

**ISTANBUL TECHNICAL UNIVERSITY**  
**COMPUTER ENGINEERING**  
DATA MINING - BLG 607

---

**CHESS PUZZLE RATING PREDICTION**

Seyyid Osman Sevgili

Instructor  
Prof. Dr. Şule Gündüz Öğüdücü

31.12.2024

# Outline

- Problem Definition
- Dataset Explanation
- Exploratory Data Analysis (EDA)
- Methodologies
- Results

# Problem Definition

## Predicting Chess Puzzle Ratings

- Enhancing user experience
- Empowering professional growth

## Glicko-2 Challenges

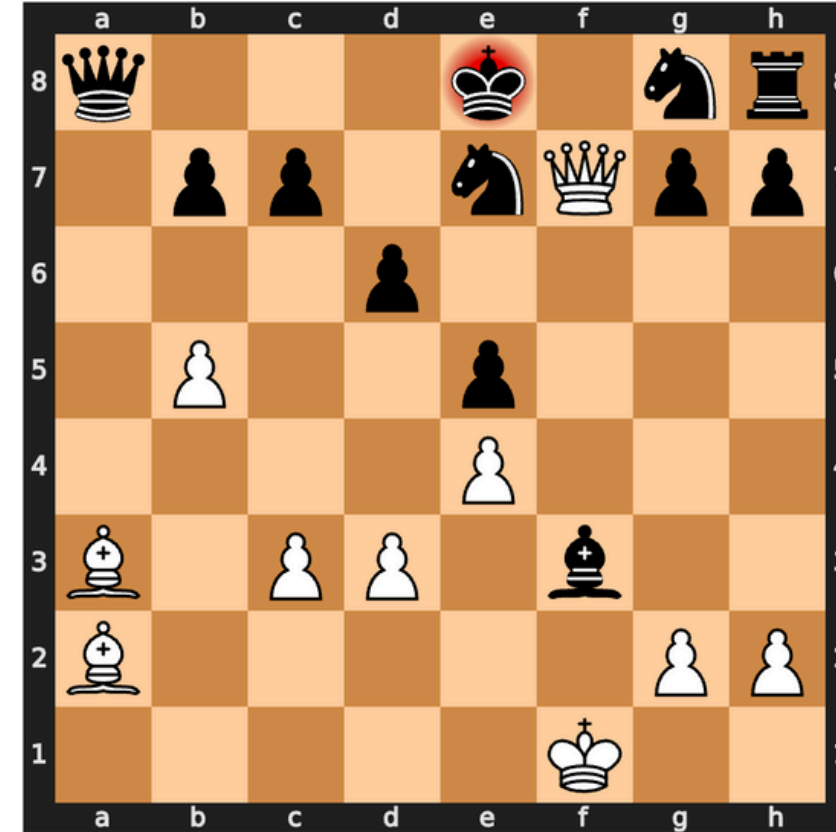
- Dependent on user performance data
- Mathematically complex



# Dataset Explanation

Field Name	Field Description	Field Type	Example Value
PuzzleId	Unique puzzle ID	string	00sHx
FEN	Board position in standard notation	string	q3k1nr/1pp1nQpp/3p4/1P2p3/4P3/B1PP1b2/B5PP/5K2 b k - 0 17
Moves	Solution to the puzzle in PGN	string	e8d7 a2e6 d7d8 f7f8
Rating	Puzzle rating	int	1760
RatingDeviation	Measure of uncertainty over difficulty	int	85
Popularity	Difference between upvotes and downvotes	int	10
NbPlays	Number of attempts at solving the puzzle	int	350
Themes	Puzzle themes	string	mates in 2 moves
GameUrl	Puzzle source link	string	<a href="https://lichess.org/game1">https://lichess.org/game1</a>
OpeningTags	Opening from which the puzzle originated	string	Sicilian Defense

