



Cryptocurrency price prediction



Preprocessing

Advanced feature engineering

- splitting dates into components (year, month, day)
- identifying quarter ends
- and creating new columns like open-close and low-high to capture critical market insights for better predictions

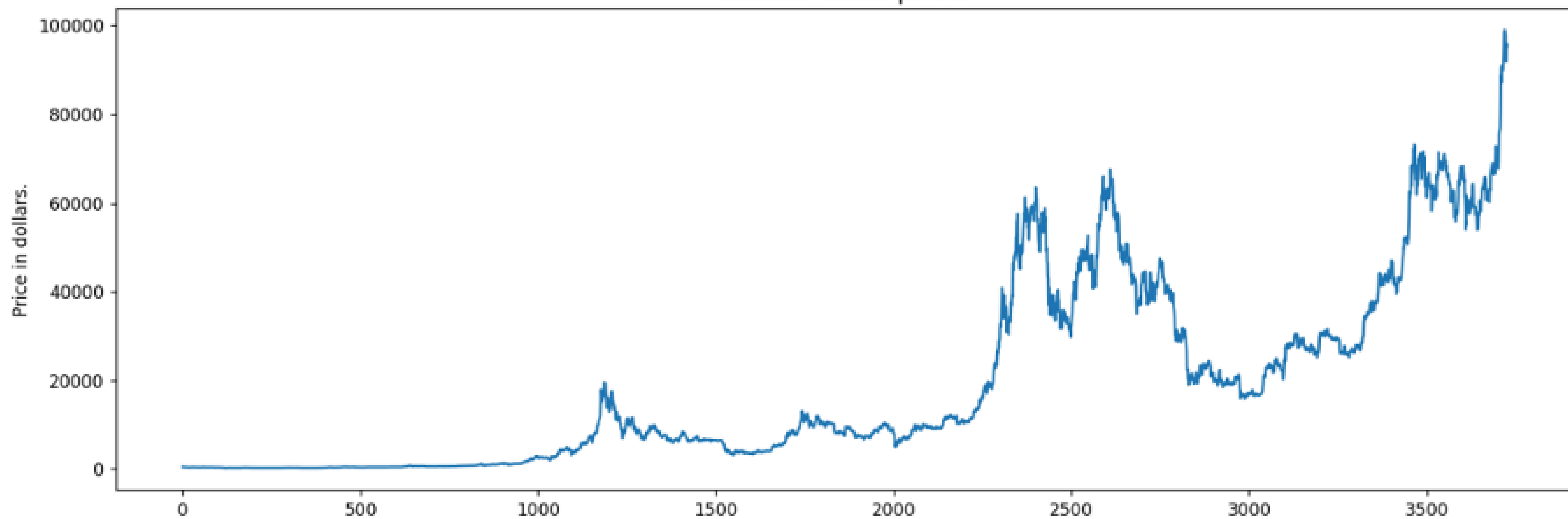
Dataset Collection + Data Pre-Processing

Collected Bitcoin data from 2014-09-17 to present

	Date	Open	High	Low	Close	Adj Close	Volume
0	2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	21056800
1	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
2	2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700
3	2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	36863600
4	2014-09-21	408.084991	412.425995	393.181000	398.821014	398.821014	26580100

