

DSCI – 6011-02

Enhanced Customer Experience: Product Complaints Identification

- ***Sravya Kaitha***
- ***Jagannath Patro Allupati***
- ***Narsi Reddy Sanivarapu***

Enhanced Customer Experience: Product Complaints Identification Using NLP

- ▮ In today's business landscape, ensuring customer satisfaction is a crucial task for the success and longevity for any business. One of the key aspect of achieving this is through addressing and resolving customer complaints effectively. This has been a problem to most businesses due to the large volume of customer feedback that have to be analyzed manually.
- ▮ Application of Natural Language Processing (NLP) thus plays an important role in helping the businesses analyze customer feedback ensuring customer satisfaction. With NLP approach, businesses can accurately categorize customer complaints and this enables them to respond promptly and appropriately to each complaints raised by customers.
- ▮ This project aims to achieve this by leveraging NLP concepts to accurately classify customer complaints.

Statement of project objective

▮ The objective of this project is to develop an NLP model to accurately identify product complaints from various customer feedback with an aim to enhance customer experience by addressing the complaints effectively.

Statement of value

▮ This project is worthwhile because it addresses the vital need for businesses to efficiently handle and respond to consumer complaints. Businesses may streamline their customer service operations, reduce response times, and ultimately increase overall customer satisfaction and retention by automating the process of identifying complaints using NLP.

Approach

- ▮ Algorithms: Bi-LSTM model for multiclass classification and DistilBERT model (from Hugging Face).
- ▮ Datasets: Consumer Complaints Dataset with 162,421 complaints for 5 products.
- ▮ Link: https://github.com/halpert3/complaint-content-classification-nlp/blob/main/project_data/complaints_processed.csv
- ▮ Models: Utilization of Bi-LSTM and DistilBERT for classification.
- ▮ Tools and Techniques: Random oversampling to address dataset imbalance.

Approach

- ▮ Trained Bi-LSTM model and DistilBERT for multiclass classification.
- ▮ Detailed documentation outlining the model's architecture, training process, and evaluation metrics.
- ▮ Source code implementing the model and preprocessing steps.
- ▮ Report summarizing findings and potential areas for improvement.

References

▮ Gao, P., Sun, N., Wang, X., Yang, C., & Zitikis, R. (2024, March 26). NLP-based detection of systematic anomalies among the narratives of consumer complaints. ArXiv.org. <https://doi.org/10.48550/arXiv.2308.11138>

▮ Peter Adebowale Olujimi, & Abejide Ade-Ibijola. (2023). NLP techniques for automating responses to customer queries: a systematic review. Discover Artificial Intelligence, 3(1). <https://doi.org/10.1007/s44163-023-00065-5>

