# INTRODUCTION

Novelty of this study can be listed as follows:

1. Comparing DoE method with KerasTuner methods in hyperparameter optimization.
2. Using hyperparameter optimizations in multistep time series prediction problems.
3. Defining custom loss and metric functions complying with portfolio optimization problem.

# METHODOLOGY

As it is illustrated in Business Process Modelling Notation (BPMN) diagram in Fig. 2.1, firstly the exchange rates are imported.

Diagram

Description automatically generated

* 1. BPMN Graph of Applied Method

## Full Factorial Design

## Response Surface Methodology

## Hyperband Tuner

## Prediction Models

### Multi-Layer Perceptron (MLP)

### Long-Short Term Memory (LSTM)

### Convolutional Encoder Decoder (Conv Enc-Dec)

### Luong’s Attention Mechanism (Luong-Att)

## Portfolio Optimization Model

Portfolio optimization problem is a special type of investment problem to select the optimal mix of opportunities that will maximize return while meeting requirements set by the investor and the market [1].

Decision variables of the optimization problem are and representing the amounts of the positions and types of the positions respectively, for the currency to open the position on time future time step and to close the position on future time step . There are 3 types of positions: buy, sell or nothing. In this context, mathematical model can be formulated as follows:

**Maximize**

**Subject to:**

|  |  |
| --- | --- |
|  | (**1**) |
|  | (**2**) |
|  | (**3**) |
|  | (**4**) |
|  | (**5**) |
|  | (**6**) |
|  | (**7**) |
|  | (**8**) |
|  | (**9**) |

First constraint ensures that investment amount is set to 0 in case position is nothing. Second constraint represents the calculation of price difference ratio where represent closing price of the time step when the position is closed. represent opening price of the time step when the position is opened. Similarly, is the ratio of spread on at the time step when the position is opened. Forth constraint demonstrates the calculation of return of investment (). Return of investment considers the ratios of price difference and spread. Fourth equation is the calculation of balances () in each time step . Fifth constraint ensures the nonnegativity of the returns. Sixth constraint calculates the balances in each time step . Seventh constraint ensures the nonnegativity of the balances. Amounts of the investments are nonnegative values. Types of the investment are -1, 0 and 1 that denote respectively to sell, nothing and to buy.

## Evolutionary Algorithms

### Genetic Algorithm (GA)

### Particle Swarm Optimization (PSO)

# APPLICATION

## Description of Data

As an application, 5 symbols of cryptocurrencies are taken into consideration: BCHUSD, BTCUSD, ETHUSD, LTCUSD and RPLUSD. The frequency of the data that is imported is 30 mins. Market data is imported for the date interval 2021-09-01 and 2022-03-10. In order to import market dataset, MetaTrader 5 [1] platform’s API [2] is used. First Prudential Markets (FP Markets) [3] is used as broker to fetch market data. Descriptive statistics of closing prices of imported datasets are demonstrated in Table 1.1.



Descriptive Statistics of Closing Prices

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BTCUSD | ETHUSD | BCHUSD | LTCUSD | RPLUSD |
| count | 6516.00 | 6516.00 | 6516.00 | 6516.00 | 6516.00 |
| mean | 49089.19 | 3563.92 | 481.58 | 161.14 | 0.93 |
| std | 8509.96 | 649.56 | 132.84 | 39.90 | 0.18 |
| min | 33098.05 | 2181.45 | 259.73 | 90.39 | 0.56 |
| 25% | 42633.72 | 3053.48 | 355.08 | 131.84 | 0.78 |
| 50% | 47299.22 | 3548.43 | 478.23 | 154.77 | 0.93 |
| 75% | 56888.63 | 4096.23 | 599.14 | 187.79 | 1.09 |
| max | 68609.11 | 4845.78 | 794.64 | 288.53 | 1.41 |

Each symbol is trained with 4 prediction models: Multi-layer perceptron (MLP), Long-short term memory (LSTM), Convolutional encoder-decoder model (Conv-EncDec) and Luong’s Attention (Luong-Att).