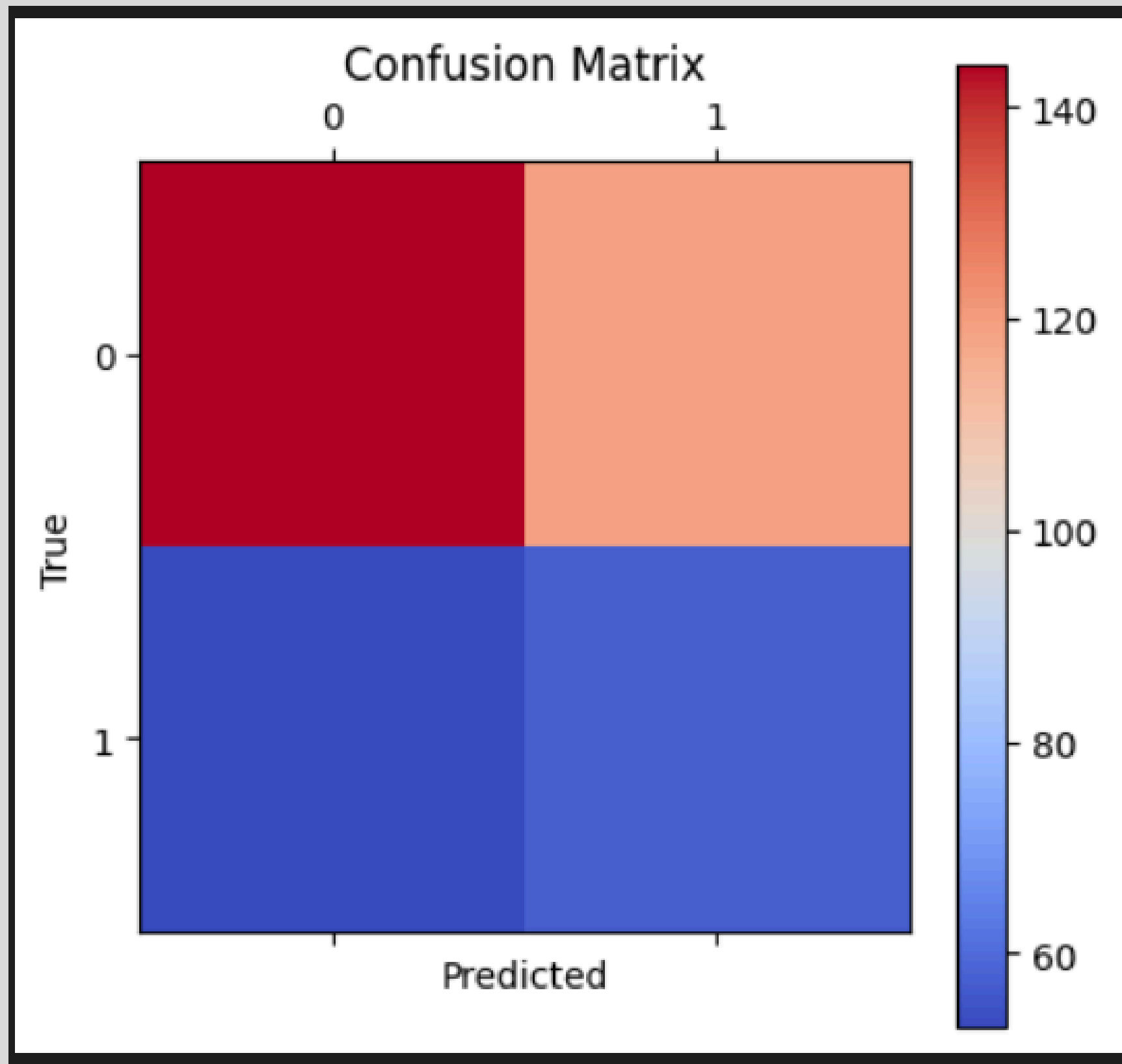
```
RangeIndex: 3726 entries, 0 to 3725
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Date        3726 non-null   object
1   Open        3726 non-null   float64
2   High        3726 non-null   float64
3   Low         3726 non-null   float64
4   Close       3726 non-null   float64
5   Adj Close   3726 non-null   float64
6   Volume      3726 non-null   int64
dtypes: float64(5), int64(1), object(1)
```

Bitcoin Close price.



	Date	Open	High	Low	Close	Volume	year	month	day	is_quarter_end	open-close	low-high	target
0	2014-09-17	465.864014	468.174011	452.421997	457.334015	21056800	2014	9	17	1	8.529999	-15.752014	0
1	2014-09-18	456.859985	456.859985	413.104004	424.440002	34483200	2014	9	18	1	32.419983	-43.755981	0
2	2014-09-19	424.102997	427.834991	384.532013	394.795990	37919700	2014	9	19	1	29.307007	-43.302979	1
3	2014-09-20	394.673004	423.295990	389.882996	408.903992	36863600	2014	9	20	1	-14.230988	-33.412994	0
4	2014-09-21	408.084991	412.425995	393.181000	398.821014	26580100	2014	9	21	1	9.263977	-19.244995	1

