

Инструкция по линейной регрессии

Для того, чтобы обучить и использовать линейную регрессию необходимо выполнить следующие шаги:

1. Создать виртуальное пространство, написав в терминале команду:

```
python -m venv venv
```

2. Установить датасет домов, выполнив в терминале команду:

```
pip install scikit-learn==1.1.3
```

3. Создать файл main.py и написать в нём следующий код:

```
from sklearn.datasets import load_boston #импорт датасета в
программу
boston = load_boston() #объявление переменной с данными
# print(boston.DESCR)

















from sklearn.model_selection import train_test_split #импорт
функции для разделения датасета на части
X_train, X_test, y_train, y_test = train_test_split(boston.data,
boston.target, test_size=0.2, random_state=42) #разделение
датасета на тестовую и обучающую части
#в соотношении 80% : 20%

from sklearn.linear_model import LinearRegression #импорт
алгоритма линейной регрессии
model = LinearRegression() #создание линейной регрессии
model.fit(X_train, y_train) #обучение линейной регрессии

house1 = [[0.00632, 18.00, 2.310, 0, 0.5380, 6.5750, 65.20
, 4.0900, 1, 296.0, 15.30, 396.90, 4.98]] #определяем
параметры дома для предсказания цены
cost_house1 = model.predict(house1) #предсказываем цену дома
```

```
print(cost_house1) #печатаем цену дома
```

4. Запустить выполнение программы
5. Загрузить файл main.py в МЭО -> 3 модуль -> Тема 3.2
Линейная регрессия -> Практическая работа. Линейная регрессия

Модуль 3. Машинное обучение	
В модуле обучающиеся знакомятся с основными понятиями машинного обучения, учатся работать с библиотекой Scikit-learn, создают и обучают простейшие модели машинного обучения для задач классификации и регрессии	
Тема 3.1. Основные понятия и задачи в машинном обучении	
 Лекция. Основные понятия и задачи в машинном обучении	<input checked="" type="checkbox"/>
 Видеоинструкция для практического задания	<input checked="" type="checkbox"/>
 Практическая работа. Основные понятия и задачи в машинном обучении	<input checked="" type="checkbox"/>
Тема 3.2. Линейная регрессия	
 Лекция. Линейная регрессия	<input checked="" type="checkbox"/>
 Видеоинструкция для практического задания	<input checked="" type="checkbox"/>
 Практическая работа. Линейная регрессия	<input checked="" type="checkbox"/>
 Самостоятельная работа. Линейная регрессия	<input checked="" type="checkbox"/>
Тема 3.3. Линейная классификация	
 Презентация. Линейная классификация	<input checked="" type="checkbox"/>
 Лекция. Линейная классификация	<input checked="" type="checkbox"/>
 Видеоинструкция для практического задания. Часть 1	<input checked="" type="checkbox"/>
 Видеоинструкция для практического задания. Часть 2	<input checked="" type="checkbox"/>
 Практическая работа. Линейная классификация	<input checked="" type="checkbox"/>
 Самостоятельная работа. Линейная классификация	<input checked="" type="checkbox"/>
Тема 3.4. Градиентный спуск	
 Лекция. Градиентный спуск	<input checked="" type="checkbox"/>
 Практическая работа. Градиентный спуск	<input checked="" type="checkbox"/>
Тема 3.5. Аттестация по итогам модуля	
 Итоговая работа по модулю 3	<input checked="" type="checkbox"/>

В датасете содержатся данные о домах в Бостоне, США. На их основе формируется цена жил. площади. Задача линейной регрессии - попытаться предсказать цену дома, на основе данных о доме:

CRIM - Уровень преступности на душу населения по городам.

ZN - Доля земель под жилую застройку, зонированных под участки площадью более 25 000 кв. футов.

INDUS - Доля акров неторгового бизнеса на город.

CHAS - Charles River фиктивная переменная(= 1, если участок граничит с рекой; 0 в противном случае).

NOX - Концентрация оксидов азота (частей на 10 миллионов).

RM - Среднее количество комнат на жилище.

AGE - Пропорция квартир, занимаемых владельцами, построенных до 1940 г.

DIS - Взвешенные расстояния DIS до пяти центров занятости Бостона.

RAD - Индекс доступности радиальных автомагистралей.

TAX - Полная ставка налога на имущество за 10 000 долларов США.

PTRATIO - Соотношение учеников и учителей по городам.

B - $1000 (B_k - 0,63)^2$, где B_k - доля чернокожих по городам.

LSTAT - % более низкий статус населения.