## Full Factorial Experiments

In each prediction model, hyperparameters that are considered in DOE are given in Table 1.2 with the levels of them. For each cryptocurrency symbol & predictive model types a full factorial design (Table 1.3) of the above-mentioned hyperparameters is created. Full factorial design consists of 16 runs with 4 repetitions to minimize the impact of random weight initializations of the predictive models.



Levels of Factors

|  |  |  |
| --- | --- | --- |
| Hyperparameter | Lower Level | Upper Level |
| Batch size | 60 | 70 |
| Number of hidden neurons | 10 | 14 |



Full Factorial Design

| Run ID | Batch Size | Number of Hidden Neurons |
| --- | --- | --- |
| 0 | 70 | 10 |
| 1 | 60 | 10 |
| 2 | 60 | 10 |
| 3 | 70 | 14 |
| 4 | 70 | 10 |
| 5 | 60 | 14 |
| 6 | 60 | 14 |
| 7 | 60 | 10 |
| 8 | 70 | 14 |
| 9 | 60 | 14 |
| 10 | 70 | 14 |
| 11 | 70 | 14 |
| 12 | 70 | 10 |
| 13 | 70 | 10 |
| 14 | 60 | 10 |
| 15 | 60 | 14 |

Full factorial design is executed for each symbol and predictive models and response values of experiments are calculated based on custom metric function. Experimental results of full factorial design for each symbol and predictive model is presented in Appendix 1, Appendix 2, Appendix 3 and Appendix 4. After completing all runs of full factorial design, a 1st order mathematical model is created. Coefficient of determination scores of 1st order mathematical models are summarized in Table 1.4.



Coefficient of Determination of 1st Order Models

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| MLP |  |  |  |  |  |
| LSTM |  |  |  |  |  |
| Conv Enc-Dec |  |  |  |  |  |
| Luong-Att |  |  |  |  |  |

Based on the first order model, steepest descent process is executed until the custom metric value starts to increase. As output of steepest descent process, factor levels where custom metric value starts to increase, are shown in Table 1.5.



Output of Steepest Descent Process

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| MLP |  |  |  |  |  |
| LSTM |  |  |  |  |  |
| Conv Enc-Dec |  |  |  |  |  |
| Luong-Att |  |  |  |  |  |

Values are presented in (Batch Size; Number of Hidden Neurons)

## Response Surface Methodology

Factor levels identified by steepest descent process are considered as central point of central composite inscribed (CCI) design. CCI design consists of 8 central points. Experiment results of CCI designs are shown in Appendix 5, Appendix 6, Appendix 7, Appendix 8 and Appendix 9 for MLP.

After execution of CCI design, a 2nd order mathematical model is created. 2nd order mathematical model consists of interaction and square effects of the factors as well as linear affects. Coefficient of determination of 2nd order models are shown in Table 1.6.



Coefficient of Determination of 2nd Order Models

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| MLP |  |  |  |  |  |
| LSTM |  |  |  |  |  |
| Conv Enc-Dec |  |  |  |  |  |
| Luong-Att |  |  |  |  |  |

Based on the 2nd order mathematical model, a grid search algorithm is applied to find the optimum hyperparameter configuration within the defined interval of CCI design. Surface plots of 2nd order models are given in Fig. 3.1, Fig. 3.2, Fig. 3.3 and Fig. 3.4.

|  |  |  |
| --- | --- | --- |
| a) | Chart, surface chart  Description automatically generated  b) | c) |
| d) | e) |  |

* 1. Response Plots for MLP: a) BCHUSD; b) BTCUSD; c) ETHUSD; d) LTCUSD; e) RPLUSD

|  |  |  |
| --- | --- | --- |
| Chart, surface chart  Description automatically generated  a) | Chart, surface chart  Description automatically generated  b) | Chart, surface chart  Description automatically generated  c) |
| Chart, surface chart  Description automatically generated  d) | Chart, surface chart  Description automatically generated  e) |  |

