SEC

SEC RESEARCH DIRECTION

EXPECTED BOND LIQUIDITY ON THAILAND BOND MARKET

SEC capital market symposium 2024

นิรันดร์ อร่ามรัตนพันธุ์
ผู้ช่วยผู้อำนวยการ
ฝ่ายกำกับธุรกิจจัดการลงทุน
สำนักงาน ก.ล.ต.

15 พฤศจิกายน 2567

Motivation of Research

The crisis transfer risks and impact to liquidity in capital market

1997

LIQUIDITY CRISIS

Thailand experienced a liquidity crisis that had severe consequences for its stock market and the broader economy.

2010

FLASH CRASH

The "Flash Crash" occurred in the US stock market, which was a sudden and severe drop in stock prices that occurred within minutes.

2019

COVID-19

Thailand experienced a liquidity crisis due to the economic impact of the COVID-19 pandemic and several mutual funds in Thailand experienced a liquidity crisis due to a sudden increase in redemption requests from investors (2019-2020)

Motivation of Research

The relationships between the Efficiency market hypothesis and the liquidity in capital market.

The price of security asset in a liquidity markets reflect all readily available information.

The securities in capital markets with highly liquidity will show significant price anomalies.

Increasing of arbitrage activity from the proposition is an improvement in liquidity which decrease frictions.

The liquidity stimulates arbitrage activity, which, in turn, enhance market efficiency.

Tetlock, P. C. (2006). "Does Liquidity Affect Securities Market Efficiency?" Working paper, Chordia 2008, Hui Guo (2017)
Chordia, R., Subrahmanyam (2008). "Liquidity and Market Efficiency." Journal of Financial Economics 87(2): 249-268

Motivation of Research

Liquidity in stock and bond is the nonlinear context of payoffs.

Considering liquidity in the stock and bond markets in a nonlinear context is essential for capturing the complexities and dynamics of financial markets that linear models might overlook. Non-linear characterization of market liquidity in the Over-the-Counter (OTC). Market conditions (Ex., total volume, number of trades..) have an asymmetric impact on the tails of the liquidity distribution.

Technology have uncovered pattern to predict further data

Machine learning and Deep learning provides an automated technique of data analysis. Specifically, is defined as a set of methods that enable to identify patterns in data, and then uncover patterns to predict the further data, or to support decision making under uncertainty.

Paolo P, Mirela S. (2024) "Speculation and Liquidity in Stock and Corporate Bond Markets" SSRN

Ruzhen Y, Ding Y, Xudong C. (2020)"Non-linear characterization and trend identification of liquidity in China's new OTC stock market based on multifractal detrended fluctuation analysis" Chaos, Solitons & Fractals Vol.139

Helena C, Stephania M, Jorge M. (2023) "Nonlinear Market Liquidity : An empirical examination" International Review of Financial Analysis Vol.87

Boumediene S, Francoise S. (2024) "Liquidity on Eurozone stock markets : A non-linear approach" Economic Bulletin Vol.44 (321-340)

Murphy, Kevin P. (2012) "Machine learning: a probabilistic perspective." MIT press

Owoc, Mieczysław L., Weichbroth, Paweł, and Żuralski, Karol (2017) "Towards better understanding of context-aware knowledge transformation", in 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), 1123-1126.

Sałabun, Wojciech, Karczmarczyk, Artur, Wątróbski, Jarosław, and Jankowski, Jarosław (2018) „Handling data uncertainty in decision making with COMET", in 2018 IEEE Symposium Series on Computational Intelligence (SSCI), 1478-1484.

Purpose of Study

TRADITIONAL STATISTICAL MODELS

Traditional statistical models **may not be able to capture the predictor and the complex relationships between the various factors that affect liquidity,** and there is a need for more sophisticated models that can provide accurate predictions.

PREDICTOR

To understand and predict liquidity from stocks and bonds' characteristic in term of risk in the Thailand Bond Market, it is Important to study "predictor" using the complex "model"

Literature review

TYPE	Subject	Authors	Proposed
Bond	Forecasting liquidity risk premia by Machine learning	Daniele Bianchi (2017)	measuring a bond risk premia with machine learning. They employ machine learning methods to revisit the debate on the presence of predictable variation in bond returns which be context of machine learning in asset pricing.
	Expected Bond Liquidity	Müller, Reichenbacher, Schuster, Uhrig-Homburg (2023)	machine learning methodology for predicting the future liquidity distribution of individual bonds in the U.S. corporate bond market and use it to compute two forward-looking illiquidity measures
	Using predictor to make liquidity	Amihud (2002)	using six measure which have common for liquidity proxies. Especially, we are using 6 measure which have common for liquidity proxies, including Bid-Ask spread, relative bid-ask spread, illiquidity ratio (AMH) Turnover ratio, Trading volume and Trading value.

What's represent "Liquidity"?

SEC

BID-ASK SPREAD*

Definition : The bid-ask spread is the difference between the highest price a buyer is willing to pay for an asset (the bid price) and the lowest price a seller is willing to accept (the ask price).

Liquidity Reflection:

Tighter Spreads: A narrow or tight bid-ask spread suggests high liquidity. This means there are many buyers and sellers, and the asset can be bought or sold with minimal price impact. In other words, you can enter or exit a position without significantly affecting the price.

Wider Spreads: A wide bid-ask spread indicates lower liquidity. Fewer buyers and sellers are active in the market, so trading can lead to more substantial price changes. This can make it more costly to enter or exit a position due to the larger difference between what you can buy and sell the asset for.

Market Depth: Liquidity is also related to market depth, which refers to the number of buy and sell orders at various price levels. A deep market with many orders at different prices generally has a tighter spread, while a shallow market with few orders can have a wider spread.

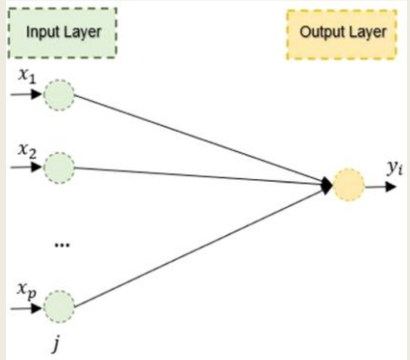
Trading Volume: Higher trading volumes usually lead to tighter spreads because the increased activity helps to balance supply and demand more effectively. Conversely, lower trading volumes can lead to wider spreads due to the lower level of trading activity.

