Отчет №5

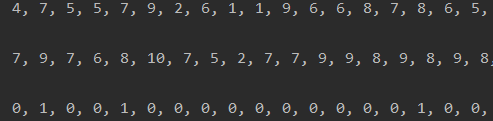
Отчет подготовил студент

2 курса группы БПИ183

Плешков Даниил

Для работы над данной задачей будет использоваться java для считывания данных и Python для вычислений.

С помощью кода, проекта Mark, считаем и представим данные в удобном виде (каждый признак с новой строки, через запятую все данные внутри одного признака).



Далее с помощью библиотек sklearn и skmetrics вычислим необходимые коэффициенты.

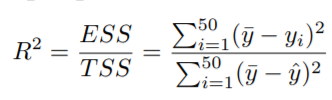
Коэффициенты для длинной регрессии:

[0.46855586, 0.10991112], константа - 3.8208926694064798

Коэффициент для короткой регрессии:

[0.47101449], константа - 3.8260869565217397

По данным можно сказать что между оценкой за курс Теории вероятностей и оценкой за контрольную работу есть некая значимая корреляция, пол же незначительно влияет на оценку за контрольную.

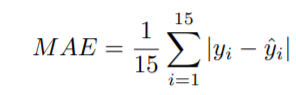
Теперь рассчитаем R2  для обеих моделей с помощью функции score типа LinearRegression

Получены следующие данные

|  |  |
| --- | --- |
| Длинная | Короткая |
| 0.17739930524874414 | 0.17714884696016764 |

В целом можно сказать, что точность модели очень мала для того, чтобы опираться на нее при расчете оценок за контрольную работу. Для увеличения точности можно добавить несколько признаков.

|  |  |
| --- | --- |
| Длинная | Короткая |
| 2.149664 | 2.151691 |

Пришло время рассчитать MAE для последних 16 наблюдений. Рассчитана она будет с использованием функции predict типа LinearRegressison. Формула вычислений следующая: 

Получены следующие данные:

Поскольку у модели длинной регрессии меньшая Mean Absolute Error, предскажу собственную оценку с помощью этой модели. Моя оценка за курс по теории вероятности: 5, пол - мужской.

Рассчитаем оценку по следующей формуле :

Mark = 3.8208926694064798 + 0.46855586\*5 + 0.10991112\*0 = 6,16.

С учетом округления : 6.

Данную оценку я считаю приемлемой.