

Оптимальная регрессионная модель

Филатович Дмитрий

28 декабря 2021 г.

1 Описание задачи

Написать программу, находящая оптимальную регрессионную модель из линейной, полиномиальной до 5 степени и экспоненциальной.

2 Ход решения задачи

1. Написать линейную модель
2. Написать полиномиальную модель
3. Написать экспоненциальную модель
4. Визуализация
5. Подсчет оптимальности моделей
6. Определение оптимальной модели

3 Решение задачи

3.1 Написание линейной модели

Для написания линейной модели я использовал `LinearRegression` из библиотеки `sklearn.linear_model`. Далее я обучаю модель с помощью `fit`.

3.2 Написание полиномиальной модели

Для полиномиальной модели я написал функцию:

```
def func(x, y, degree):  
    poly = PolynomialFeatures(degree=degree)  
    X_poly = poly.fit_transform(x)  
    poly.fit(X_poly, y)  
    model = LinearRegression()  
    model.fit(X_poly, y)  
    return model.predict(poly.fit_transform(x))
```

Рис. 1: Код полиномиальной модели

Здесь я использовал `PolynomialFeatures` из библиотеки `sklearn.preprocessing`. Её мы преобразуем аргументы, и применяем `LinearRegression`.

3.3 Написание экспоненциальной модели

Для экспоненциальной я преобразую аргументы по формуле e^x . После этого опять применяем `LinearRegression`.

3.4 Визуализация

```
plt.figure(figsize=(15,10))
plt.subplot(2,3,1)
plt.scatter(x,y,c='blue')
plt.plot(x.to_numpy(),model.predict(x),c='red')
plt.subplot(2,3,2)
plt.scatter(x,y,c='blue')
plt.plot(x.to_numpy(),func(x,y,2),c='green')
plt.subplot(2,3,3)
plt.scatter(x,y,c='blue')
plt.plot(x.to_numpy(),func(x,y,3),c='pink')
plt.subplot(2,3,4)
plt.scatter(x,y,c='blue')
plt.plot(x.to_numpy(),func(x,y,4),c='brown')
plt.subplot(2,3,5)
plt.scatter(x,y,c='blue')
plt.plot(x.to_numpy(),func(x,y,5),c='black')
plt.subplot(2,3,6)
plt.scatter(x,y,c='blue')
plt.plot(x.to_numpy(),model2.predict(xexp),c='yellow')
```

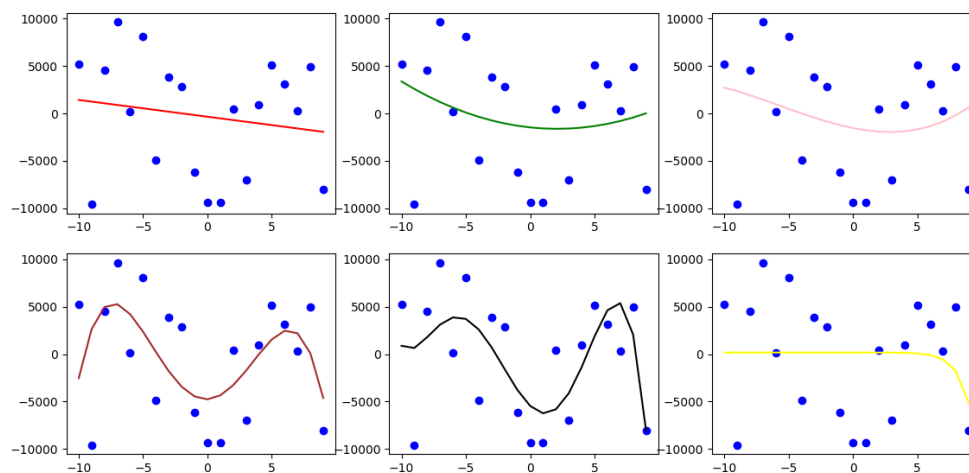


Рис. 2: Визуализация

3.5 Определение оптимальности моделей

Оптимальность моделей определялась с помощью метрики r^2 score

3.6 Определение оптимальной модели

Оптимальная модель определяется сравнением метрик `r2score` каждой модели

4 Вывод

Программа, определяющая оптимальную регрессионную модель написана. Полный код находится в папке с отчетом