"Options, Futures, and Other Derivatives" by John C. Hull, "Principles of Corporate Finance" by Richard A. Brealey, Stewart C. Myers, and Franklin Allen, "Liquidity and Market Depth" by Acharya, V. V., & Pedersen, L. H. (2005). Journal of Financial Economics., "The Determinants of Bid-Ask Spread in Emerging Markets" by Chen, N., & Zhang, L. (2004). Emerging Markets Review., "Financial Econometrics: Models and Methods" by Christian Gourieroux and Joann Jasiak

Methodology

Summary of model architecture

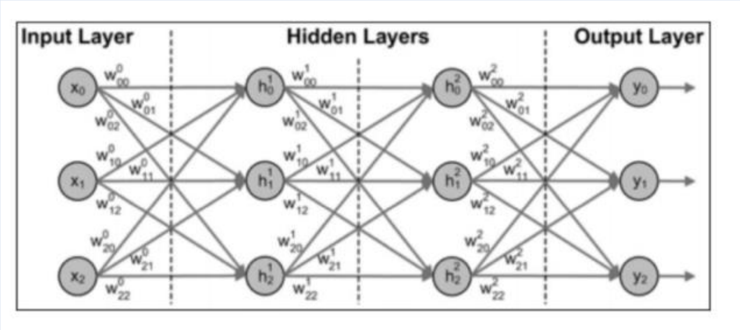
Multiple Linear Regression (MLR)



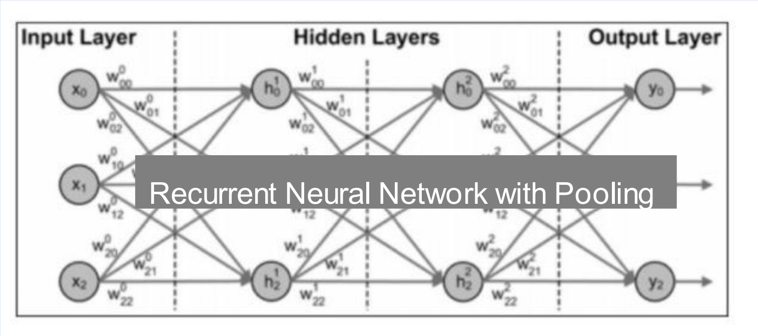
$$Yhat = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} w_0 + w_1 x_1^1 + w_2 x_1^2 + \cdots + w_D x_1^D \\ w_0 + w_1 x_2^1 + w_2 x_2^2 + \cdots + w_D x_2^D \\ \vdots \\ w_0 + w_1 x_N^1 + w_2 x_N^2 + \cdots + w_D x_N^D \end{bmatrix}$$

Machine and Deep learning

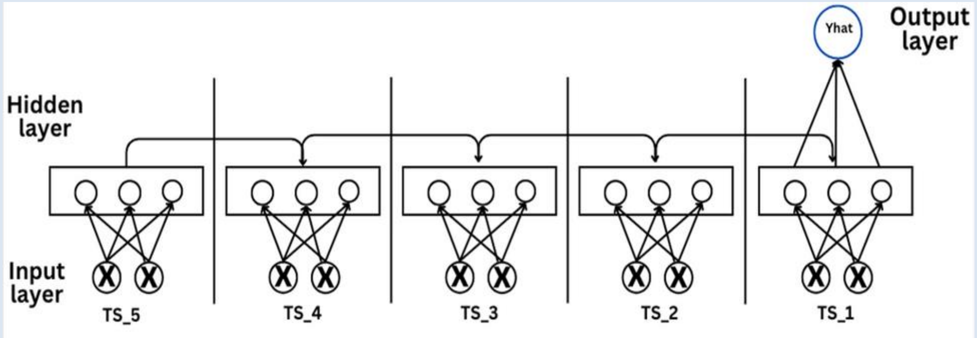
Multi-layer Perceptron (MLP)



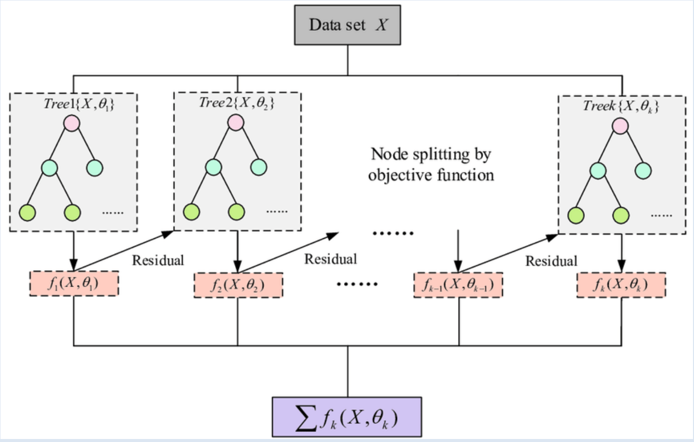
Mixed Deep Learning (MDL)



Long Short-Term Memory (LSTM)



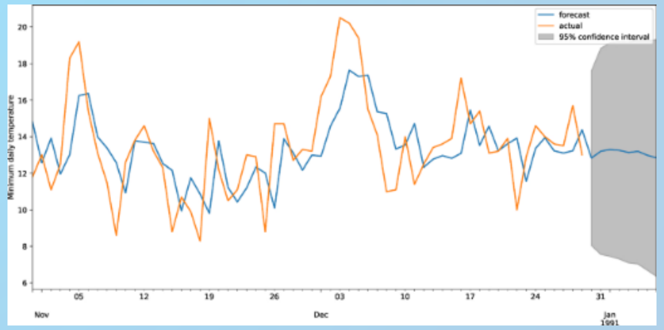
Extreme Gradient Boosting (XGB)

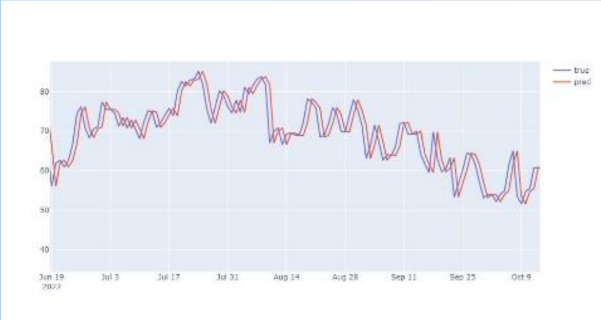


Methodology

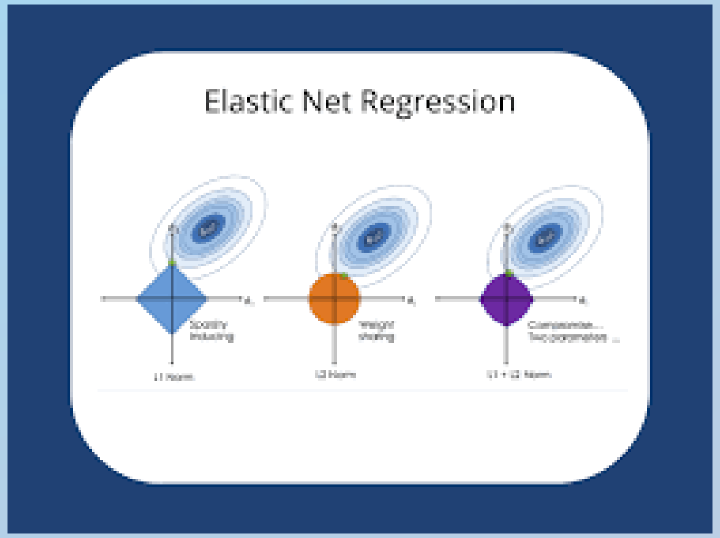
Other model for comparing the performance