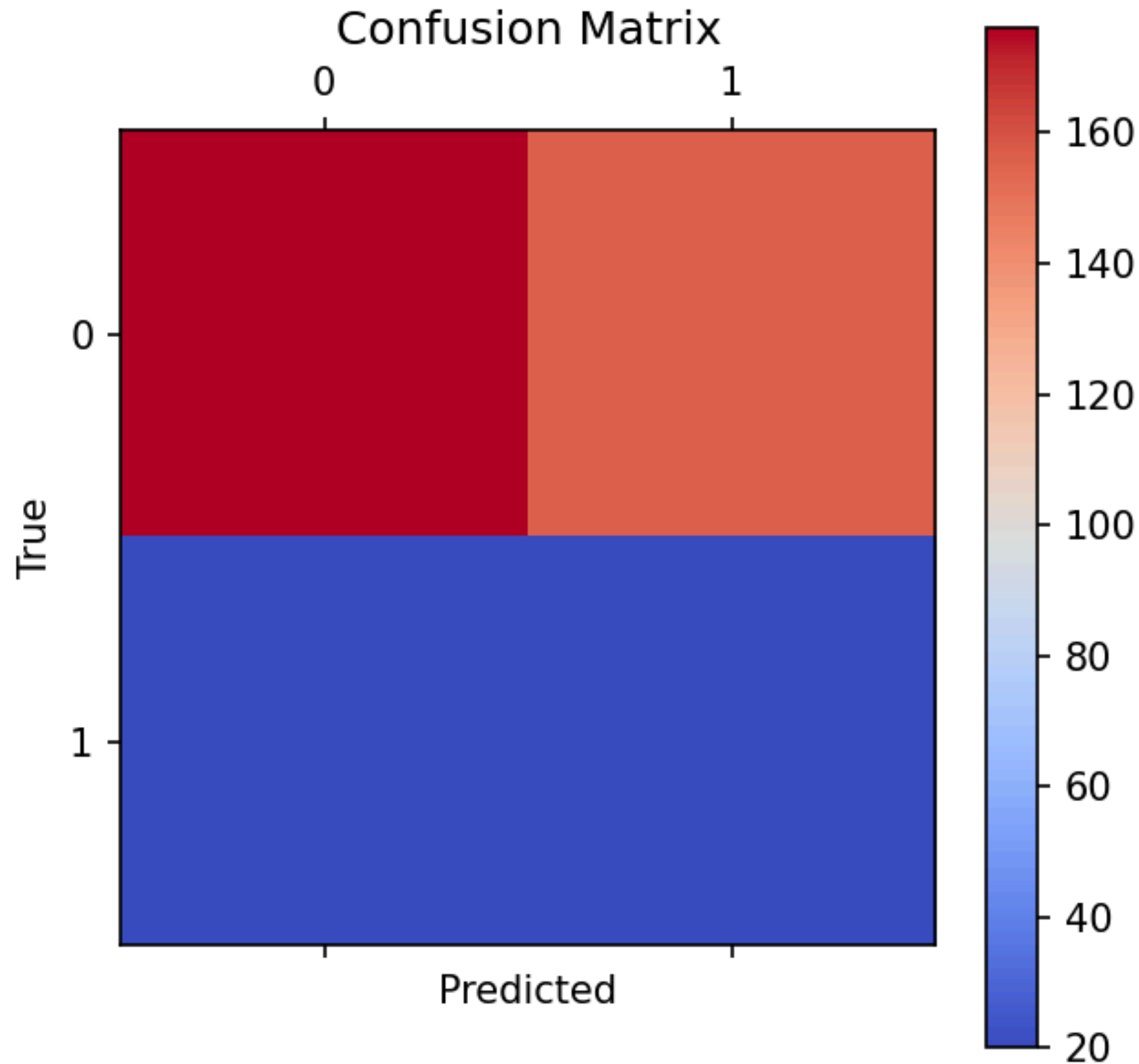
Logistic regression



```
Confusion Matrix:  
[[ 57 119]  
 [ 53 144]]
```

