Exploiting Betting Odds Using Machine Learning

WONG Wing Keung (1155093416)

WONG Ching Yeung Wallace (1155093534)

Main Objective

Using Machine Learning to beat the odds

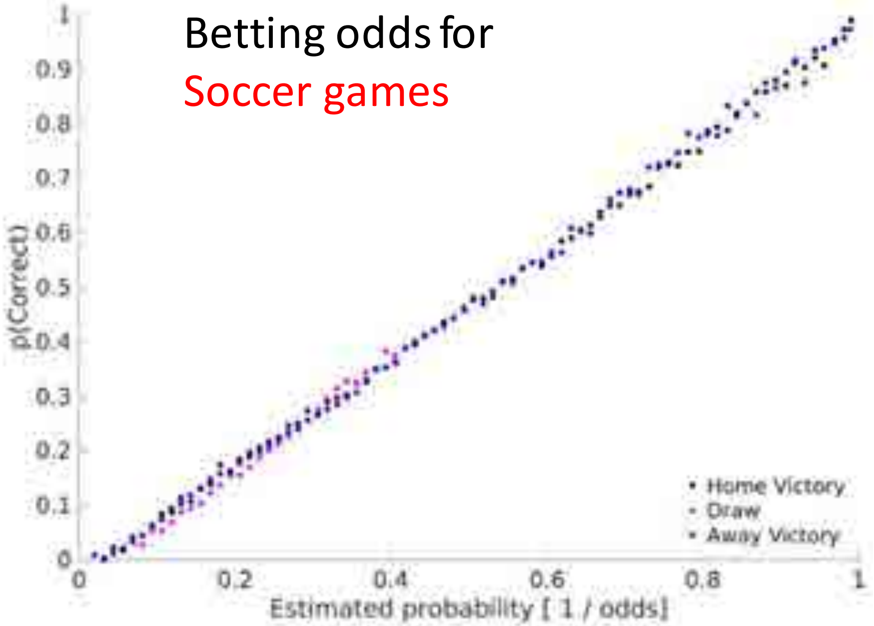
Our Solution

- Easy to obtain data
- Optimized for betting

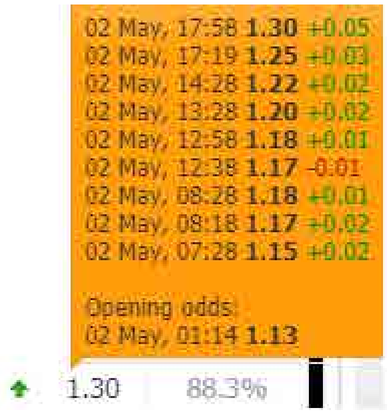
Easy to Obtain Data

- Use Betting odds as features
  - Betting odds are accurate in general
  - Odds are widely available

Betting odds for
Soccer games



- Use Odds movements as features
  - Odds movements are not noises but informative
  - Odds change from time to time
  - Odds are becoming more accurate in general



More accurate

Less accurate

Optimize for Betting

- **Kelly Formula** – a simple and optimal betting strategy
  - Relates probability of winning and payoffs (odds) together
  - **Compute the optimal bet size that maximizes the growth of profit**

- Payoffs are known, probability is unknown

➢ Build machine learning model
to map features into probability

Model Requirements

1. Activation and Loss Function
2. Ensemble model

1. Activation and Loss Function

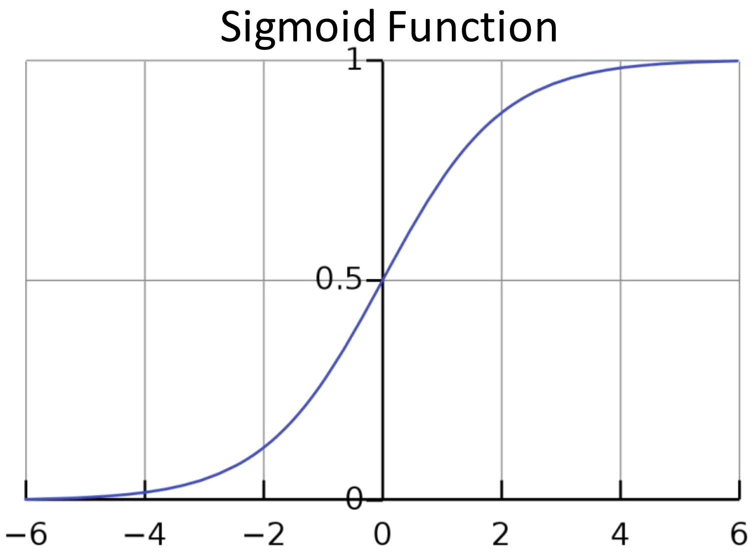
2. Ensemble model

- Objective: Mapping features into probability

- Activation in output layer: Sigmoid
  - To produce values ranging from 0 – 1

- Loss Function in training: Binary Cross Entropy
  - The optimal loss function to use with Kelly Betting

Have been shown in our report



$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1-y_i) \cdot log(1-p(y_i))$$

