

Лабораторная работа #6

Источник данных на PySpark

В рамках данной работы было разработано взаимодействие между моделью и источником данных.

Реализация флуктуирующей работы модели и источника данных (main.py).

В качестве источника данных используется локальный Hadoop File System

В результате обучения модели была получена следующая метрика : Silhouette with squared euclidean distance =
0.9999993029872573