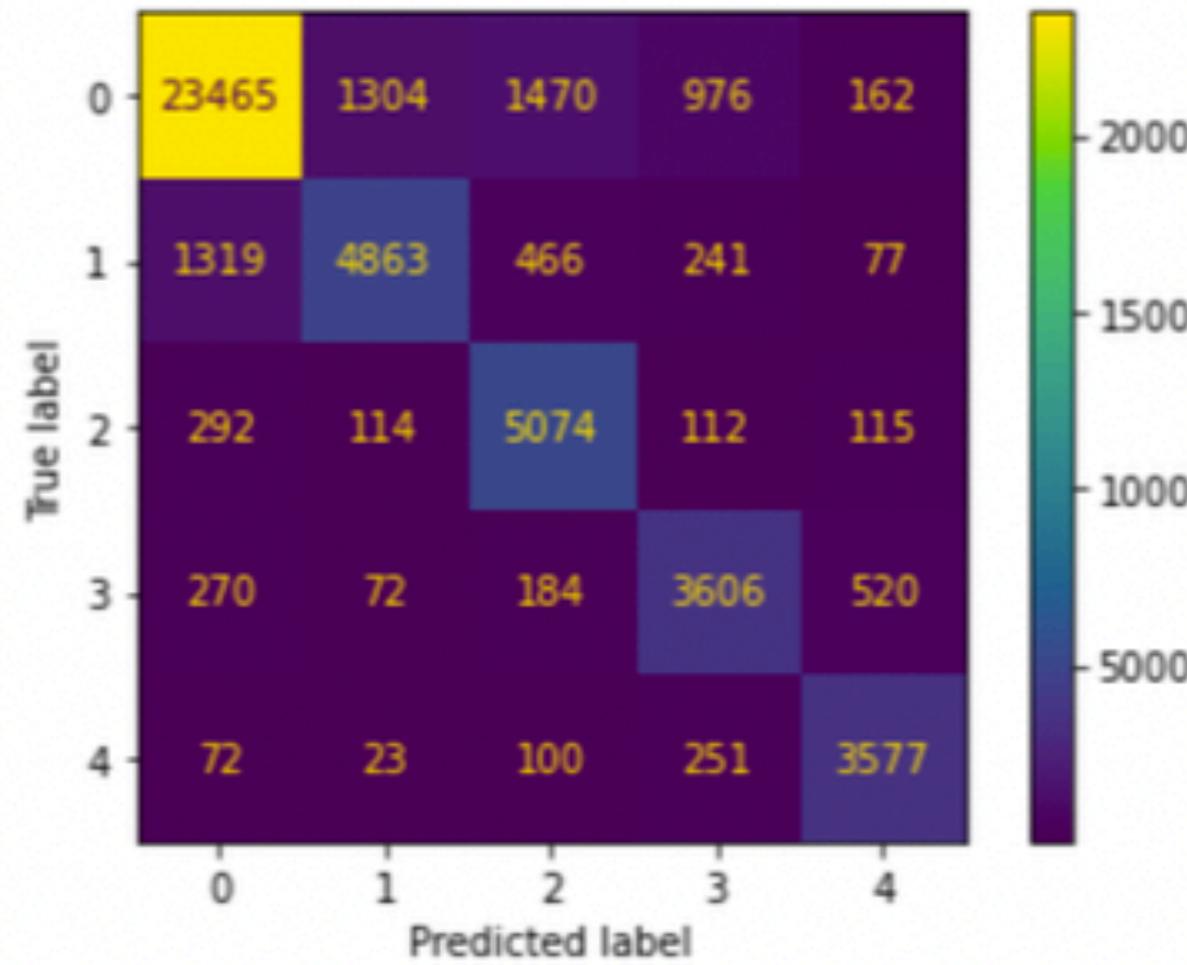


Figure 3: Confusion matrix for Multinomial Naive-Bayes Classifier on test data.

