

```
# Importing necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error

# Load the dataset
file_path = 'employee-processed.csv' # Replace with your file path
df = pd.read_csv(file_path)

# Drop non-feature columns (id, EmpID)
df_cleaned = df.drop(columns=['id', 'EmpID'])

# Handle any missing values by filling with mean
df_cleaned = df_cleaned.fillna(df_cleaned.mean())

# Split the data into features (X) and target variable (y)
# We'll assume "JobSatisfaction.1" is the target variable
X = df_cleaned.drop(columns=['JobSatisfaction.1'])
y = df_cleaned['JobSatisfaction.1']

# Normalize the feature data
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Split data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2,
random_state=42)

# Initialize the linear regression model
lin_reg = LinearRegression()

# Train the model
lin_reg.fit(X_train, y_train)

# Make predictions on the test set
y_pred = lin_reg.predict(X_test)

# Evaluate the model's performance using mean squared error
mse = mean_squared_error(y_test, y_pred)
print(f"Mean Squared Error on test data: {mse}")
```