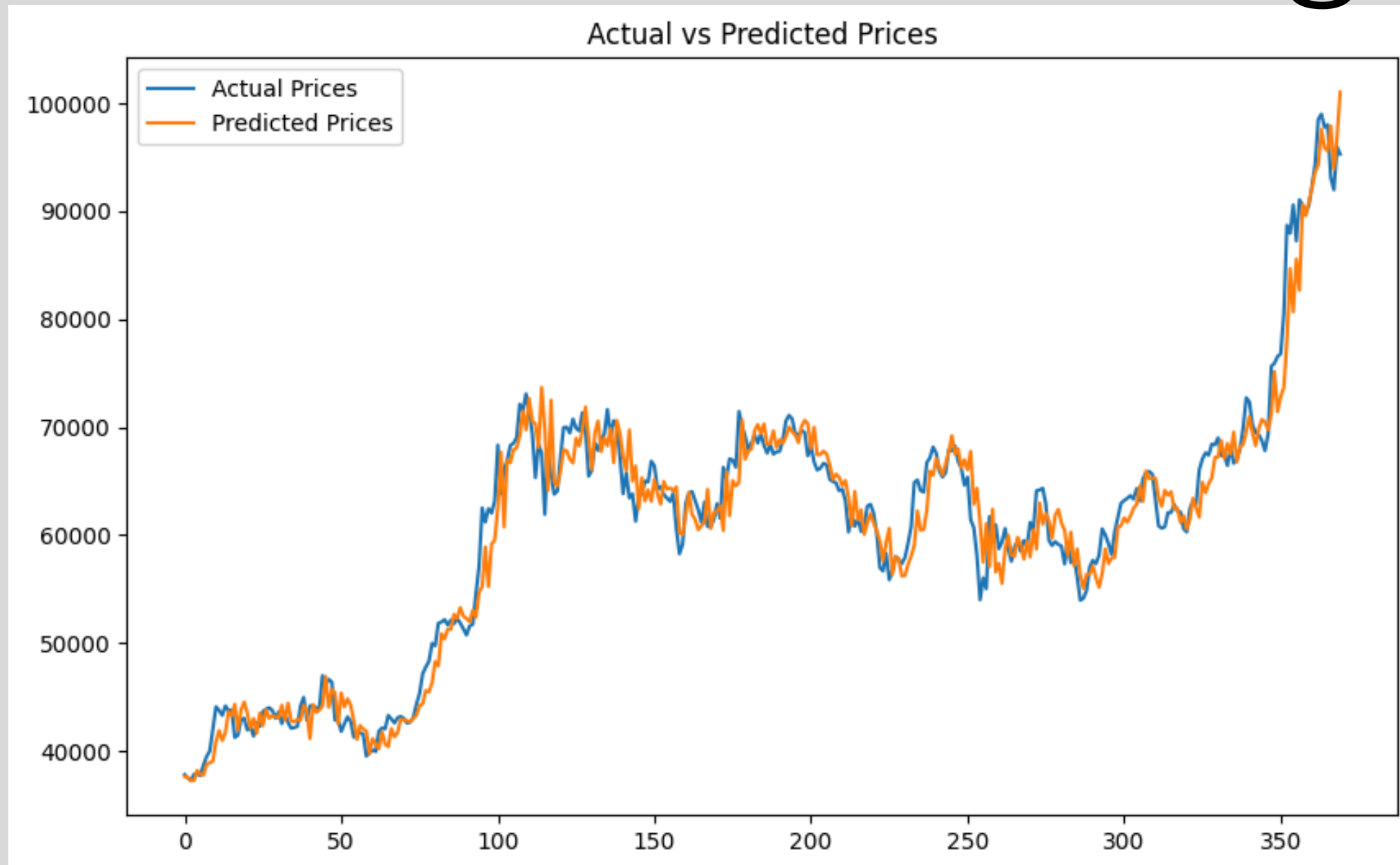
Logistic Regression - Training Accuracy: 0.826
Logistic Regression - Validation Accuracy: 0.789

SVM Classifier



```
SVM Training Accuracy: 0.820  
SVM Validation Accuracy: 0.775  
Confusion Matrix:  
[[176 156]  
 [ 21  20]]
```