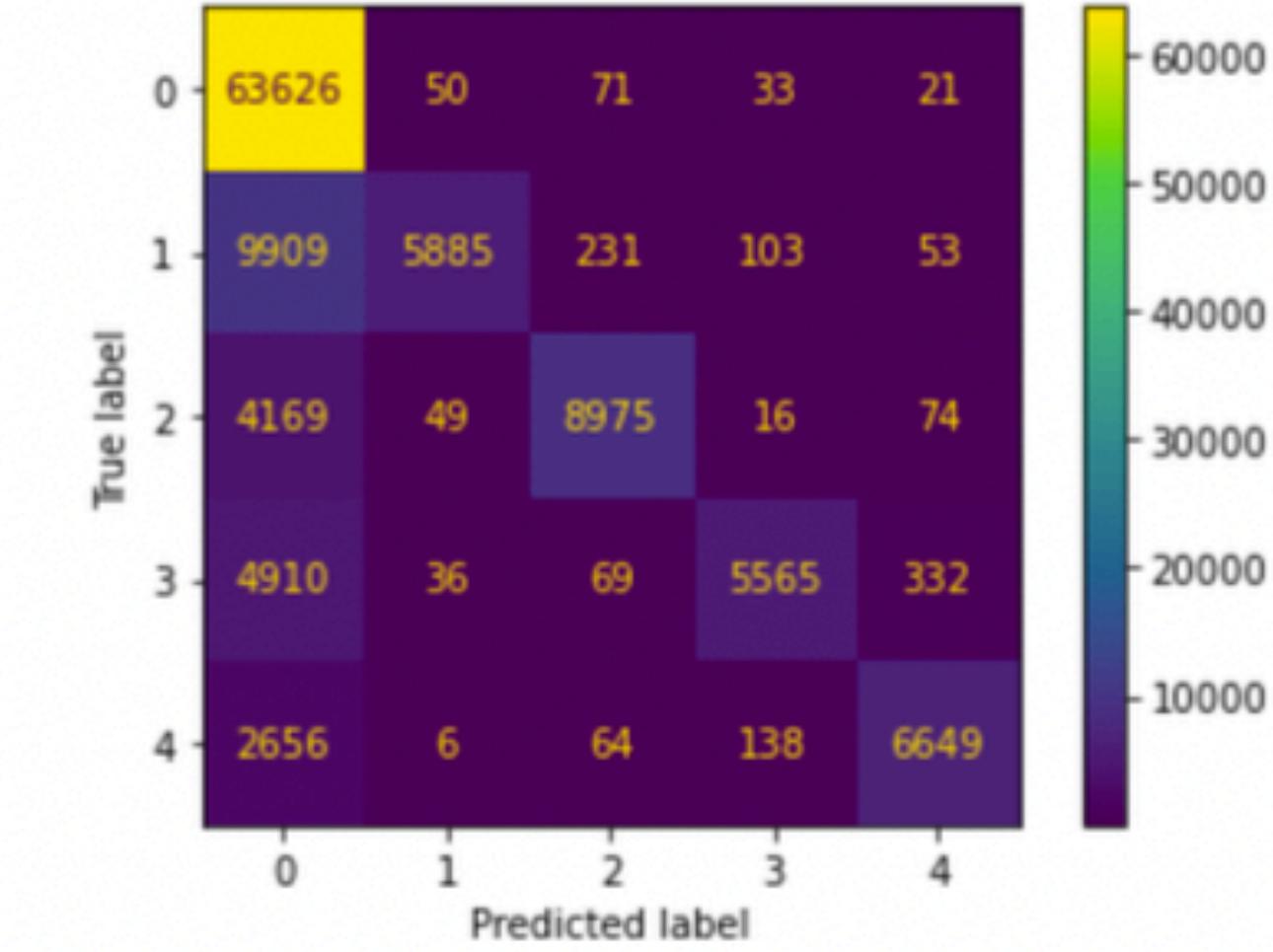


Figure 4: Confusion matrix for Random Forest Classifier on test data.

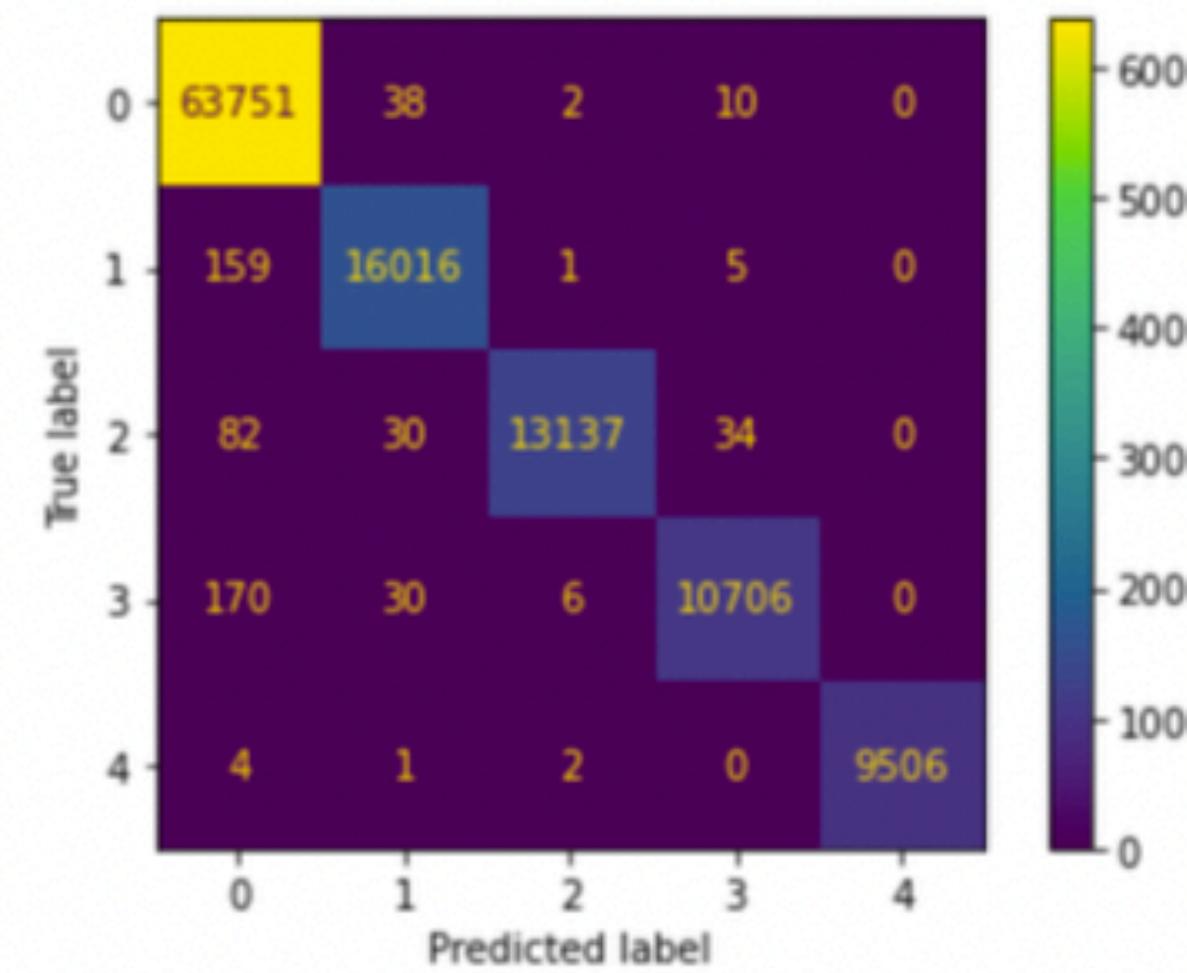


Figure 5: Confusion matrix for kNN algorithm on training dataset. As you can see, we have high accuracy here, actually 99%, but that's because of overfitting.

