Machine Learning Challenge: Cross E-Shop Session Conversion Prediction

Task description

Three independent e-commerce companies have agreed to share anonymized user behavior data to support the development of advanced machine learning models for purchase prediction.

Each company operates its own online store and collects detailed user interaction logs, including:

- Session starts (shop_visit)
- Product views (product_view)
- Add-to-cart actions (add_to_cart)
- Purchases (purchase)
- Other browsing activities (idle, etc.)

All the logs are collected into one file (file *training_data.xlsx*) and, for your convenience, summarized into each user session start and purchase existence summary (file training_results.xlsx).

The companies also share product catalogs containing descriptive information about the products available in their stores (files *products_eshop_1.xlsx, products_eshop_2.xlsx, products_eshop_3.xlsx*).

The objective is to investigate whether combining behavioral knowledge from multiple e-shops can improve purchase prediction performance for a target e-shop.

You can use the data or ignore it, based on your preferences for feature engineering. The goal is to build a machine learning model that predicts whether a user session in e-shop 1 will result in a purchase.

To test, you will have a set of logs from the e-shop (testing_data.xlsx) and a result file (testing_results.xlsx), illustrating the final result – user IP address, session start time, and final purchase status.

For every session, the prediction target is “purchase”:

1 = purchase occurred during the session

0 = no purchase occurred

Evaluation Criteria

The challenge aims to identify not only the most accurate predictive model but also the most practical and trustworthy solution for deployment in a real-world e-commerce environment.

You will therefore be evaluated using a combination of the following criteria:

1. Predictive performance (40%)

We will evaluate the submitted model on its ability to correctly predict whether a user session will result in a purchase. The primary ranking score will be the F1-Score, while Precision and Recall will distinguish closely performing submissions.

2. Prediction time and computational efficiency (20%)

In practical business environments, predictions often need to be generated in near real-time. Therefore, we will also consider the efficiency of the proposed solution. You should provide model inference time (how much time will be needed to output the file with prediction results)

3. Explainability and business interpretability (40%)

A highly accurate model is often difficult to adopt if its decision-making process cannot be understood. You are therefore encouraged to provide a clear explanation of: why the model predicts a purchase or non-purchase; which behavioral signals contribute most to the prediction; which user actions increase or decrease the likelihood of conversion; how information obtained from multiple e-shops contributes to the decision process; etc.

Explainability can be implemented for each prediction, the model, or the situation itself. The most meaningful explainability will be prioritized.

The organizing committee reserves the right to perform qualitative assessment of the explainability component to ensure that explanations are meaningful, consistent, and useful for business decision-makers.

Winner Selection

The objective of this challenge is not only to identify the most accurate purchase prediction model for e-shop 1, but also to encourage innovative use of shared e-commerce data, efficient machine learning solutions, and transparent decision-making mechanisms that can be trusted and adopted in real-world business environments.

Each solution will be tested using new *testing_data.xlsx* and *testing_results.xlsx* files. Those will include new records for e-shop 1, with more recent event logs. Your code must read from testing_data.xlsx and write to testing_results.xlsx; the output will later be compared with the real data to estimate accuracy metrics. Estimate prediction time based on the run timestamps. While the explainability has to be presented during the pitch.