Finance BSc Thesis: The Impact of Political Indicators on Asset Pricing Predictions: An LSTM-based approach for ASEAN Markets

A Thesis Presented to the Tilburg School of Economics and Management,

Department of International Business Administration,

Tilburg University

By

Michaïl Eric Maréchal

U627441/SNR 2081750

m.e.marechal@tilburguniversity.edu

24th May 2024

Supervisor: M. Stam

m.h.a.stam@tilburguniversity.edu

Number of Words: 7797

Abstract

The present study aims to assess the significance of Political Indicators in predicting the returns of ASEAN major Indices. To achieve this, a Long Short Term Memory (LSTM) model is constructed to predict the index returns of major ASEAN countries (Indonesia, Singapore, Malaysia, Philippines and Thailand). The predictors used include political and economic factors, as well as lagged returns from 1 to 3 periods. Two models are then executed, one including and one excluding political indicators. The results indicate that the model including political factors performs better across all three metrics and suggests that political indicators are important in predicting equity market returns in Southeast Asia.

Permutation Feature Importance is used to compute the contribution of each political indicator towards the outcome. The results show that Political Stability, Control of Corruption and Government Effectiveness have a significant contribution to the prediction model. The procedure is repeated for each country to observe if there are any differences in which factors display significance and to what extent. To evaluate the significance of the results, an ANOVA test was performed and concluded that Thailand and Indonesia showed the most sensitivity to Political Indicators. In Thailand's case, it is a vicious cycle of military coups, political corruption and an inconsistent legal framework that contributed to this situation. For Indonesia, political violence at the start of the century coupled with corruption has had detrimental effects, but recent political actions are bringing the country on the right path.

Table of Contents

I. Introduction ... 4

II. Literature Review ... 7

III. Data and Methodology ... 11

1. Data ... 11

2. Methodology ... 13

2.1 Processing data ... 13

2.2 Model: LSTM (Long Short-Term Memory) ... 13

2.3 Performance Evaluation ... 16

2.4 Permutation Feature Importance ... 17

IV. Results ... 18

1. Political Indicators Significance ... 18

2. Overall Feature Importance of Political Indicators ... 18

2.2 Variation Across Countries ... 20

2.2.1 ANOVA Results ... 22

V. Conclusion ... 25

VI. References ... 27

1. Academic: ... 27

2. Databases: ... 29

3. Other: ... 30

VII. Appendix: ... 31

I. Introduction

Asset pricing is the area of financial research which explores the factors that determine the prices and returns of financial assets. Some of the most major contributions were made during the second half of the 20th century. Among them is the famous Capital Asset Pricing Model (CAPM), resulting from the combined works of Treynor (1961), Sharpe (1964), Lintner (1965) and Mossin (1966). Its innovativeness was in explaining equity returns as exposure to market risk, expressed as Beta.

The creation of stock databases and the growth in computational power made it possible to test empirically the CAPM model (Dimson & Mussavian, 1999). Based on empirical evidence Fama & French (1992) came up with two additional factors. The three factors model offered a better explanation of stock returns and errors of the CAPM model. It was later completed by the five factors model introduced by Fama & French (2015), identifying two additional factors.

However, Gu et al. (2020) point out an important limitation of linear models, which is their linearity and, as a consequence, the impossibility of capturing non-linear relationships. The further development of computational capabilities, coupled with more comprehensive data availability enabled the use of more complex statistical models, known as Machine Learning, over the last two decades. Those models and algorithms allow to handle more data, more flexibility and most importantly to capture non-linear relationships between variables. The use of regression trees and neural networks significantly contributed to increase the predictivity of stock returns and the identification of new hidden factors.

The growing interest in Emerging markets in the context of Asset Pricing has been fuelled by numerous reasons including the search for diversification, availability of data and the unique characteristics of emerging markets (volatility, liquidity issues, institutions and political environment). A particularly representative context for that is ASEAN. It is a political and economic union of Southeast Asian nations. The region is one of the most populous in the world with 671 million inhabitants[1] and has a significant growth potential. This can be illustrated by the fact that the region achieved a compounded annual growth rate of 4,8% from 2013 to 2023 compared to 1,9% for the European Union.[2] Improvement in life quality has been a driving factor for growth and the region is yet to become an international financial centre, with Singapore as its main pillar. The micro-state already being ranked third globally as a financial centre and experts expect it to gain traction due to the development of the rest of the region.[3] The union includes the following countries: Brunei Darussalam, Cambodia, Lao PDR, Myanmar, Indonesia, Malaysia, Philippines, Singapore, Thailand and Vietnam.[4] The present work will only include: Indonesia, Malaysia, Philippines, Singapore and Thailand for data availability reasons.

[1] (ASEAN, 2023)
[2] (ASEAN Economic Data | Data| World Economics, n.d.)
[3] (The Global Financial Centres Index 34, n.d.)
[4] (ASEAN, n.d.-b)

Nonetheless, those countries are a good representation of the region's economic activity as they represent 85,4%[1] of ASEAN's GDP.

Building on the growing interest for ASEAN countries in Asset Pricing and the better performance offered by Machine Learning, this work's main contribution will be to investigate the intersection of the above two and the significance of the political factors. Specifically, this paper aims to answer the question of to what extent political factors contribute to enhance the accuracy of asset pricing models using machine learning in ASEAN countries, and which political factors are most significant in this context? Additionally, how does the significance of political factors on prediction outcomes vary across different ASEAN countries?

The political factors are intended to be measured using the Worldwide Governance Indicators published on a yearly basis by the World Bank. Political factor is measured through six indicators: Voice & Accountability, Political Stability, Government Effectiveness, Regulatory Quality, Rule of Law and Control of Corruption. Political factors have not yet been investigated in the context of Machine Learning in Asset Pricing. The relevance of Political factors is supported by Teeramungcalanon et al. (2020). In their research, they found that political factors play a crucial role in determining the ability of a country to attract Foreign Direct Investment (FDI) as a fraction of GDP. This was further supported by Mai (2023), who showed a significant connection between political stability and stock returns. Moreover, general governance had also a positive impact on the stock market.

The objective of this study is to predict the returns of the five major ASEAN indices spanning from 2018 to 2022, employing data from the training period of 2000 to 2017. The inputs for the model will be the lagged returns from 1 to 3 periods, in addition to political and economic indicators. The Machine Learning model selected for this work is the Long-Short Term Memory (LSTM). It is a recurrent neural network (RNN) and it was chosen for its ability to keep long and short term memory. Moreover, it also offers strong abilities to deal with non-linear data and time-series which are of interest for this research (Mehtab et al., 2021). To evaluate the different research objectives of the study, two LSTM models will be constructed: one including political factors and a second one without the political factors. The performances of these two models will be then compared to assess the significance of political factors for Asset Pricing in ASEAN countries. Finally, the importance of the different political indicators overall and across countries will be determined using Permutation Feature Importance.

This research paper aims to contribute to the literature in two ways. Firstly, by examining the potential increase in prediction ability of a Machine Learning model considering the political and economic factors. The factors have already been studied in different contexts and they are supported to be significant (Irshad, 2017; Mai et al., 2023; Teeramungcalanon et al., 2020). However, all the previous studies were only limited to spotting a correlation between the two phenomena and did not intend to incorporate the political factor as one of the predictors to determine individual countries' stock market returns. Furthermore, those studies used "traditional" linear regression. Therefore, this research paper, leveraging on Machine Learning's capacity to capture non-linear relationships and deliver improved performance compared to classical models (Gu et al., 2020), would propose a new implementation and methodology incorporating political factors. The reason why this factor has not been previously researched in Machine Learning models for Asset Pricing could be explained for several reasons. Preceding studies mostly focused on analysing a single country (ex: USA) or a relatively homogenous region from a political perspective (ex: EU). In that context, relatively little political divergences can be observed and hence may not be relevant.