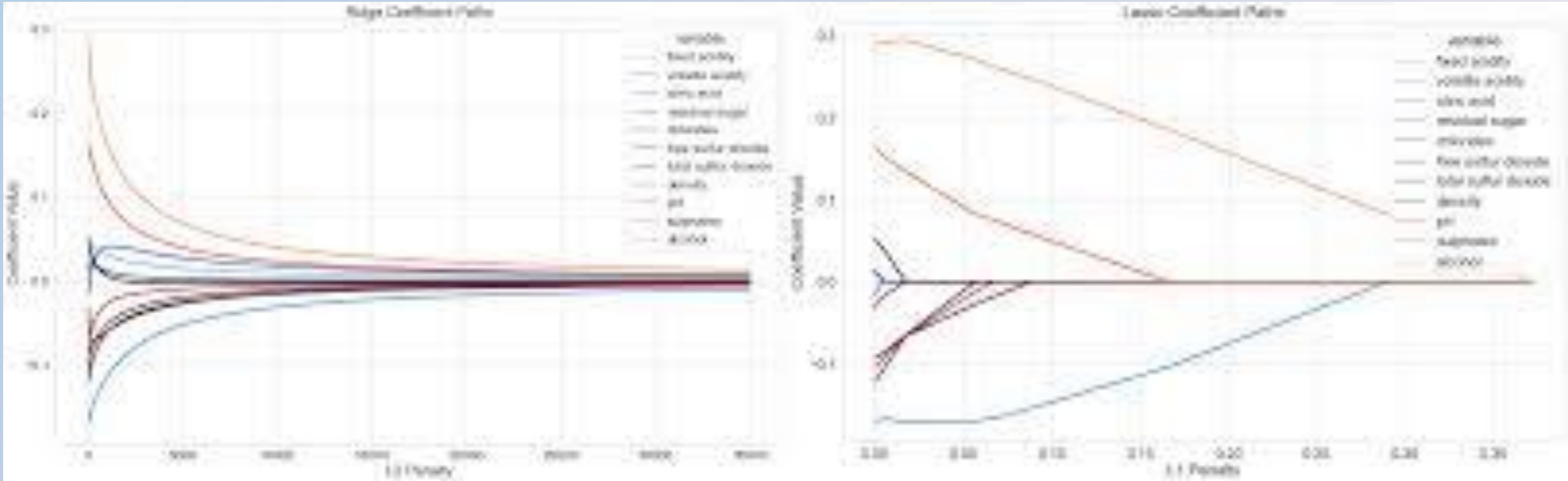
Naive Regression





Elastic net regression





Data description

Data description and preparation for corporate bonds

The daily of transaction data of corporate bonds issuing and appear in TBMA between 2015-2024. The bonds which issue by issuers cover every sector in market. The result of bond represent 1) Sectors 2) Duration 3) Credit rating which each individual of bond under each type have different liquidity.

The predicted and predictor
(i) Lifetime (TTM) (Chordia,T. 2005)
(ii) Age (Warga, Arthur)
(iii) Outstanding (Edwards 2007, Jankowitsch 2011, Bao 2011)
(iv) Credit rating (Mahanti2008, Hotchkiss 2017)
(v) Trading value (Brennan 1998)
(vi) Turnover ratio (Amihud,1997)
(vii) AMH (Amihud, 2002)
(viii) SPREAD (Amihud Y. , 1980)

Training & Testing Process (80:20)

Training : JAN 2015 – FEB 2022	Testing : MAR2022 - JAN 2024

Methodology

Model Evaluation

MEAN SQUARED ERROR (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

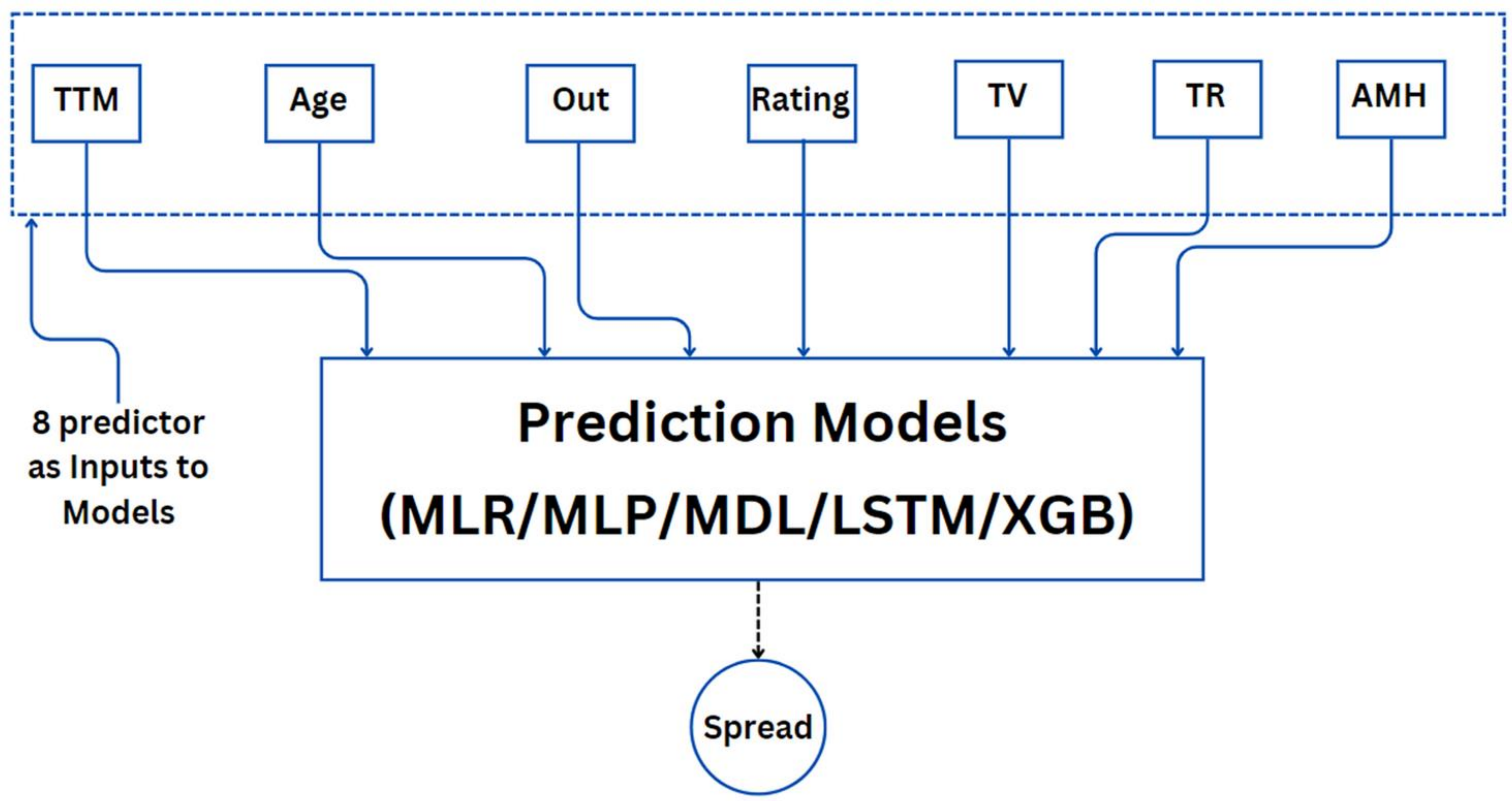
MEAN ABSOLUTE ERROR (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MEAN ABSOLUTE PERCENT ERROR

