**Отчёт по лабораторной работе №3**

**Никита Чучелов, Карев Артём, Давыдов Дмитрий, Зубарева Алина, Бусыгина Екатерина**

1. Сформировали Тестовые, Тренировочные и контрольные выборки на основе РЕАЛЬНОГО Dataset, полученного на экспериментальных данных

df = pd.read\_csv('input.txt', header=None, delimiter = ",")

array = df.values

X = df.iloc[:, 2:11].values

y = df.iloc[:,11].values

X\_train,X\_test,y\_train,y\_test=train\_test\_split(X,y,train\_size=0.2,random\_state=10)

1. Сформировали Тестовые, Тренировочные и контрольные выборки на основе синтезированного Dataset, полученного на экспериментальных данных

df2 = pd.read\_csv('df1.txt', header=None, delimiter = " ")

X2 = df2.iloc[:, 0:9].values

y2 = df2.iloc[:, 9].values

X2\_train,X2\_test,y2\_train,y2\_test=train\_test\_split(X2,y2,train\_size=0.2,random\_state=10)

1. Реализовать на Python любые три на выбор метода для решения задачи. Обучить машину обучения по всем методам. Метриками измерить функционал качества.

*Метод случайный лес*

model = RandomForestRegressor()

model.fit(X\_train, y\_train)

RandomForestRegressor()

xs1=model.score(X\_test, y\_test)

print(xs1)

xs2=model.score(X2\_test, y2\_test)

print(xs2)

*Метод ближайших соседей*

model2= KNeighborsRegressor()

model2.fit(X\_train, y\_train)

KNeighborsRegressor()

xs4=model2.score(X2\_test, y2\_test)

xs3=model2.score(X\_test, y\_test)

print(xs3)

print(xs4)

*Метод экстра деревья*

model3 = ExtraTreesRegressor()

model3.fit(X\_train, y\_train)

ExtraTreesRegressor()

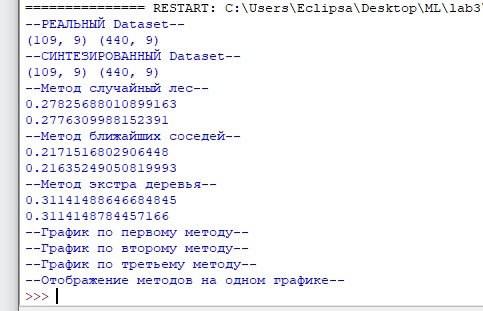
xs5=model3.score(X\_test, y\_test)

print(xs5)

xs6=model3.score(X2\_test, y2\_test)

print(xs6)

1. Метод экстра деревьев лучше других методов, потому что разница между ошибками минимальна.



1. Сравнение результатов:

|  |  |  |
| --- | --- | --- |
|  | Реальный Dataset | Cинтезированный Dataset |
| Случайный лес | 0.278256880 | 0.227630998 |
| Ближайших соседей | 0.217151680 | 0.216352490 |
| Экстра деревья | 0.311414886 | 0.311414878 |