The second way this study aims to contribute to the literature is by developing an Asset Pricing model using Machine Learning in the context of ASEAN countries. ASEAN stocks have not yet received full attention for Asset Pricing research up to this point, the only relevant studies that come close are Purnaningrum & Fariana (2022) which focuses on predicting major indices in ASEAN over one year after the Coronavirus and Jordan (2016) which focuses on more conventional methods to predict equity betas for 500 random US and ASEAN stocks. Other studies that focused on Southeast Asia scrutinised the equities of specific countries, such as Malaysia in Theang et al. (2019). This study aims to contribute to further investigate the case of ASEAN, as it offers great economic perspectives.[5]

[5] (KPMG, 2017)

II. Literature Review

Asset Pricing is the area of financial research which explores the factors that determine the prices and returns of financial assets. The CAPM model is a foundational concept in Asset Pricing and has continued to be used. The model is the result of the combined works of Treynor (1961), Sharpe (1964), Lintner (1965) and Mossin (1966). Their work has also been largely influenced by Markowitz (1952). Markowitz in his work explored the relationship between returns and volatility, which led to the mean-variance portfolio optimisation and modern portfolio theory. His paper also pointed out the need to diversify away the idiosyncratic risk, which is the risk associated with an individual security, by investing in multiple securities, preferably ones not correlated to one another. In that way, he argues that the market should not give any premium returns for bearing idiosyncratic risk as it can be diversified away by holding multiple securities. Based on this conception the CAPM model argues that the only factor influencing equity returns is exposure to market risk, as it is the only risk that cannot be diversified away. This exposure to market risk is captured in the term *beta* in the formula.

The creation of stock databases and the growth of computational power enabled to test empirically the CAPM model (Dimson & Mussavian, 1999). Based on empirical evidence Fama & French (1992) came up with two additional factors. The three factors model was a combination of the market risk factor, small minus big (SMB) and high minus low (HML). Small minus big represented the empirical findings that small capitalisation companies tend to outperform big capitalisation companies. In addition, high minus low stands for the observation that companies with a higher book-to-market ratio perform better compared to those with a low one. Those additions resulted in a better predictive ability and explained part of the errors observed in the CAPM model, referred to as *alphas* in the literature. The model was later completed furthermore with the inclusion of two new factors robust minus weak (RMW) and conservative minus aggressive (CMA) (Fama & French, 2015). Robust minus weak represents the observation that companies with robust or high operating profitability perform better than those with weaker profitability. Moreover, conservative minus aggressive illustrates that on average companies with a more conservative investment approach outperform their peers with more aggressive strategies. This final state of the Fama & French model continues to contribute to our understanding of the factors driving equity returns and it remains an essential linear model in the field of Asset Pricing.

Researchers are constantly looking for new factors that could contribute to a better comprehension of stock price dynamics. In that context, the political factor has already been investigated in various settings. Teeramungcalanon et al. (2020) investigated its importance for Foreign Direct Investments (FDI) for ASEAN and the Korean economies using the Worldwide Governance Indicators. They concluded by suggesting that a stable political atmosphere helps build trust for foreign companies and a strong rule of law has an important role in attracting FDI. Mai et al. (2023) also point out the role of political factors on the stock market performance. The study focused specifically on Pakistan and used the Worldwide Governance Indicators, with an emphasis on political stability for their regression. The results showed a significant link between political stability and good stock market performance. Furthermore, all indicators have been found to be significant, as they foster better governance in general. The researchers conclude by underlying the importance of diversification, in particular for emerging markets, to mitigate the risk related to political instability. Irshad (2017) came to the same conclusion regarding the negative impact of political instability on stock market performances. Moreover, the study used slightly different measures to express political stability and incorporated other elements into their regression. Their results showed that inflation also has a significant negative impact and on the contrary industrial output and industrial exports have a significant positive effect on equities' return. They concluded by stressing the need for political stability, controlled inflation and stimulated industrial output and exports for a better performing stock market.

However, the main limitation of 'traditional' models and the ones presented above is their linearity and by definition their incapacity to capture non-linear relations (Gu et al., 2020). The further development of computational capabilities, coupled with more comprehensive data availability enabled the use of more complex statistical models, known as Machine Learning. Those models offer the advantages of, first, being able to capture non-linear patterns which traditional models are unable to do, and thus offer the possibility to discover new "hidden" significant parameters for Asset Pricing which do not follow linear patterns. Secondly, Machine Learning models can handle large amounts of data and variables. Many variables can be the source of issues for traditional models. However, Machine Learning models can accommodate this by only retaining the relevant predictors from a large set, leading to improved predictive accuracy without manual intervention. Thirdly, it offers greater flexibility. This translates into the model's ability to adapt to different types of variables (categorical or nominal) and adjust to new data without external intervention. Finally, studies have demonstrated the significant performance gain of Machine Learning models over traditional counterparts. This resulted in increased stock return prediction accuracy and generated economic gains for investors, demonstrating its relevance for finance.

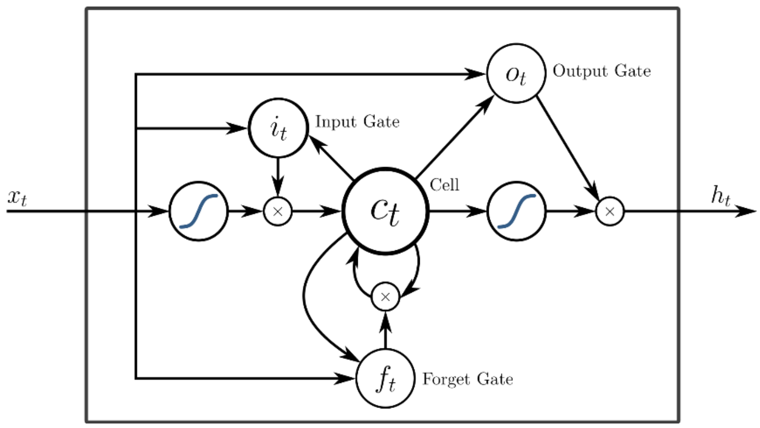
The field of Machine Learning has elaborated various models, each with their specificities and applications. There exist many categories of Machine Learning algorithms, but one of interest for this study and most employed is Supervised Learning. This approach tries to learn patterns in input-output pairs, enabling it to predict an output for a given input. Additionally, the data must be split into a training and test set, and it requires some level of supervision. This category incorporates notable models such as Linear Regression, Regularisation, Decision trees & Random Forests and Neural Networks (Mahesh, 2020).

Regularisation is a linear model which was built to better manage numerous predictors and avoid overfitting, this is achieved by adding a penalty term to the loss function[6]. Decision trees, including Random Forests, use a divide-and-conquer approach, splitting the training data into subsets (leaves) based on features and values. Predictions are made based on these subsets, with Random Forests creating multiple trees and averaging results for robustness (Kubat et al., 1998). Neural networks represent some of the most complex machine Learning Models. They have demonstrated impressive performances and a wide range of applications (pattern recognition, prediction or natural language processing). Its structure is meant to replicate the one of a human brain. It is made of a network of layers of neurons: input, output and a number of so called "hidden layers" determined by the complexity of the model. All those layers are interconnected by weights, which the model calculates to optimise its predictions.

A model particularly suited for this research is the Long Short Term Memory (LSTM), which is a recurrent neural network (RNN). Literature has pointed out the shortcomings of 'simple' RNN, which can arise when the long-term memory starts to become more distant. In that case, two types of errors can occur known as gradient vanishing and gradient growth (explosion) (Hochreiter & Schmidhuber, 1997). Considering this issue LSTM was designed with the intention to correct those errors. LSTM architecture is made of three gates: forget gate, input gate and output gate. Those components are responsible for deciding how much of the inputted data should be used for updating the cell state. The cell state represents the long-term memory, whereas the hidden states are the short term or working memory. All together those components decide what should be the output. Its unique features have granted it increased performances compared to other models, in particular, basic RNNs and made it extremely performant to deal with time-series data, which is the case in this study (Mehtab et al., 2021).

Figure 1

Architecture of LSTM



[6] *What Is Regularization and How Does It Help to Avoid Overfitting in a Machine Learning Model?* | Kaggle, n.d.

Note. The figure illustrates the connections between the key components of LSTM's architecture: forget gate, input gate, cell state and output gate. From "*Long short-term memory*", by Wikipedia contributors. (2024, March 16). Wikipedia.

The applications of Machine Learning in finance have not been curtained to return prediction. Academic research touched upon a wider range of areas of implementation, including portfolio construction. Ma et al. (2021) in their study compared the performances of various Machine Learning and Deep Learning models at selecting stocks before forming portfolios. The portfolios were then formed using Mean-Variance and predicted stock returns in the optimisation process. This is known as Mean-Variance with Forecasting (MVF). They found that Random Forest presented the best performance. Furthermore, their results suggest that the models studied present significantly better performances than traditional linear models. Additionally, an area which also has gained from the use of Machine Learning models is the treatment of missing financial data. Treatment of missing data is particularly important in finance, as the discipline often utilises times-series and inappropriate use of future data points to fill-in anterior missing data can cause "look-ahead biases" (Freyberger et al., 2022). To resolve this issue Guilin (2021) suggests splitting data into two categories short term and long-term missing data. The first category would be treated by a linear model, such as LASSO and long-term missing data would be treated with a light gradient boost. Alternatively, Stam (2022) argues for the use of Random Mice Forest (RMF) as a method for data treatment.