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

Structure of model



Empirical Results

SEC

compare with all model

compare with naive and enet

Summary model : Sector comparative analysis of MAPE, MAE and MSE

Models	Metrics	Performance	Argo	Bank	COMM	CONMAT	ENERGY	FINANCIAL	FOOD	HEALTH	ICT	PETRO	PKG	PROP	SA	SPV	TOURRISM	TRANS
MLR	MAPE	error test	0.25	3.00	0.32	10.57	0.91	1.09	0.72	1.60	0.59	0.51	0.24	0.78	0.00	0.27	0.19	1.56
	MAE	error test	30.51	19.66	25.52	26.49	22.33	15.13	25.23	32.71	31.67	27.99	24.55	19.19	0.00	28.24	32.27	25.01
	MSE	error test	2378.59	844.12	1689.88	1202.29	1020.46	436.67	1635.68	1188.25	2024.06	2059.92	617.33	877.56	0.00	1384.88	2468.93	1170.37
MLP	MAPE	error test	9.07E-02	1.55E+00	3.11E+13	6.66E+00	5.88E+12	7.16E-01	2.38E+13	3.81E+00	1.29E+00	3.47E+13	1.05E-01	4.37E+12	1.61E+14	2.49E-01	1.18E+14	7.47E-01
	MAE	error test	1.29E+01	2.85E+01	2.47E+01	2.98E+01	2.24E+01	1.16E+01	2.68E+01	8.12E+01	8.56E+01	2.06E+01	1.09E+01	1.46E+01	3.57E-02	2.60E+01	3.43E+01	2.89E+01
	MSE	error test	2.68E+02	2.08E+03	1.88E+03	3.21E+03	1.38E+03	2.47E+02	3.26E+03	8.57E+03	3.57E+04	1.32E+03	1.30E+02	7.12E+02	1.31E-03	7.53E+02	2.81E+03	3.34E+03
MDL	MAPE	error test	3.50E-01	3.74E+00	8.59E-01	1.81E+01	1.46E+00	2.72E+00	1.08E+00	1.31E+00	1.04E+00	7.06E-01	1.80E-01	1.44E+00	0.00E+00	2.52E-01	1.12E-01	3.23E+00
	MAE	error test	5.20E+01	3.53E+01	3.64E+01	4.72E+01	2.69E+01	4.53E+01	3.66E+01	2.47E+01	6.58E+01	3.81E+01	1.86E+01	2.82E+01	0.00E+00	2.60E+01	1.93E+01	5.68E+01
	MSE	error test	5.98E+03	2.90E+03	3.69E+03	3.25E+03	1.46E+03	6.16E+03	2.44E+03	8.25E+02	1.08E+04	6.63E+03	5.13E+02	1.25E+03	0.00E+00	1.45E+03	9.18E+02	4.69E+03
LSTM	MAPE	error test	8.48E-01	3.07E+00	2.53E+13	1.21E+01	9.78E+12	8.13E-01	2.25E+13	2.86E-01	7.14E-01	5.12E+13	5.23E-01	1.14E+13	7.96E+13	9.06E-01	8.76E+13	9.41E-01
	MAE	error test	1.12E+02	5.17E+01	6.20E+01	4.11E+01	3.86E+01	1.10E+01	5.01E+01	7.36E+00	5.50E+01	2.77E+01	5.42E+01	3.24E+01	1.77E-02	9.03E+01	6.09E+01	4.85E+01
	MSE	error test	1.32E+04	7.22E+03	8.53E+03	4.91E+03	3.00E+03	1.67E+02	6.60E+03	7.91E+01	6.53E+03	2.32E+03	2.94E+03	3.97E+03	4.26E-04	8.36E+03	7.65E+03	7.69E+03
XGB	MAPE	error test	7.25E-02	1.37E+00	2.30E-01	9.35E+00	7.71E-01	7.07E-01	3.69E-01	1.09E+00	2.98E-01	1.15E-01	6.92E-06	4.63E-01	0.00E+00	8.66E-02	1.24E-01	7.94E-01
	MAE	error test	1.01E+01	1.55E+01	1.51E+01	1.79E+01	1.54E+01	8.22E+00	1.28E+01	2.23E+01	1.57E+01	7.24E+00	7.17E-04	8.60E+00	0.00E+00	9.59E+00	2.09E+01	1.20E+01
	MSE	error test	2.13E+02	7.50E+02	9.88E+02	5.18E+02	5.76E+02	1.22E+02	3.53E+02	8.74E+02	5.53E+02	1.42E+02	6.79E-06	1.42E+02	0.00E+00	1.88E+02	1.13E+03	2.61E+02
NAVIE	MAPE	error test	1.55E-01	3.26E+00	7.55E-01	2.24E+01	1.71E+00	2.87E+00	1.22E+00	2.31E+00	1.19E+00	4.25E-01	5.15E-02	1.51E+00	0.00E+00	1.97E-01	6.92E-02	3.45E+00
	MAE	error test	1.91E+01	2.60E+01	2.72E+01	5.19E+01	2.66E+01	4.48E+01	3.16E+01	4.81E+01	5.29E+01	2.41E+01	5.33E+00	2.60E+01	0.00E+00	2.13E+01	1.16E+01	4.73E+01
	MSE	error test	8.29E+02	1.35E+03	1.80E+03	4.01E+03	1.13E+03	3.92E+03	1.46E+03	2.43E+03	5.82E+03	1.21E+03	2.84E+01	9.28E+02	0.00E+00	5.17E+02	3.77E+02	2.99E+03
ENET	MAPE	error test	1.80E-01	2.76E+00	3.83E-01	7.83E+00	8.13E-01	1.44E+00	8.44E-01	1.90E+00	6.55E-01	1.95E-01	1.29E-01	8.25E-01	0.00E+00	2.99E-01	1.39E-01	1.80E+00
	MAE	error test	2.18E+01	1.79E+01	2.49E+01	2.65E+01	1.67E+01	2.24E+01	2.25E+01	3.86E+01	3.52E+01	1.40E+01	1.33E+01	1.59E+01	0.00E+00	3.20E+01	2.33E+01	2.77E+01
	MSE	error test	1.24E+03	8.31E+02	1.77E+03	1.10E+03	5.11E+02	8.51E+02	9.27E+02	1.62E+03	2.76E+03	5.46E+02	1.83E+02	4.41E+02	0.00E+00	1.14E+03	1.42E+03	1.21E+03

Empirical Results

Summary the result of machine learning and deep learning model classify by Sector