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3888765 entries, 0 to 3888764
Data columns (total 10 columns):
#   Column          Dtype
---  ---
0   PuzzleId        object
1   FEN              object
2   Moves           object
3   Rating          int64
4   RatingDeviation int64
5   Popularity      int64
6   NbPlays         int64
7   Themes          object
8   GameUrl         object
9   OpeningTags     object
dtypes: int64(4), object(6)
memory usage: 296.7+ MB
```



# Exploratory Data Analysis (EDA)

## UNIVARIATE ANALYSIS

- RatingDeviation < 107.5
- NbPlays > 10
- Popularity > 65.5

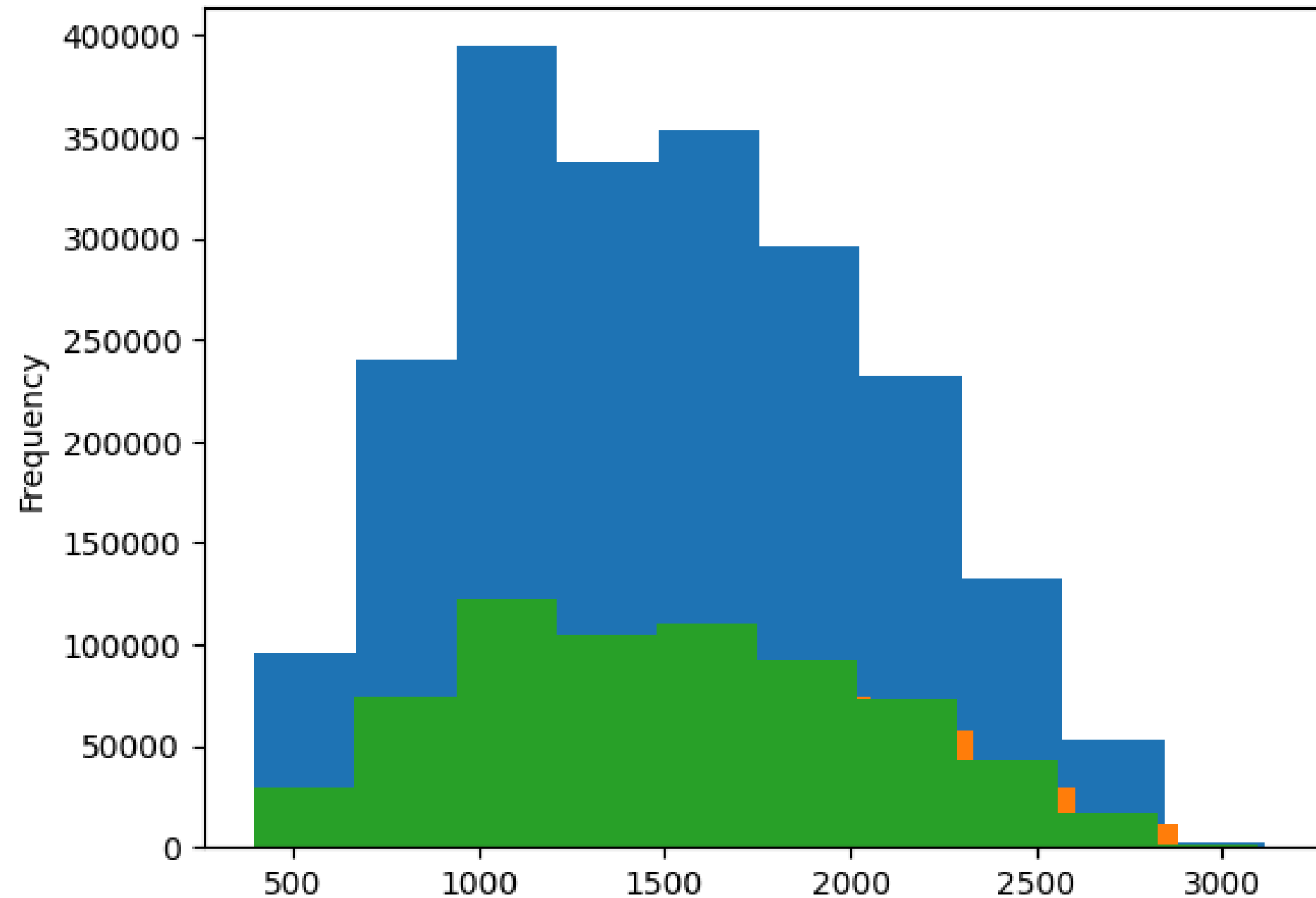
## MULTIVARIATE ANALYSIS

- MoveLength Rating Correlation

## OUTLIER DETECTION

- Puzzles are shared one move earlier

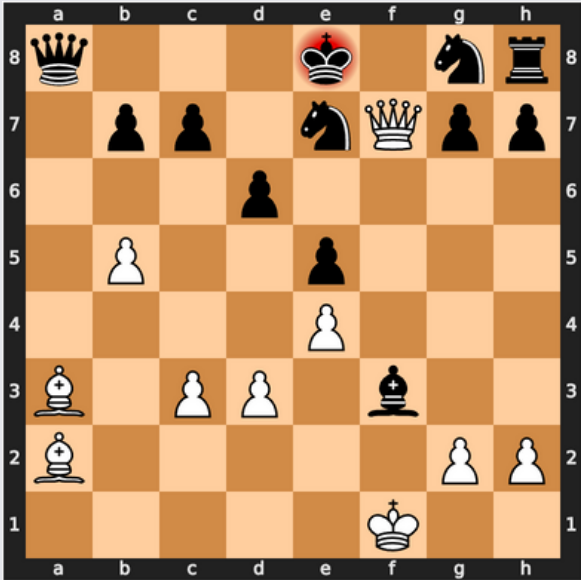
# Experiment Setup



BLUE -> Train (2137801 rows)

GREEN -> Test (668064 rows)

ORANGE -> Validation (534451 rows)



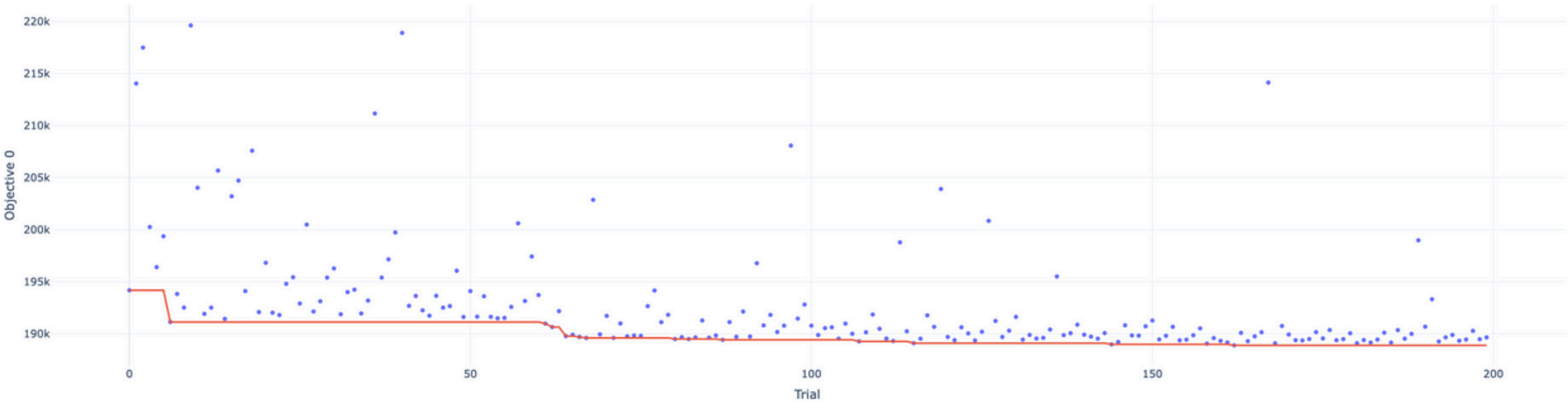
# Methodology 1 Flatten

114 0 0 0 0 0 0 107 112 112 ...

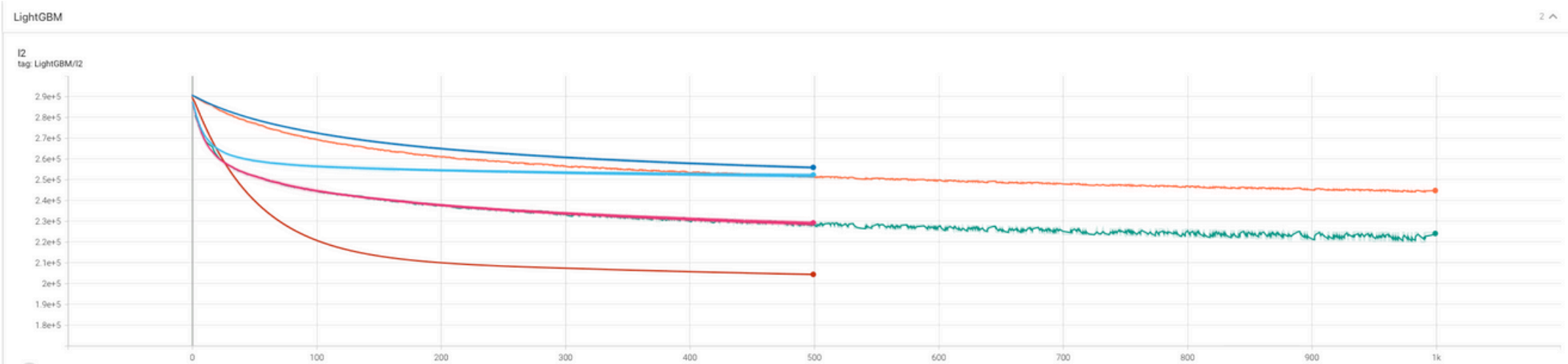
