

Anastasija Leskiv



Amazon Reviews

Natural language processing

Project overview

I build a model which would help manufacturers to make their products better by incorporating customer feedback and suggestions.

Models will determine whether customers are satisfied or dissatisfied with the product based on their reviews. I would show my client which are the most helpful reviews that we have to pay attention to.



Business and Data Understanding

For the best result I used :

- Data Collection
- Data cleaning
- Exploratory Data Analysis
- Modeling
- Evaluation
- Code Quality
- Visualization

For my client I checked :

- What are the most helpful reviews
- How reviews can help manufactures
- What should we look at when analyzing reviews
- What products have the most reviews



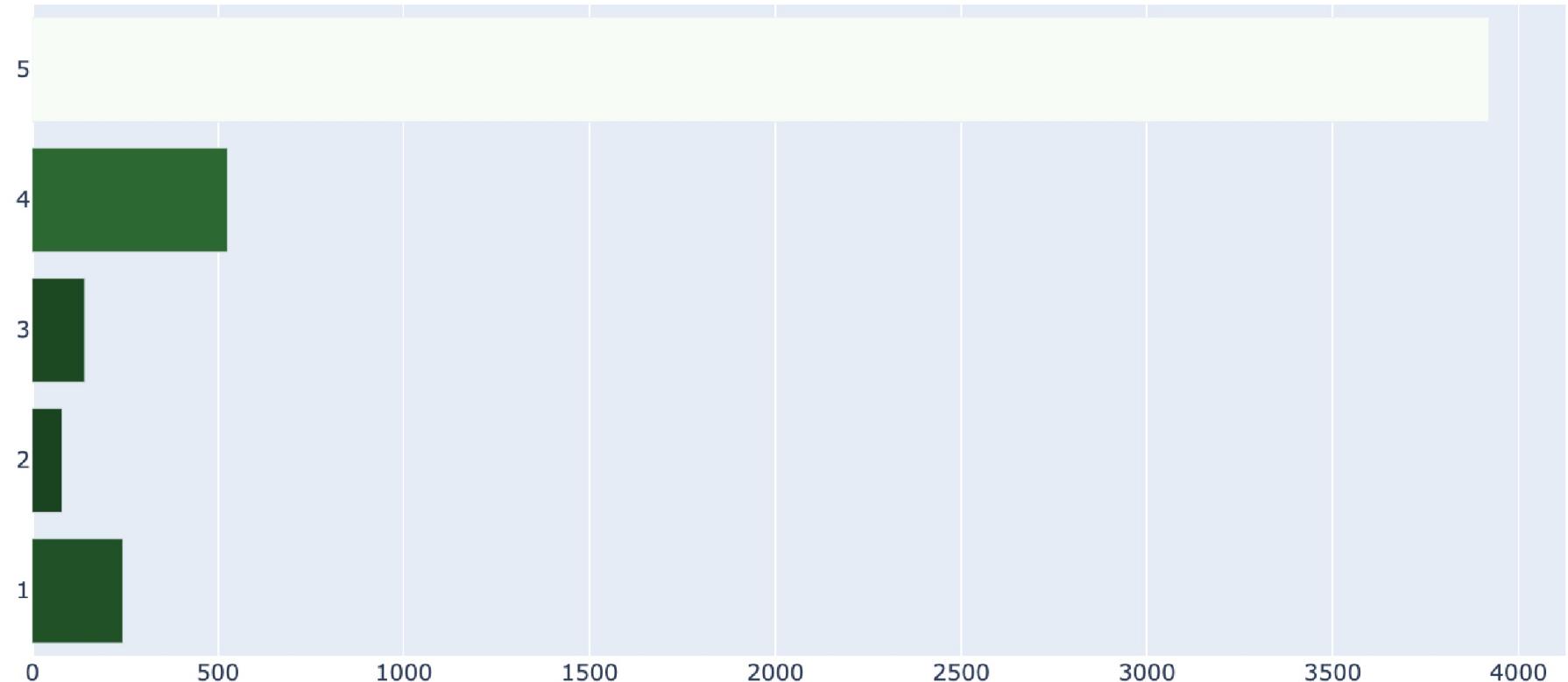


Data

- User Name
- Product Rating
- Evaluation Summary
- Evaluation Time
- Number of days since assessment
- Product Rating