- **Sector-specific insights,**
  - Argo, XGB and MLP perform well in this sector, indicating manageable volatility and relatively predictable patterns.
  - Bank, XGB and MLP shows strong performance
  - COMM, XGB and MLR provide good accuracy
  - CONMAT, MLP and XGB capture sector-specific complexities effectively
  - ENERGY, XGB models perform well, suggesting the efficacy
  - Financial, XGB and MLP robustness to outliers and Elastic Net's balance of linear and regularization.
  - FOOD, XGB predictions are accurate, indicating a relatively predictable sector.
  - HEALTH, LSTM and XGB shows better predictability
  - ICT, XGB and MLR showing predictable patterns and benefiting from ensemble and regularized methods.
  - PETRO, XGB low prediction errors indicating less volatility or more consistent patterns.
  - PKG, XGB complex models perform well, indicating a stable and predictable sector.
  - PROP, extremely low prediction errors suggest very predictable patterns, especially for XGB
  - SA, highly predictable sector, benefiting from all modeling approaches
  - SPV, Low errors with XGB and MDL, indicating the advantage of advanced ensemble methods in this sector
  - TOURISM, MDL and XGB performance, suggesting stable patterns within the tourism sector.
  - TRANS, MLP and XGB shows lower errors, highlighting its robustness in capturing transportation sector complexities

- **Model selection,** <u>XGB is preferred for its robust performance and lower errors and generally performs</u> the best in terms of lowest mean and min errors across all metrics and phases including appears to be the most reliable model with the lowest errors.
- **General Insights,** <u>XGB, dominates in most sectors with the lowest MAPE MAE and MSE,</u> indicating its strong performance across diverse and complex datasets.

Empirical Results

Summary model : Duration comparative analysis of MAPE, MAE and MSE

Models	Metrics	Performance	1-5	5-10	10-15	15-20
MLR	MAPE	error test	2.96E+00	1.06E+00	5.17E-01	5.90E-01
	MAE	error test	1.95E+01	2.79E+01	4.32E+01	2.17E+01
	MSE	error test	8.69E+02	1.63E+03	2.78E+03	8.94E+02
MLP	MAPE	error test	2.77E+12	1.10E+00	5.94E-01	4.25E-01
	MAE	error test	1.65E+01	3.42E+01	5.83E+01	2.88E+01
	MSE	error test	1.04E+03	6.63E+03	8.39E+03	1.31E+03
MDL	MAPE	error test	4.99E+00	1.83E+00	7.10E-01	4.46E-01
	MAE	error test	3.21E+01	4.46E+01	5.82E+01	3.50E+01
	MSE	error test	2.23E+03	4.41E+03	5.65E+03	2.82E+03
LSTM	MAPE	error test	1.37E+12	9.57E-01	9.18E-01	7.93E-01
	MAE	error test	2.79E+01	4.20E+01	1.26E+02	8.96E+01
	MSE	error test	2.71E+03	4.69E+03	1.83E+04	9.71E+03
XGB	MAPE	error test	2.18E+00	6.64E-01	4.21E-01	4.21E-01
	MAE	error test	9.90E+00	1.58E+01	2.40E+01	1.90E+01
	MSE	error test	2.19E+02	6.07E+02	1.27E+03	6.64E+02
NAVIE	MAPE	error test	5.64E+00	2.13E+00	5.43E-01	3.78E-01
	MAE	error test	3.17E+01	4.10E+01	3.54E+01	1.94E+01
	MSE	error test	1.81E+03	2.73E+03	1.97E+03	6.03E+02
ENET	MAPE	error test	2.55E+00	1.08E+00	4.25E-01	4.95E-01
	MAE	error test	1.84E+01	2.62E+01	3.40E+01	2.02E+01
	MSE	error test	6.47E+02	1.44E+03	1.72E+03	6.47E+02

Empirical Results

Summary the result of machine learning and deep learning model classify by duration

- **Duration-specific insights,**
  - 1-5 years, XGB stands out with the lowest MAPE, MAE, and MSE values, indicating superior performance for short-duration predictions.
  - 5-10 years, XGB maintains its strong performance with low MAPE, MAE, and MSE values.
  - 10-15 years, XGB remains the best performer with consistently low error metrics.
  - 15-20 years, XGB continues to excel with low MAPE, MAE, and MSE.

- **Overall conclusion**, Short Durations, XGB is the most accurate, followed by ENET and MLR. MLP and LSTM have notable accuracy issues. Medium Durations, XGB continues to lead in accuracy. ENET and MLR show stable performance, whereas MLP and LSTM struggle with increasing duration. Long Durations, XGB remains the most reliable, but MLP and LSTM perform poorly, showing a significant drop in accuracy.

- **General Insights,** XGB, dominates in most sectors with the lowest MAPE MAE and MSE, indicating its strong performance across diverse and complex datasets.

Empirical Results

Summary model : Credit rating comparative analysis of MAPE, MAE and MSE

Models	Metrics	Performance	AAA	AA	A	BBB
MLR	MAPE	error test	2.28E+00	8.93E-01	1.95E+00	1.48E-01
	MAE	error test	1.73E+01	2.90E+01	2.42E+01	2.40E+01
	MSE	error test	5.78E+02	1.72E+03	1.36E+03	1.43E+03
MLP	MAPE	error test	1.39E+00	7.80E-01	8.73E+12	7.25E+13
	MAE	error test	1.91E+01	3.13E+01	2.42E+01	5.15E+01
	MSE	error test	8.11E+02	2.49E+03	2.37E+03	2.13E+04
MDL	MAPE	error test	3.24E+00	1.52E+00	3.49E+00	3.55E-01
	MAE	error test	2.27E+01	4.54E+01	3.74E+01	5.14E+01
	MSE	error test	8.35E+02	3.61E+03	3.21E+03	8.09E+03
LSTM	MAPE	error test	2.41E+00	7.72E-01	1.08E+13	6.00E+13
	MAE	error test	2.59E+01	5.93E+01	4.31E+01	7.57E+01
	MSE	error test	1.61E+03	7.38E+03	5.02E+03	1.16E+04
XGB	MAPE	error test	1.30E+00	5.60E-01	1.53E+00	7.47E-02
	MAE	error test	1.30E+01	1.81E+01	1.25E+01	9.82E+00
	MSE	error test	3.51E+02	7.63E+02	4.54E+02	3.87E+02
NAVIE	MAPE	error test	3.19E+00	1.66E+00	4.10E+00	2.97E-01
	MAE	error test	2.10E+01	3.92E+01	3.25E+01	4.13E+01
	MSE	error test	6.27E+02	2.04E+03	1.84E+03	5.20E+03
ENET	MAPE	error test	2.28E+00	8.37E-01	1.68E+00	1.77E-01
	MAE	error test	1.62E+01	2.47E+01	2.20E+01	2.64E+01
	MSE	error test	4.21E+02	1.09E+03	1.00E+03	2.04E+03

Empirical Results

Summary the result of machine learning and deep learning model classify by Credit rating

- **Credit rating-specific insights,**
  - Higher Credit Ratings (e.g., AAA, AA), XGB's lower error metrics across these higher ratings suggest that it is proficient at predicting high creditworthiness with accuracy.
  - Lower Credit Ratings (e.g., A, BBB), XGB’s consistent low error metrics for lower ratings suggest it is effective at differentiating between lower creditworthiness levels.