	square_1	square_2	square_3	square_4	square_5	square_6	square_7	square_8	square_9	square_10	...	square_55	square_56	square_57	square_58	square_59
0	114	0	0	0	0	0	0	107	112	112	...	80	80	0	0	0
1	0	0	0	0	0	114	107	0	0	112	...	80	80	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
3	114	0	0	113	114	0	107	0	98	0	...	0	80	82	78	0
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3340311	0	0	0	114	0	0	0	107	112	0	...	80	80	0	0	0
3340312	0	0	0	0	0	81	0	0	112	112	...	0	0	0	0	75
3340313	114	0	0	0	107	98	0	114	112	112	...	80	0	82	78	0
3340314	114	0	0	113	0	114	107	0	0	0	...	80	80	82	66	0
3340315	0	0	0	0	0	0	0	0	112	0	...	0	0	0	0	0

3340316 rows × 64 columns

Optuna

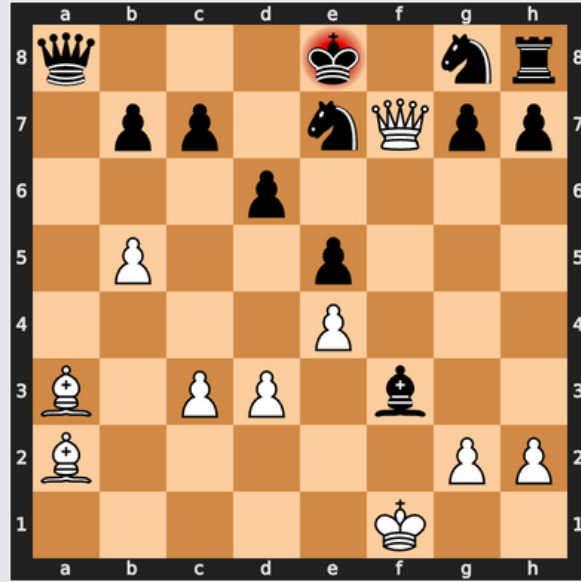


Tensorboard

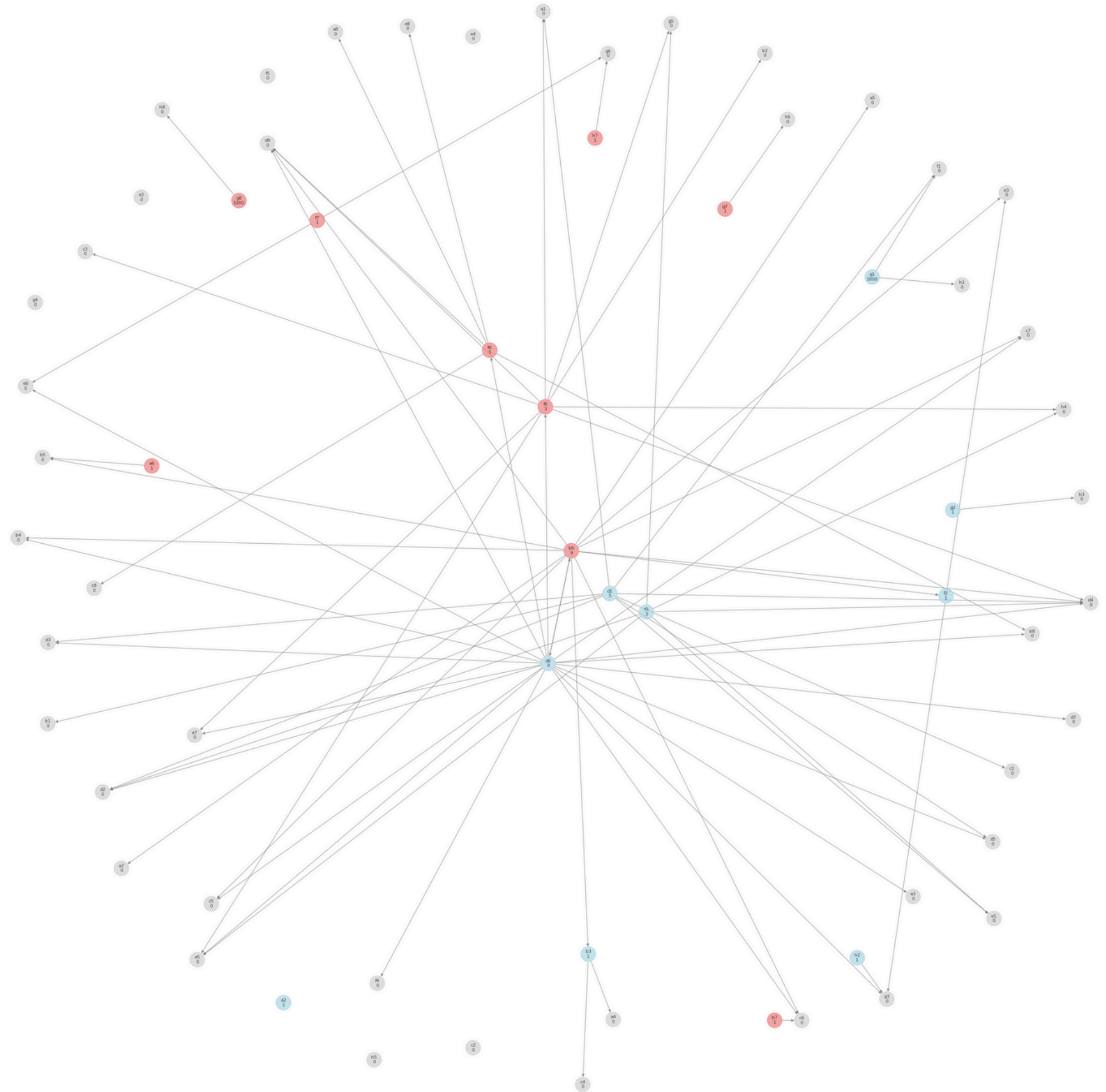
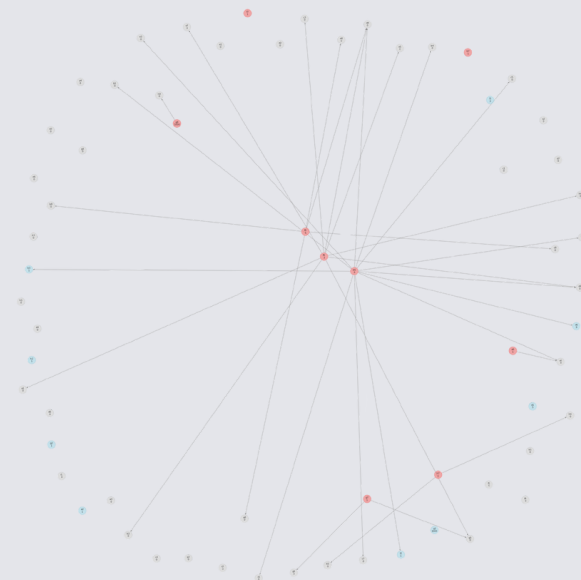


Test MSE: 176159





# Methodology 2 Graph





# Engineered Features



## ADVANTAGE

- Material advantage
- Central control



## SAFETY

- King safety
- Total threats



## MOBILITY

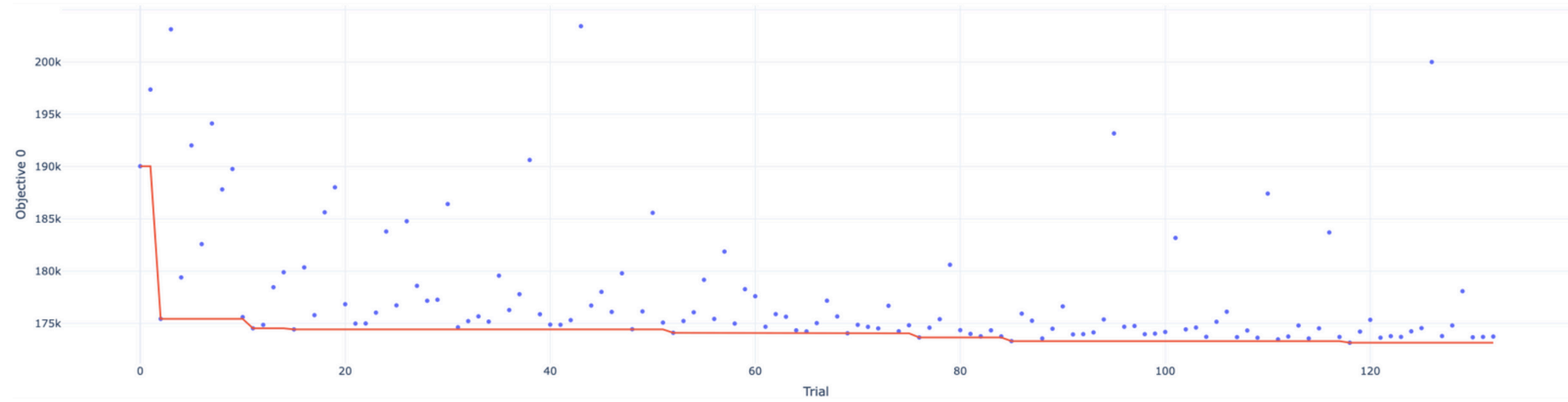
- Piece mobility
- Connectivity

$f(x)$

## OTHERS

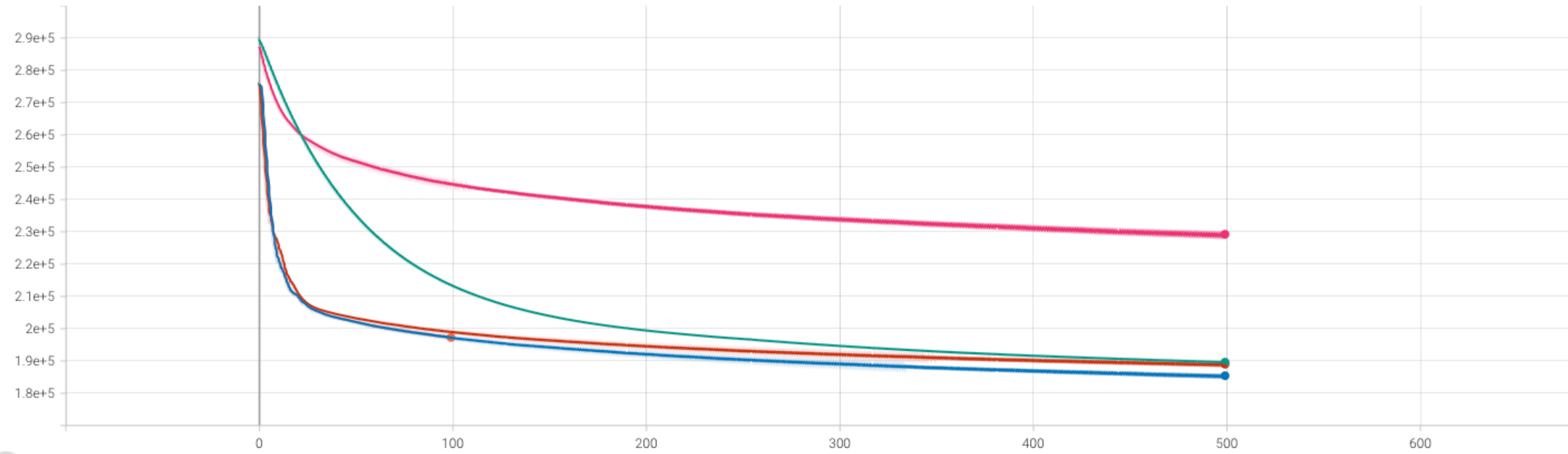