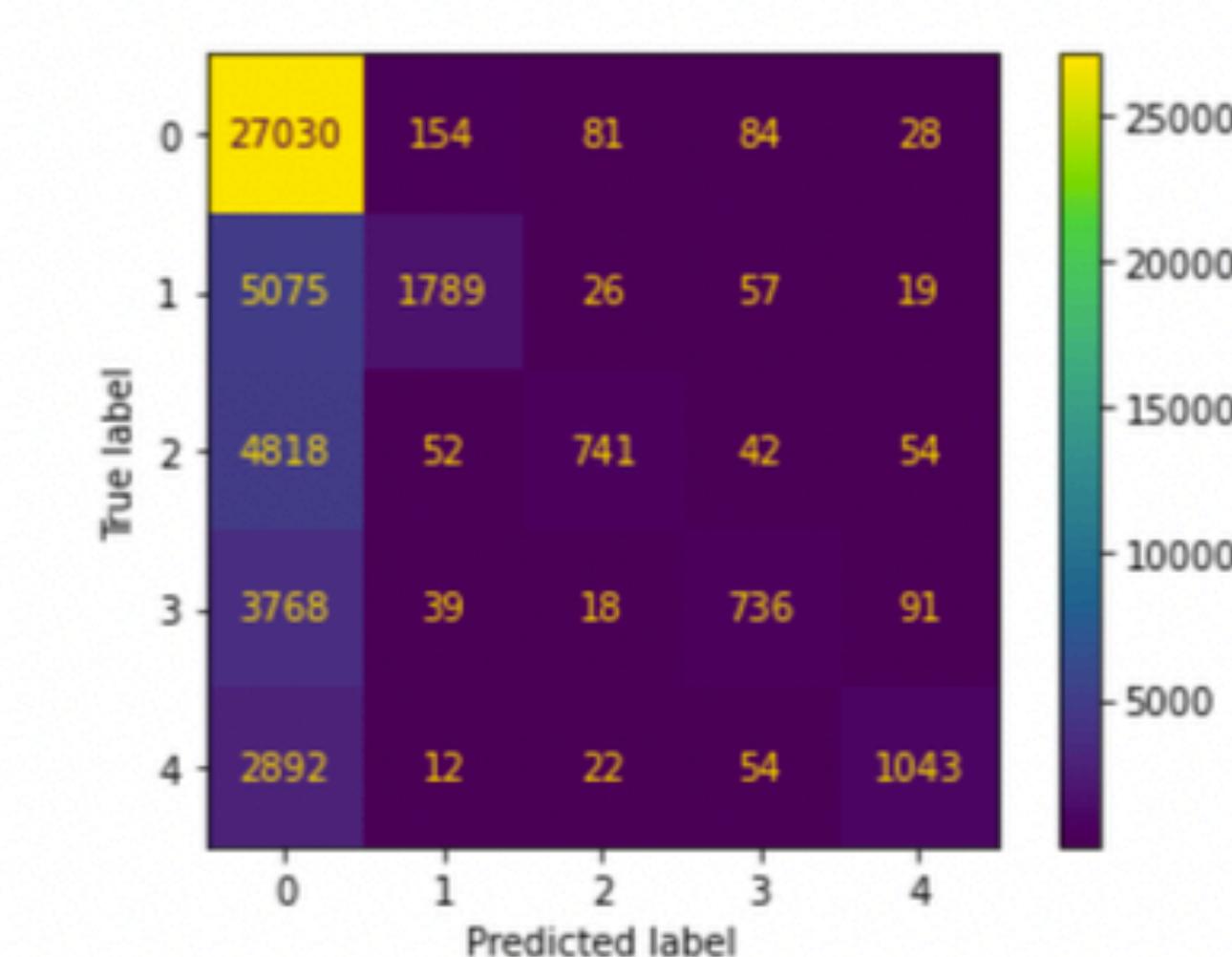


Figure 6: Confusion matrix for kNN algorithm on test dataset. The algorithm severely favours class 0 for all entries and has measly low accuracy of 64%.



# Nested Models Tuning

- Two models
  1. Predict class ‘credit reporting’
  2. Predict one of the other four classes
- RandomizedSearchCV followed by GridSearchCV
- Tuned Model 1: 90% (train accuracy)
- Tuned Model 2: 87% (train accuracy)
- Nested Model: 87% (train accuracy)
- Nested Model: 84% (test accuracy)