III. Data and Methodology

1. Data

The data used for this paper comes from multiple databases and covers the timeframe from 2000 until 2022. Due to data availability constraints only the five following ASEAN countries will be studied: Indonesia (IDN), Singapore (SGP), Philippines (PHL), Malaysia (MYS) and Thailand (THA). Nonetheless, those countries are a good representation of the region's economic activity as they represent 85,4%1 of ASEAN's GDP.

The first source of data is the World Indices database offered by Wharton Research Data Services. Their services will be used to access the monthly returns including dividends for the concerned countries. The data obtained spans from 8/1999 to 12/2022, in order to compute the lagged returns from 1 to 3 periods for the relevant timeframe. The dataset is entirely complete and does not require further manipulation.

Table 1

Descriptive Statistics for Monthly Returns Including Dividends

Country	*Observations*	*Mean*	*Median*	*Min*	*Max*	*STDV*
IDN	276	1,11%	1,43%	-31,30%	19,68%	0,0610
MYS	276	0,58%	0,78%	-14,52%	13,74%	0,0409
PHL	276	0,72%	1,19%	-22,34%	15,47%	0,0544
SGP	276	0,48%	0,74%	-25,80%	25,47%	0,0513
THA	276	0,70%	0,84%	-29,54%	24,91%	0,0598

Additionally, the Key Indicators Database from the Asian Development Bank will be providing the economic indicators about ASEAN countries. The database holds a total of 775 variables updated yearly for most Asian countries. The data is sorted into several major categories and subcategories. Major categories are: People, Economy & Output, Energy & Electricity, Money & Finance & Prices, Government & Governance, Globalization, Transport & Communication, Environment and Sustainable Development Goals (SDGs). After selecting relevant factors and ruling out variables with excessive missing data (>20%), this resulted in 60 variables across 23 years and 5 countries. Finally, missing data (4,71%) will be filled using Rolling Mice Forest (Stam, 2022). This method of imputation offers good performance in filling incomplete financial databases, while respecting the chronological nature of financial data and applying a Rolling Window Approach to avoid "look-ahead biases". This is accomplished by only using the information available to that point in time for imputing and no posterior data points which could bring potential leakage from future data.

Moreover, the measurement of the Political factor will be made using the Worldwide Governance Indicators database from the World Bank. Since 1996, the database has annually reported six indicators: Voice and Accountability (VAA), Political Stability and Absence of Violence/Terrorism (POL), Governance Effectiveness (GOV), Regulatory Quality (RQ), Rule of Law (ROL) and Control of corruption (COC)[7]. Each metric is reported on a scale from -2,5 (weak governance) to 2,5 (strong governance). This database is entirely complete, except for the year 2001 as before 2002 the data was released every two years. The missing year is addressed by simple mean imputation. Apart from that no further processing is required as the indicators are reported on a comparable scale.

Table 2

Descriptive Statistics for Political Indicators

	Observations	*Mean*	*Median*	*Min*	*Max*	*STDV*
COC	*115*	*0,13*	*-0,37*	*-1,14*	*2,3*	*1,0636*
GOV	*115*	*0,65*	*0,26*	*-0,6*	*2,47*	*0,8818*
POL	*115*	*-0,29*	*-0,5*	*-2,1*	*1,6*	*1,0113*
RQ	*115*	*0,48*	*0,17*	*-0,87*	*2,25*	*0,8187*
ROL	*115*	*0,21*	*-0,01*	*-0,91*	*1,84*	*0,8241*
VAA	*115*	*-0,18*	*-0,11*	*-1,05*	*0,47*	*0,3138*

The data collected spans from 2000 and 2022. It will split into two sets: 2000 to 2017 for training and 2018 to 2022 will be kept hidden and used for test purposes. Cross-validation will be performed using the training set to tune the hyperparameters and come at a lower cost of "lost" training data for validation. In that perspective, 5-fold cross-validation is chosen for computational reasons (Bergmeir & Benítez, 2012). Besides that, no issues are expected regarding the fact that our test data includes the COVID-19 years, as our training data includes financial crises (2007-2009). Therefore, the model should identify the patterns and react accordingly. Additionally, for the model to be judged efficient it should be able to perform in various market conditions and the 2018 to 2022 timeframe allows to observe its behaviours in the event of a crisis, recovery and expansion.

[7] Refer to the Appendix for the World Bank's definitions of the Political Indicators.

2. Methodology

2.1 Processing data

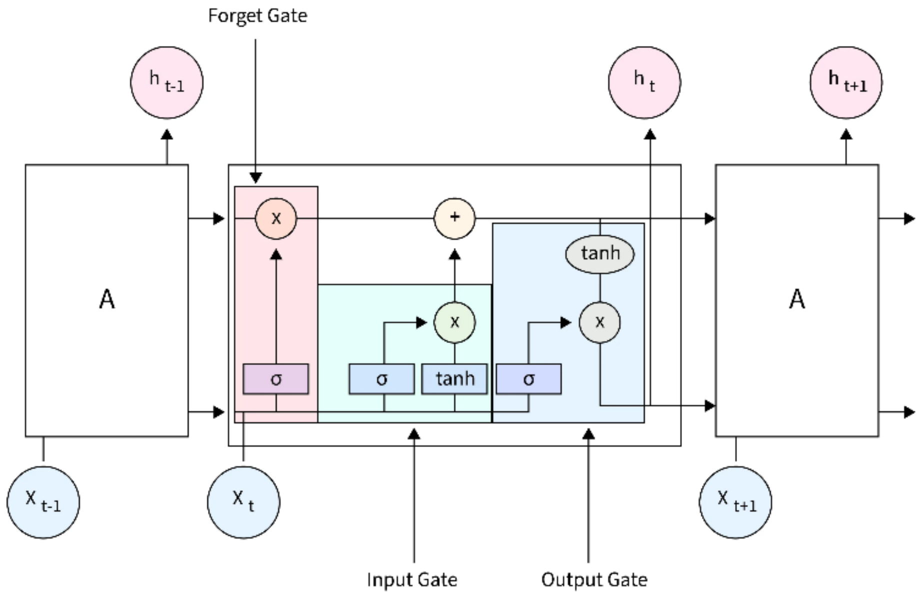
The first step will be to prepare the data for further processing. Firstly, missing data will have to be treated. The missing data in the economic factors dataset will be filled using Rolling Mice Forest (Stam, 2022). Additionally, the missing year, 2001, for the political indicators is retrieved using mean imputation. Regarding the indices' returns, the lagged returns from 1 to 3 periods are calculated and the data used for that purpose before 2000 is disposed of. Finally, the three datasets are manipulated to ensure they have compatible formats and merged into one long format dataset with each sheet representing a country by its ISO code.

2.2 Model: LSTM (Long Short-Term Memory)

The model selected for this study is the Long Short-Term Memory (LSTM) model, which is a recurrent neural network (RNN). Its strong abilities to deal with non-linear data and times-series through its long term memory make it well suited for this study. RNNs are a type of neural network made to recognise patterns in sequences of data. Its main specificity lies in the fact that contrary to Feedforward Neural Networks (FNN), the model is able to keep a certain form of memory as the information propagates through the network. This is because the information does not only flow in one direction through the hidden layers, like in an FNN, but instead, the model has loops, that in turn create internal states by passing several times the data and this enables the model to have a memory. However, RNNs can face severe issues when the long-term memory starts to become more distant. Errors start to arise during backpropagation (the process of updating the weights of the various states). In that case, two types of errors can occur known as gradient vanishing and gradient growth (explosion) (Hochreiter & Schmidhuber, 1997). Considering this issue LSTM was designed with the intention to correct those errors. The uniqueness of its long-term memory enables it to be very efficient at dealing with time-series data. This makes it a relevant choice for this study given the time-series nature of the political and economic factors. Additionally, it exhibits increased performance compared to other models, in particular, basic RRNs and displays the ability to capture non-linear patterns (Mehtab et al., 2021).