Binary Cross-Entropy / Log Loss

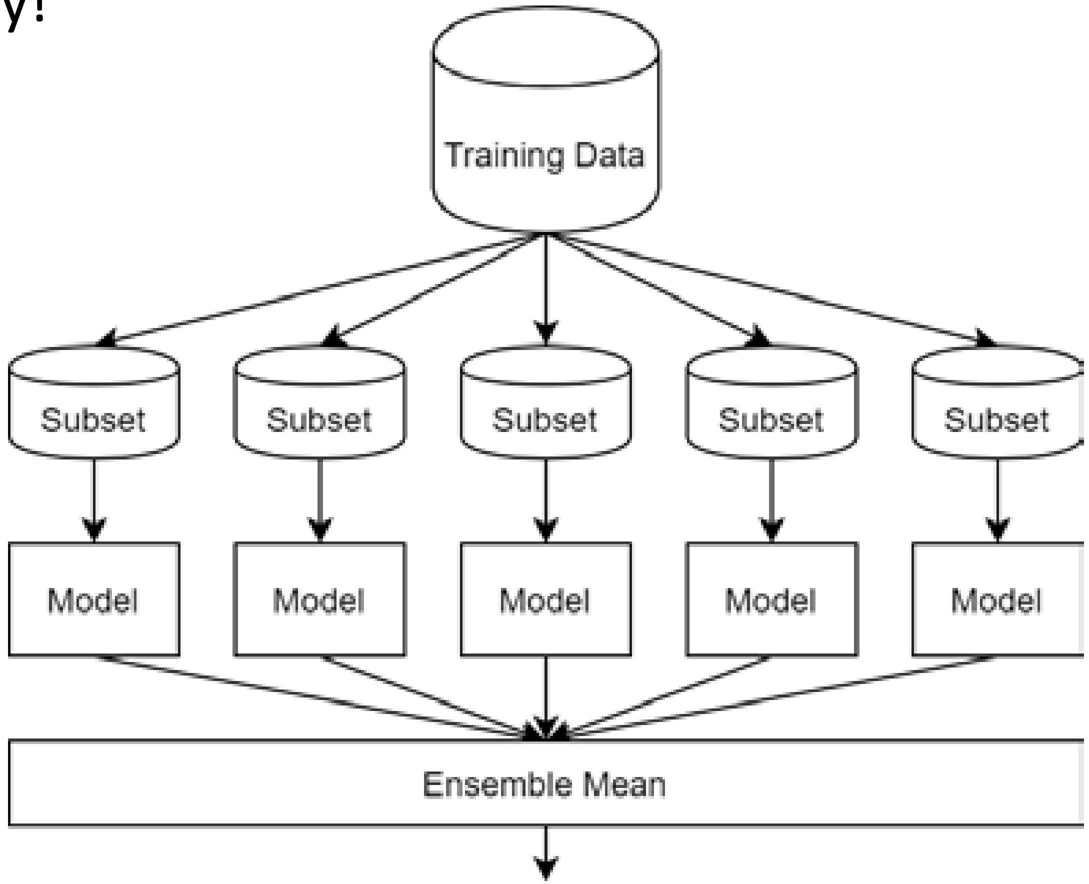
1. Activation and Loss Function

2. Ensemble model

- Kelly Formula assumes an unbiased probability
  - Overestimation causes bankruptcy quickly!

- To improve robustness
  - Train multiple models
  - Group them into an ensemble
  - Output Ensemble Mean

- 300 – 1000 models in an ensemble



Betting strategies

- Kelly & Fractional Kelly
- Improved Kelly

Betting strategies

- Kelly Betting (Full Kelly):
  - Use Kelly Formula to compute the optimal bet size
- Fractional Kelly:
  - Multiply a fraction (e.g. 20%, 30%) to the Kelly Formula
- Improved Kelly:
  - Solve for the optimal fraction under given uncertainty estimation

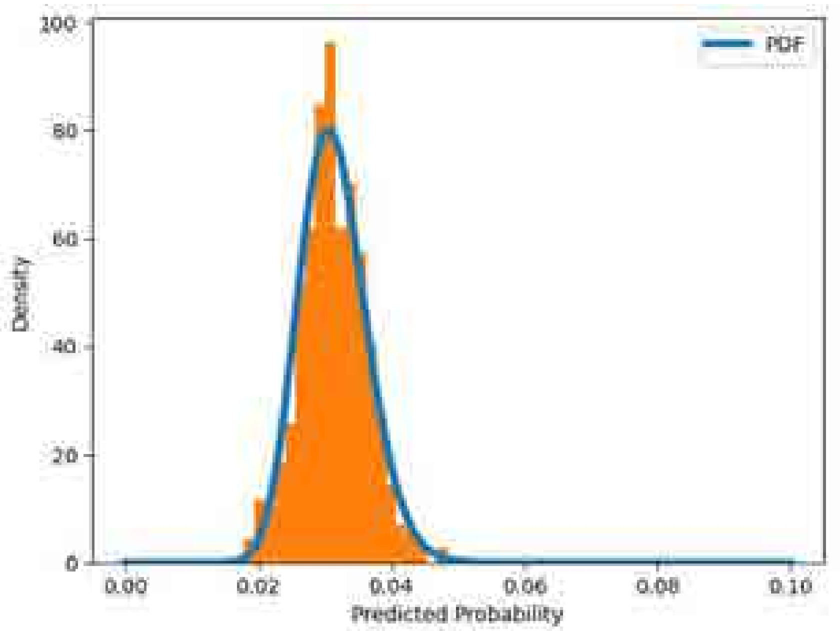
	Input	Output
Kelly Formula	Deterministic Probability, Odds	Optimal bet
Improved Kelly	PDF of Probability, Odds	Optimal fraction + Optimal bet

PDF of Probability: Uncertainty in prediction

Improved Kelly

	Input	Output
Original Kelly	Predicted Probability, Odds	Optimal bet
Improved Kelly	PDF of Probability, Odds	Optimal fraction + Optimal bet

- Assume the **uncertainty = ensemble variance**
- We obtain the PDF by performing Beta fit on ensemble members' predictions.