- **Overall conclusions**, XGB shows robust performance across different credit rating indices, making it the preferred model for comprehensive credit rating predictions. MLP and LSTM might require additional refinement to improve consistency across indices, while MDL and NAVIE may not be as effective for detailed credit rating tasks.

- **General Insights,** XGB, dominates in most sectors with the lowest MAPE MAE and MSE, indicating its strong performance across diverse and complex datasets.

Empirical Results

Output parameter of model : Average bias and weight of the equation MLR model for predicting the spread.

$$\hat{y} = w_0 + w_1x_1 + w_2x_2 + \ldots\ldots + w_px_p$$

Weight output / coefficient of Variables	Average
Bias (interception (w_0))	80
AMH (w_1)	-0.64
Credit rating (w_2)	-0.50
Issue size (w_3)	0
Issue term (w_4)	0
Trading value (w_5)	147.57
Term to maturity (w_6)	7.80
Turnover ratio (w_7)	147.57
Volume (w_8)	295.83

Empirical Results

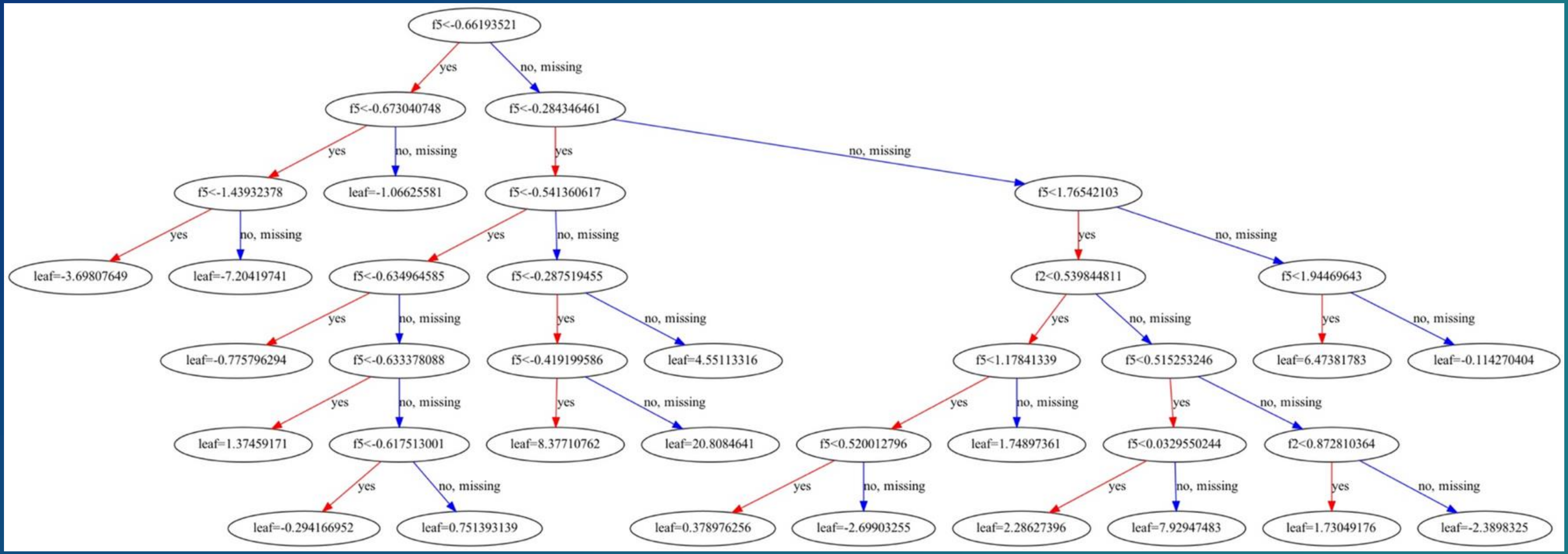
Output parameter of model : Average bias and weight of the equation MLP model for predicting the spread.

$$Z^{[1]} = XW^{[1]} + b^{[1]} \qquad A^{[1]} = ReLU\left(Z^{[1]}\right) \qquad Z^{[3]} = A^{[2]}W^{[3]} + b^{[3]} \qquad \hat{y} = A^{[3]}W^{[out]} + b^{[out]}$$

Weight output / coefficient of Variables	MLP	MDL	LSTM
Input layer Hidden layer 1 $(W^{[1]})$	-0.0004125	-0.000353906	0.535931719
Hidden layer 1 (BIAS) $(b^{[1]})$	0.168631875	0.0035125	0.000346582
Hidden layer 2 $(W^{[2]})$	0.027387083	0.016515	0.358102271
Hidden layer 2 (BIAS) $(b^{[2]})$	0.016515	0.003870792	-0.008585352
Hidden layer 3 $(W^{[3]})$	0.050803125	0.028845479	0.076608617
Hidden layer 3 (BIAS) $(b^{[3]})$	0.063284688	0.013106875	0.025435506
Output layer $(W^{[out]})$	0.119960313	0.02665107	0.138594431
Output layer (BIAS) $(b^{[out]})$	0.32075	0.0358125	0.23485

Empirical Results

Output parameter of model : XGB the output show in decision tree which represent the regression for predicting the spread.



Discussion

Comparative Analysis of Difference Models

- The methods proposed in this paper (MDL/LSTM/XGB) attained a lower Mean Squared Error (MSE) compared to the others approach.

Authors	Years	Models	MSE (%)
Marcel Müller, Michael Reichenbacher, Philipp Schuster, Marliese Uhrig-Homburg Invalid source specified.	2023	Extreme Gradient Boosting	0.604
this research	2024	Extreme Gradient Boosting	0.0288
		Multi-layer perceptron	0.3602
		Mixed Deep Learning	2.5246
		Long-term Short-term Memory	29.3

Conclusion and Policy implication

- **Importance of Liquidity:** Liquidity enables easier trade execution and reduces costs; illiquidity increases risk, requiring a liquidity premium.

- **Key Contribution:** Empirical validation of model structure, addressing liquidity's role in investment decisions and cost of capital reduction.

- **Dynamic Market Environment:** The research underscores the importance of adapting to dynamic market conditions by focusing on liquidity metrics such as trading volume, turnover ratio, and spread to mitigate liquidity risk.

- **Key Liquidity Proxies for Bonds:** Volume, turnover ratio, Amihud illiquidity, issue size, term to maturity, credit rating, and spread.

- **Impact of Bid-Ask Spread:** The bid-ask spread was identified as a critical factor influencing both liquidity and the cost of trading, offering important insights for market participants.

- **Predictive Model for Capital Market in Thailand:** Developed for forecasting individual bond liquidity using deep learning and machine learning techniques.

- **Deep Learning Models Applied:** XGB and LSTM models provided the most accurate liquidity predictions (lowest MAPE, MAE, MSE).

- **Predictive Accuracy:** The model's predictive accuracy was confirmed with deep learning techniques (XGB and LSTM), highlighting the potential of advanced AI methods in financial forecasting.