Figure 2

Structure of LSTM as part of a network

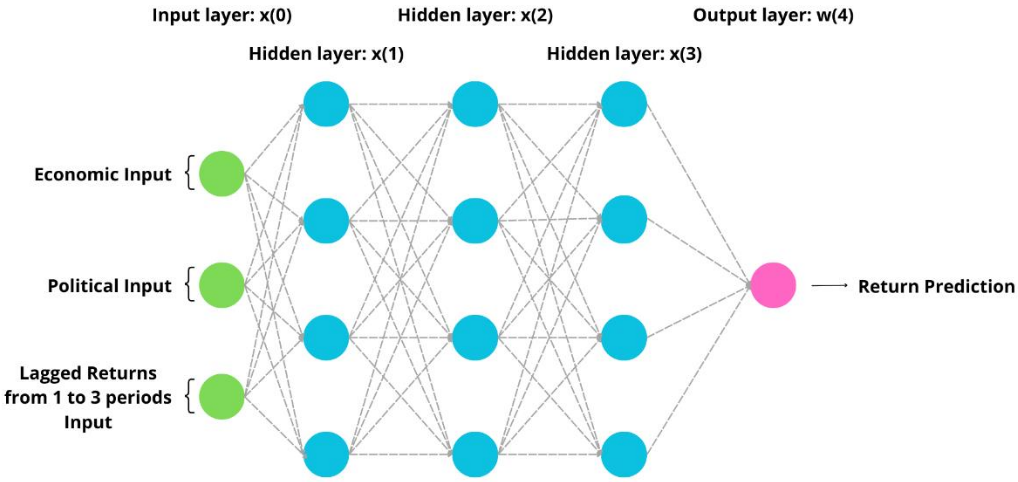


Note. This figure displays the different components of the LSTM model and the functions inside them. Moreover, it highlights the cell's location within a network of similar cells which constantly update each other. From *What is LSTM? - Introduction to Long Short-Term Memory - Scaler Topics*, by Lawrence, S. J. (2023, April 18), Scaler Topics. https://www.scaler.com/topics/deep-learning/lstm/

The present LSTM model developed for this work forecasts the monthly returns including dividends of the major ASEAN indices based on the following inputs (or predictors): the lagged returns from 1 to 3 periods, economic and political indicators. The dependent (or target) variable, in our case monthly returns with dividends of the major ASEAN indices, is paired with the predictors for that period. The data is divided into training and test sets. During the training, in this case, the years 2000 to 2017, the model will learn by observing the relation between the predictors and the target variable. The model determines the parameters and weights that give the most optimal output given the loss function. Once it is done it is evaluated on the test set, years 2018 to 2022, which acts as new data, that the model has not seen and therefore, learned from it. This way it gives a good representation of how well the model has understood the dynamics at play in the predictors and how well it can predict unseen/future data with new values for the predictors.

Figure 3

Visual representation of the developed LSTM model



Note. The figure offers a good way to visually represent the model with its inputs (Economic, Political and lagged returns) and output (Predicted Returns). However, the model offers only a simplified description of the actual model. The actual model has between 2 to 4 hidden layers with 4 to 128 nodes on each layer depending on the value chosen by the model for the hyperparameters.

To effectively obtain the desired outcome and avoid overfitting particular attention will be granted to the different parameters and hyperparameters of the model. The selection of hyperparameters is comprised of the following elements: the number of hidden layers, the number of neurons in each layer, dropout rate, learning rate, batch size and epoch size. The selection was made based on Mao et al. (2022) and the suggested values for each hyperparameter can be found in the table below:

Table 3

Suggested values for Hyperparameters

Hyperparameter	**Suggested Values**
Number of hidden Layers	[2, 3, 4]
Number of neurons in each layer	[4, 8, 16, 32, 64,128]
Dropout rate	[0.05, 0.1, 0.15, 0.2, 0.25]
Learning rate	[0.01, 0.001, 0.0001]
Batch size	[1, 5, 10, 20]
Epoch size	[50,100, 200, 300, 400, 500]

Note. From *Selection of precise long short term memory (LSTM) hyperparameters based on particle swarm optimization.* By Mao et al. (2022).

The suggested values for the hyperparameters are selected during tuning by either random search or grid search. For this work random search is preferred over grid search for its ability to find equally good or even better models, and it requires less computational power (Bergstra & Bengio, 2012). This will be achieved with the use of cross-validation. This method was selected over simply splitting the training data into two parts for training and validation, as in that case the data used for validation is in some way "lost" as it cannot be used to train the model. Instead, cross-validation offers a solution which allows to use the validation data for training purposes, by splitting the data into k-fold groups and alternatively using them for validation. For this model, 5-fold cross-validation was selected for computational reasons (Bergmeir & Benítez, 2012).

2.3 Performance Evaluation

To answer the research question and assess the significance of the contribution of Political Indicators in predicting the monthly returns of major ASEAN indices, two models are constructed: one including political indicators and the second one excluding political factors. Each model will be presented with the same set of hyperparameters and suggested values. Then the model's performance will be compared on three dimensions: MSE, RMSE and MAPE.

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

$$RMSE = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}(|y_i - \hat{y}_i|)^2\right]}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| * 100$$

Where:

y_i: $Original\ times\ series,$

$\hat{y}_i$: $Predicted\ time\ series\ computed\ from\ the\ model,$

n: Number of observations.

MSE and RMSE calculate the mean square of error and for the latter its root. Those metrics offer a good way to evaluate the models' prediction accuracy as they are outlier-sensitive. This means that they will react exponentially negatively to large prediction errors. This is relevant in the context of stock price predictions as even a few large prediction errors can cause a lot of troubles for an investor. MAPE is a relative metric, which expresses the average value of relative error as a percentage of the actual data. It is particularly well suited for evaluating the forecasting accuracy of a model and returns a value in percentage, which can be easily compared (da Silva et al., 2022). The smaller the value of MSE, RMSE and MAPE, the better the model performed. The model with the smallest MSE, RMSE and MAPE would be considered the best (Hum Nath Bhandari, 2022).

2.4 Permutation Feature Importance

Complex Machine Learning models can often result in black-boxes, predictors are inputted and something gets out of it and miraculously works. However, the importance and contribution of each feature remain uncertain and unobservable. For that reason, statisticians have come up with various models to evaluate the contribution of each feature towards the final output. In the context of Neural Networks and LSTM, a method commonly used is Permutation Feature Importance. It was introduced by Breiman (2001)and consists of randomly shuffling several times the values of each feature separately and measuring its impact on the model's error (MSE). The reasoning behind this is that this way the connection between the feature and the target is broken. Therefore, if the feature was important to the model then the error (MSE) is expected to increase and oppositely remain or decrease if the feature was irrelevant. For this work, the Permutation Feature Importance will be performed ten times for each input. Finally, the mean change in MSE for all the Political Indicators overall will be extracted. The procedure will be repeated by grouping each country to get the country-specific results. Those results will allow to first see if the same Political Indicators are significant for all the countries and second if for certain countries, they are more significant towards the final outcome.

To ensure the differences in mean change of MSE observed between countries are the result of an underlying different sensitivity to political indicators and not statistical randomness an Analysis of Variance (ANOVA) will be performed. This will be done over the scores of the ten iterations of Permutation Feature Importance for each country. This procedure is relevant as it allows to test for statistically significant differences between the means of multiple groups. In case of significant results (p-value < 0,05), Tukey's Honestly Significant Difference test will be used as a Post-Hoc analysis (Tukey, 1949). This will help determine which specific pairs of countries have significantly different mean changes in MSE expressed in terms of p-value, which grants ease of interpretation. This procedure assumes homogeneity of variance and independence of observations.

IV. Results

1. Political Indicators Significance

The significance of Political Indicators is assessed by comparing the results of two LSTM models, with and without the Political Indicators, in addition to 60 Economic factors and lagged returns from periods 1 to 3. The two models were presented with the same set of values to choose for the hyperparameters, and the selected values can be found for both models in the Appendix under Tables 6 and 7. Finally, they were compared across three metrics: Mean Square Errors (MSE), Root Mean Square Errors (RMSE) and Mean Absolute Percentage Error (MAPE). The table below expresses the results for the two models.