Proposed Models

1. Closing Model
2. Continuous Model

Overview

1. Closing Model
   - Model that **gives predictions only** just before closing
     - Closing (收盤): the moment that bookmakers do not accept bets anymore

2. Continuous Model
   - Model that **gives predictions** continuously until closing

Overview

- Select an odds considering period
    - Every period is in 1-minute interval

- Use the odds-implied probabilities in that period to form features
    - For example: 0 min – 29 min before closing
    - Odds-implied probabilities: [P(t=0min), P(t=1min), ..., P(t=28min), P(t=29min)]

1. Closing Model
2. Continuous Model

- **Regression-based Closing Model (Term 1)**
- **LSTM-based Closing Model**
- **LSTM-based Continuous Model**
- **Convolution-based Continuous Model**

Application in Horse Racing – Forming Records

- Aim: Predict Pr(Win) for each horse

- Pick an Odds Considering Period
  - Sequence of Odds-Implied Probabilities (P)

- Create a record for each horse:
  - lastP, minP, maxP ….
  - Coefficients from Polynomial Regression

Application in Horse Racing – Forming Records

- **Training Set:** Data from 2017/01/01 – 2018/12/31 (19647 records)
  - 2-Year Data
- **Testing Set:** Data from 2019/01/01 – 2019/12/31 (9827 records)
  - 1-Year Data

Application in Horse Racing – Results:

- Kelly Betting
  - Initial capital: $10000
  - Betting against the highest closing odds among ≈ 10 bookmakers

- We tried models with different configurations
- The best model is �

- Problem:
  - The returns are very sensitive to the choice of degree in polynomial regression

Model	Return
0-39-4deg	21287
0-39-6deg	20941
0-39-8deg	**24307 (highest)**
0-39-10deg	19255
0-39-12deg	17903
0-39-14deg	17472
0-39-16deg	10185
0-39-18deg	213

Application in Horse Racing – Improvement: LSTM-based

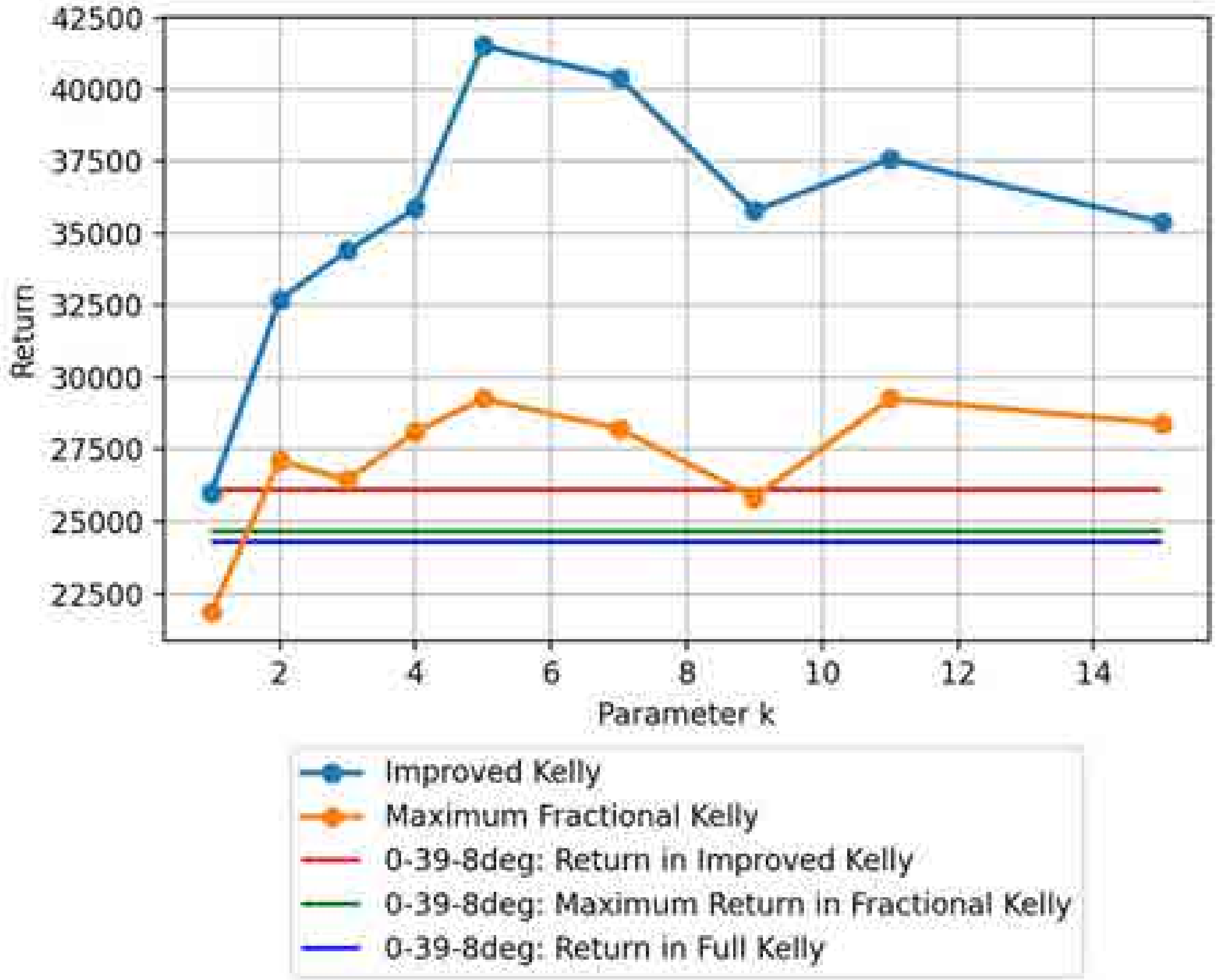
Application in Horse Racing – Improvement: LSTM-based