- First played peace
- Game phase
- Game status
- ....

Optuna



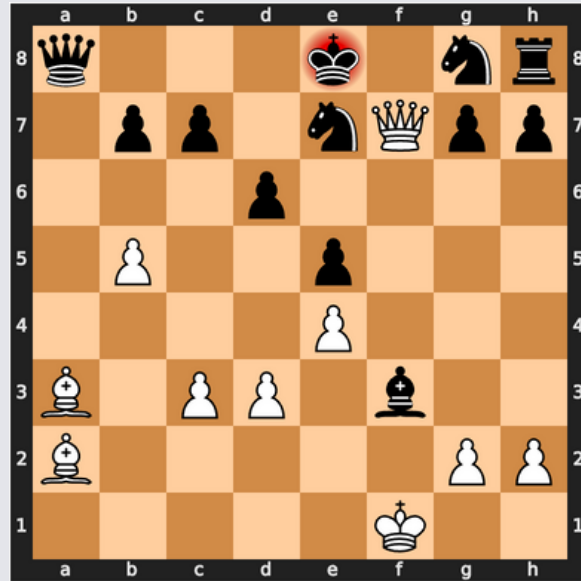
LightGBM

l2  
tag: LightGBM/l2

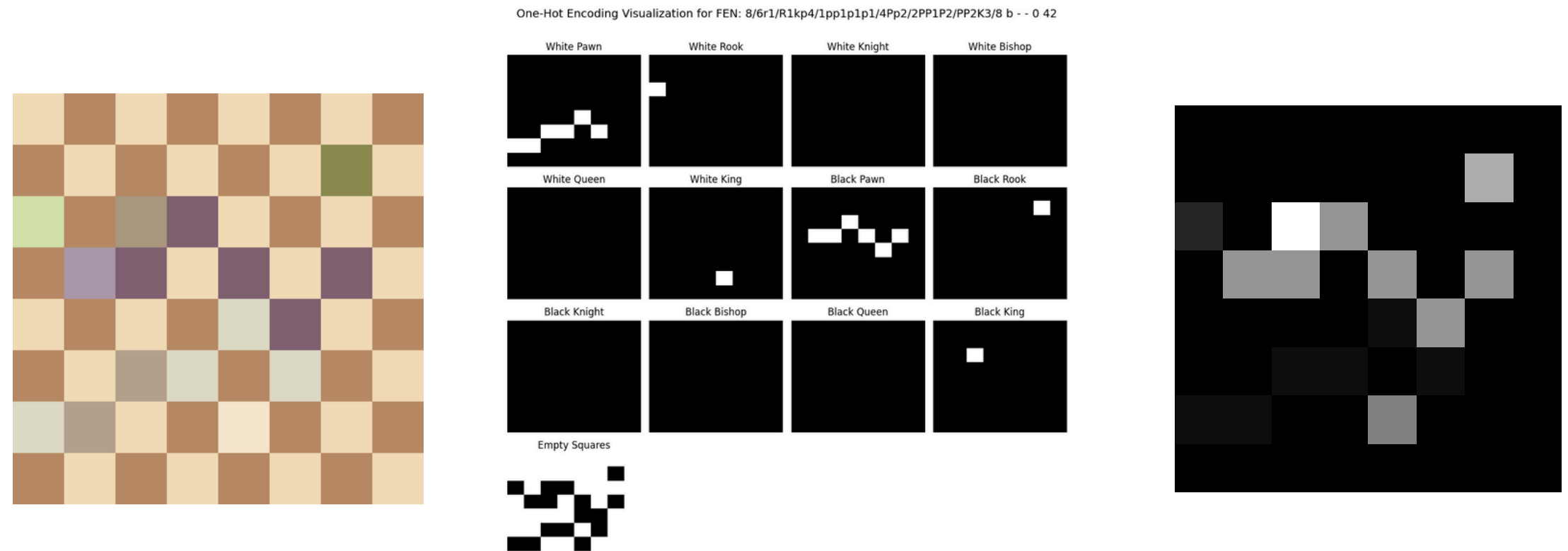
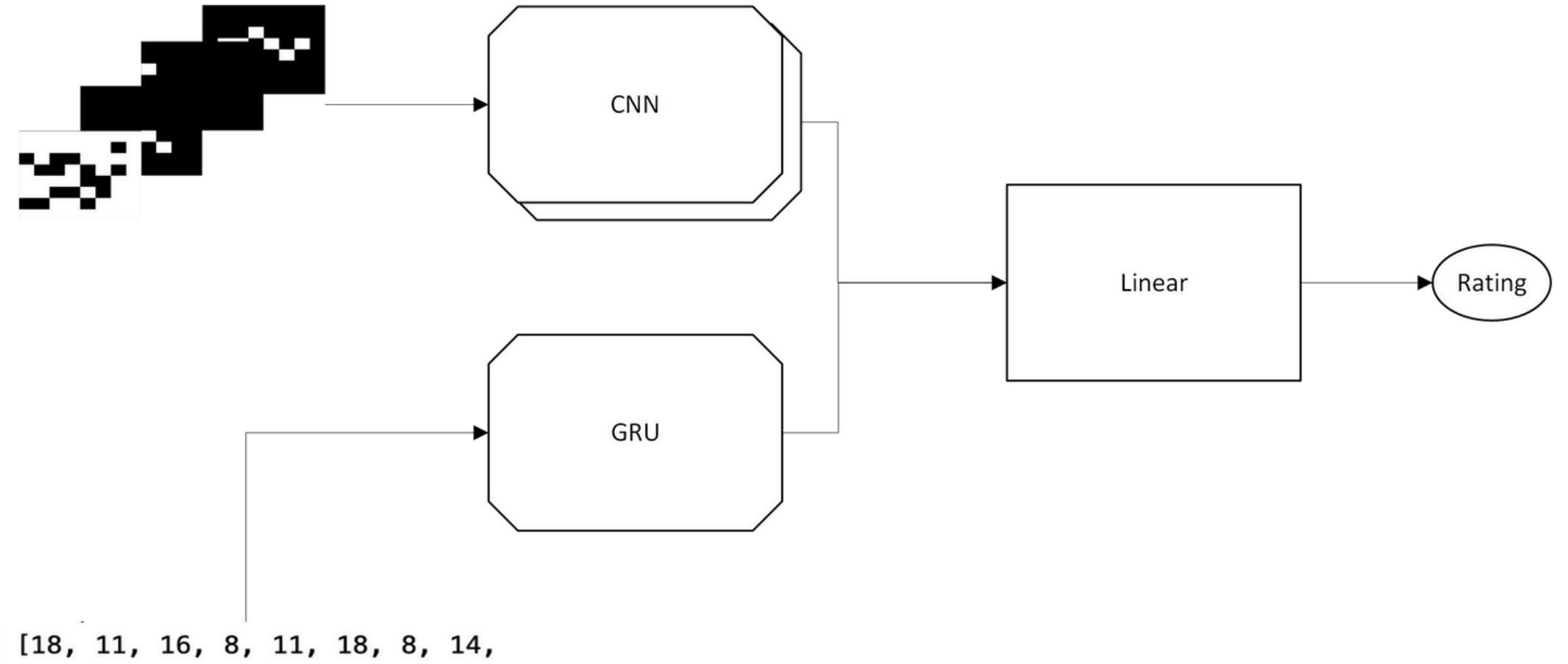


Tensorboard

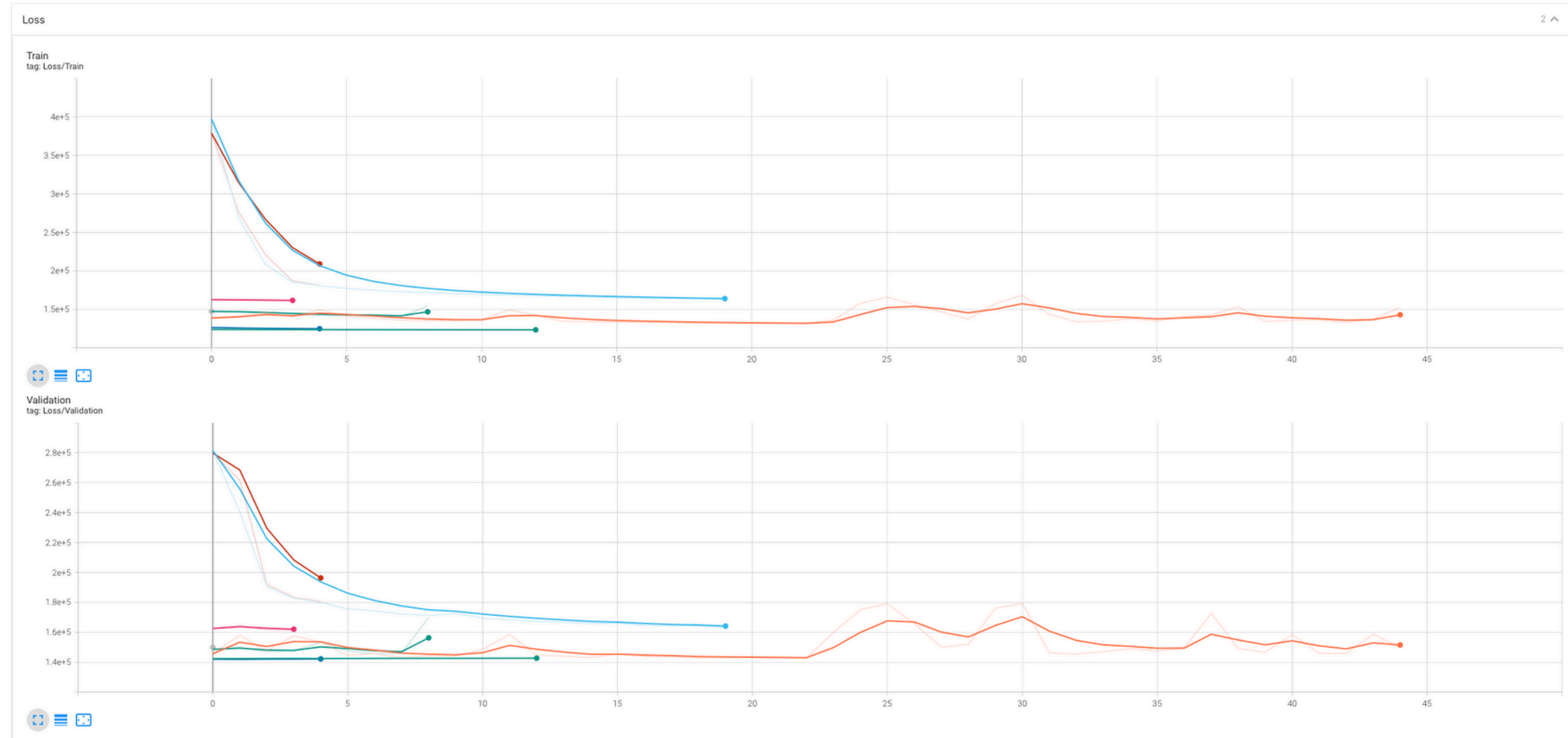
Test MSE: 163668



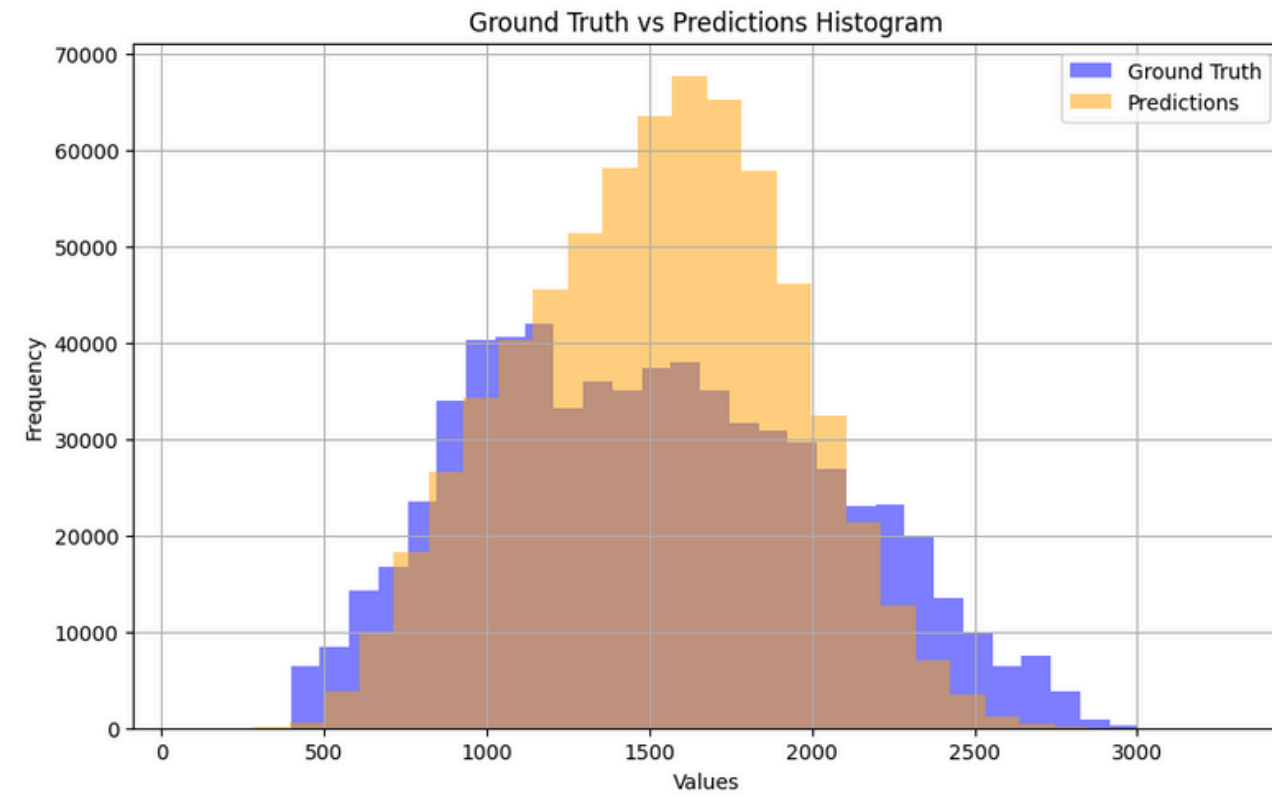
# Methodology 3 CNN-GRU



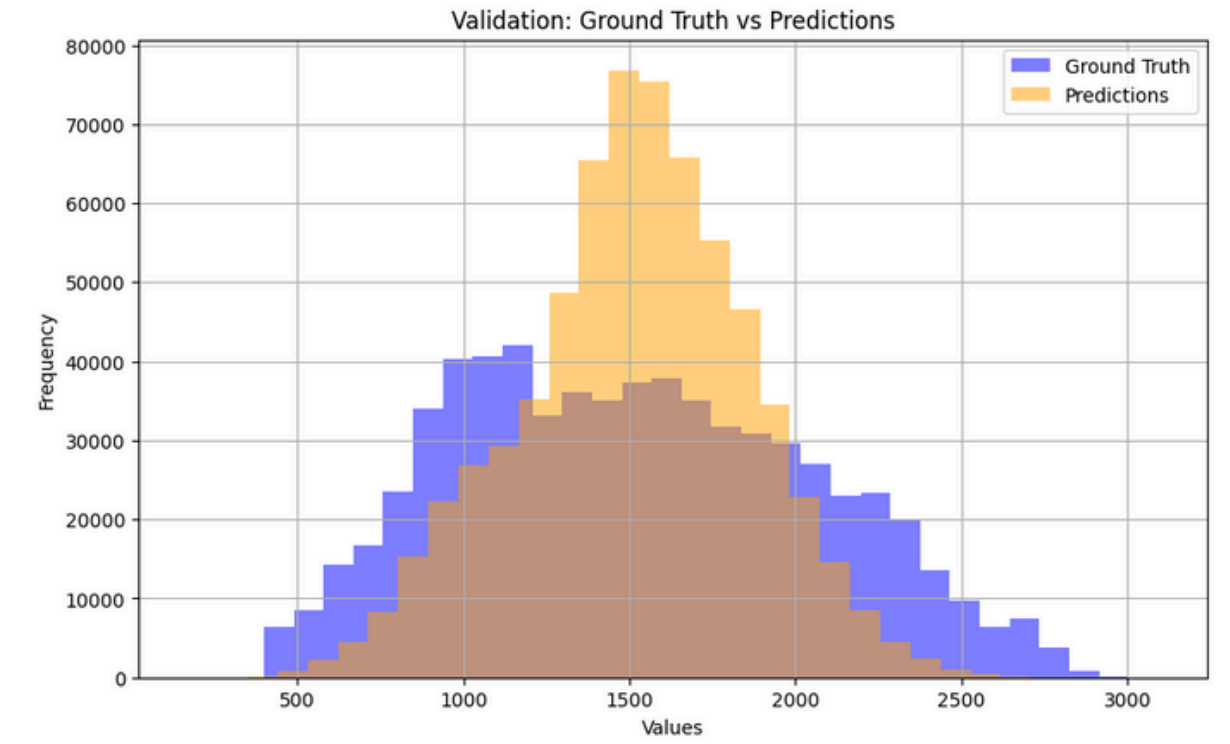
# Tensorboard



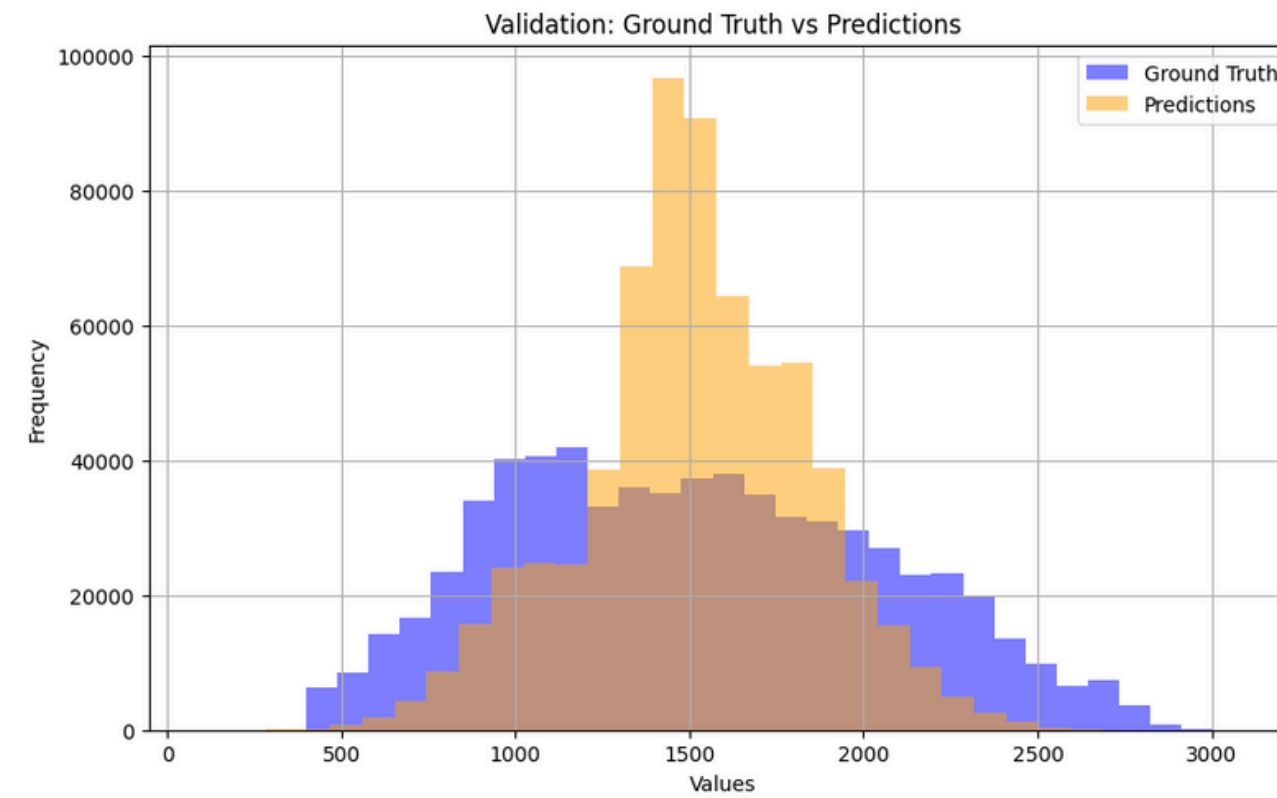
# Results



CNN-GRU



LGBM-Graph



LGBM-Flattened