Linear regression



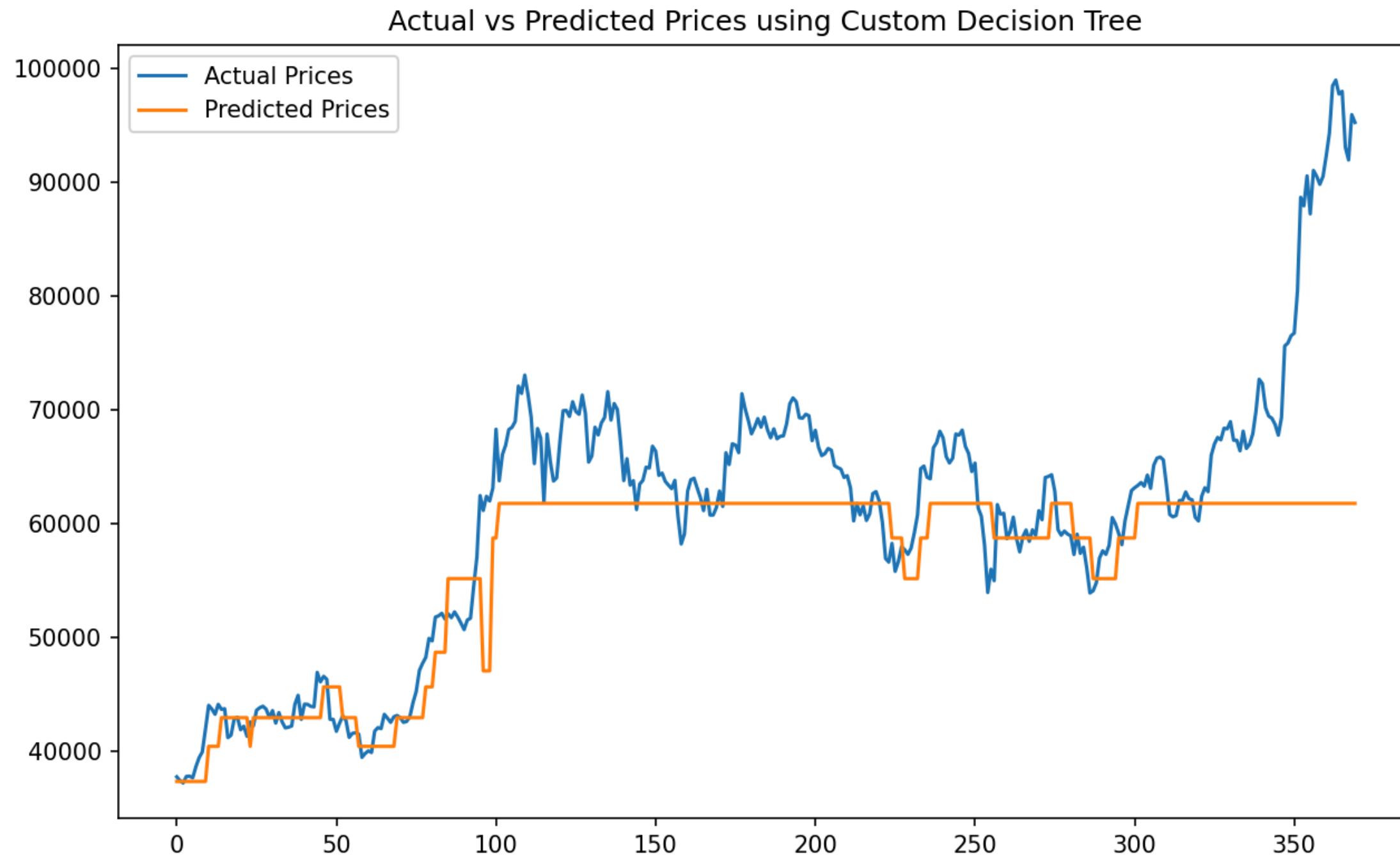
Linear Regression Metrics:

Training MAE: 560.0379, Validation MAE: 1817.0227

Training RMSE: 1122.2195, Validation RMSE: 2503.5600

Training R2: 0.9951, Validation R2: 0.9583

Decision Tree



```
Custom Decision Tree Regressor:  
Training MAE: 693.71  
Validation MAE: 5005.09  
Training RMSE: 1266.85  
Validation RMSE: 8371.29  
Training R^2: 0.99  
Validation R^2: 0.53  
Predicted next close price: 40475.59
```

THANK YOU