

Temat

"Nasz serwis zarabia wtedy, gdy ludzie słuchają muzyki. Jeśli nie wiemy ile czasu będą słuchali jej w przyszłości, to trudno nam rozliczać się z artystami i negocjować z nimi stawki"

Kontekst

Serwis streamingowy muzyki "Pozytywka"

Zadanie biznesowe

Przewidywanie czasu odsłuchu każdego artysty przez użytkowników w następnym tygodniu.

Biznesowe kryterium sukcesu

Wyższa dokładność niż statyczne podejście - średnia z poprzednich czterech tygodni.

Zadanie modelowanie

Przygotowanie modelu analizującego szeregi czasowe:

- przewidującego czas odsłuchu w następnym tygodniu na podstawie danych historycznych

Dane do modelowania

- sesje użytkowników
- lista użytkowników
- lista artystów
- lista utworów

Analityczne kryterium sukcesu

Dokładność na **zbiorze testowym** tworzonego modelu większa niż dokładność na **zbiorze testowym** modelu naiwnego

Model naiwny

Model naiwny: *średnia z poprzednich czterech tygodni*

Miara dokładności

Miarą dokładności jest RMSE (Root Mean Square Error) - jest to funkcja obliczająca różnicę między predykcją a prawdziwym wynikiem, błąd jest tego samego wielkość rzędu co różnica wyników. Czym mniejszy błąd tym wyższa dokładność. Jednostką miary dokładności będzie jednostka czasowa (np. sekundy, godziny).