The University of Hong Kong

FITE4801 Final Year Project

AI-based Real-time and Early Detection of Financial Institution Vulnerabilities in Hong Kong

Interim Report

Student name	**UID**
Lee Man Kit	3035783637
Chan Chun Ho	3035741029
Chen Hsin-mei	3035685847
Chang Hsi-yun	3035729588

Supervisor

Dr. Liu Qi

21 January 2024

Abstract

With heightened regulation standards after the 2008 global financial crisis and the ever-changing financial landscapes in recent years, regulators around the world have started applying supervisory technology (SupTech) to their supervisory frameworks for enhanced supervision effectiveness and efficiency. Hong Kong suffers from not only a time-consuming and lengthy risk assessment method, but also the lack of adoption of SupTech solutions. The aim of this project is to develop a SupTech application for early detection and forecast financial vulnerabilities in local banks in an attempt to address these problems. This will be achieved through the computation of relative vulnerability index (RVI) on a peer group basis, alongside with the prediction of risk profiles using machine learning and natural language processing techniques, and the development of a dashboard providing key financial metrics. Currently, the project is halfway to completion, with the processing and model training on textual data still ongoing. The immediate next steps are model finetuning and front-end development.

Acknowledgements

We would like to express our special thanks to our project supervisor Dr. Liu Qi for his support and guidance. His valuable input has guided us to overcome various difficulties along the way and shaped the project to what it is today. This project would not have been the same without his guidance. We would also like to take this opportunity to thank each team member Lee Man Kit, Chan Chun Ho, Chen Hsin-mei and Chang Hsi-yun for their unique contribution and devotion to the project.

Table of Contents

List of Figures

List of Tables

Abbreviations

AI	artificial intelligence
API	application programming interface
BIS	Bank for International Settlements
CAMEL	capital adequacy (C), asset quality (A), management (M), earnings (E), liquidity (L)
ESMA	European Securities and Markets Authority
FSB	Financial Stability Board
GDP	gross domestic product
HKMA	Hong Kong Monetary Authority
IMF	International Monetary Fund
LSTM	long short-term memory
ML	machine learning
NLP	natural language processing
NLTK	Natural Language Toolkit
NN	neural network
OLS	ordinary least squares
P2P	peer-to-peer
RMSE	root mean squared error
RVI	relative vulnerability index
SupTech	supervisory technology
VADER	Valence Aware Dictionary and Sentiment Reasoner

1. Introduction

1.1 Background

Supervisory technology (SupTech) refers to the use of advanced technology to support supervisory authorities and regulators in supervision (Broeders & Prenio, 2018; Financial Stability Board [FSB], 2020). The SupTech trend is driven by two main factors: increased regulation standards following the 2008 global financial crisis and the recent changes in financial landscapes (Buckley et al., 2019; FSB, 2020). Ever since the 2008 financial crisis, regulatory authorities have been seeking more effective and efficient supervision methods to keep up with the raised international regulation standards (Dordevic et al., 2021). This has underlined the importance of adopting a data-driven approach in risk reporting and assessment (Dordevic et al., 2021). Additionally, financial activities being increasingly digitalized, paired with new business forms, such as cryptocurrency exchange platforms and peer-to-peer (P2P) lending, being introduced, has led to the creation of previously unknown types of financial risks, posing a potential threat to global financial stability (Buckley et al., 2019).

Given that SupTech often involves the use of machine learning (ML) and artificial intelligence (AI), the additional benefits of adopting SupTech may include: the ability to collect, store and analyse large amounts of data, identification of hidden patterns in datasets and real-time monitoring of financial risks (FSB, 2020). As a result, an increasing number of regulatory authorities worldwide have begun developing SupTech solutions and adopting them into their regulatory frameworks (European Securities and Markets Authority, 2019; FSB, 2020).