- **Practical Applications:** The model supports decision-making in bond markets and portfolio management based on liquidity forecasts.

Conclusion and Policy implication

- **Advanced Models for Liquidity Prediction:**
  - Liquidity prediction in capital markets requires advanced models like LSTM and XGB.
  - These models provide more accurate forecasts, especially in high-liquidity risk sectors.
  - Prioritizing such models aids in better liquidity management and financial stability.
- **Importance for Policymakers:**
  - Accurate predictive models are essential for informed decision-making in capital markets.
  - Policymakers should focus on enhancing the accuracy of these models for better economic outcomes.
- **Integration of Predictive Analytics:**
  - Advanced predictive models offer opportunities to enhance economic stability and inform policy decisions.
  - Sector-specific interventions, improved data quality, and regulatory support reduce liquidity risks.
- **Using Predictive Models for Capital Market Regulation:**
  - Predictive models transform liquidity risk management, enhancing market stability and investor confidence.
  - Advanced machine learning techniques like LSTM and XGB help predict liquidity risks and prevent disruptions.
- **Enhanced Market Surveillance:**
  - Utilize predictive models to monitor liquidity and detect early warning signs of market stress.
  - Enable timely interventions to prevent market crises.
- **Risk Management Frameworks:**
  - Tailored regulatory measures should target sectors with high liquidity risks (e.g., Financials, Industrials, Services).
  - Use predictive insights in stress testing to assess financial resilience under adverse conditions.

Conclusion and Policy implication

- **Optimized Regulatory Policies:**
  - Leverage predictive models to ensure regulations are data-driven and dynamically adaptable.
  - Flexible policies allow for responsive regulation based on predictive analysis.

- **Improved Transparency and Confidence:**
  - Promote transparency by communicating the use of predictive models in regulatory processes.
  - Strengthen investor protection by mitigating liquidity risks through predictive analytics.

- **Strategic Actions for Regulatory Implementation:**
  - Encourage adoption of predictive models within regulatory agencies.
  - Develop sector-specific regulations to address unique liquidity risks and market vulnerabilities.

- **Impact on Market Stability and Investor Protection:**
  - Predictive models enable better surveillance, risk management, and policymaking.
  - This approach supports a resilient, transparent financial system and promotes sustainable market growth

Key Takeaways

- **Portfolio Development:** Liquidity predictions can be integrated into portfolio management strategies, enhancing decision support systems for both institutional and retail investors.

- **Implications for Policy and Regulation:** The findings suggest that enhancing market liquidity could have broader implications for regulatory policies, potentially leading to a more robust and efficient capital market structure in Thailand.

SEC capital market symposium 2024

Literature review

TYPE	Subject	Authors	Proposed
Stock	Prediction liquidity by Machine learning	Michael ReichenbAcher, Philip Schuster, Marliese Uhrig-Homburg (2020)	prediction models to estimate the bond liquidity by machine learning with the transparency of a simple linear model.
	Forecasting liquidity risk by Deep learning Model	Pham, Marcin, Katarzyna, Artur and Wieslawa (2020)	Deep learning forecast liquidity Vietnamese stock market which be the most modern technique to forecast the future direction and developed a predictive model to forecast liquidity.
	Forecasting signal of risk premium by Machine learning	Sebastiao (2017)	Machine Learning model to build forecasts of the signal of the risk premium which they used combinations of ML model like Regression Trees, Random Forests and Support Vector Machine, using financial information set

Research GAP

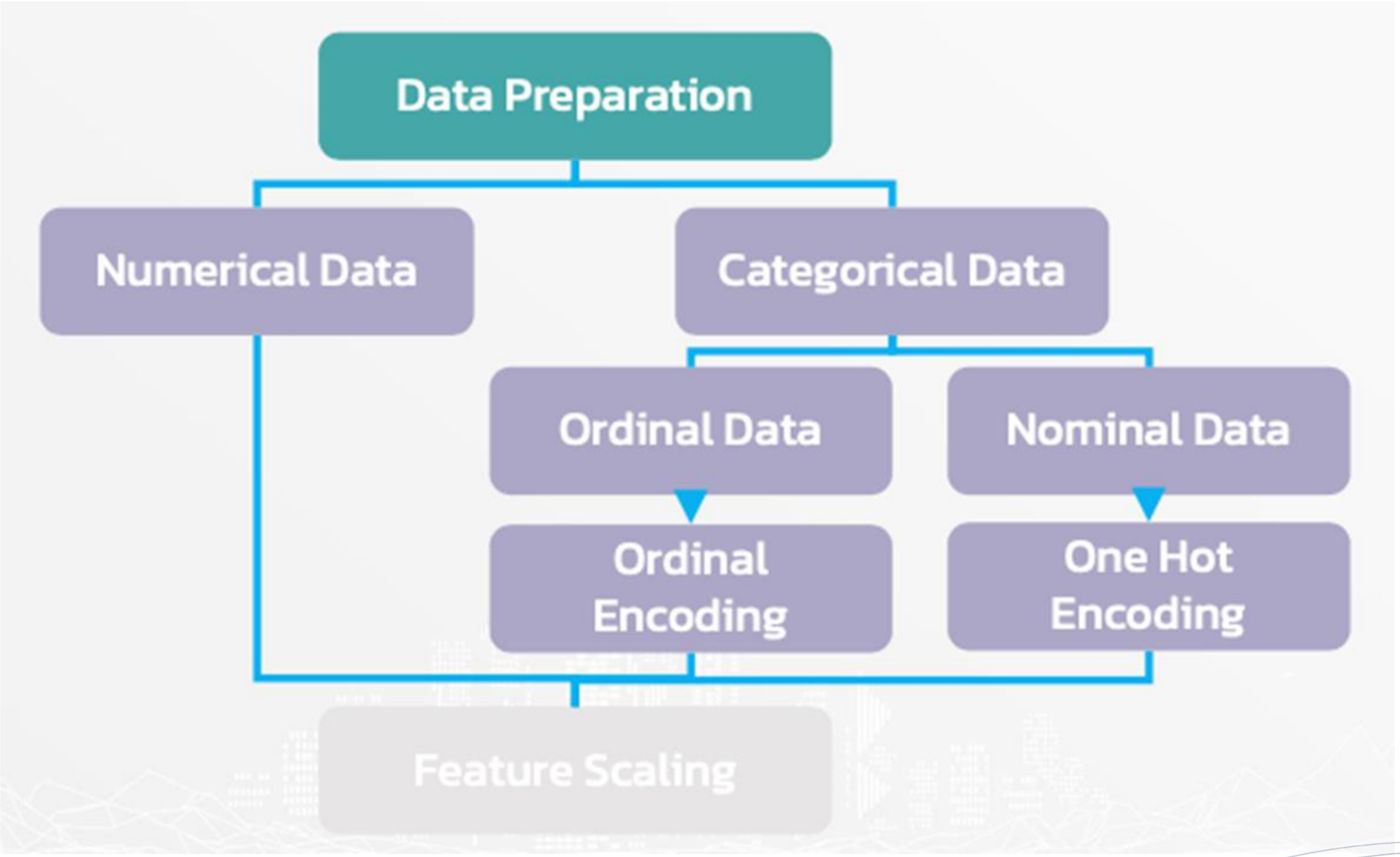
I. Liquidity is a complex concept, and defining it clearly presents a significant challenge.