Application in Horse Racing – Results

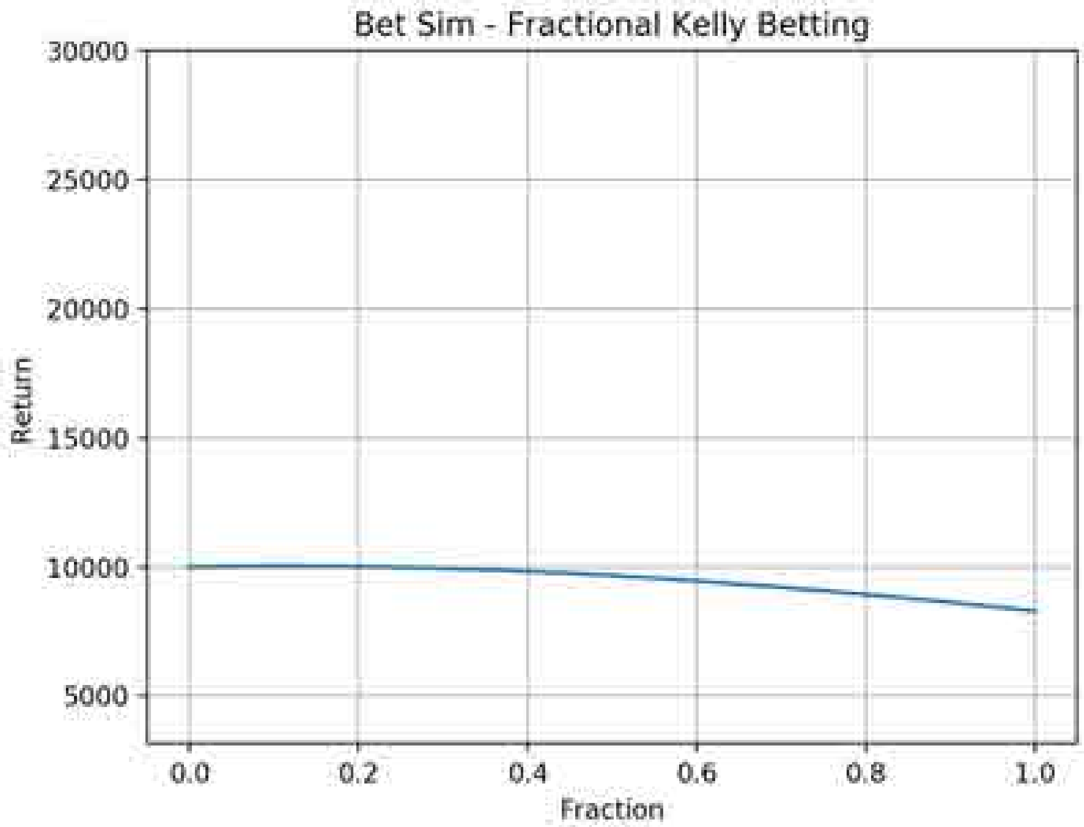
- Best regression-based model
  - 0-39-8deg

- Initial capital: $10000
- Betting against the highest closing odds among ≈ 10 bookmakers

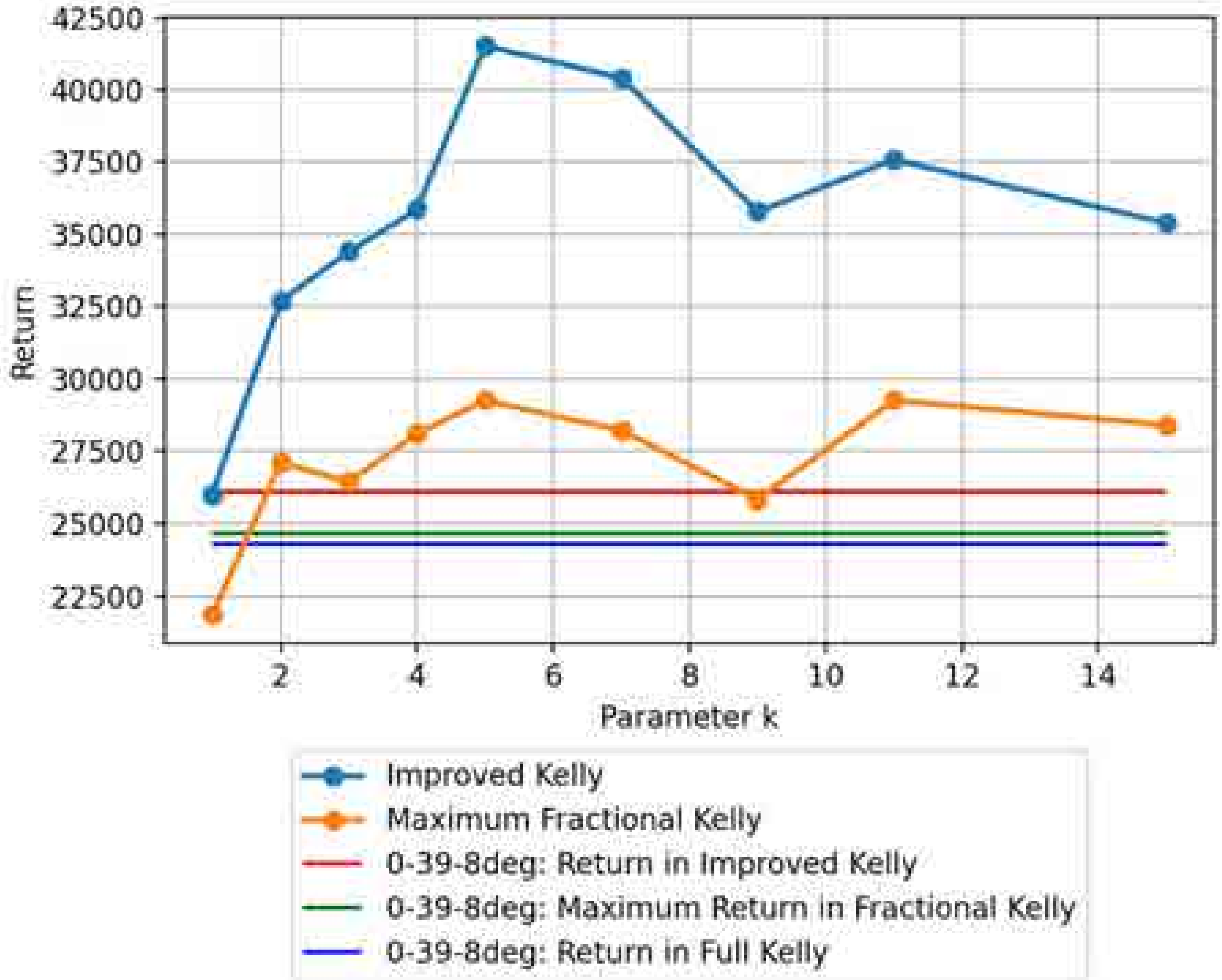
- Every LSTM-based model can outperform the Regression-based



