Lab Assignment-3 (Machine Learning)

```
import pandas as pd
import numpy as np
from sklearn.model_selection import KFold, train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.linear_model import LinearRegression
from sklearn.decomposition import PCA
from sklearn.metrics import r2_score

# Q1: K-Fold Cross Validation (USA House Price Dataset)
# data = pd.read_csv('USA_House_Price.csv')
# X = data.drop('price', axis=1)
# y = data['price']
# scaler = StandardScaler()
# X_scaled = scaler.fit_transform(X)

# kf = KFold(n_splits=5, shuffle=True, random_state=42)
# for train_idx, test_idx in kf.split(X_scaled):
#     X_train, X_test = X_scaled[train_idx], X_scaled[test_idx]
#     y_train, y_test = y.iloc[train_idx], y.iloc[test_idx]
#     model = LinearRegression()
#     model.fit(X_train, y_train)
#     y_pred = model.predict(X_test)
#     r2 = r2_score(y_test, y_pred)

# Q2: Validation Set with Gradient Descent
# (Custom gradient descent implementation skipped in skeleton)

# Q3: Car Price Prediction Dataset
# car_df = pd.read_csv('imports-85.data', names=[...])
# Preprocessing: handle missing values, encoding, scaling
# Train/Test split, linear regression, PCA

# ------------------------------------------------
# Name: Aaditya Khanna
# Roll No: 102483002
# Sub Group: 3C53
# ------------------------------------------------
```