Final Project:

XRP prediction with Machine Learning

CS 414: Introduction to Machine Learning

Omar Drira, Michael Nelson

May 2023

Table of contents

I . Introduction

A. Background information on the problem

B. Brief overview of the machine learning techniques used

C. Objectives of the project

A. Data collection and preparation

B. Exploratory data analysis

C. Feature selection and engineering

III. Model Selection and Evaluation

A. Description of the chosen models and their parameters

B. Model evaluation metrics

C. Results of model training and testing

IV. Discussion of Results

A. Explanation of the rationale behind the chosen models

B. Interpretation of the results obtained

V. Conclusion

A. Summary of the key findings and insights

B. Limitations and future directions of the study

VI. References

A. List of sources cited in the report

I . Introduction

A. Background information on the problem

The purpose of this project is to forecast the XRP cryptocurrency's closing prices. Cryptocurrency markets are notorious for their volatility, making them an intriguing and difficult subject for price prediction. Individual and institutional investors can benefit from accurate price predictions for decision-making purposes.

B. Brief overview of the machine learning techniques used

In this study, we employ a Long Short-Term Memory (LSTM) neural network, a form of Recurrent Neural Network (RNN) that is particularly well-suited to time series data. Because LSTM networks can learn long-term dependencies, they are an excellent candidate for predicting XRP prices.

C. Objectives of the project

The main goal is to create and train an LSTM model to forecast XRP cryptocurrency closing prices and evaluate its performance using various measures.

II. Data Analysis

A. Data collection and preparation

The dataset contains daily XRP/USD prices from the past (2017 to 2023). The dataset was collected from (*XRP USD (XRP-USD) price history & historical data*). It has options like Date, Open, High, Low, Close, and Volume. The data is broken down into three sets: training, validation, and testing.

B. Exploratory data analysis

To acquire a better understanding of the price trends in each set, the training, validation, and testing sets are represented using line plots.

C. Feature selection and engineering

For the first model, using a window of 5 days prior data (Open, High, Low, Close, and Volume), the LSTM model predicts the next day's closing price. To improve training efficiency, the input features are normalized using a zero-base normalization or a min-max normalization.

For the second model, we added a new attribute to the dataset which is the price movement/difference. We compute this by taking the current closing price and subtracting the previous day's closing price. This LSTM model predicts the next day's price movement using a window of 5 days prior data (Open, High, Low, Close, and Volume) plus the added price movement.

The third model predicts the next day's closing price using a window of 5 prior days' closing prices only. This allows us to use its own predictions to predict the next days' closing prices.

III. Model Selection and Evaluation

A. Description of the chosen models and their parameters

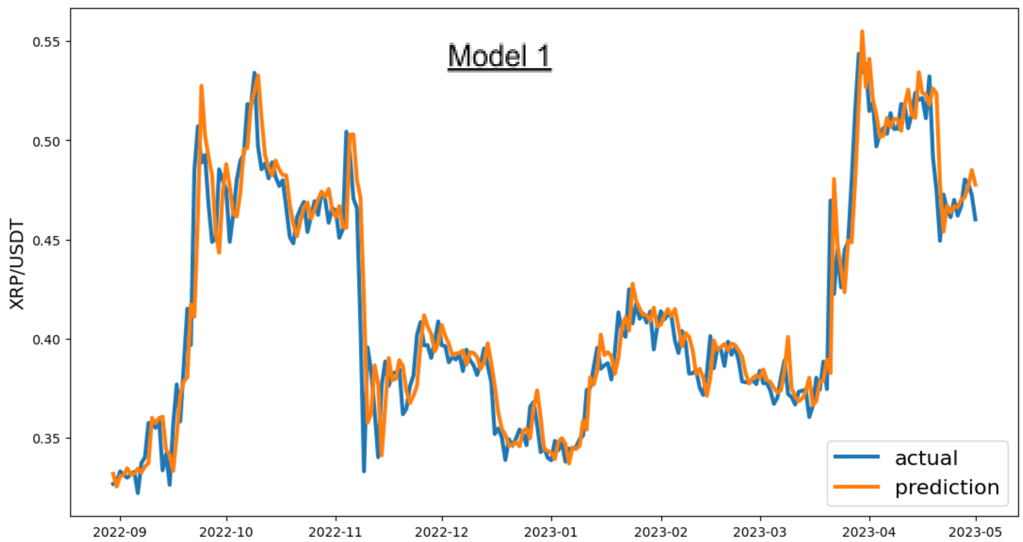
To prevent overfitting, an LSTM model with 50 neurons is utilized, along with a dropout layer 0.15. The loss function is the mean squared error (MSE), and the Adam optimizer is used for training. The three models have the same parameters.

