An Approach to Analyzing Text Customers’ Reviews with a Rating

Andrei Plotnikov, Oleg Betin, Anatoly Truba, Grigorii Voloshin, Evgeniia Dusaeva, Andrei Andrunik
We presented with predicted rating values that users did not specify when they posted on the site. The dataset contains 10441 reviews about Russian Agricultural Bank, generated by users between March 2006 and December 2020. The analyzed data were obtained using raw data and applying XGBoost methods and neural networks to predict those reviews that users did not post. Scientists can use a research and dataset that determines the tone of the review, their subjectivity, reasons for customer satisfaction, and dissatisfaction. Most reviews have negative sentiment because users like to shared negative experience, also because borrowers (as deptors) of bank loans identify with the difficult life situation despite of causes of depts.
The silhouette plot for the various clusters

The visualization of the clustered data

Cluster label

Feature space for the 2nd feature

The silhouette coefficient values

Feature space for the 1st feature
Thank you for your attention!