ZADANIE REKRUTACYJNE

W ubezpieczeniach, podobnie jak w wielu innych branżach, jednym z celów budowania modeli predykcyjnych jest optymalizacja baz marketingowych. Poniższe zadanie dotyczy tego właśnie problemu- zostało przygotowane w oparciu o publiczne dane dotyczące kampanii marketingowej w banku, z modyfikacjami.

Zadanie polega na przygotowaniu modelu predykcyjnego, który umożliwi wyselekcjonowanie klientów z najwyższym prawdopodobieństwem kupna produktu w ramach prowadzonej kampanii marketingowej.

Dane znajdują się  w załączeniu:

* current\_campaign.csv – w tym pliku znajdują się dane dotyczące bieżącej kampanii, w oparciu o które powinien zostać zbudowany model,
* previous\_campaigns.csv – to dane z poprzednich kampanii, które mogą posłużyć do wzbogacenia zbioru treningowego,
* data\_description.txt – to opis danych.

Oczekiwany wynik zadania:

1. Zakładając, że dane w plikach current\_campaigns.csv i previous\_campaigns.csv to dane w dwóch tabelach na bazie danych (o nazwach data.current\_campaign i data.previous\_campaign odpowiednio), proszę przygotować kod sql, który posłuży do dołączenia do danych z bieżącą kampanią informacji dotyczących danego klienta z poprzednich kampanii:

* liczba kontaktów w ramach poprzednich kampanii,
* wynik ostatniego kontaktu w poprzednich kampaniach,
* odstęp pomiędzy datą pierwszego kontaktu w bieżącej kampanii a datą ostatniego kontaktu w poprzednich kampaniach .

Samo łączenie danych na cele późniejszego modelowania może być wykonane w Pythonie, natomiast to podzadanie dotyczy umiejętności pisania kodu sql. Jeśli zna Pan/Pani PostgreSQL, proszę napisać kod w tej składni. Jeśli nie- może być to dowolny inny wariant. Kod proszę przesłać w pliku .sql lub .txt.

1. Kod w Pythonie (plik .py lub notebook .ipynb) wykorzystany do zbudowania modelu.
2. Podsumowanie procesu budowy modelu i wyników w formie kilku slajdów (.pptx) lub notebooka (html), z założeniem, że materiał ten miałby być omawiany z biznesem.

Powodzenia!