1.2 Current Situation in Hong Kong

Currently, the Hong Kong Monetary Authority (HKMA) adopts an integrated supervisory framework, where the CAMEL rating system is incorporated into the original risk-based approach (HKMA, 2022). There are five elements being measured in the CAMEL rating system: capital adequacy (C), asset quality (A), management (M), earnings (E) and liquidity (L) (HKMA, 2022). Each CAMEL element can take a score ranging from 1 to 5, where 1 stands for "of the least supervisory concern" and 5 for "of the most supervisory concern" (HKMA, 2022). The ratings of the five elements are assigned based on the risk profile of a financial institution. Determining the CAMEL scores requires the HKMA to go through a lengthy multi-step process, which includes several prudential meetings, on-site visitations and off-site examinations (HKMA, 2022). The HKMA has put forward a general outline of its three-year SupTech journey in its newsletter published in 2019 (HKMA, 2021a). The outline specifies milestones for each of the three phases and the proposed technologies to be implemented (HKMA, 2021a).

1.3 Problems Identified

Two major limitations were found in the supervisory framework currently implemented by the HKMA: the lengthy risk assessment process and the lack of SupTech solutions. While the risk assessment approach is comprehensive, it is rather time consuming. It is estimated that running an entire risk assessment cycle will take up to months or even a year. This may hinder the HKMA from acting promptly in the face of potential financial problems, thus lowering the forward-looking capability of the HKMA. Moreover, according to the HKMA, they are still in the early stages of the SupTech journey (HKMA, 2021a). In this highly digitalized age, it is crucial for supervisors to utilize SupTech in order to keep up with the large amount of data created every day and the ever-changing financial landscapes. The lack of SupTech adoption by the HKMA may cause its supervision to be less effective, leading to the formation of a regulatory gap.

1.4 Objectives, Scope and Deliverables

The main aim of this project is to address the problems mentioned above by developing a SupTech application to achieve early detection and forecast of financial institution vulnerabilities in Hong Kong using ML and natural language processing (NLP) techniques. There are three main components to the SupTech application:

1. **Real-time monitoring of financial data**: The latest financial news, numerical data and risk assessment reports are delivered by the web-based dashboard in real-time. This allows for real-time detection and monitoring of financial institution vulnerabilities.

2. **CAMEL-based risk profile prediction**: ML techniques are employed to predict the risk profiles of financial institutions more accurately. This provides insights into the general trend of the financial performance of each financial institution, thus strengthening the forward-looking capabilities of regulators.

3. **Relative vulnerability index (RVI)**: This is a relative risk rating system. The RVIs are computed on a group basis and serve as an indicator of the relative amount of financial risk a financial institution is exposed to as compared to other financial institutions in the group. This facilitates comparison between financial institutions and quick identification of those that are more at risk of failure.

The ultimate goal of this project is to provide an additional tool for the HKMA, offering faster results, more accurate predictions and real-time financial data, to assist them in risk assessment. In the case that the project is successful, it is hoped that the final product can help to refine the supervisory framework in Hong Kong, and thus closing regulatory gaps. The main deliverables of this project will be the trained models resulting from ML and the web-based dashboard.

1.5 Outline

The remaining report is organized as follows. Section 2 provides a detailed description of the methodology of the project. Section 3 states the current status of the project and future plans. Section 4 reports and discusses the results. Finally, Section 5 puts forward the conclusion of the report.

2. Methodology

2.1 CAMEL-based Risk Profile Prediction

To assign supervisory ratings for early detection of vulnerabilities, the risk profile of a financial institution is instrumental. In this project, financial ratios are considered as proxies for the risk profile of financial institutions. This is found on the fact that financial ratios have long been crucial indicators to assess the risks attendant on financial institutions (Van Greuning & Bratanovic, 2009). Furthermore, in linkage with the CAMEL rating system in the current supervisory framework, it is crucial for the financial ratios to quantify risks relating to the five areas of the supervisory rating. Thus, five financial ratios ("representative ratios") were selected to represent each area of the CAMEL rating, of which the specifics are summarized as follows:

- **Capital adequacy**: The common equity tier 1 capital ratio was chosen to assess the solvency risk of financial institutions because it is prioritized by the HKMA in supervision for high capital strength (HKMA, 2020). A higher value of the ratio suggests the financial institution has better capital adequacy, and hence lower solvency risk.

