به نام خدا

کوییز 1 – نام نا خانوادگی : نمره :

**سوال ۱: مراحل اصلی یک پروژه ماشین‌لرنینگ را نام ببرید و هر کدام را به اختصار توضیح دهید.**

**سوال ۲: کد زیر چه کاری انجام می‌دهد؟**

data = pd.read\_csv("Bengaluru\_House\_Data.csv")

data.head()

* الف) دیتاست را بارگذاری می‌کند و ۵ سطر اول آن را نمایش می‌دهد.
* ب) دیتاست را بارگذاری می‌کند و اطلاعات کلی آن را نمایش می‌دهد.
* ج) دیتاست را بارگذاری می‌کند و ستون‌های آن را حذف می‌کند.
* د) دیتاست را بارگذاری می‌کند و مقادیر گم‌شده را پر می‌کند.

**سوال ۳: تفاوت بین loc و iloc در پانداس چیست؟**

* الف) loc بر اساس نام ستون و iloc بر اساس شماره ستون کار می‌کند.
* ب) loc بر اساس شماره ستون و iloc بر اساس نام ستون کار می‌کند.
* ج) هر دو بر اساس نام ستون کار می‌کنند.
* د) هر دو بر اساس شماره ستون کار می‌کنند.

**سوال ۴: کد زیر چه کاری انجام می‌دهد؟**

filtered\_data = data[data['price'] > 100]

* الف) سطرهایی که قیمت آن‌ها بیشتر از ۱۰۰ است را فیلتر می‌کند.
* ب) سطرهایی که قیمت آن‌ها کمتر از ۱۰۰ است را فیلتر می‌کند.
* ج) ستون price را حذف می‌کند.
* د) مقادیر گم‌شده در ستون price را پر می‌کند.

**سوال ۵: هدف از استفاده از data.info() در پانداس چیست؟**

* الف) نمایش تعداد سطرها و ستون‌ها، نوع داده‌ها و تعداد مقادیر غیرتهی.
* ب) نمایش مقادیر منحصربه‌فرد در هر ستون.
* ج) نمایش مقادیر گم‌شده در هر ستون.
* د) نمایش مقادیر تکراری در دیتافریم.

**سوال ۶: کد زیر چه کاری انجام می‌دهد؟**

data['bhk'] = data['size'].apply(lambda x: int(str(x).split(' ')[0]))

* الف) ستون size را به عدد تبدیل می‌کند.
* ب) ستون size را به دو بخش تقسیم می‌کند و بخش اول را به عنوان bhk ذخیره می‌کند.
* ج) ستون size را حذف می‌کند.
* د) مقادیر گم‌شده در ستون size را پر می‌کند.

**سوال ۷: تفاوت بین MinMaxScaler و StandardScaler چیست؟**

* الف) MinMaxScaler داده‌ها را به بازه [۰, ۱] نرمال‌سازی می‌کند، اما StandardScaler داده‌ها را استاندارد می‌کند (میانگین صفر و واریانس یک).
* ب) MinMaxScaler داده‌ها را استاندارد می‌کند، اما StandardScaler داده‌ها را نرمال‌سازی می‌کند.
* ج) هر دو روش داده‌ها را به بازه [۰, ۱] نرمال‌سازی می‌کنند.
* د) هر دو روش داده‌ها را استاندارد می‌کنند.

**سوال ۸: هدف از استفاده از train\_test\_split چیست؟**

* الف) تقسیم داده‌ها به دو بخش آموزش و تست.
* ب) تقسیم داده‌ها به دو بخش ویژگی‌ها و برچسب‌ها.
* ج) تقسیم داده‌ها به دو بخش نرمال‌سازی شده و استاندارد شده.
* د) تقسیم داده‌ها به دو بخش مقادیر گم‌شده و مقادیر پر شده.

**سوال ۹: کد زیر چه کاری انجام می‌دهد؟**

X = pd.get\_dummies(df, drop\_first=True)

* الف) تبدیل متغیرهای کیفی به متغیرهای عددی با استفاده از One-Hot Encoding.
* ب) تبدیل متغیرهای عددی به متغیرهای کیفی.
* ج) حذف ستون‌های تکراری از دیتافریم.
* د) پر کردن مقادیر گم‌شده در دیتافریم.

**سوال ۱۰: هدف از استفاده از plt.hist() در کتابخانه matplotlib چیست؟**

* الف) رسم نمودار خطی.
* ب) رسم نمودار هیستوگرام.
* ج) رسم نمودار پراکندگی.
* د) رسم نمودار میله‌ای.