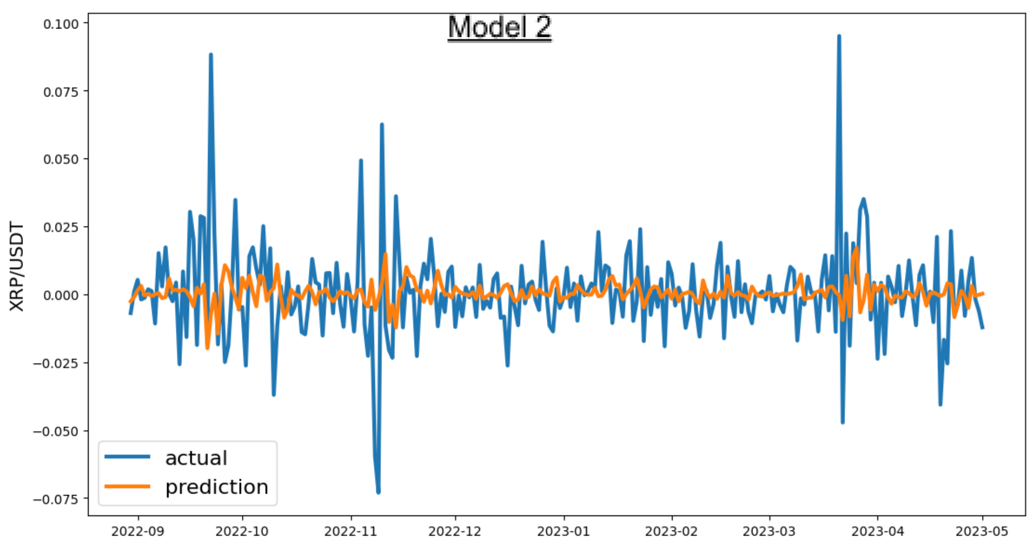
B. Model evaluation metrics

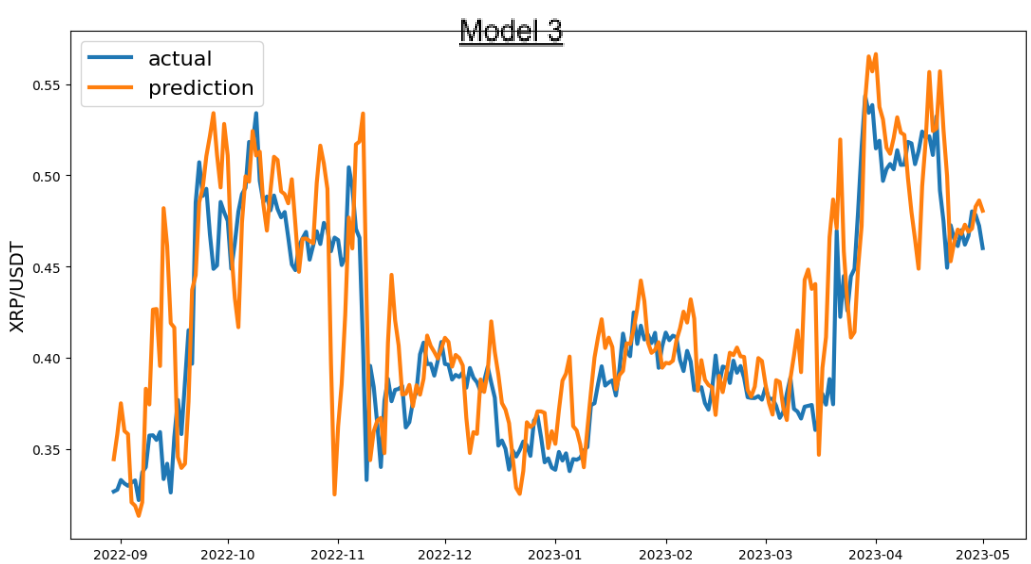
Metrics like mean absolute error (MAE), mean squared error (MSE), and R-squared (R2) are used to assess the model's performance. In addition, trading simulations and visuals are used to give insight into the model's performance.