- Number of people who didn't support the comment and didn't find it helpful
- The number of times the evaluation was found useful
- Number of votes given to the evaluation
- Score average rating

Ratings



Most of the reviews are 5 stars - 3921

4gb device cardthings
reliable Length
or reliable worked think
works Packaging used stuff
dtype SD object storage
anything amount name
expected anything class 16GB
sprung want mini
fast space Sandisk adv Good

The most used words in this dataset



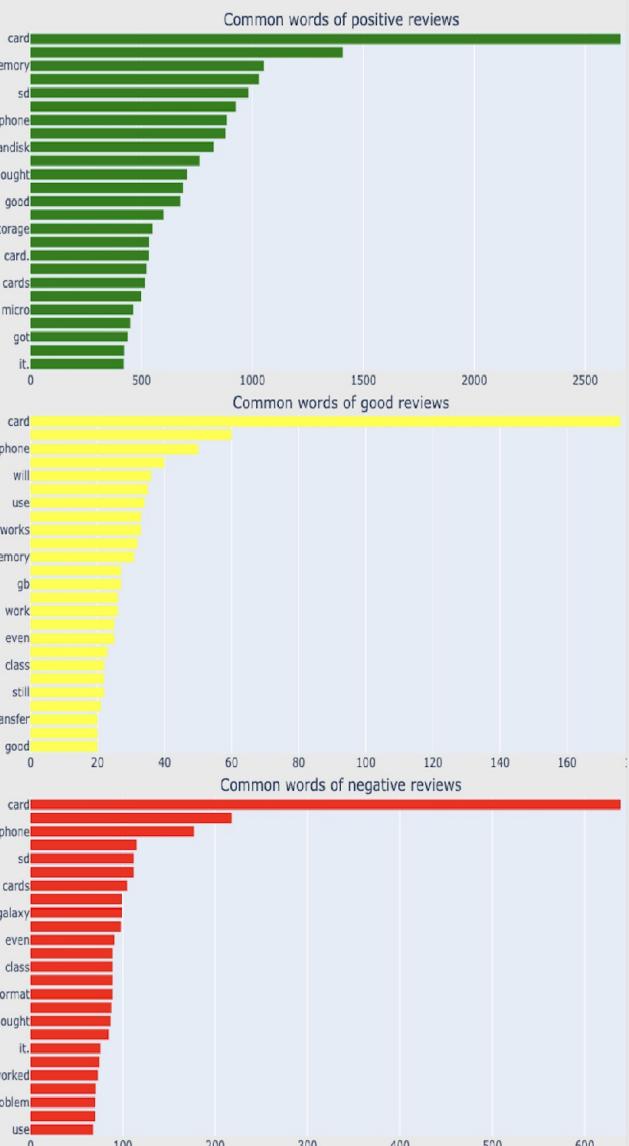
Common words in positive reviews

Common words in negative reviews

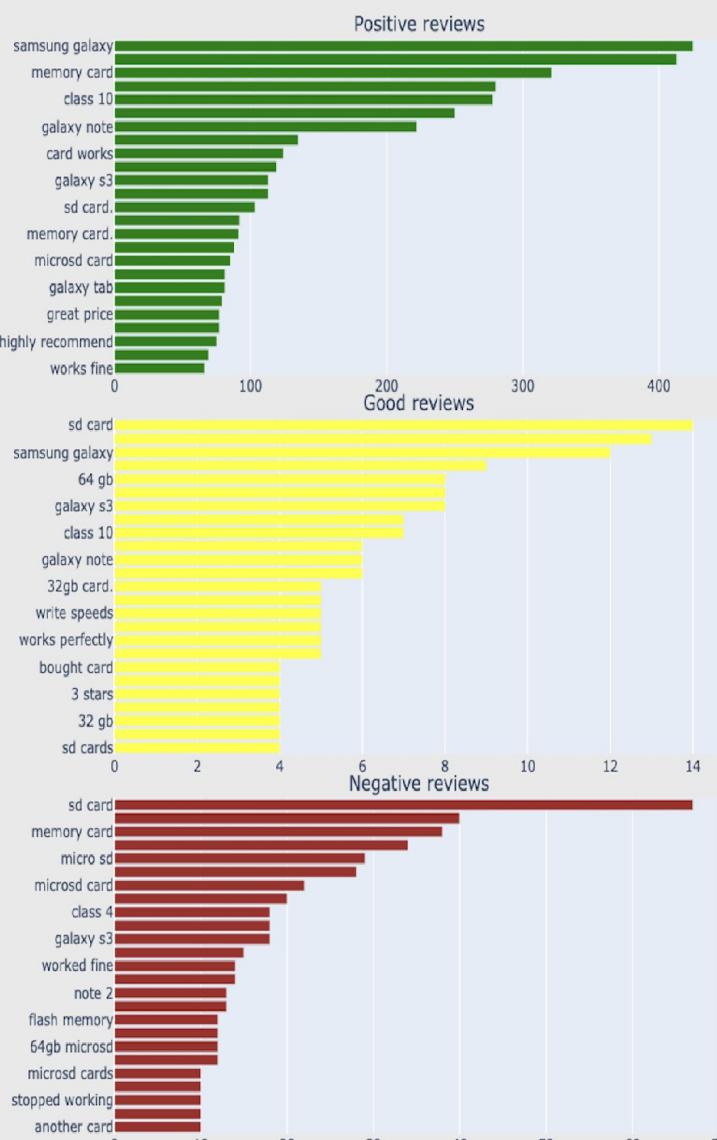


Common words in good (3stars) reviews

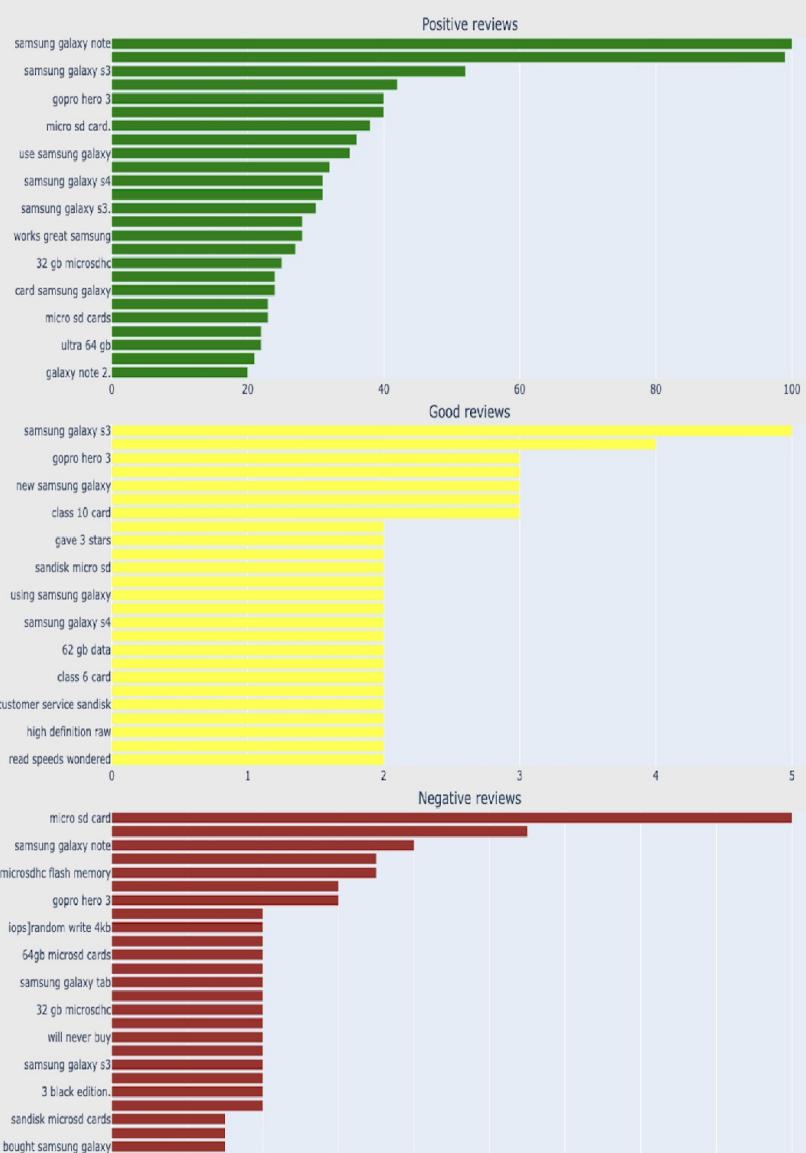
Words Plots

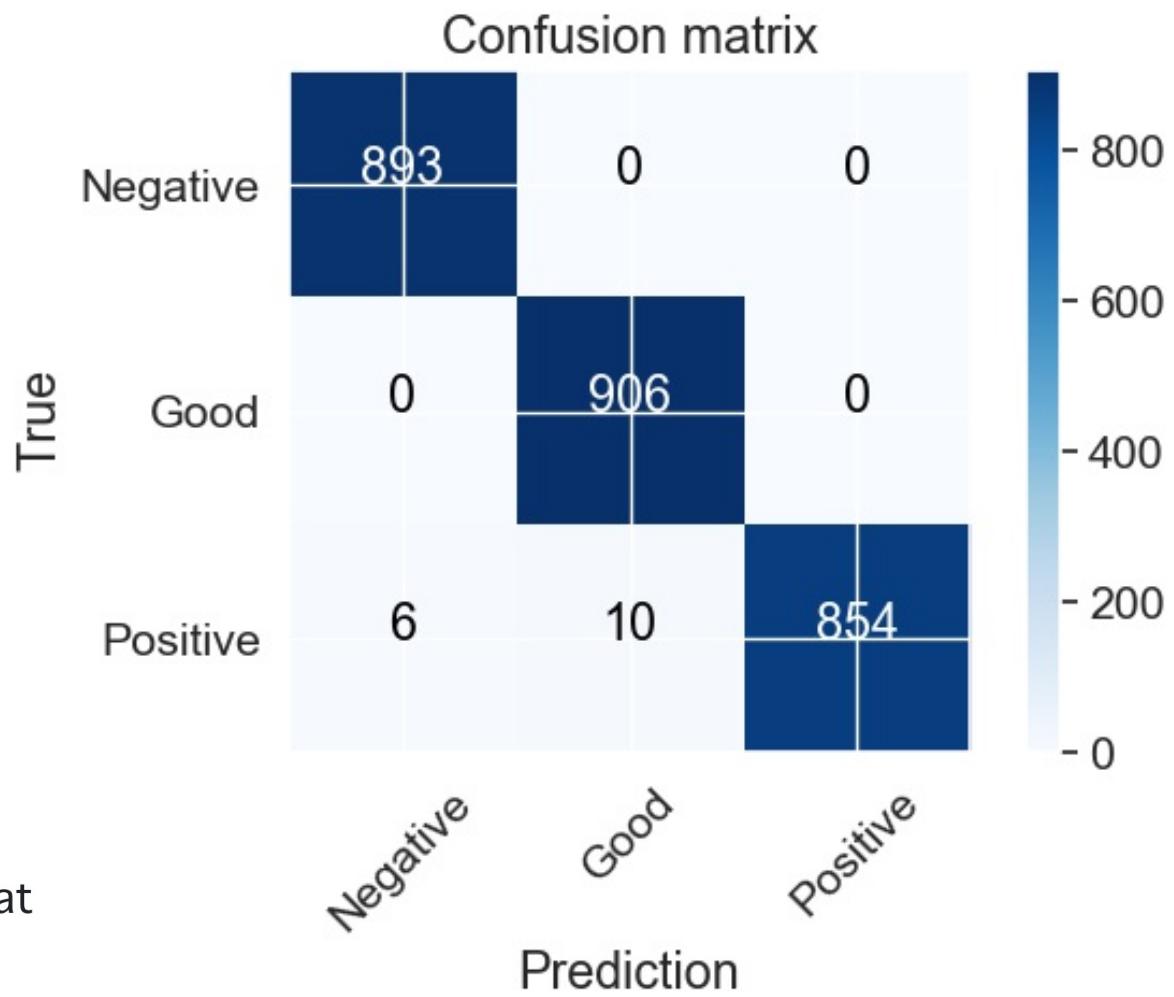


Two words review Plots

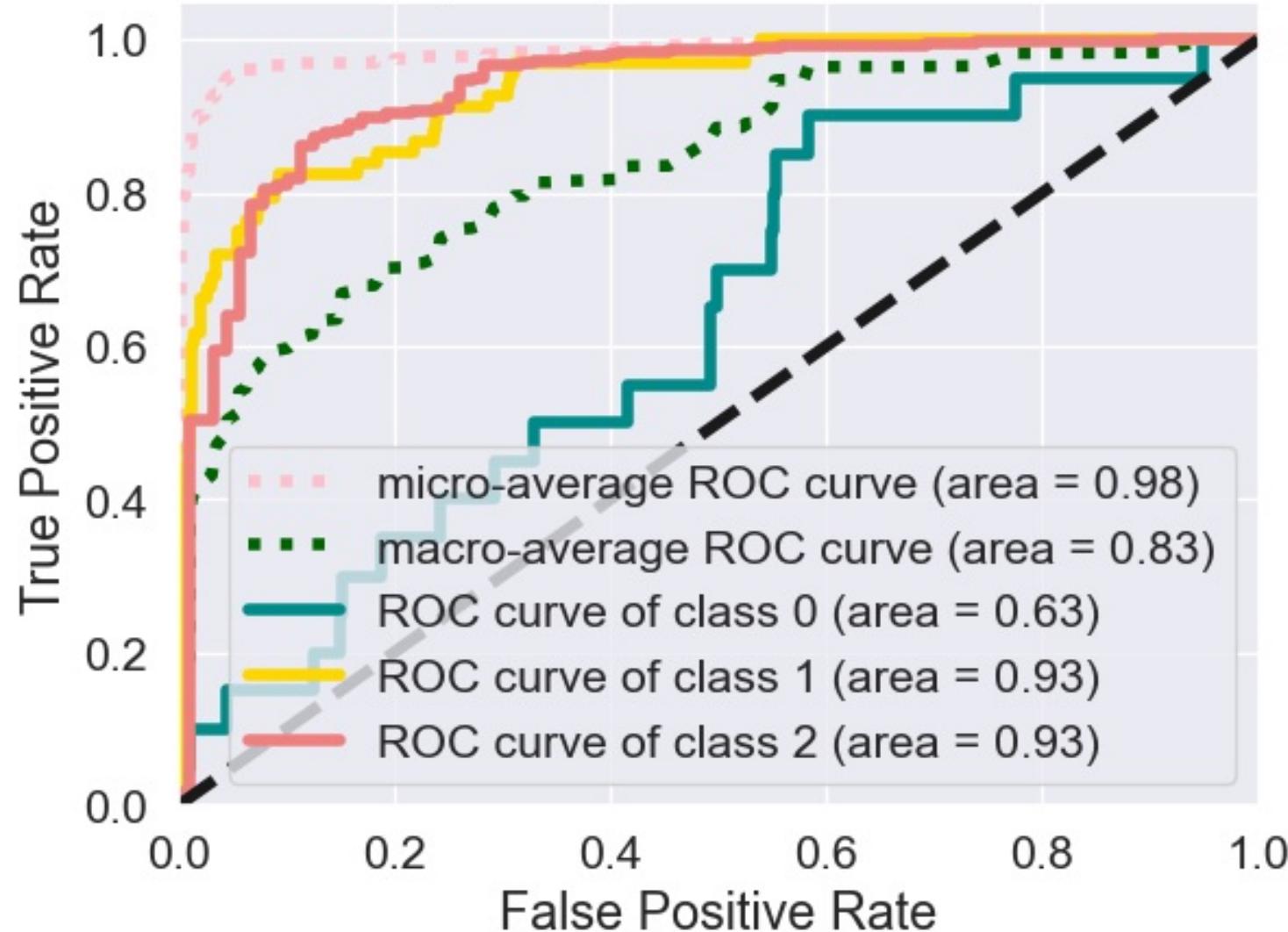


Three words review Plots





Receiver operating characteristic to multi-class



It's great to see that class 2 and class 1 have been classified effectively, as indicated by their high area under the curve (AUC) values. This suggests that the model performs well in distinguishing these classes.

These evaluations and visualizations are fundamental for making informed decisions about the model's performance and tuning its behavior to meet specific objectives

Conclusion

- Recognizing that many good reviews were criticism or feedback from buyers is a valuable insight.
- This feedback can be shared with sellers to help them improve their products.
- Good(3 stars) reviews are the most valuable.
- Most of the reviews were on electronics such as phones, Samsung galaxy, GoPro, cards.



*Sentiment analysis using sentiment word
dictionaries has low reliability when the
number of positive and negative words is
small.*

Recommendations

Pay attention to good reviews(3 stars reviews)

next steps



*Explore 3 stars reviews more to have
specific recommendations to my client.*

Thank You!

Anastasiia Leskiv

[Linkedin](#)

Email: prokopivanastasiia@gmail.com

[GitHub](#)