|  |  |  |
| --- | --- | --- |
|  |  | TECHNICAL TASK  YASA-1 LLC |

16/04/2024

# TASK SETTING

Dear representatives of the X team, I am writing to you on behalf of YASA-1. My name is Vladyslav Viacheslavovych, I am the director of the analytical department of YASA-1. We are a large e-commerce platform for posting and selling a variety of products. Following a recent audit of our South American branch, I was given the responsibility to help my colleagues optimize their operations. You were recommended to me as experts in data analytics and demand forecasting. We face several challenges, namely:

## Task 1. Forecasting

The main goal of this block is to implement a demand forecasting system for a short-term period (14 days) 7 days from the last date in the data for all product groups. We rely on the transaction data provided to you. It is important to make sure the system can produce a forecast for new products (which have little or no training data). We would like to see 2 approaches for comparison: using machine learning and classical time series forecasting.

The technical solution must be justified by comments regarding the appropriateness of a particular method or approach.

The result: a file (.py/.ipynb) with a technical solution for forecasting. This functionality should be able to calculate the result for a future period (scoring).

**Task 2. Analysis of sellers and products**

In this block, I would like to see analytics of sellers and products presented on our marketplace. You have been provided with the following data: *sellers, products* – based on them, you need to build an analytical report, including: the sellers with the biggest/smallest turnover, leaders/outsiders in sales in each area. For products, we are interested in turnover analytics, which products sell best in each category, as well as the dependence of product weight on turnover and price. It is also interesting to see the segmentation of sellers and products with textual conclusions regarding the business value of using these results.

The result: a visual representation of the results (.ipynb file with visualization or BI report) of analytics and the code that generated it.

**Task 3. Analysis of product semantics**

Our database contains information about numerical product ratings with text comments. I hear that modern algorithms can read and understand these reviews. Let’s implement the functionality for classifying positive/negative comments and consolidate the analytics. I’m interested in the correlation of text comments with a numerical rating (1-5), products with the best/worst reviews, sellers who only collect negative feedback.

I would also like to highlight such an entity as price in the comment text, so that the analytical department looks through this information for mentions of the competitors’ prices.

The result: a visual representation of the results (.ipynb file with visualization or BI report) of analytics and the code that generated it. Also, code that can take a text comment as input and classify the semantics as positive, negative or neutral.

**Contact**: ?