Table 4

Results of model comparison

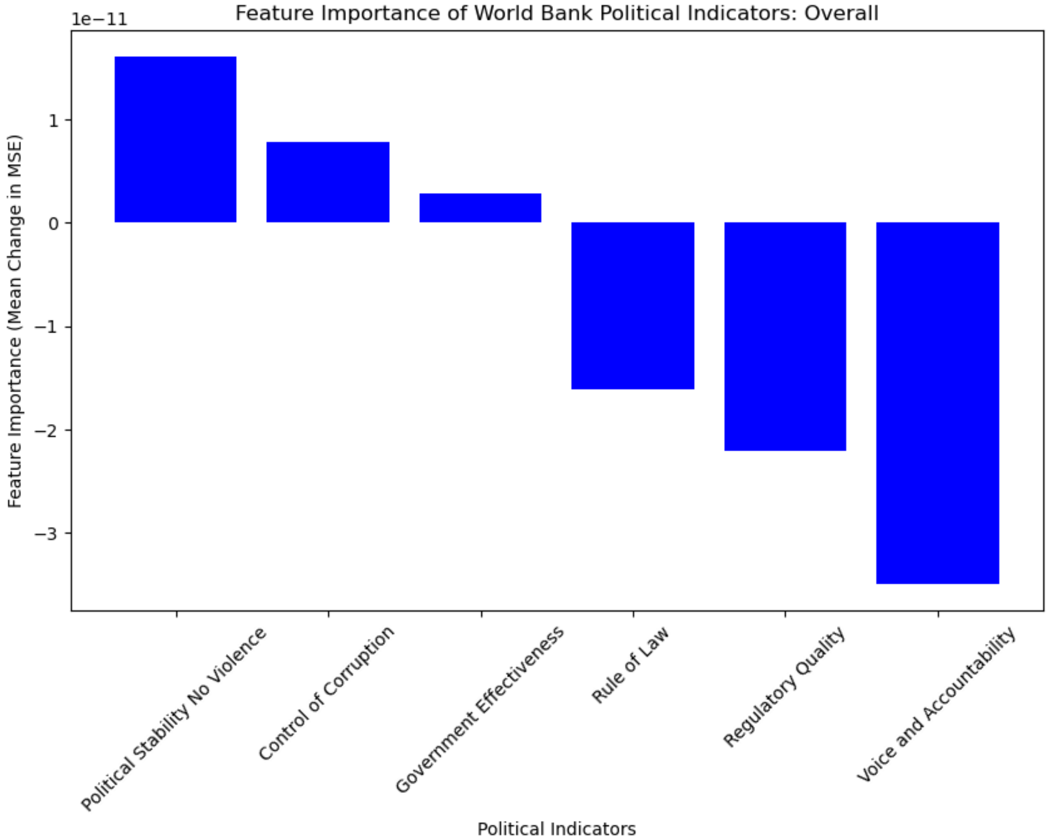
	With Political Indicators	*Without Political Indicators*
MSE	0,0019	0,0034
RMSE	0,0431	0,0585
MAPE	2,1%	5,57%

From the table above it can be clearly observed that the model with the Political Indicators obtained better performance across all metrics, as defined in the methodology. This outcome goes in line with what has already been investigated by Teeramungcalanon et al. (2020), regarding the significance of Political Indicators to attract Foreign Direct Investment (FDI), or by Mia et al. (2023) and Irshad (2017) regarding their impact on the Pakistani stock market returns. The obtained results support that Political Indicators are also relevant in predicting the returns of ASEAN indices.

2. Overall Feature Importance of Political Indicators

Following the insights that Political Indicators play a significant role in predicting the returns of ASEAN indices, a Permutation Feature Importance test is performed ten times to assess the overall contribution and relevance of each of the six factors across all the countries. The results are expressed in the figure below in terms of mean change in MSE. A positive value expresses that the feature contributes to the model's accuracy, whereas a negative or zero value suggests that this value does not bring prediction power.

Figure 4

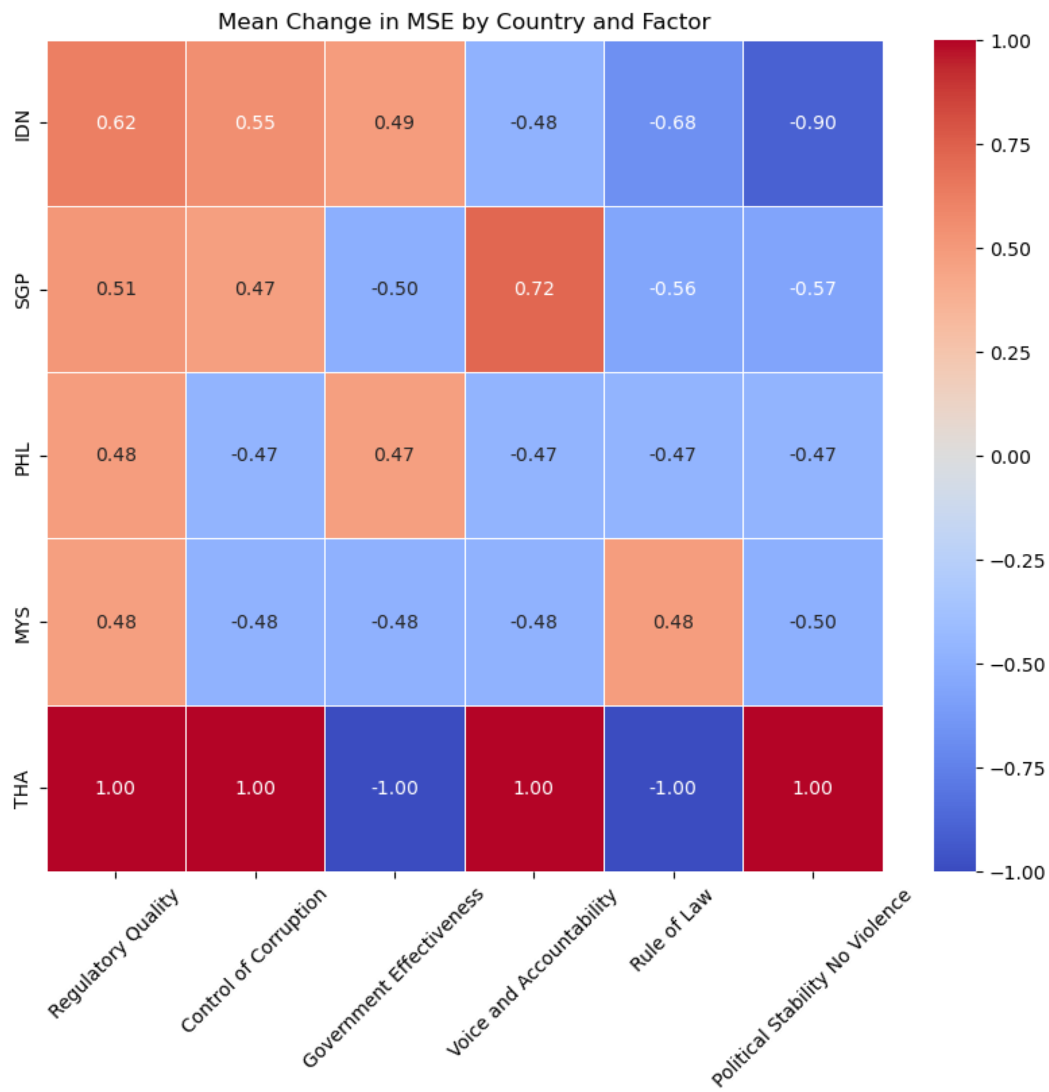


From the figure above it can be observed that three Political Indicators contribute to the model's overall prediction across countries. These are: Political Stability, Control of Corruption and Government Effectiveness. Conversely, Rule of Law, Regulatory Quality and Voice & Accountability did not show importance for the model when predicting returns of the major ASEAN indices. Teeramungcalanon et al. (2020), observed that Political Stability and Rule of Law were found to be positively significant in predicting FDI for ASEAN and Korean economies, whereas Voice and Accountability had a significant negative effect. Moreover, Mai et al. (2023) found in their paper support for the same Indicators in addition to Government Effectiveness when trying to predict the returns of the Pakistani stock market, but all had a positive impact. The results obtained above only share Political Stability which was found to be important among the three studies. The differences can be explained by the different nature between predicting index returns and FDI. Furthermore, another explanation could be the employment of non-linear models for this study, as opposed to a simple linear model for the aforementioned studies. LSTM's ability to capture non-linear interactions and observe patterns across time steps led the model to better grasp the interactions at play and their significance.