Application in Horse Racing – Results- By Luck? NO



Using the closing-odds-implied probabilities as estimations **IS NOT profitable**



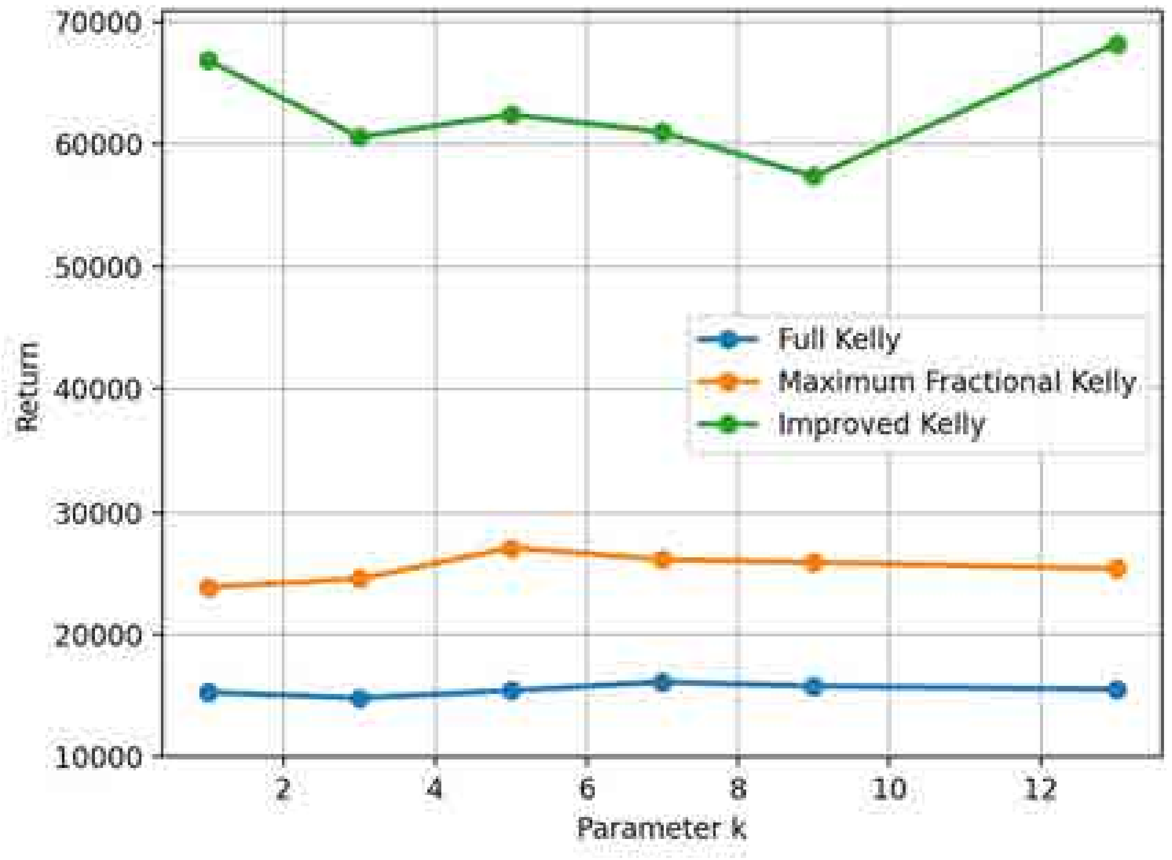
Application in Soccer

- Tested in the Over/Under (入球大細) market
  - Guess the **total goal** is Over(大) or Under(細) a Line
  - Models predict the probability of Over
- Using � to compute features
  - Odds Considering Period: �
- Data
  - Training Set: games before 2019/07/01 (18847 lines)
  - Testing Set: games from 2019/07/01 – 2020/03/08 (8567 lines)

