تمرین 1: پیش‌بینی قیمت خانه‌ها

توضیحات:

- داده‌ها: استفاده از داده‌های [California Housing Dataset](https://scikit-learn.org/stable/modules/generated/sklearn.datasets.fetch\_california\_housing.html) (یا هر دیتاست مشابه).

- هدف: پیش‌بینی میانگین ارزش خانه‌ها (`MedHouseValue`) بر اساس ویژگی‌های مختلف مانند درآمد متوسط (`MedInc`), سن خانه‌ها (`HouseAge`), تعداد اتاق‌ها (`AveRooms`), و غیره.

- مراحل:

1. داده‌ها را بارگذاری کنید و به صورت آموزش و آزمون تقسیم کنید.

2. یک مدل رگرسیون خطی ساده ایجاد کنید.

3. ضرایب مدل را تفسیر کنید (هر ویژگی چقدر تأثیرگذار است؟).

4. معیارهای ارزیابی مانند `MSE` (Mean Squared Error) و `R²` را محاسبه کنید.

---

تمرین 2: رگرسیون خطی برای پیش‌بینی فروش\*\*

توضیحات:

- داده‌ها: یک دیتاست مصنوعی یا واقعی شامل اطلاعات فروش محصولات (مانند قیمت، تبلیغات، و تخفیف‌ها).

- هدف: پیش‌بینی میزان فروش بر اساس عوامل مختلف مانند قیمت محصول، هزینه تبلیغات، و تخفیف‌های ارائه شده.

- مراحل:

1. داده‌های مصنوعی ایجاد کنید.

import numpy as np

np.random.seed(42)

X = np.random.rand(100, 3) \* 100 # 3 ویژگی: قیمت، تبلیغات، تخفیف

y = 50 + 2 \* X[:, 0] - 1.5 \* X[:, 1] + 3 \* X[:, 2] + np.random.randn(100) \* 10

2. یک مدل رگرسیون خطی بسازید و آن را آموزش دهید.

3. ضرایب مدل را تفسیر کنید و تأثیر هر ویژگی را بررسی کنید.

4. نمودار واقعی و پیش‌بینی‌شده را رسم کنید.

---

تمرین 3: مقایسه رگرسیون خطی با روش‌های دیگر

توضیحات:

- داده‌ها: استفاده از داده‌های [Ames Housing Dataset](https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data).

- هدف: مقایسه عملکرد رگرسیون خطی با مدل‌های دیگر مانند رگرسیون لاسو (Lasso Regression) و رگرسیون ریج (Ridge Regression).

- مراحل:

1. داده‌ها را پیش‌پردازش کنید (مدیریت مقادیر گمشده، استانداردسازی ویژگی‌ها، و غیره).

2. یک مدل رگرسیون خطی ساخته و آن را آموزش دهید.

3. مدل‌های Lasso و Ridge را نیز آموزش دهید و ضرایب آن‌ها را با رگرسیون خطی مقایسه کنید.

4. معیارهای ارزیابی (`MSE`, `MAE`, `R²`) را برای هر سه مدل مقایسه کنید.

---

تمرین 4: رگرسیون خطی برای پیش‌بینی درآمد

توضیحات:

- داده‌ها: یک دیتاست شامل اطلاعات افراد مانند سن، تحصیلات، تجربه کاری، و درآمد.

- هدف: پیش‌بینی درآمد افراد بر اساس ویژگی‌های داده شده.

- مراحل:

1. داده‌های مصنوعی ایجاد کنید:

import numpy as np

np.random.seed(42)

age = np.random.randint(20, 60, 100)

education = np.random.randint(1, 20, 100)

experience = np.random.randint(0, 30, 100)

income = 1000 + 200 \* age + 150 \* education + 50 \* experience + np.random.randn(100) \* 500

data = pd.DataFrame({'Age': age, 'Education': education, 'Experience': experience, 'Income': income})

2. یک مدل رگرسیون خطی بسازید و آن را آموزش دهید.

3. ضرایب مدل را تفسیر کنید و ببینید کدام ویژگی بیشترین تأثیر را دارد.

4. نمودار واقعی و پیش‌بینی‌شده را رسم کنید.

---

تمرین 5: رگرسیون خطی برای پیش‌بینی مصرف انرژی

توضیحات:

- داده‌ها: استفاده از داده‌های [Energy Efficiency Dataset](https://archive.ics.uci.edu/ml/datasets/energy+efficiency)

که شامل اطلاعات مربوط به ساختمان‌ها و مصرف انرژی آن‌ها است.

- هدف: پیش‌بینی میزان مصرف انرژی گرمایشی و سرمایشی بر اساس ویژگی‌های مختلف مانند مساحت، جهت‌گیری، و عایق‌بندی.

- مراحل:

1. داده‌ها را بارگذاری و پیش‌پردازش کنید.

2. یک مدل رگرسیون خطی برای پیش‌بینی مصرف انرژی بسازید.

3. ضرایب مدل را تفسیر کنید و ببینید کدام ویژگی‌ها بیشترین تأثیر را دارند.

4. معیارهای ارزیابی (`MSE`, `R²`) را محاسبه کنید و نتایج را تحلیل کنید.