2.2 Variation Across Countries

Now addressing Permutation Feature Importance per country for Political Indicators. The results obtained showed a lot of variations which made it not feasible to create a visually comprehensible figure. For that reason, the following figure was scaled. This was achieved by initially taking the absolute values of the results, then robust scaling was applied to them, subsequently, the outputs were processed by a sigmoid function. Finally, their original signs were restored. These manipulations resulted in more visually comparable values. Consequently, the following figure should be interpreted bearing in mind those manipulations. To explore the unprocessed values for Permutation Feature Importance by country across all Indicators, please refer to Table 8 in the Appendix.

Figure 5



The following results can be interpreted in the following way. Values in the range between -0,6 and 0,6 suggest that this factor does not show more or less significance than on the aggregate level for the specific country. Alternatively, values <-0,6 suggest that the particular indicator is found to be less significant for this country and >0,6 indicates the indicator is more important for the specific country. From the figure above it can be directly noticed that the country that expresses the most extreme results is Thailand, always scoring -1 or 1 on all the factors. This is particularly remarkable as the data was processed using a sigmoid function returning values of 0 and 1[8], where 1 stands for very large values and 0 for very small values. This showcases extreme difference in the scales of the results obtained, given that it is a logistic function. The next country that seems to be demonstrating some deviating results is Indonesia, scoring on three factors above |0,6| and with a result of -0,9 for Political Stability and No Violence. Lastly, Voice and Accountability show stronger importance for Singapore.

Turning to indicators displaying compelling results. Firstly, Control of Corruption was also found significant on an overall basis for the LSTM model. Only two countries deviate from that on an individual case: the Philippines and Malaysia. This dimension represents a reality for emerging countries, which can often be prone to suffer more from corruption as they do not have the institutions to fight it and the living standards remain comparatively low. Its presence consequently can disturb the natural flow of business and be detrimental to the overall economy. Secondly, Voice & Accountability was found significant by the two previous studies using linear regression, however in the case of Teeramungcalanon et al. (2020) this significance was negative. This was argued also on based on Guerin & Manzocchi (2009) and Jadhav (2012) that public opinion on trade policies can be detrimental. Nonetheless, for LSTM it did not appear as relevant. When looking at a per country basis, it can be observed that it displays strong importance for Singapore and Thailand. Finally, Regulatory Quality was not found significant by any other study and on an aggregate level, yet it shows relevance for all countries when taken individually. This highlights the complexity of the interactions at play and demonstrates that some features can get diluted on an aggregate basis and emerge on an individual basis. Those results also point out the role of Regulations in offering a good framework for business conduct, by promoting the development of the private sector and ensuring property rights.

[8] The values include negative numbers as the sigmoid function is applied to their absolute values and the signs are then retrieved.

2.2.1 ANOVA Results

To assess the significance of the above-mentioned results a one-way ANOVA was conducted to analyse the effects of 'Country' and 'Factor' on the mean change in MSE[9]. The corrected model showed significance at F (29, 270) = 50,553, p<0,001[10] and explained 84,4% of the variance, with an adjusted R^2 of 0,828. More specifically, the 'Country' effect did not show significance, F (4, 270) = 0, p = 1, this implies that no country when averaging across all factors consistently shows a higher or lower mean change in MSE. Nonetheless, this does not mean that no country displays a significant deviation from the others. As in this case a country could have more sensitivity to political factors and obtain results way above and below the mean than other countries (which would imply more sensibility), but as this metric averages out the results across all factors it would not show significance on aggregate basis. Moreover, 'Factor' did show relevance at F (5, 270) = 54.503, p < .001 and the interaction effect ('Country' * 'Factor') was also found significant at F (20, 270) = 55.164, p < .001. The fact that the interaction term presents significance suggests that certain countries could showcase meaningful differences compared to other countries for specific factors. To further elaborate Tukey's HSD Post Hoc test is performed and the following results are obtained:

Table 5

Results of Tuckey HSD Post Hoc Test

					95% Confidence Interval	
(I) Country	(II) Country	Mean Difference	Std. Error	Sig.	Lower Bound	Upper Bound
THA	MYS	-2,57E-06	3,38E-07	0,000***	-3,49E-06	-1,64E-06
	PHL	-2,56E-06	3,38E-07	0,000***	-3,49E-06	-1,63E-06
	SGP	-2,58E-06	3,38E-07	0,000***	-3,51E-06	-1,65E-06
	IDN	-1,42E-06	3,38E-07	0,000***	-2,35E-06	-4,93E-07
MYS	PHL	5,49E-09	3,38E-07	1,000	-9,23E-07	9,34E-07
	SGP	-1,66E-08	3,38E-07	1,000	-9,45E-07	9,12E-07
	IDN	1,14E-06	3,38E-07	0,007**	2,16E-07	2,07E-06
PHL	SGP	-2,21E-08	3,38E-07	1,000	-9,50E-07	9,06E-07
	IDN	1,14E-06	3,38E-07	0,008**	2,10E-07	2,07E-06
SGP	IDN	1,16E-06	3,38E-07	0,006**	2,32E-07	2,09E-06

[9] The full table of results can be found under Table 9 in the Appendix.
[10] F (X, Y) = Z, p < p-value, Where F stands for F-test, X for the degrees of freedom and Y the number of observations. Z represents the F value, and the significance value is expressed.

The table above presents the significance level for specific pairs of countries concerning their different mean changes in MSE for Permutation Feature Importance. It can be noted that the two countries that exhibit significance are Thailand ($p < 0,001$) and Indonesia ($p < 0,05$). This is consistent with the visual observation that could be made from the figure above, but those results bring empirical evidence that they are statistically different. This suggests that those countries are more sensitive to political indicators and those vary significantly away from the overall results. Alternatively, the three other countries studied do not indicate significant deviation from the overall results.

Those results could be explained by the fact that Thailand and Indonesia have been navigating since the early 2000s tumultuous political environments. At the paramount of this Thailand suffered two military coups, in 2006 and 2014, which are part of a history marked by already 13 coups d'état since constitutional law was introduced in 1932 (Pongsudhirak, 2014). Thai politics is characterised by a fragile trinity between the monarchy, bureaucracy and military. All the political struggle in Thailand resides in the opposition between the so-called Red and Yellow factions. This opposition can be transcribed respectively as liberal and Western-endorsed notions against a more traditional view which fights for the nation's homogeneity, referred to as "Thai-ness" and its three pillars of nation-religion-monarchy. The two coups were the result of a corrupted bureaucracy and were massively supported by the Yellow faction. After seizing power, the military sought to fight corruption in the administration and put in place a new constitution. When the military decides it is done "cleaning" the system, it organises new elections and leaves power but keeps in place some of its representatives. Thailand's political situation is stuck in a cycle of coups, introducing a new constitution inevitably leading politicians to become corrupted and seeding the ground for a new coup. In conclusion, Thailand has not been able to make much progress in any direction because it finds itself stuck between the military, corrupted politicians and unsteady constitutions, consequently leading to inconsistent regulatory frameworks (Harding & Leelapatana, 2019).

On the other side, Indonesia entered the 21st century with some turmoil. The country was part of the Asian Miracle, which describes the rapid economic development of eight eastern Asian countries. However, the Asian Financial Crisis of 1997 led to a severe economic downturn, extreme inflation and consequently devaluation of the local currency. Those events were coupled with social unrest, which resulted in President Suharto's resignation from power, Indonesia's second leader and dictator. The severity of the crisis was furthermore fuelled by a weak banking system, lack of accountability and corruption[11]. In 1999 the first free election was held. Additionally, the ambient instability created a favourable ground for secessionist and religious extremist groups. Resulting in numerous terrorist attacks across the country. Those groups might have been instrumentalised by foreign interests and the situation was resolved with the cooperation of other ASEAN states (Singh, 2004). In 2004 the

[11] (Singh, n.d.)

Corruption Eradication Commission (KPK) was established in an attempt to combat corruption and improve regulatory quality. However, the results are still mitigated to this day, as the country still lacks strong institutions to effectively address the issue (Dick & Mulholland, 2016). In recent years under the rule of the incumbent President, Joko Widodo, stability has been restored and associated with the realisation of economic potential. This was most notably observed in improved government efficiency, in particular regarding infrastructure development, which had to compensate for previous underinvestment in that domain. This was accomplished through a shift in fiscal policy to encourage infrastructure projects, with a focus on maritime infrastructure to benefit from the country's geographic location and multiplication of Private-Public-Partnership (PPP) as a new source of financing (Salim & Negara, 2018). In conclusion, those could be some of the reasons explaining the increased sensitivity to political indicators for Thailand and Indonesia.