# Results

Rank <span>↑↓</span>	Team Name <span>↑↓</span>	Preliminary Score <span>↑↓</span>	Final Score <span>↑↓</span>	Submissions <span>↑↓</span>
1	bread emoji	49141.5359	<b>104540.656891</b>	47
2	anansch	58810.4586	<b>120682.234604</b>	38
3	ousou	69890.9227	<b>123103.229717</b>	50
4	Andryyyyyy	61381.3812	<b>129245.229228</b>	56
5	ToDoFindATeamName	65136.8232	<b>132631.424731</b>	53
6	alexmolaz	74378.0110	<b>137839.668622</b>	22
7	dymitr	69202.5691	<b>141488.501466</b>	50
8	Feiwyth	70792.7182	<b>146729.234115</b>	6
9	NxGTR	73832.3591	<b>150757.042522</b>	35
10	BigData2024	74135.4586	<b>154905.759042</b>	51

## Competition Results





# Evaluation

## POSITIVE

- I utilized both tree based and nn models.
- Graph representation for feature engineering
- Custom architecture CNN-GRU

## NEGATIVE

- Low performance
- Huge data requires huge! hardware

## OTHER APPROACHES

- Vision transformer based architectures. 150k as a final score.

## FUTURE WORK

- Apply move to board and feed new board after each move to CNN.
- Ensemble architectures can be used.
- Move encoding can be improved.





Thank you for listening