BI-LSTM Model Evaluation

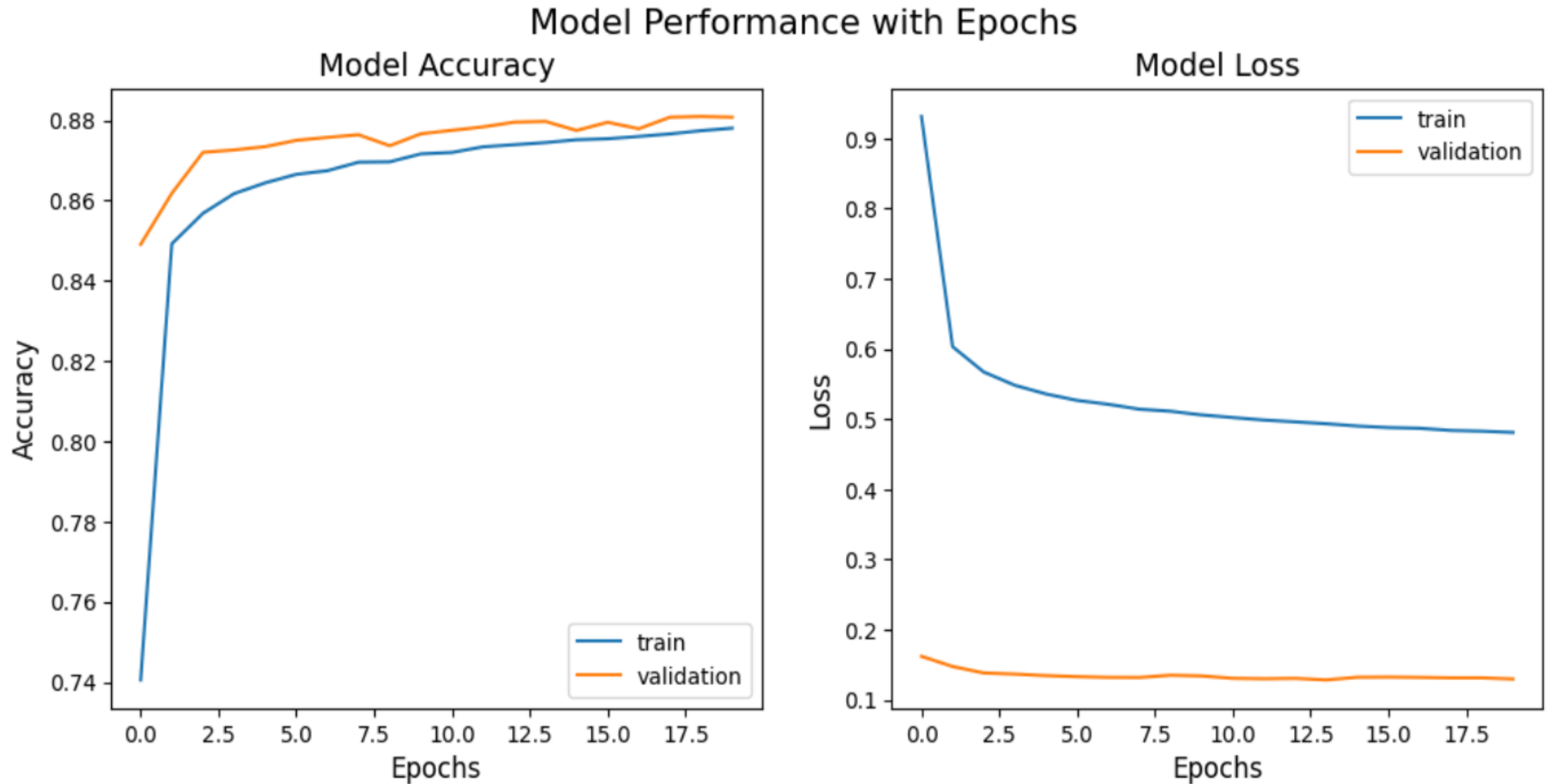
1. Model performance

```
precision: [0.77858002 0.92771216 0.81839791 0.86786704 0.86336996]
recall:    [0.81537492 0.92898209 0.76122043 0.82490785 0.86495413]
fscore:    [0.79655279 0.92834669 0.78877434 0.84584233 0.86416132]
support:   [ 3174 18263  4523  3798  2725]
```