V. Conclusion

In conclusion, by comparing two LSTM models across 3 metrics it was determined that Political Indicators show evidence of contributing to better prediction of the returns of major ASEAN indices. Furthermore, with the use of Permutation Feature Importance it was observed that Political Stability, Control of Corruption and Government Effectiveness were the Political Indicators with an impactful contribution towards the model's final output. Interestingly Regulatory Quality did show significance on an aggregate basis but turned out to be significant on a per country basis for all. This highlights the difficulty of accurately capturing all the active connections among variables and the dilution that can occur on an aggregate basis. Additionally, the analysis revealed substantial differences among countries in terms of the importance they assigned to each indicator. The variations were observed based on the specific dimensions that showed significance, and sometimes did not align with the aggregate results. Moreover, the degree to which each indicator contributed to the prediction varied largely across countries. The countries that demonstrated the most sensitivity to Political Indicators after the ANOVA test using Tukey's HSD were Thailand followed by Indonesia. Both countries have experienced in their recent histories, tumultuous political environments, most likely explaining their increased sensitivity to political indicators. Thailand seems to be trapped in a cycle of social unrest and corruption leading to military coups. In the aftermath, a new constitution is introduced and the government falls back into its bad ways (Pongsudhirak, 2014). Regarding Indonesia, the country experienced a severe economic downturn and extreme political violence at the beginning of the century, which led to a lot of transformations. In recent years, the country has been doing well under the Presidency of Joko Widodo, partly due to his infrastructure investment plan and is on a good path (Salim & Negara, 2018). This stresses the differences between countries and how the importance of Political Indicators can vary.

In general, this study contributes to emphasise the role of political influence on the economy and resultingly on the stock markets of countries. This brings to light the importance of good governance for economic success. Political Stability inspires confidence in investors which makes them less likely to withdraw their investments from a given country. This follows from the perceived continuity and certainty that will follow. Moreover, Control of Corruption represents a real challenge for emerging countries as they lack the institutions to combat it and poor standards of living make it more appealing. However, it has detrimental impacts on the economic activity of the region. ASEAN countries, familiar with the topic, have nonetheless put in place efforts to reduce this issue and those initiatives' importance is supported by the results obtained. Government Effectiveness and its policies are also major drivers of economic performance. Policies that foster investment in public infrastructures or services coupled with stability and ensuring predictability are well received by markets (Mai et al., 2023). On the contrary, uncertainty produced by sudden changes in regulations leads to more uncertainty and volatility which has negative effects on the stock market. Drafting and executing good policies can be a difficult job, requiring nuance as they are not always straightforward and subject to other influences. Finally, Regulatory Quality interestingly displays significance for all countries on a country per country basis, but not overall. Reasons for it could be its dilution on an aggregate level as its relevance is only applicable to its local context, due to all the nuances and differences that exist. Moreover, this once again puts forward the importance of having a clear and consistent regulatory environment, which in turn leads to less volatility for business. Policies that strive to promote private sector development contribute largely to that.

Overall, this study contributes to emphasise the importance for governments and policymakers to encourage stability and good governance. This can be achieved by implementing policies which promote transparency and stability while ensuring consistency to wave uncertainties. Poor management of those can lead investors to withdraw their investments from a given region as they perceive rising risks and not adequate returns to compensate. Investors should also monitor political developments when constituting and reviewing their portfolios. Diversification and hedging should consequently be the main pillars of a well-managed portfolio to mitigate the downside. This also reveals itself to be relevant for international companies which should put in place efficient risk management, as well as diversify their presence across politically sensitive regions.

Finally, this study aims to be a building block for future research interested in asset pricing for ASEAN. Firstly, a relevant direction could be to investigate the prediction power of political factors as standalone informational. This builds upon the empirical evidence provided by Chen et al. (2023) to suggest that for economic factors time-series offer better quality predictions, whereas the cross-sectional approach makes the economic information harmful to the model. In the second place, it would be insightful to construct the same model, but instead of predicting ASEAN major stock markets, predict their constituents using firm-specific factors. However, this would require complete data and handling well the high dimensionality. Lastly, another relevant direction to investigate as the region continues to develop, and more data becomes available is to incorporate the remaining countries part of ASEAN. It would be particularly interesting to observe the results for countries such as Laos and Vietnam as those countries still have a communist government in place. Additionally, Brunei can also be relevant as the country still has an absolute monarchy in power.

VI. References

1. Academic:

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, *26*(10), 1340–1347.

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, *191*, 192–213.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, *13*(2).

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Chen, L., Pelger, M., & Zhu, J. (2023). Deep learning in Asset Pricing. *Management Science*.

da Silva, D. G., Geller, M. T., Moura, M. S., & Meneses, A. A. (2022). Performance evaluation of LSTM neural networks for consumption prediction. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, *2*.

Dick, H., & Mulholland, J. (2016). The politics of corruption in Indonesia. *Geo. J. Int'l Aff.*, *17*, 43.

Dimson, E., & Mussavian, M. (1999). Three centuries of asset pricing. *Journal of Banking and Finance*, *23*(12), 1745–1769.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Freyberger, J., Höppner, B., Neuhierl, A., & Weber, M. (2022). *Missing data in asset pricing panels* (w30761). National Bureau of Economic Research.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Guerin, S. S., & Manzocchi, S. (2009). Political regime and FDI from advanced to emerging countries. *Review of World Economics*, *145*, 75–91.

Guilin, H. (2021). Missing Data Filling Method Based on Linear Interpolation and Lightgbm. *1754(1)*.

Harding, A. J., & Leelapatana, R. (2019). Constitution-making in 21st-century Thailand: The continuing search for a perfect constitutional fit. *The Chinese Journal of Comparative Law*, *7*(2), 266–284.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Hum Nath Bhandari, R. K. C. K., Binod Rimal, Nawa Raj Pokhrel, Ramchandra Rimal, Keshab R. Dahal. (2022). Predicting stock market index using LSTM. *Machine Learning with Applications*, *9*, 100320-.

Irshad, H. (2017). Relationship among political instability, stock market returns and stock market volatility. *Studies in Business and Economics*, *12*(2), 70–99.

Jadhav, P. (2012). Determinants of Foreign Direct Investment in BRICS Economies: Analysis of Economic, Institutional and Political Factor. *Procedia – Social and Behavioral Sciences*, *37*, 5–14.

Jordan, F. (2016). Back to the future betas: Empirical Asset Pricing of us and southeast asian markets. *International Journal of Financial Studies*, *4*(3), 15–15.

Kubat, M., Bratko, I., & Michalski, R. (1998). A review of machine learning methods. In *Machine learning and data mining: Methods and applications* (pp. 3–69). Springer.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economic Statistics*, *13–37*.

Ma, Y., Han, R., & Wang, W. (2021). Portfolio optimization with return prediction using deep learning and Machine Learning. *Expert Systems with Applications*, *165*.

Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, *9*(1), 381–386.

Mai, Z., Nawaz Saleem, H., & Kamran, M. (2023). The relationship between political instability and stock market performance: An analysis of the MSCI index in the case of Pakistan. *Plos One*, *18*(10), e0292284.

Mao, Y., Pranolo, A., Wibawa, A., Utama, A., Dwiyanto, F., & Saifullah, S. (2022). Selection of precise long short term memory (LSTM) hyperparameters based on particle swarm optimization. *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 1114–1121.

Markowitz, H. M. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91.

Mehtab, S., Sen, J., & Dutta, A. (2021). Stock price prediction using Machine Learning and LSTM-based deep learning models. In *Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, Chennai, India, October 14–17, 2020, Revised Selected Papers 2* (pp. 88–106). Springer Singapore.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica 34 (4)*, 768–783.

Pongsudhirak, T. (2014). Thailand: The Coup and its consequences. In R. Huisken, O. Cable, D. Ball, A. Milner, R. Sukma, & Y. Wanandi (Eds.), *CSCAP REGIONAL SECURITY OUTLOOK 2015* (pp. 46–48). Council for Security Cooperation in the Asia Pacific. http://www.jstor.org/stable/resrep22263.14

Purnaningrum, E., & Fariana, R. (2022). Dynamic ensemble time series for prediction major indices in asean. *Indonesian Journal of Social Research (Ijsr)*, *4*(1), 54–62.

