实验练习11 参考答案

```python
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import PolynomialFeatures
from sklearn.model_selection import train_test_split
from sklearn.pipeline import Pipeline
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
scale = MinMaxScaler()

data = pd.read_csv('data/bike.csv')
```

```python
data = data.drop(columns=['id'])
```

```python
data = data.loc[data['city']==1]
data = data.drop(columns=['city'])
```

```python
data.loc[data['hour']>=19,'hour']=0
data.loc[data['hour']<=5,'hour']=0
data.loc[data['hour']!=0,'hour']=1
```

```python
y = data['y'].to_numpy()
y = y.reshape(len(y),1)
data = data.drop(columns=['y'])
```

```python
data = data.to_numpy()
```

```python
x_train,x_test,y_train,y_test = train_test_split(data,y,test_size=0.2)
```

```python
x_train = scale.fit_transform(x_train)
x_test = scale.fit_transform(x_test)
y_train = scale.fit_transform(y_train)
y_test = scale.fit_transform(y_test)
```

```python
model = LinearRegression()
model.fit(x_train,y_train)
```

```python
y_predict = model.predict(x_test)
```

```python
print(mean_squared_error(y_test,y_predict)**0.5)
```