Jakub Zieliński, 299041

Jakub Hebda, 283484

Sprawozdanie IUM

Zadanie 03 wariant 02:

"Wygląda na to, że nasze firmy kurierskie czasami nie radzą sobie z dostawami. Gdybyśmy wiedzieli, ile taka dostawa dla danego zamówienia potrwa – moglibyśmy przekazywać tę informację klientom."

Etap 1:

Definicja problemu biznesowego:  
 Oszacowanie czasu dostawy na podstawie firmy kurierskiej, użytkownika (miejsce zamieszkania), ceny, produktu, kategorii kupowanego produktu oraz daty zakupu.

Zdefiniowanie zadania/zadań modelowania:

Założenia:

Kryteria sukcesu:

Analiza danych:  
 Z pliku sessions.jsonl zostały wybrane te rekordy, które kończą się zakupem. Połączyliśmy je z odpowiadającymi im rekordami z plików purchases.jsonl, products.jsonl i users.jsonl, obliczyliśmy czas dostawy oraz usunęliśmy kolumny, które naszym zdaniem nie noszą dodatkowej informacji. Powstałą tabelę zapisaliśmy do pliku data/processed/sessions.jsonl.

Todos:

* poprawa modelu biznesowego, czyli jak będzie wyglądał produkt, który klient będzie mógł użyć do rozwiązania swojego problemu.
* Czy to zadanie regresji, klasyfikacji , itd., czyli czy szukamy klasy tzn. dzień, 2 lub tydzień, czy po prostu długosći okresu, czyli ciągłej. Jaki jest kwant predykcji?
* Zastanowić się, czy uwzględnić weekendny przy wyliczaniu rzeczywistego czasu dostawy.