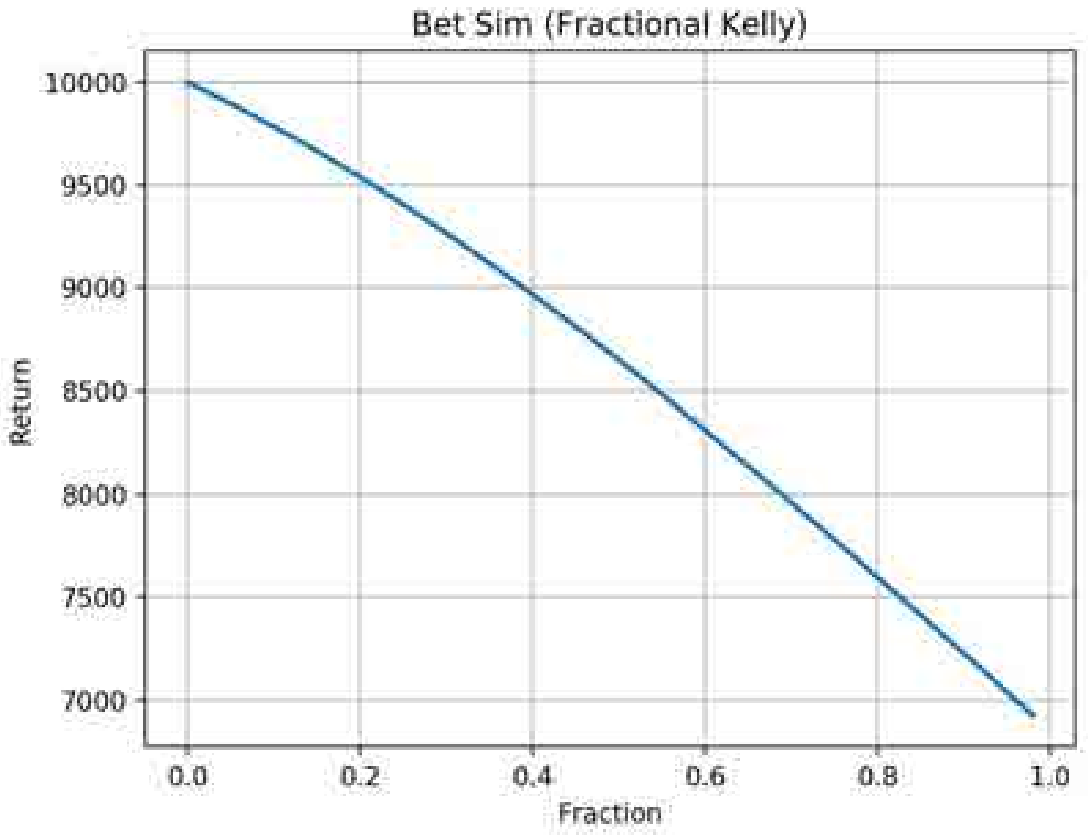
Application in Soccer – Results

- Kelly Betting
  - Initial capital: $10000
  - Betting against the highest closing odds

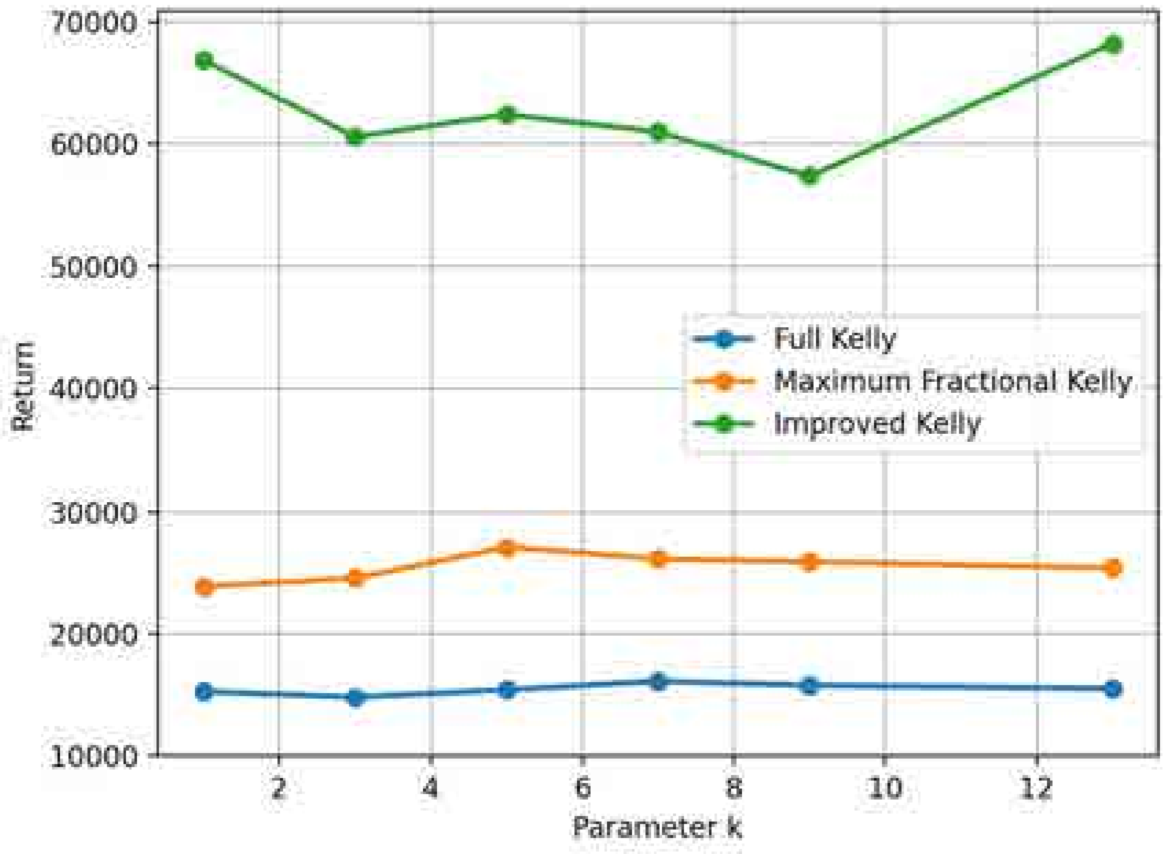
- All models are profitable



Application in Soccer – Results- By Luck? NO



Using the closing-odds-implied probabilities as estimations **IS NOT profitable**



1. Closing Model

2. Continuous Model

- **Regression-based Closing Model**
- **LSTM-based Closing Model**
- **LSTM-based Continuous Model**
- **Convolution-based Continuous Model**

Application in Horse Racing – LSTM-based Continuous Model

- Similar structure to the LSTM-based Closing Model
- Continuous Prediction can be achieved by replicating data records
- Model:
  - Give minute-by-minute predictions for period 0 min – 29 min before closing