* 1. Response Plots for LSTM: a) BCHUSD; b) BTCUSD; c) ETHUSD; d) LTCUSD; e) RPLUSD

|  |  |  |
| --- | --- | --- |
| Chart, surface chart  Description automatically generated  a) | Chart, surface chart  Description automatically generated  b) | Chart, surface chart  Description automatically generated  c) |
| Chart, surface chart  Description automatically generated  d) | Chart, surface chart  Description automatically generated  e) |  |

* 1. Response Plots for Conv Enc-Dec: a) BCHUSD; b) BTCUSD; c) ETHUSD; d) LTCUSD; e) RPLUSD

|  |  |  |
| --- | --- | --- |
| Chart, surface chart  Description automatically generated  a) | Chart, surface chart  Description automatically generated  b) | Chart, surface chart  Description automatically generated  c) |
| Chart, surface chart  Description automatically generated  d) | Chart, surface chart  Description automatically generated  e) |  |

* 1. Response Plots for Luong-Att: a) BCHUSD; b) BTCUSD; c) ETHUSD; d) LTCUSD; e) RPLUSD

Optimum configurations that are identified via RSM are given in Table 1.7.



Optimum Configurations via RSM

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| MLP |  |  |  |  |  |
| LSTM |  |  |  |  |  |
| Conv Enc-Dec |  |  |  |  |  |
| Luong-Att |  |  |  |  |  |

Values are presented in (Batch Size; Number of Hidden Neurons)

## Hyperband

Optimum configurations handled via Keras Tuner are shown in Table 1.8.



Optimum Configurations via Hyperband

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| Hyperband | MLP |  |  |  |  |  |
| LSTM |  |  |  |  |  |
| Conv Enc-Dec |  |  |  |  |  |
| Luong-Att |  |  |  |  |  |

## Portfolio Optimization

# RESULTS AND CONCLUSION

Following steps can be considered as further research areas:

1. Social media and financial news dataset can be processed via using natural language processing techniques.
2. Statistical process control of prediction results can be applied to monitor the performance of prediction. In case the model starts to predict with error higher than a threshold, alerts can be generated to perform whole processes.
3. A strategy should be built on during the prediction and optimization durations.

appendixes

[Appendix 1](#_Toc100440779)

[Experiments of Full Factorial Design for MLP](#_Toc100440780)

[Appendix 2](#_Toc100440781)

[Experiments of Full Factorial Design for LSTM](#_Toc100440782)

[Appendix 3](#_Toc100440783)

[Experiments of Full Factorial Design for Conv-EncDec](#_Toc100440784)

[Appendix 4](#_Toc100440785)

[Experiments of Full Factorial Design for Luong-Att](#_Toc100440786)

[Appendix 5](#_Toc100440787)

[CCI Experiments for MLP/BCHUSD](#_Toc100440788)

[Appendix 6](#_Toc100440789)

[CCI Experiments for MLP/BTCUSD](#_Toc100440790)

[Appendix 7](#_Toc100440791)

[CCI Experiments for MLP/ETHUSD](#_Toc100440792)

[Appendix 8](#_Toc100440793)

[CCI Experiments for MLP/LTCUSD](#_Toc100440794)

[Appendix 9](#_Toc100440795)

[CCI Experiments for MLP/RPLUSD](#_Toc100440796)



Experiments of Full Factorial Design for MLP

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Run ID | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| 0 |  |  |  |  |  |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 6 |  |  |  |  |  |
| 7 |  |  |  |  |  |
| 8 |  |  |  |  |  |
| 9 |  |  |  |  |  |
| 10 |  |  |  |  |  |
| 11 |  |  |  |  |  |
| 12 |  |  |  |  |  |
| 13 |  |  |  |  |  |
| 14 |  |  |  |  |  |
| 15 |  |  |  |  |  |



