

Projekt z eksploracji danych - wybór zbioru danych

Cezary Żak & Bartłomiej Guz

Zbiór danych:

Dane zawierają informacje o samochodach osobowych wystawionych na sprzedaż za pośrednictwem platformy otomoto.pl z okolic Lublina (w promieniu 50 km).

Baza zawiera 1248 obserwacji oraz 10 parametrów.

Parametry:

Sprzedający opisują swoją ofertę przy pomocy wielu zmiennych, jednak wiele z nich jest opcjonalna. Z tego powodu w bazie znajdują się jedynie najważniejsze z nich:

- marka pojazdu
- model pojazdu
- rok produkcji
- przebieg
- pojemność skokowa silnika
- rodzaj paliwa
- moc
- typ nadwozia
- kolor
- cena

Cel projektu

Głównym celem projektu będzie stworzenie modelu, który szacować będzie cenę pojazdu na podstawie dostępnych parametrów. Taki model może być użyteczny na kilka sposobów:

- dla osób które nie mają pojęcia o rynku samochodowym, a chcą sprzedać auto
- dla osób które chcą kupić pojazd a zarazem nie chcą dać się oszukać
- dla osób które szukają pojazdów w celach inwestycyjnych (auta sprzedawane po zaniżonych cenach)