```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures
from sklearn.metrics import r2_score

from google.colab import files
uploaded = files.upload()
# Load the dataset from the CSV file
data = pd.read_csv('dataset.csv')
X = data[['horse_age']]
y = data['winner']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
# Apply polynomial regression
poly = PolynomialFeatures(degree=2)
X_train_poly = poly.fit_transform(X_train)
X_test_poly = poly.transform(X_test)
model = LinearRegression()
model.fit(X_train_poly, y_train)
y_pred = model.predict(X_test_poly)
# model's accuracy
r2 = r2_score(y_test, y_pred)
print('R2 score: ', r2)
# Predict the winner of a new horse race
new_race = np.array([[4]]) # horse_age
new_race_poly = poly.transform(new_race)
winner = model.predict(new_race_poly)
print('Predicted winner: ', winner)
```

```
Choose Files dataset.csv
• dataset.csv(text/csv) - 53 bytes, last modified: 3/24/2023 - 100% done
Saving dataset.csv to dataset (3).csv
R2 score:  -1.0
Predicted winner:  [1.]
/usr/local/lib/python3.9/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but PolynomialFeatures was
  warnings.warn(
```

Colab paid products - Cancel contracts here

✓ 8s completed at 6:23 AM