- **Asset quality**: The non-performing loan coverage ratio was selected to evaluate the credit risk exposed by financial institutions for its common usage by central banks in reporting the asset quality (European Central Bank, 2023; HKMA, 2023). A smaller value of the ratio implies the financial institution has better asset quality and less credit risk.

- **Management**: The cost-to-income efficiency ratio was used to reflect the management capability of financial institutions relating to their operational efficiency. It has been widely used and preferred by analysts in banks and credit agencies (European Central Bank, 2010). The lower the ratio value, the better management capability of the financial institution.

- **Earnings**: The return on risk-weighted assets ratio was used to gauge the profitability of financial institutions. As risk-weighted assets are considered, this metric is robust against excessive risk-taking and thus is more consistent in the long run (European Central Bank, 2010). A larger ratio value means better risk-adjusted returns for the financial institution.

- **Liquidity**: The liquid assets to short-term funding ratio was chosen to monitor the liquidity risk of financial institutions. This ratio was considered to resemble the regulatory ratio, i.e., the liquidity coverage ratio, which are not applicable to all financial institutions in Hong Kong due to differentiated regulatory requirements (HKMA, 2019). A larger value of the ratio would imply higher liquidity of the financial institution.

More importantly, given that foreseeability is a cornerstone in banking supervision for early detection, the representative ratios will be forecasted to simulate forward-looking risk profiles. As such, these representative ratios denote the dependent variables for five separate models. In each model, the ratio will be predicted by independent variables using the past representative ratio and macro-factors which comprise five macroeconomic indicators, i.e., real gross domestic product (GDP), inflation rate, unemployment rate, 3-month interbank borrowing rate, stock market index, and one bank confidence index. These macroeconomic indicators were chosen because they are frequently used by regulators to perform stress-testing, an analysis to test whether financial institutions can withstand stressed market scenarios (Bank of England, 2022; Federal Reserve Board, 2022). In addition, the bank confidence index was intended to gain useful information on the market sentiment of financial institutions. Considering the above, the representative ratio can be formulated by a function of independent variables and the error term in the equation below:

$$Y_{it} = f(Y_{it-s}, X_{it-s}, u_{it}) \quad \textbf{(2.1)}$$

where Y_{it} denotes the representative ratio of financial institution i at time t; Y_{it-s} is the s-period lagged representative ratio; X_{it-s} is a $K \times 1$ vector of s-period lagged macro-factors; and u_{it} is the error term at time t.

2.1.1 Data Collection and Processing

Numerical data

To collect numerical data on financial information, Moody's Analytics Orbis Bank Focus was utilized (Moody's Analytics, 2023). Although the database could trace back annual data from the past 30 years in the search setting, the actual data available in the database for local financial institutions was found to be much less, with only 8 to 9 years on average for all financial information. This could be problematic for time-series forecasting due to insufficient amount of data for satisfactory prediction accuracy. Hence, it was decided to pool time-series data from different financial institutions to form a panel dataset so that the issue of limited data availability can be addressed (Cheng, 2007). Furthermore, to provide a fair amount of data for model training, data on the representative ratios were further collected from four other jurisdictions, i.e., Canada, China, Singapore, and the United States, which are subject to comparable supervisory frameworks based on their high degree of compliance with the Basel core principles for effective banking supervision and consistent implementation of the Basel III framework (Bank for International Settlements [BIS], 2023; Garcia & de Mendonça, 2023). Following the decision on the selected jurisdictions, annual data on the macroeconomic indicators in the past 10 years were also sourced from public repositories of the jurisdictions respectively.