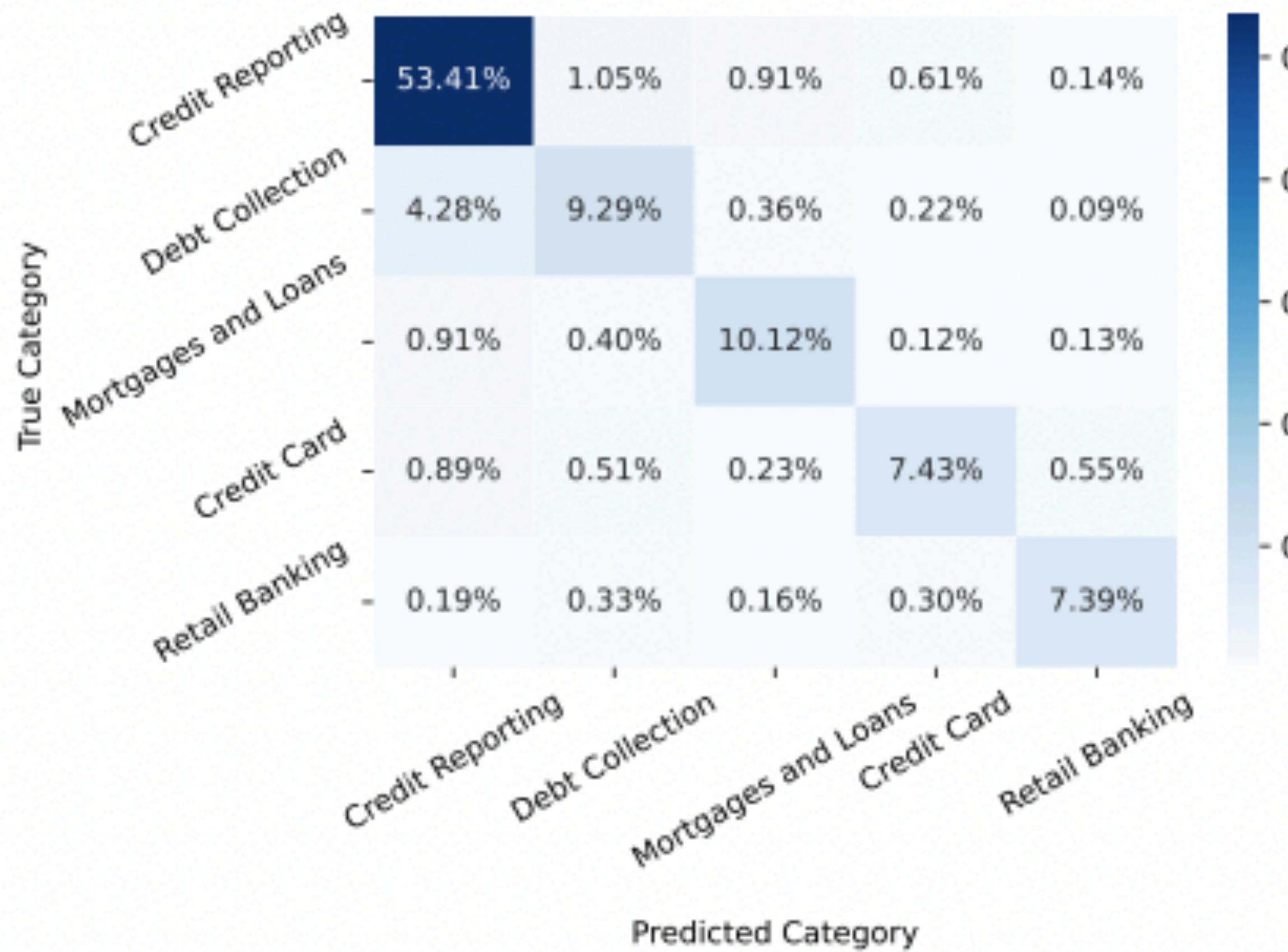


Figure 7: Confusion matrix for nested model on training dataset. The model performs great for except for debt collection, many of which are classified as credit reporting.