Basic Features		Sequence of odds-implied probability P_{avg}				
lastP	Minute before closing	$P_{avg}(t_0)$	$P_{avg}(t_1)$	$P_{avg}(t_2)$	…	$P_{avg}(t_{29})$
$P_{avg}(t_0)$	0	0.1080	0.1101	0.1117	…	0.1378
$P_{avg}(t_1)$	1	-1	0.1101	0.1117	…	0.1378
$P_{avg}(t_2)$	2	-1	-1	0.1117	…	0.1378
…	…	…	…	…		…
$P_{avg}(t_{29})$	29	-1	-1	-1	…	0.1378

Records formed by a single horse

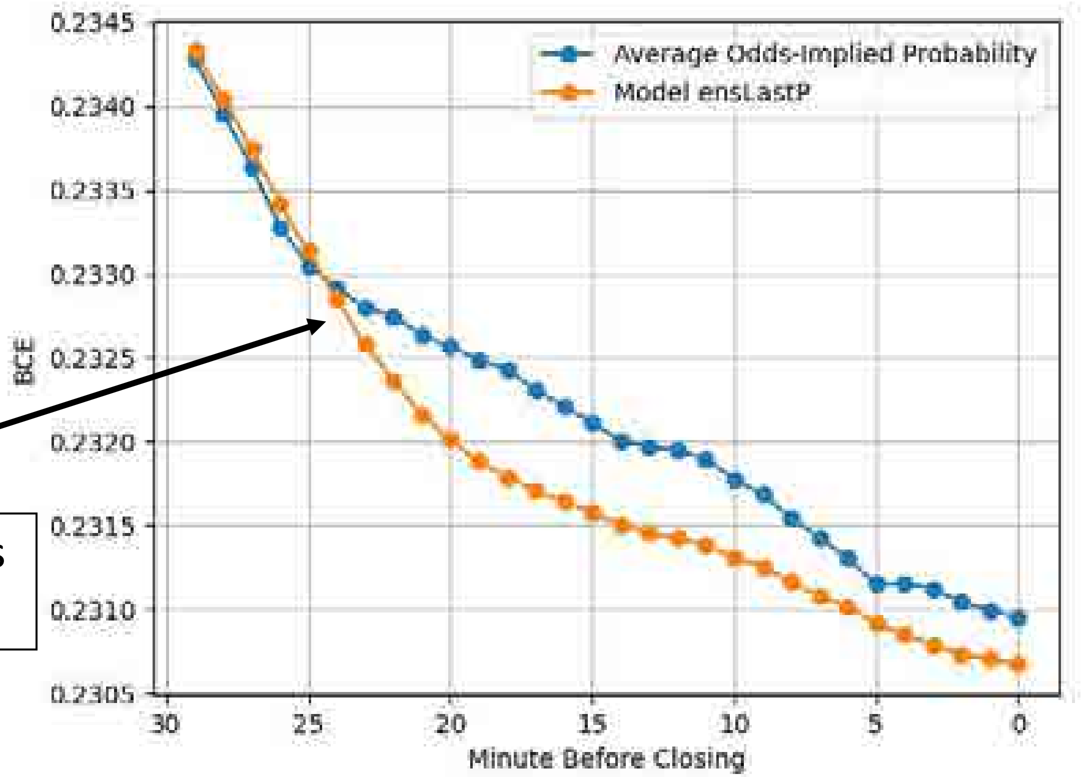
�

Application in Horse Racing – Results

- Since it is difficult to achieve optimal in continuous betting
  - We compute the Binary Cross Entropy on the testing set instead

- Binary Cross Entropy
  - How good the probability estimations are

- **Lower is Better**



Start to outperform the odds at 24min before closing

Limitation

- Data size grows with number of timesteps
  - Original: 19647 records
  - After replicating records for every horse: 19647 x 30 = 589 410
- Not suitable for long period prediction
  - For example: Soccer games
    - Bookmakers offer odds several days or even a month before kickoff