Experiments of Full Factorial Design for LSTM

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Run ID | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| 0 |  |  |  |  |  |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 6 |  |  |  |  |  |
| 7 |  |  |  |  |  |
| 8 |  |  |  |  |  |
| 9 |  |  |  |  |  |
| 10 |  |  |  |  |  |
| 11 |  |  |  |  |  |
| 12 |  |  |  |  |  |
| 13 |  |  |  |  |  |
| 14 |  |  |  |  |  |
| 15 |  |  |  |  |  |



Experiments of Full Factorial Design for Conv-EncDec

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Run ID | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| 0 |  |  |  |  |  |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 6 |  |  |  |  |  |
| 7 |  |  |  |  |  |
| 8 |  |  |  |  |  |
| 9 |  |  |  |  |  |
| 10 |  |  |  |  |  |
| 11 |  |  |  |  |  |
| 12 |  |  |  |  |  |
| 13 |  |  |  |  |  |
| 14 |  |  |  |  |  |
| 15 |  |  |  |  |  |



Experiments of Full Factorial Design for Luong-Att

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Run ID | BCHUSD | BTCUSD | ETHUSD | LTCUSD | RPLUSD |
| 0 |  |  |  |  |  |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |
| 5 |  |  |  |  |  |
| 6 |  |  |  |  |  |
| 7 |  |  |  |  |  |
| 8 |  |  |  |  |  |
| 9 |  |  |  |  |  |
| 10 |  |  |  |  |  |
| 11 |  |  |  |  |  |
| 12 |  |  |  |  |  |
| 13 |  |  |  |  |  |
| 14 |  |  |  |  |  |
| 15 |  |  |  |  |  |



CCI Experiments for MLP/BCHUSD

| Run ID | Batch Size | Number of Hidden Neurons | Response |
| --- | --- | --- | --- |
| 0 |  |  |  |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |
| 7 |  |  |  |
| 8 |  |  |  |
| 9 |  |  |  |
| 10 |  |  |  |
| 11 |  |  |  |
| 12 |  |  |  |
| 13 |  |  |  |
| 14 |  |  |  |
| 15 |  |  |  |



CCI Experiments for MLP/BTCUSD

| Run ID | Batch Size | Number of Hidden Neurons | Response |
| --- | --- | --- | --- |
| 0 |  |  |  |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |
| 7 |  |  |  |
| 8 |  |  |  |
| 9 |  |  |  |
| 10 |  |  |  |
| 11 |  |  |  |
| 12 |  |  |  |
| 13 |  |  |  |
| 14 |  |  |  |
| 15 |  |  |  |



CCI Experiments for MLP/ETHUSD

| Run ID | Batch Size | Number of Hidden Neurons | Response |
| --- | --- | --- | --- |
| 0 |  |  |  |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |
| 7 |  |  |  |
| 8 |  |  |  |
| 9 |  |  |  |
| 10 |  |  |  |
| 11 |  |  |  |
| 12 |  |  |  |
| 13 |  |  |  |
| 14 |  |  |  |
| 15 |  |  |  |



CCI Experiments for MLP/LTCUSD

| Run ID | Batch Size | Number of Hidden Neurons | Response |
| --- | --- | --- | --- |
| 0 |  |  |  |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |
| 7 |  |  |  |
| 8 |  |  |  |
| 9 |  |  |  |
| 10 |  |  |  |
| 11 |  |  |  |
| 12 |  |  |  |
| 13 |  |  |  |
| 14 |  |  |  |
| 15 |  |  |  |



CCI Experiments for MLP/RPLUSD

| Run ID | Batch Size | Number of Hidden Neurons | Response |
| --- | --- | --- | --- |
| 0 |  |  |  |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |
| 7 |  |  |  |
| 8 |  |  |  |
| 9 |  |  |  |
| 10 |  |  |  |
| 11 |  |  |  |
| 12 |  |  |  |
| 13 |  |  |  |
| 14 |  |  |  |
| 15 |  |  |  |