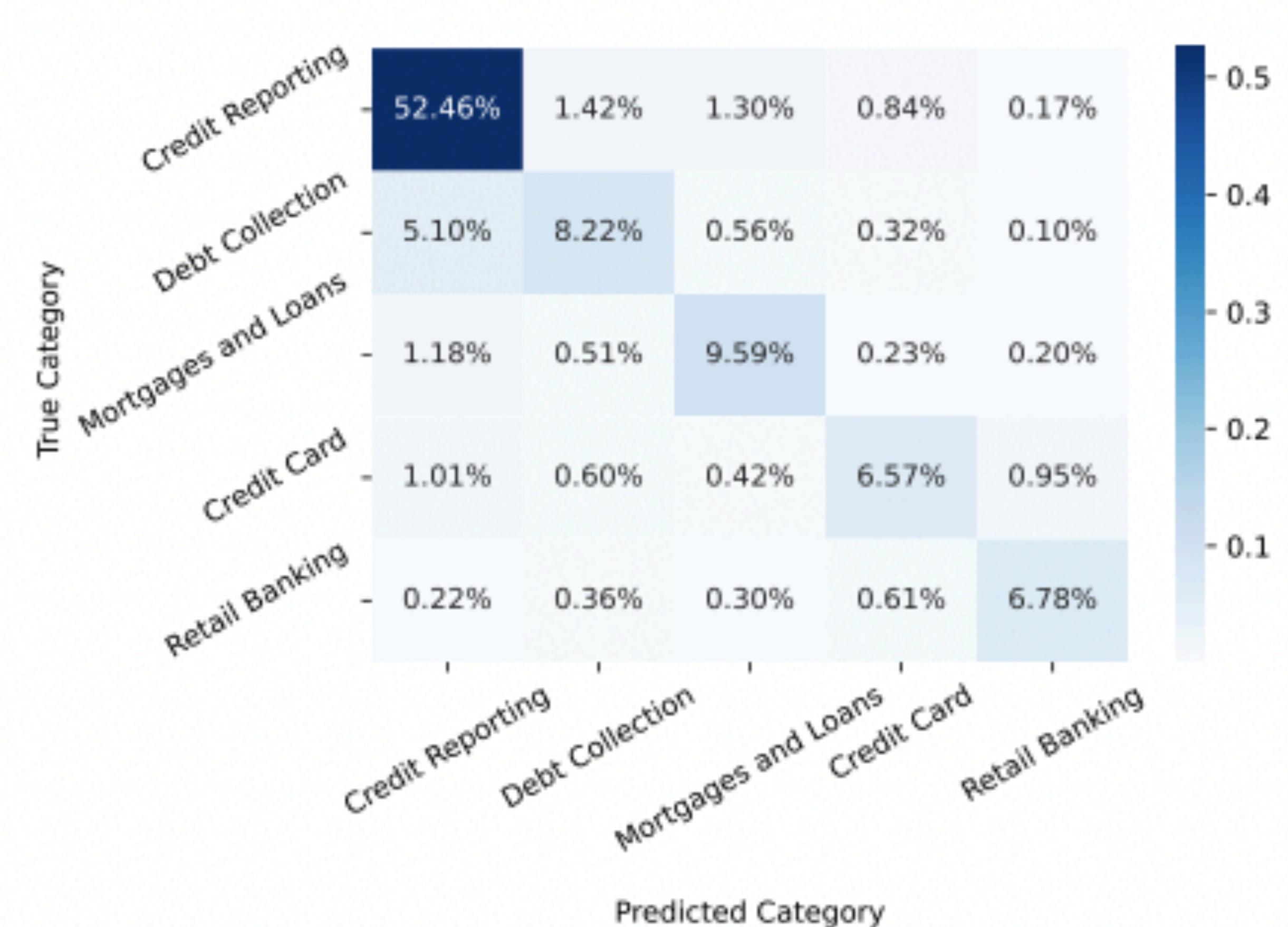


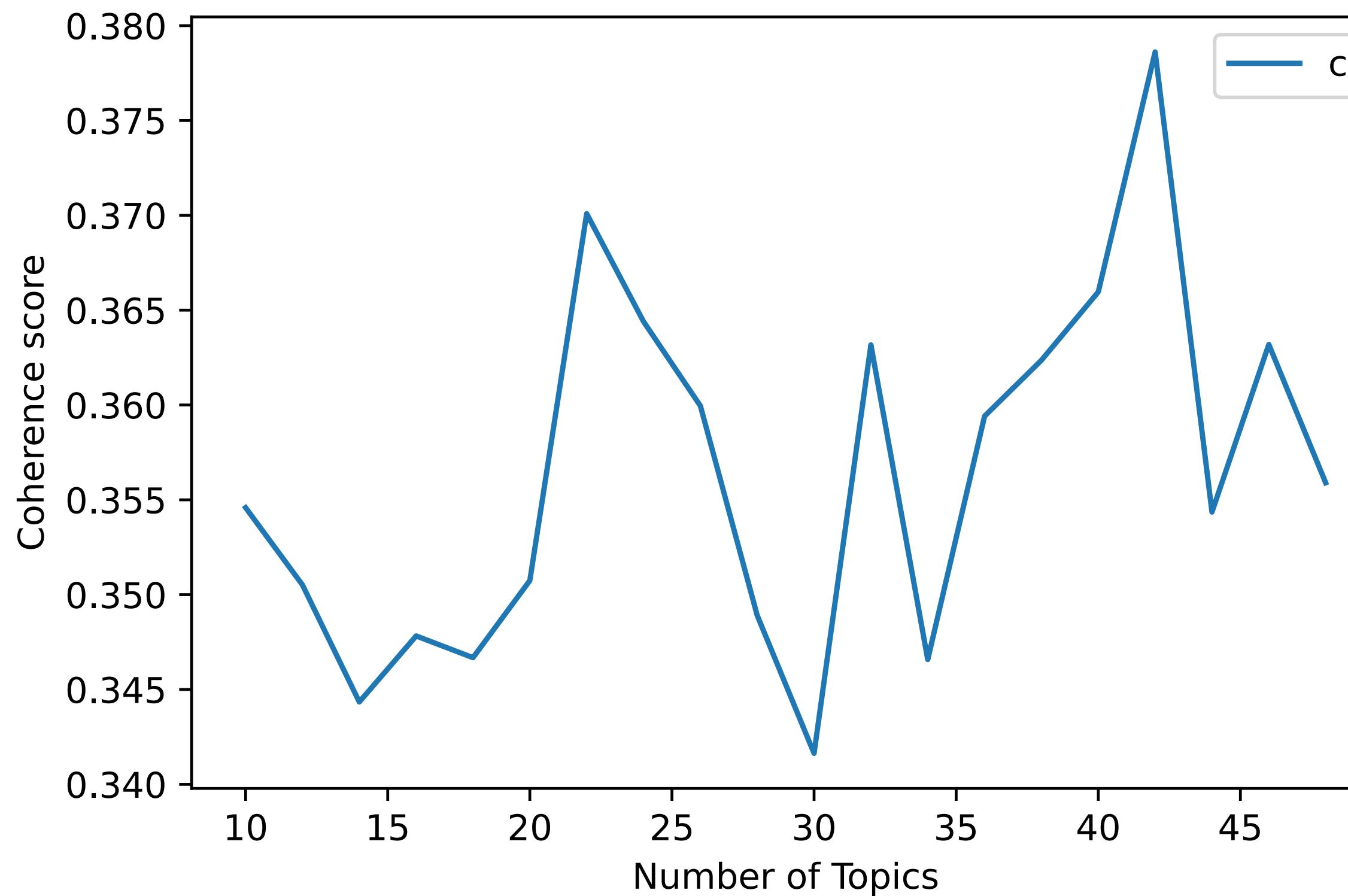
Figure 8: Confusion matrix for nested model on test dataset. The training model accuracy can very well be translated to test model accuracy.

