

Листок 04. Метод k-ближайших соседей (k-Nearest Neighbors, k-NN)

Упражнение 1. Для набора данных `sleep75` рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
sleep	totwrk, age, south, male

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	totwrk	age	south	male
0	2160	32	1	0
1	1720	24	0	1
2	2390	44	0	1

вычислите прогноз **sleep** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 2. Для набора данных `sleep75` рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
sleep	totwrk, age, south, male, smsa, yngkid, marr

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	totwrk	age	south	male	smsa	yngkid	marr
0	2150	37	0	1	1	0	1
1	1950	28	1	1	0	1	0
2	2240	26	0	0	1	0	0

вычислите прогноз **sleep** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 3. Для набора данных **wage2** рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
wage	age, IQ, south, married, urban

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	age	IQ	south	married	urban
0	36	105	1	1	1
1	29	123	0	1	0
2	25	112	1	0	1

вычислите прогноз **wage** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 4. Для набора данных **wage2** рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
log(wage)	age, IQ, south, married, urban

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	age	IQ	south	married	urban
0	36	105	1	1	1
1	29	123	0	1	0
2	25	112	1	0	1

вычислите прогноз **wage** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 5. Для набора данных **wage1** рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
wage	exper, female, married, smsa

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	exper	female	married	smsa
0	5	1	1	1
1	26	0	0	1
2	38	1	1	0

вычислите прогноз **wage** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 6. Для набора данных **wage1** рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
log(wage)	exper, female, married, smsa

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	exper	female	married	smsa
0	5	1	1	1
1	26	0	0	1
2	38	1	1	0

вычислите прогноз **wage** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 7. Для набора данных Labour рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
output	capital, labour

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	capital	labour
0	2.970	85
1	10.450	60
2	3.850	105

вычислите прогноз **output** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 8. Для набора данных Labour рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
log(output)	log(capital), log(labour)

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	capital	labour
0	2.970	85
1	10.450	60
2	3.850	105

вычислите прогноз **output** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 9. Для набора данных **Labour** рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
output	capital, labour, wage

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	capital	labour	wage
0	2.970	85	36.98
1	10.450	60	33.82
2	3.850	105	40.23

вычислите прогноз **output** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE

Упражнение 10. Для набора данных **Labour** рассмотрим задачу прогнозирования для переменных

зависимая/target	объясняющая/предикторы/features
log(output)	log(capital), log(labour), log(wage)

1. Подгоните на исходном датасете модель kNN-регрессии с параметрами

№	k	веса
1	5	uniform
2	5	distance
3	10	uniform
4	10	distance

2. Рассмотрим трёх людей с характеристиками

index	capital	labour	wage
0	2.970	85	36.98
1	10.450	60	33.82
2	3.850	105	40.23

вычислите прогноз **output** по каждой модели

3. На обучающей выборке вычислите метрики подгонки: R^2 , MSE, MAE, MAPE, RMSE