

Проект 2: Прогнозирование цен на жилье

Описание задачи: Модель для предсказания цен на жилье на основе характеристик (площадь, комнаты и т.д.).

Стек технологий: Python 3.x, pandas, numpy, scikit-learn, matplotlib, seaborn (для визуализации).

Начальный каркас кода:

Импорт библиотек

```
import pandas as pd Данные
```

```
import numpy as np Массивы
```

```
from sklearn.model_selection import train_test_split Разделение
```

```
from sklearn.linear_model import LinearRegression Регрессия
```

```
from sklearn.metrics import mean_squared_error Оценка
```

```
import matplotlib.pyplot as plt Визуализация
```

Шаг 1: Загрузка данных (используйте Boston Housing с `sklearn.datasets.load_boston()`)

```
from sklearn.datasets import fetch_california_housing
```

```
data = fetch_california_housing(as_frame=True)
```

```
df = data.frame
```

Шаг 2: Предобработка

```
X = df.drop('MedHouseVal', axis=1) Признаки
```

```
y = df['MedHouseVal'] Цена
```

Шаг 3: Разделение

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Шаг 4: Обучение

```
model = LinearRegression()  
model.fit(X_train, y_train)
```

Шаг 5: Предсказание и оценка

```
predictions = model.predict(X_test)  
print("MSE:", mean_squared_error(y_test, predictions))
```

Визуализация

```
plt.scatter(y_test, predictions)  
plt.xlabel('True Values')  
plt.ylabel('Predictions')  
plt.show()
```

TODO: Добавьте нормализацию (StandardScaler), попробуйте
RandomForestRegressor

Направления для дальнейшего развития: Добавить полиномиальную регрессию,
визуализацию с seaborn, веб-интерфейс с Flask.