# Complaint Topics Areas of Concern



# Credit Card

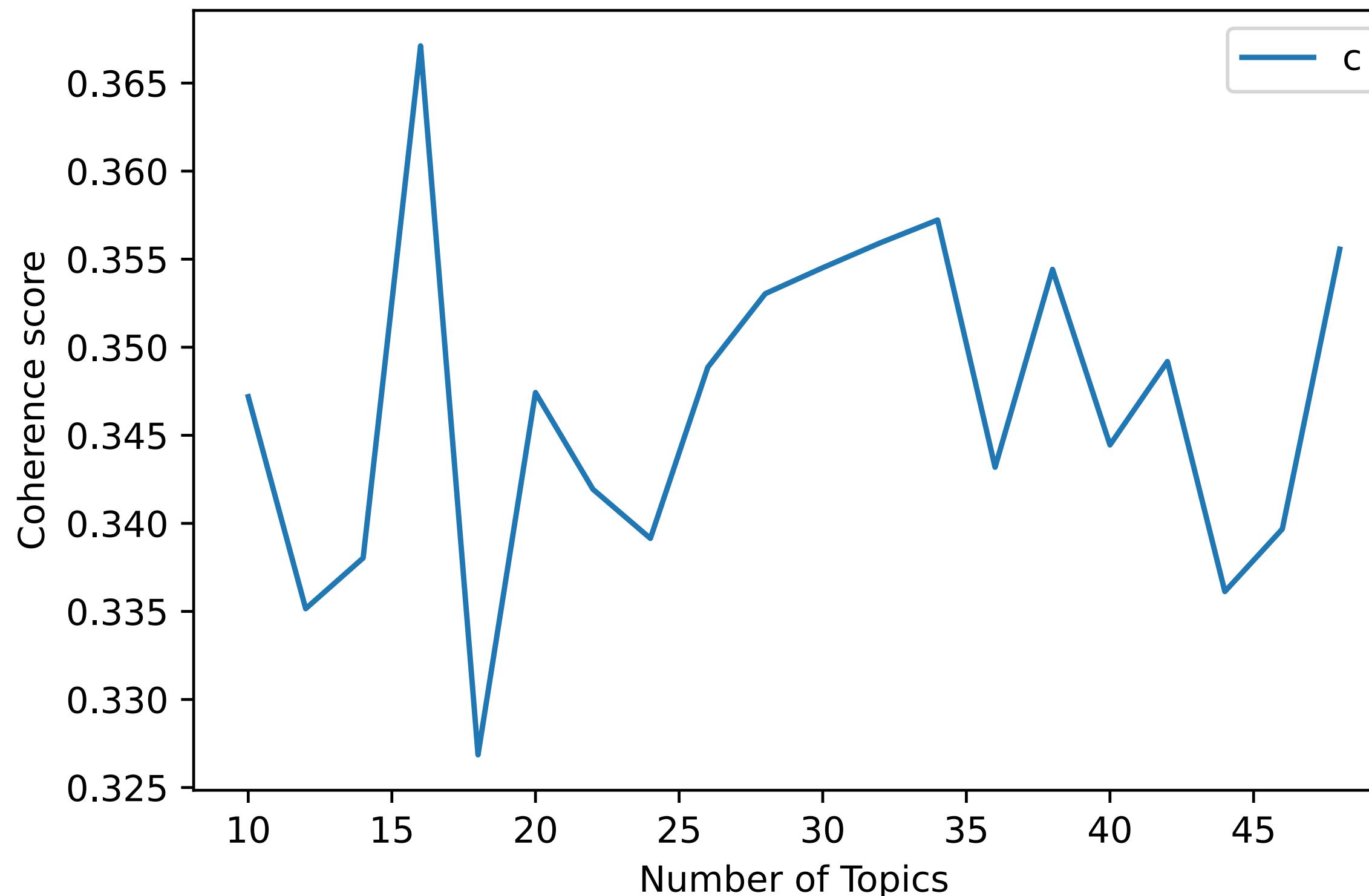
## 42 Topics



- Apple Pay
- Reward points; airline miles
- Amazon; ‘Didn’t call back’
- Late Fees
- Dispute