Salim, W., & Negara, S. D. (2018). Infrastructure Development under the Jokowi Administration: Progress, Challenges and Policies. *Journal of Southeast Asian Economies*, *35*(3), 386. https://doi.org/10.1355/ae35-3e

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425–442.

Singh, B. (2004). The challenge of militant Islam and terrorism in Indonesia. *Australian Journal of International Affairs*, *58*(1), 47–68. https://doi.org/10.1080/1035771032000184764

Stam, M. H. (2022). *Asset Pricing and the applications of Machine Learning in missing data treatment* [PhD Thesis]. Tilburg University.

Teeramungcalanon, M., Chiu, E., & Kim, Y. (2020). Importance of political elements to attract FDI for ASEAN and Korean Economy. *Journal of Korea Trade*, *24*(8), 63–80.

Theang, K., Zhang, L., Chee Ghee, T., & Rangel, G. (2019). Multifactor Asset Pricing model: Contemporary evidence from Malaysia. *SSRN. 2024*.

Treynor, J. L. (1961). *Toward a theory of market value of risky assets*.

Tukey, J. W. (1949). Comparing Individual Means in the Analysis of Variance. *Biometrics*, *5*(2), 99–114.

2. Databases:

Daniel Kaufmann and Aart Kraay (2023). Worldwide Governance Indicators, 2023 Update (www.govindicators.org), Accessed on 10/19/2023.

Key Indicators Database – Asian Development Bank. (n.d.). Key Indicators Database – Asian Development Bank.

Monthly World Indices. (2024, April 8. Wharton Research Data Services. https://wrds-web.wharton.upenn.edu/wrds/)

3. Other:

ASEAN Economic Data | Data| World Economics. (n.d.). World Economics. https://www.worldeconomics.com/Regions/ASEAN/

ASEAN. (2023). ASEAN Key Figures - 2023. In Asean.org. https://asean.org/serial/asean-key-figures-2023/

ASEAN. (n.d.-b). Member States - ASEAN. ASEAN.org. https://asean.org/member-states/

International Monetary Fund. (2019, November 1). ABOUT OAP. IMF. https://www.imf.org/en/Countries/ResRep/OAP-Home/oap-about

KPMG. (2017). Politics, Power, and Change: What's Next for ASEAN. In KPMG. https://kpmg.com/mm/en/home/insights/2017/09/politics-power-and-change-what-next-for-asean.html

Lawrence, S. J. (2023, April 18). What is LSTM? - Introduction to Long Short-Term Memory - Scaler Topics. Scaler Topics. https://www.scaler.com/topics/deep-learning/lstm/

Molnar, C. (2023, August 21). *8.5 Permutation Feature Importance | Interpretable Machine Learning*. https://christophm.github.io/interpretable-ml-book/feature-importance.html#theory-3

Singh, U. B. (n.d.). Indonesia: From economic crisis to political turmoil. Strategic Analysis: A Monthly Journal of the IDSA. https://ciaotest.cc.columbia.edu/olj/sa/sa_nov00siu01.html#txt10

The Global Financial Centres Index 34. (n.d.). https://www.longfinance.net/publications/long-finance-reports/the-global-financial-centres-index-34/

What is Regularization and how does it help to avoid overfitting in a machine learning model? | Kaggle. (n.d.). https://www.kaggle.com/discussions/general/384887

Wikipedia contributors. (2024, March 16). *Long short-term memory*. Wikipedia. https://en.wikipedia.org/wiki/Long_short-term_memory#:~:text=March%202022),and%20other%20sequence%20learning%20methods.

VII. Appendix:

Definitions of the Political Indicators according to the World Bank.

Voice and Accountability: Reflects perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.

Political Stability No Violence: Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism.

Government Effectiveness: Reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.

Regulatory Quality: Reflects perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.

Rule of Law: Reflects perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.

Control of Corruption: Reflects perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.

Table 6

Hyperparameters selected for LSTM with Political Factors

Hyperparameter	**Suggested Values**
Number of hidden Layers	[3]
Number of neurons in each layer	[32]
Dropout rate	[0.05]
Learning rate	[0.0068]
Batch size	[10]
Epoch size	[100]

Table 7

Hyperparameters selected for LSTM without Political Factors

Hyperparameter	**Suggested Values**
Number of hidden Layers	[2]
Number of neurons in each layer	[16]
Dropout rate	[0.1]
Learning rate	[0.00028]
Batch size	[10]s
Epoch size	[100]

Table 8

Result of Permutation Feature Importance by Country and Political Indicator

	IDN	*SGP*	*PHL*	*MYS*	*THA*
RQ	0,000000123354	0,000000029394	0,000000004571	0,000000003613	0,000005047762
COC	0,000000061713	0,000000000923	-0,000000000037	-0,000000004315	0,000035247273
GOV	0,000000011578	-0,000000025231	0,000000001191	-0,000000010598	-0,000015548798
VAA	-0,000000007020	0,000000216832	-0,000000000713	-0,000000007571	0,000029429217
ROL	-0,000000170381	-0,000000071036	-0,000000000048	0,000000006772	-0,000037062322
POL	-0,000000471756	-0,000000077250	-0,000000000405	-0,000000020745	0,000020498322

Note. This table presents the unscaled results for Permutation Feature Importance which are expressed in terms of mean change in MSE over iterations. The scale of the results varies largely across countries, with Thailand scoring the range of 10^{-6}-10^{-5}, Indonesia 10^{-8}-10^{-7} and the other countries 10^{-11}-10^{-8}. Consequently, those scores were scaled to ensure easier visual comparison.

Table 9

Results of the ANOVA test

Source	*Sum of Squares*	*df*	*Mean Square*	*F*	*Sig.*	*Partial Eta Squared*
Corrected Model	5,024E-9[a]	29	1,732E-10	50,553	0,000	0,844
Intercept	1,635E-10	1	1,635E-10	47,724	0,000	0,150
Country	0,000	4	0,000	0,000	1,000	0,000
Factor	9,338E-10	5	1,868E-10	54,503	0,000	0,502
*Country * Factor*	3,781E-09	20	1,890E-10	55,164	0,000	0,803
Error	9,252E-10	270	3,427E-12			
Total	6,112E-09	300				
Corrected Total	5,949E-09	299				

Note. This table expresses the output of the ANOVA test. It can be observed that the overall model referred to as the corrected model shows significance and the results can be further analysed by group, including the interaction term. To get a better understanding of the pairwise differences in mean Tukey's HSD Post Hoc test is also performed and presented in the "Results" section.

Limitations:

Regardless of all the effort that this work required, it still suffers from some limitations. As previously mentioned, Machine Learning can be very data and computationally demanding. Consequently, the present research had limited means to calculate heavy models. Considering this, the model could have potentially attained better results with lower values for batch size (<10) and higher epochs (>100). However, with the models already taking 25-30 minutes to run on a personal computer it was not feasible to test with more demanding values, which could potentially contribute to enhance the model's ability to learn and generalise to unseen data. Additionally, for the model the random search only had 50 iterations. Bergstra & Bengio (2012), suggest using a value in a range of 100 and 1000 depending on the model's complexity, the number of hyperparameters and the spacing between the suggested values. Increasing the number of iterations of random search would significantly increase the chances of identifying the most efficient hyperparameter values and result in better overall performance. Finally, the interpretation of the significance of the different variables could have been made more understandable and reliable using PIMP permutation importance (Altmann et al., 2010). This method builds on Permutation Feature Importance and additionally returns the p-value of each variable. To attain this, it computes the Permutation Feature Importance scores for permuting the target variable (instead of features) to construct a null distribution. This means that the relation between the predictors and the target is broken, and the outcomes can be attributed to random chance. Subsequently, the scores for Permutation Feature Importance are calculated as usual (permuting the variables' values multiple times) and compared against the distribution to assess the p-value. However, to achieve this, it is necessary to perform a lot of permutations to create a representative distribution and, when attempted, required too much computational power. Furthermore, in this study Permutation Feature Importance was performed with 10 permutations, a higher number would ensure better robustness.

Lastly, using more frequently updated data could potentially yield better results, as the model would have more variations to analyse. Political and economic indicators were only updated yearly. Consequently, using more frequently updated variables would make the model more sensitive reflecting more accurately the dynamics at play. This should lead to better performance due to the reduction in the lag between the occurrence of an event and how it is reflected in the data.