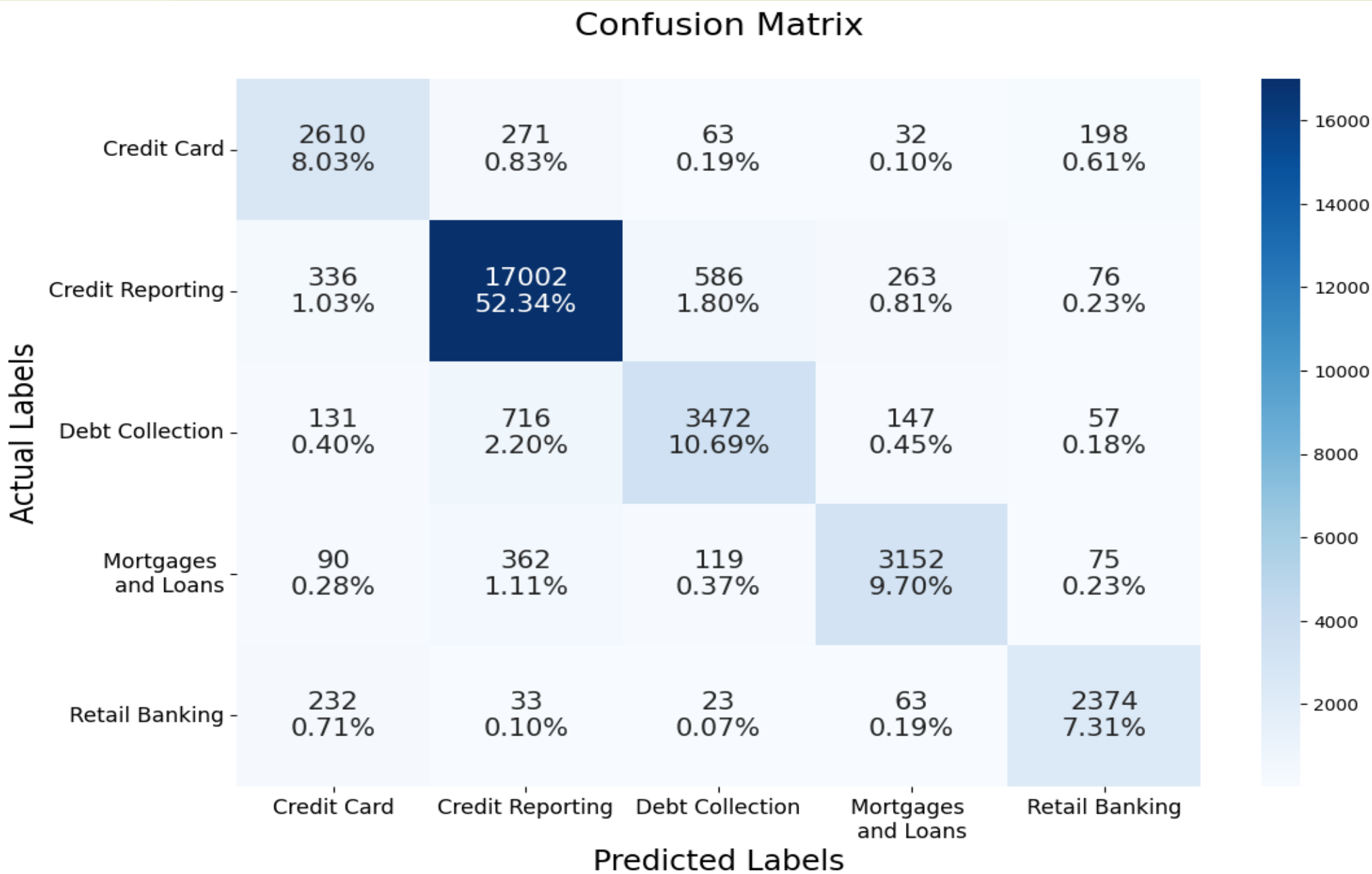
```
#####
```

	precision	recall	f1-score	support
0	0.78	0.82	0.80	3174
1	0.93	0.93	0.93	18263
2	0.82	0.76	0.79	4523
3	0.87	0.82	0.85	3798
4	0.86	0.86	0.86	2725
micro avg	0.89	0.88	0.88	32483
macro avg	0.85	0.84	0.84	32483
weighted avg	0.89	0.88	0.88	32483
samples avg	0.88	0.88	0.88	32483

2. Model Performance with Epochs



3. Confusion Matrix



DistelBERT Model [From Hugging face]

Accuracy: 79.00%

	precision	recall	f1-score	support
0	0.63	0.64	0.63	439
1	0.65	0.66	0.66	269
2	0.87	0.89	0.88	2236
3	0.73	0.73	0.73	422
4	0.72	0.65	0.68	634
accuracy			0.79	4000
macro avg	0.72	0.71	0.72	4000
weighted avg	0.79	0.79	0.79	4000



THANK YOU !