# Retail Banking

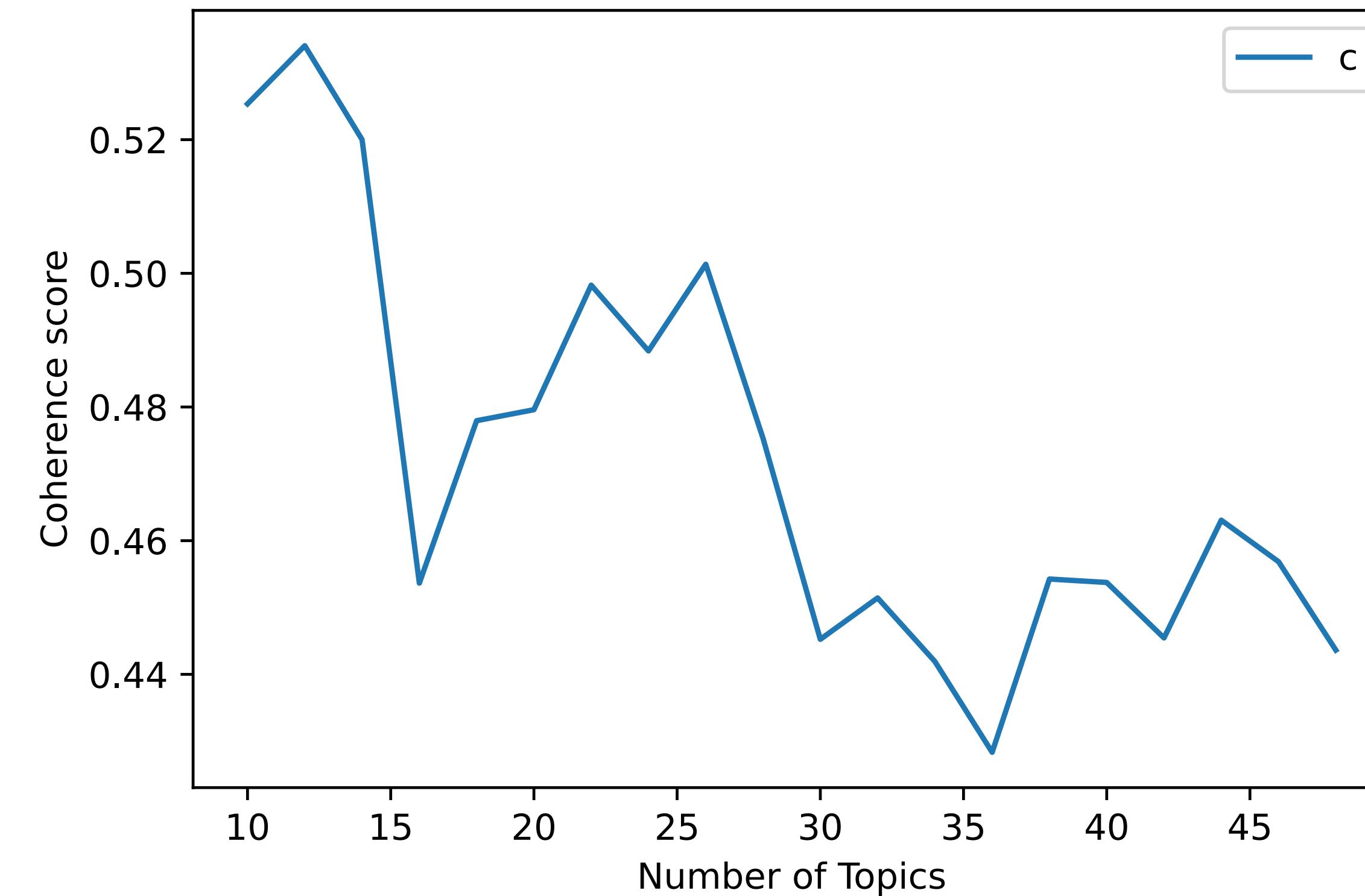
## 16 Topics



- Airline Ticket Refunds
- Amount Deducted, Merchant Didn't Receive
- Fraudulent Transaction Investigation
- Credit Score
- Account Never Opened