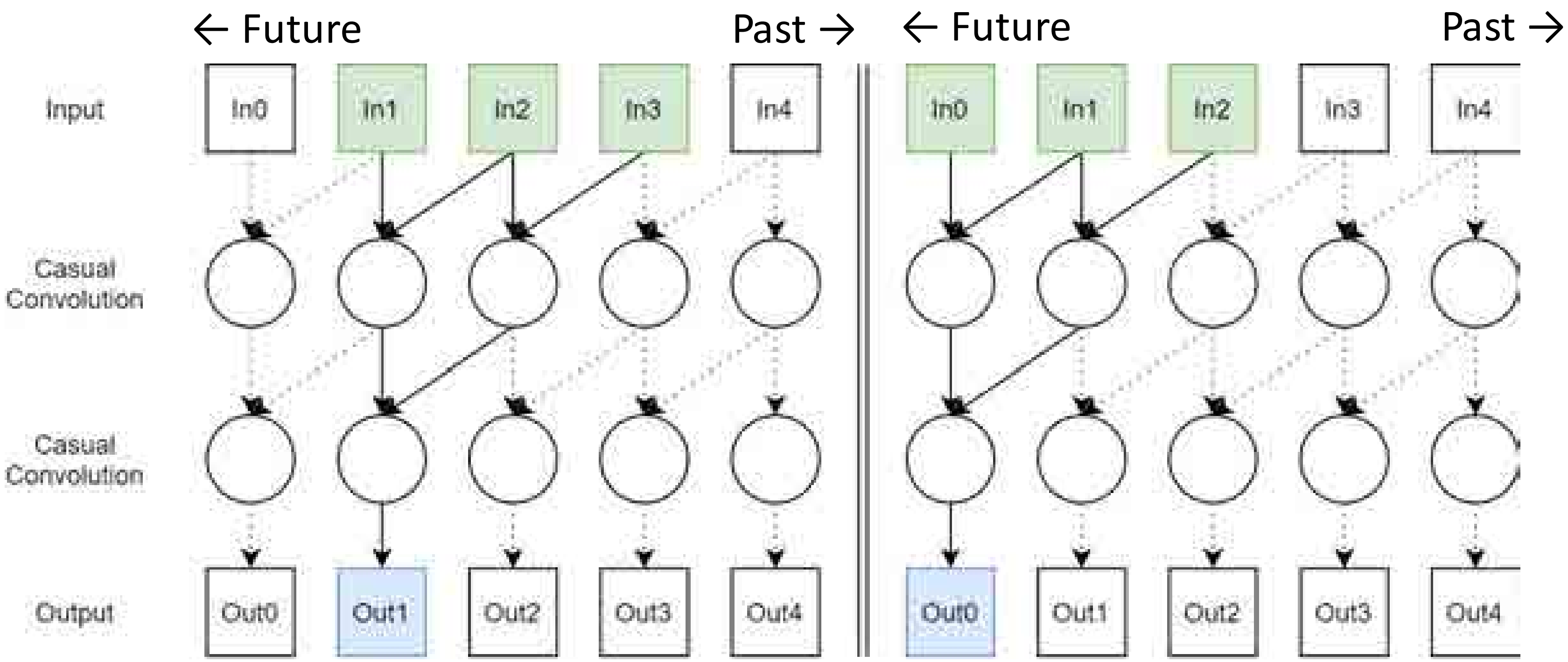
Convolution-based

- Designed for long period prediction

- Input **a sequence** of odds-implied probability
- output **a sequence** of predicted probability with time dependency preserved

Convolution-based

- Casual Convolution
  - Preserves time dependency in a Sequence
  - Example: **Window Size = 2**



Application in Soccer – Data & Model

- Over/Under (入球大細)
    - Predict Pr(Over) minute by minute in period 0 – 1439 mins before closing

Application in Soccer – Results: BCE Test

�

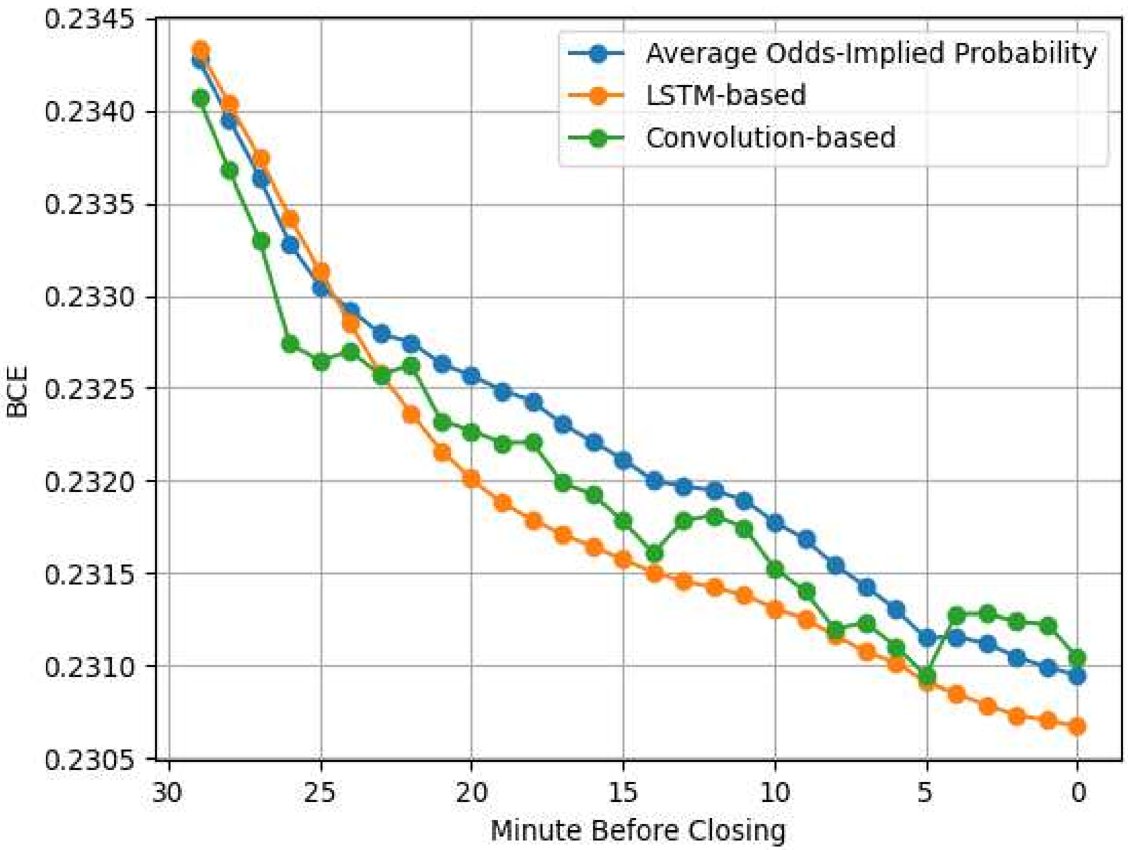
1. Regression-based Closing Model
2. LSTM-based Closing Model
3. LSTM-based Continuous Model
4. Convolution-based Continuous Model

- **Smaller window is better**

Application in Horse Racing – Results – BCE Test

- Which one is better?
  - At Closing: LSTM-Based
  - At Beginning: Convolution-Based

- Both models show they are possible to outperform the odds



Conclusion

- **We created the following models in this project**
    1. Regression-based Closing Model
    2. LSTM-based Closing Model
    3. LSTM-based Continuous Model
    4. Convolution-based Continuous Model
- **Above models are shown to be potentially profitable and able to outperform the betting odds**
- **Using to odds to beat the odds!**