

kmeans-spark

This project shows example of training K-means model using spark.

Dataset

OpenFoodFacts dataset consists of the descriptions of different food products.
More info could be found [here](#)

Data preparation

Data was preprocessed with removing of unimportant features and null columns filling.

Project structure

1. Research notebook
2. Preprocessor
3. Model trainer