Извештај за задатак 1 - једнострука линеарна регресија

Даница Газдић

SV 12/2020Милош Обрадовић

SV 55/2020

## Проблем

У овом извештају се решава проблем једноструке линеарне регресије применом алгоритма *normal equation* у комбинацији са нормализацијом *z-score*.

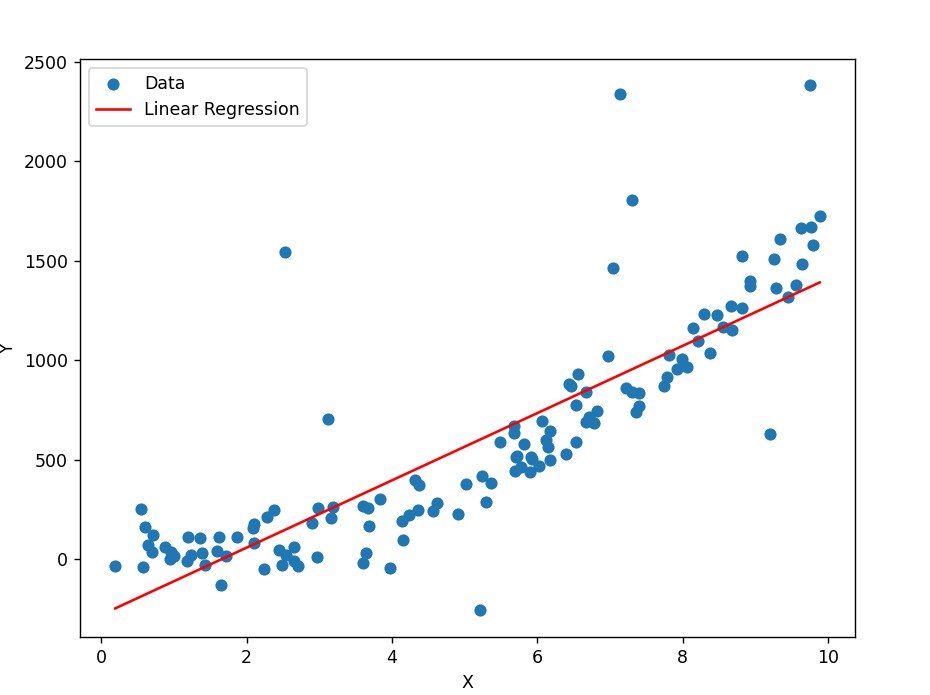
## Испробани алгоритми

С обзиром да је скуп података имао дужину од само 160, а *normal equation* најбоље ради над малим скуповима података, *stochastic* и *batch gradient descent* нису узети у обзир.

Пре примене алгоритма скуп података је поде љен на тренинг и тестни скуп у размери 80-20%. Затим су над тренинг скупом података израчунати коефицијенти за *z-score* нормализацију, помоћу којих је нормализован и тестни скуп. Нормализацијом је осигурано брже конвергирање алгоритма. Након тога је примењен алгоритам *normal equation* са циљем да се нађу оптимални коефицијенти праве. *Outllier-*и су остављени у скупу података у оригиналном облику.

## Остварени резултати

Метрика која је коришћена је *root mean square error - RMSE* и приступ проблему описан у овом извештају је дао задовољавајући резултат од 175,618 *RMSE*, као и задовољавајуће фитовану праву [Слика 1]. С обзиром да је овај метод дао задовољавајуће решење, друге имплементације нису испитиване.



Слика Резултат једноструке линеарне регресије применом normal equation алгоритма и z-score нормализације