# Credit Reporting

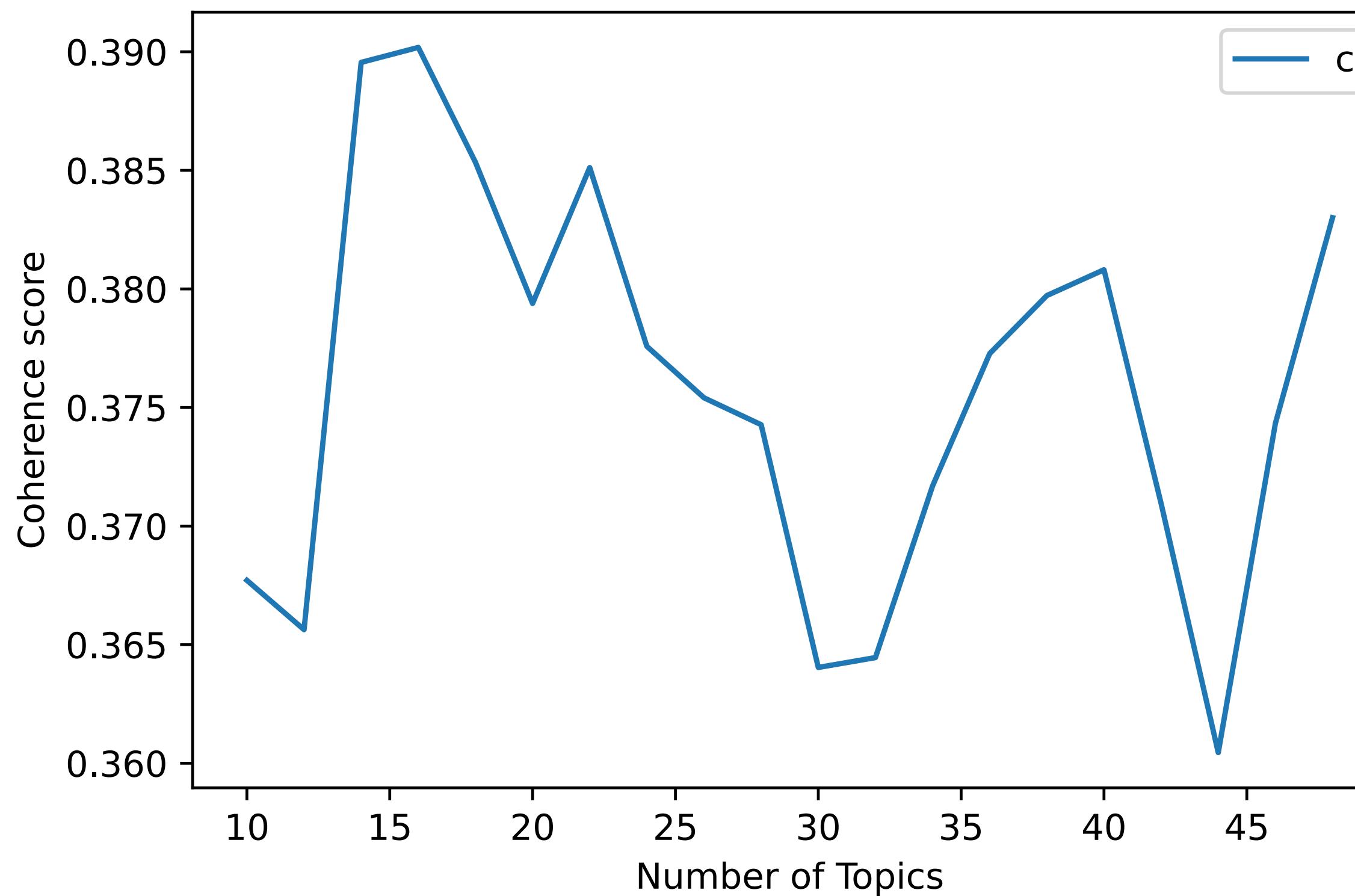
## 16 Topics



- Inaccurate Credit Reporting
- Vehicle Approval
- Account Never Opened
- Account Closed; Not Removed From Report
- Data Breach

# Mortgages and Loans

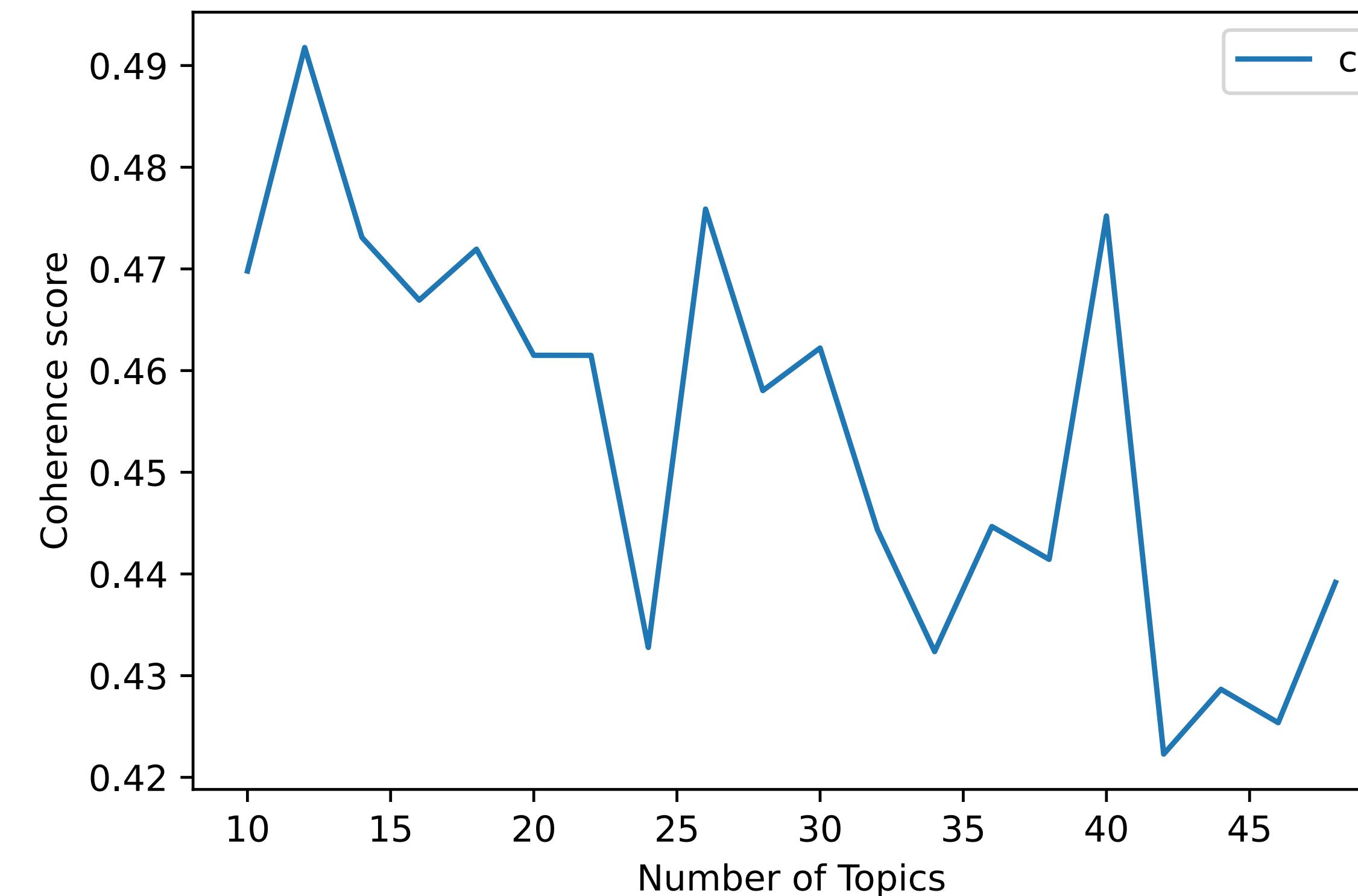
## 18 Topics



- Bankruptcy
- Car Dealership
- Rushmore, Pennymac, Caliber Home Loans, Wells Fargo
- Contracts
- Fairness

# Debt Collection

## 14 Topics



- Communication violation (FDCP Act)
- Full Settlement Not Updated
- Legal Notice
- Attorney and Court
- Car Vehicle Finance from Santander Bank

# Concluding Remarks

- Product complaints have cost financial institutions over \$187 billion since 2009. That doesn't include the cost of lost opportunities
- Predictive model with nested architecture is best in predicting complaint category
  - 87% training accuracy; 83% test accuracy
- Topic model with LDA to identify five areas of concern for each product category



[Scan QR code or](#)  
[click here for](#)  
[feedback please!](#)