

Pojekt IUM – 2021L

Kontekst

W ramach projektu wcielamy się w rolę analityka pracującego w firmie „eSzopping” – sklepu internetowego z elektroniką i gramami komputerowymi. Praca na tym stanowisku nie jest łatwa – zadanie dostajemy w formie enigmatycznego opisu i to do nas należy jego doprecyzowanie tak, aby dało się je zrealizować. To oczywiście wymaga zrozumienia problemu, przeanalizowania danych, czasami negocjacji z szefostwem. Poza tym, oprócz przeanalizowania zagadnienia i przygotowania modeli, musimy przygotować je do wdrożenia produkcyjnego – zakładając, że w przyszłości będą pojawiać się kolejne ich wersje, z którymi my lub nasi następcy będą eksperymentować.

Jak każda szanująca się firma internetowa, eSzopping zbiera dane dotyczące swojej działalności – jest to:

- baza użytkowników,
- katalog produktów,
- historia sesji,
- dane dotyczące wysyłki zakupionych produktów.

Zadania

1. “Niektórzy klienci nie mogą zdecydować się na zakup oglądając produkt na stronie – ale pewnie gdybyśmy wiedzieli, które to są osoby, to odpowiednia zniżka skłoniłaby ich do zakupu.”
2. “Są osoby, które wchodzą na naszą stronę i nie mogą się zdecydować, którym produktom przyjrzeć się nieco lepiej. Może dało by się im coś polecić?”
3. “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.”
4. „Dobrze byłoby wiedzieć czy dana sesja użytkownika zakończy się zakupem. Dzięki temu nasi konsultanci będą mogli baczniej przyglądać się tym sesjom i szybciej rozwiązywać potencjalne problemy.”
5. “Mamy co prawda dodatkowe benefity dla naszych najlepszych klientów, ale może dałoby się ustalić kto potencjalnie jest skłonny wydawać u nas więcej?”

Zasady

- 1) Projekt **realizujemy w parach** – każda para otrzymuje do realizacji jedno z zadań z listy powyżej.
- 2) **Zgłoszenia do zadań proszę wysyłać mailowo podając co najmniej dwa wybrane tematy** – w odpowiedzi otrzymają Państwo identyfikator zbioru danych.

- 3) W ramach projektu trzeba dostarczyć:
- a) etap 1 – możliwe, że będzie konieczne wykonanie więcej niż jednej iteracji tego etapu – pierwsza powinna zakończyć się do **2021.04.18 (0-20 pkt)**:
 - i) definicję problemu biznesowego, zdefiniowanie zadania/zadań modelowania i wszystkich założeń, zaproponowania kryteriów sukcesu),
 - ii) analizę danych z perspektywy realizacji tych zadań (trzeba ocenić, czy dostarczone dane są wystarczające – może czegoś brakuje, może coś trzeba poprawić, domagać się innych danych, ...),
 - b) etap 2 – do **2021.06.06**:
 - i) dwa modele: model bazowy (najprostszy możliwy dla danego zadania) i bardziej zaawansowany model docelowy, oraz raport z pokazujący proces budowy modelu i porównujący wyniki **(0-15 pkt)**,
 - ii) implementację aplikacji (w formie mikroservisu), która pozwala na **(0-15 pkt)**:
 - (1) serwowanie predykcji przy pomocy danego modelu,
 - (2) realizację eksperymentu A/B – w ramach którego porównywane będą oba modele i zbierane dane niezbędne do późniejszej oceny ich jakości,
 - iii) materiały pokazujące, że implementacja działa.
- 4) Rozwiązania należy przesłać mailem:
- a) raporty/dokumentację w formie PDF lub **notatników [jupyter](#) (do których zachęcam)**,
 - b) implementację – w pliku **zip**,
 - c) proszę nie wysyłać formatu **rar** – serwer PW często oznacza takie wiadomości jako spam,
- 5) Zachęcam do korzystania **konsultacji** (MS Teams) w przypadku jakichkolwiek wątpliwości – najpóźniej do **2021.05.28r** – ostatni tydzień przed terminem traktujemy jako „finalny sprint”, podczas którego wszystkie ustalenia są już zamrożone.