II. Liquidity is defining the related determining factors and formulating an appropriate functional form to approximate and predict its value is a difficult and intricate task.

III. No paper that applies quantitative models and machine learning models to predict liquidity and liquidity premium in the Thai capital market (both equity and bond markets).

Methodology

Data preparation process



All of data stock are Numerical Data which can process to feature scaling by doesn't convert encoding

Methodology

Feature Scaling

I.Check and manage the outlier of variables of data

II.ChII.Rescaling the numerical data of variables to a common scale. This is done to eliminate any inherent differences in the scales of the variables, which could otherwise result in biased models.

III.Min-max normalization scales the data to
a range between 0 and 1, where the minimum value of the data
is mapped to 0 and the maximum value is mapped to 1. The formula for min-max normalization is:

$$Xnorm = \frac{(X - Xmin)}{(Xmax - Xmin)}$$

Data description

Variables	Definition
Illiquidity Ratio (AMH)	Occurs to using price which sensitivity to gauge the impact or influence of order flow on price.
Credit rating	Evaluation of the creditworthiness of an individual, business, or government, typically expressed as a letter grade.
Issue size	Total value or amount of securities, such as bonds, shares, or other financial instruments, that are being offered for sale in a particular issuance.
Issue term	length of time for which a financial instrument, such as a bond, loan, or other debt security, is issued before it matures and is due for repayment.
Trading value	Quantity of trade value in the time of a specified period
Term to maturity	remaining time period until a financial instrument, such as a bond, loan, or other debt security, reaches its maturity date, at which point the principal amount is due to be repaid to the investors or lenders.
Turnover ratio	Total trading values of a stock over a specified timeframe divided by market capitalization.
Volume	Represent the number of trades

Data description

Variables of Model

Asset	Type	Data	Variables
Bonds	Independent (predictor)	Transaction	AMH, Credit rating, Issue size, Issue term, Trading value, Term to maturity, Turnover ratio, Volume
	Dependent (predicted)	Transaction	Spread

Methodology

EXtreme Gradient Boosting (XGB)

State model

$$\hat{y}_i = \sum_j \theta_j x_{i,j},$$

Objective Function : Training Loss + Regularization

$$obj(\theta) = L(\theta) + \Omega(\theta)$$

$$L(\theta) = \sum_i (y_i - \hat{y}_i)^2$$

Decision Tree Ensembles : CART

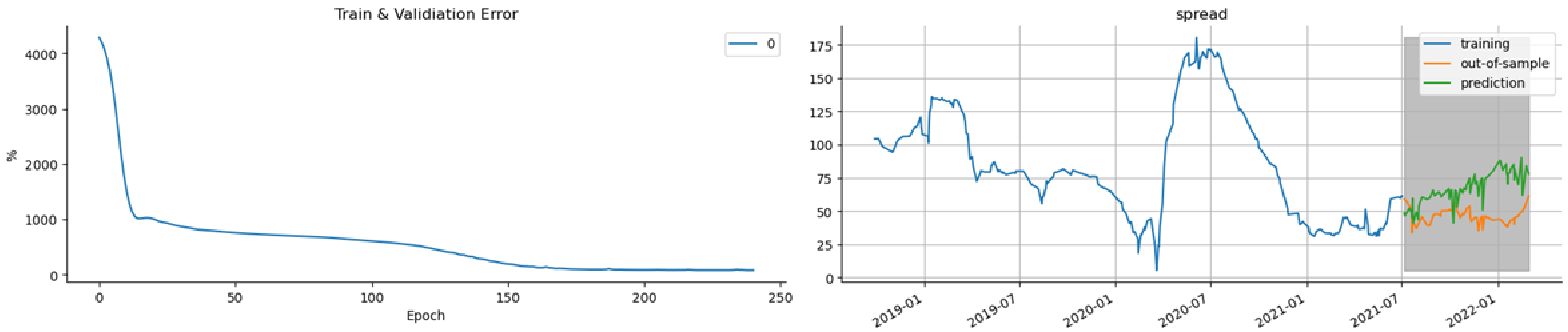
$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathrm{F}$$

$$obj(\theta) = \sum_i^n l(y_i, \hat{y}_i) + \sum_{i=1}^t \omega(f_i)$$

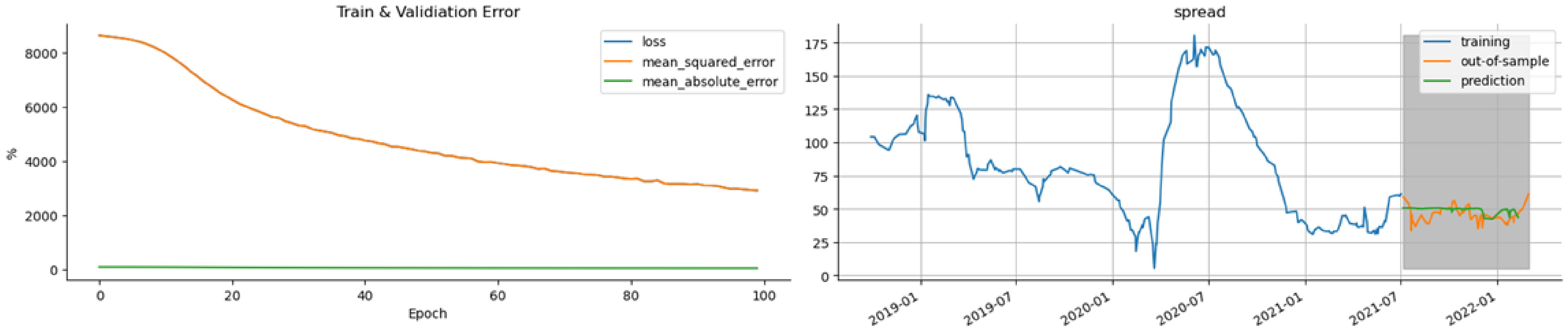
Empirical Results

Best value of performance metric in MLP, MDL and LSTM model has been achieved in approximal 100 epochs.

TBEV223A : MLP - Results | Test MAE = 21.341



TBEV223A : LSTM - Results | Test MAE = 5.713



TBEV223A : LSTM - Results | Test MAE = 5.713

Structure of model

Input variables

$$X = X^{TTM}, X^{Age}, X^{Out}, X^{rating}, X^{TV}, X^{TR}, X^{AMH}$$

Input layer (for deep learning models)

$$y = dense\left(X^{TTM}, X^{Age}, X^{Out}, X^{rating}, X^{TV}, X^{TR}, X^{AMH}, neurons, relu\right)$$

Output variables (SPREAD) (for deep learning models)

$$y(output) = f\left(y^{1}, y^{2}, y^{3}, y^{4}, \ldots.., y^{L}\right)$$