C. Results of model training and testing

To visualize the model's performance, the training and validation losses are shown in the notebook. Line plots are used to compare the model's predictions to the actual prices in the test set.


Model 1
XRP/USDT
0.55
0.50
0.45
0.40
0.35
actual
prediction
2022-09
2022-10
2022-11
2022-12
2023-01
2023-02
2023-03
2023-04
2023-05


Model 2
XRP/USDT
0.100
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
actual
prediction
2022-09
2022-10
2022-11
2022-12
2023-01
2023-02
2023-03
2023-04
2023-05


Model 3
XRP/USDT
0.55
0.50
0.45
0.40
0.35
actual
prediction
2022-09
2022-10
2022-11
2022-12
2023-01
2023-02
2023-03
2023-04
2023-05

IV. Discussion of Results

A. Explanation of the rationale behind the chosen models

The first model predicts the closing price of XRP. The second model predicts the difference in the price from one day to the next. It was chosen to reveal the flaws in the way the first model makes its predictions. The last model uses its own previous predictions to predict the next days' closing price so it is more suitable for real-world applications.

B. Interpretation of the results obtained

In the first model, it appears to predict the prices very accurately but when we explore deeper into the LSTM model we see that it's using the previous 5 days to predict the next 1 day. This continuous structure of using the testing data to predict the next day in the testing data is inherently flawed and we can even see in the graph the predictions are basically the same as the actual data but delayed by one day.

In order to confirm this is the case we created a new model that predicts the difference in price from one day to the next. In this way if our model can actually predict the prices accurately than we should see good results here. However, Model 2 clearly does not predict the price difference very well and confirms our suspicions that the structure of the first model is flawed.

In order to combat the bias from using the last 5 days of test data to predict the next day we built a 3rd model that uses the last 5 predictions to predict the next day. In this way we are not looking at any of the test data when making predictions. As expected this model is much less accurate than the 1st but is a much better application for real-world predictions. Additionally, this model actually made more money than the first in our trading simulations.

V. Conclusion

A. Summary of the key findings and insights

People should be skeptical of the stock market and cryptocurrency predictions using the LSTM model as if implemented poorly as they can look promising but could be very misleading. Based on the results, using previous predictions to predict the future is a viable strategy in creating LSTM-based cryptocurrency predictions.

B. Limitations and future directions of the study

Future research could concentrate on optimizing the 3rd model's hyperparameters, enhancing feature engineering, and adding more features in order to get better results for XRP price predictions.

VI. References

A. List of sources cited in the report

"XRP USD (XRP-USD) Price History & Historical Data." *Yahoo! Finance*, 1 May 2023, finance.yahoo.com/quote/XRP-USD/history/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAN0zcmPjImjUGUcCB1U5lHAdq8irWE5g5tUmGttUR1Bqz0quZ9wzduSOzJ1kew83No60CP2B39UmymPpXaRu1bohhH7xHTFZ6Q9Kp51_gmH7EWLOB41JXZqFHDbuJnQSQ6jq--zEoSk3tnOeipxMCsLFyf0O49SPMrdCL6Q4mnNB.

Emirhanai. "Emirhanai/Cryptocurrency-Prediction-with-Artificial-Intelligence: First Version.. Cryptocurrency Prediction with Artificial Intelligence (Deep Learning via LSTM Neural Networks)- Emirhan Bulut." *GitHub*, github.com/emirhanai/Cryptocurrency-Prediction-with-Artificial-Intelligence. Accessed 10 May 2023.

Chat.Openai.Com, chat.openai.com/.