After gathering the numerical data, data cleansing was performed by standardizing the data format and removing irrelevant data such as duplicated values and financial institutions with missing observations. More critically, considering that the data covers financial institutions from other jurisdictions, this can run the risk of using time-series data that are uncommon among local financial institutions, thereby distorting the original intention of supplementing local data. To minimize the discrepancies between local and non-local data, similarity search was applied to time-series data on the representative ratios. In particular, the Mahalanobis distance approach was adopted in lieu of conventional Euclidean distance approach since research had shown that time-series similarity search was more accurate when the data distribution was considered (Arathi & Govardhan, 2015). To illustrate, the Mahalanobis distance can be computed as

$$d_M = \sqrt{(x_i - \mu_{local})' \Sigma_{local}^{-1} (x_i - \mu_{local})} \qquad \textbf{(2.2)}$$

where x_i is a vector of time-series data on financial institution i, μ_{local} and Σ_{local} are the parameters of the distribution of the local data to be estimated. In this light, the gaussian mixture model[1] was employed to evaluate the mean μ_{local} and covariance matrix Σ_{local} of the distribution of the local data. With these parameters, the Mahalanobis distance between every financial institution and the centroid of the local data was computed. Only non-local financial institutions

[1] The Gaussian Mixture Model provides the flexibility of adjusting the structure of the covariance matrix. This allows high-level data distribution modelling and prevents excessive removal of financial institutions from the dataset.

with a Mahalanobis distance smaller than a threshold, i.e., the maximum Mahalanobis distance between time series data of local financial institutions and their centroid, were kept in the dataset. Eventually, the processed dataset of financial information was combined with the macroeconomic dataset and converted into a panel data format. This led to an 8-year panel dataset of 16 indicators[2], spanning from 2015 to 2022 for 2,401 financial institutions, i.e., 19,208 bank-year observations.

Textual data

Textual data, in the form of newspaper articles, was sourced from the Guardian's Open API, and the New York Times' Developer API. Articles relating to the banks on the dataset were retrieved from the two free APIs, and were given a sentiment score using dictionary-based sentiment analysis, following text data normalization (removal of stopwords). The dictionary used in this analysis was the Loughran-McDonald Master Dictionary with Sentiment Word Lists (hereafter "Dictionary"), and the sentiment score for each article was calculated as follows:

$$\text{sentiment} = (\text{positive words}/\text{length of article}) - (\text{negative words}/\text{length of article}) \quad \textbf{(2.3)}$$

where positive words and negative words refer to words with a positive value in the "Positive" and "Negative" columns of the Dictionary respectively.

Each bank is given a sentiment score based on the arithmetic mean of the sentiment scores of the articles relating to it. The arithmetic mean was chosen as the aggregate for ease of computation and to preserve the magnitude of sentiment scores.

[2] The 16 indicators include 5 CAMEL ratios and 5 macroeconomic indicators for the risk profile prediction, as well as 6 balance sheet items for the RVI.

2.1.2 Model Implementation for Risk Profile Prediction

In banking supervision, econometric models have been extensively used by financial regulators to evaluate risk trends and perform macro stress tests on financial institutions. The HKMA has utilized panel data model to study the trend of capital adequacy ratio of local banks and performed macro stress tests for credit risk (HKMA, 2023; Wong et al., 2005). Nonetheless, given the rapidly changing business environment, a more robust method could be warranted to effectively capture the future risk trends of financial institutions. Indeed, recent research has looked into ML models for predicting financial ratios. For instance, Park et al. (2021) and Petropoulos et al. (2019) both utilized linear regression model and deep neural network to predict the capital adequacy ratio for banks. They all demonstrated that ML models produced better prediction accuracy than the econometric models. However, they only predicted financial ratios representing a single element of the CAMEL rating. From this, it would be useful to examine the application of machine learning models in predicting other financial ratios. Therefore, it is considered to implement both the econometric and machine learning models to forecast the five representative ratios.

To evaluate the model performance, a common metric, i.e., the root mean squared error (RMSE), is utilized given better interpretability and its ability to impose bigger penalty on larger errors (Muharemi et al., 2018). Moreover, considering that predicting the representative ratios of local banks is the primary objective, an indicator variable $\mathbb{I}(i \in HK)$ can be added in conjunction with the RMSE as with the one suggested by Liu et al. (2018) in their evaluation criterion, i.e.,

$$\mathrm{RMSE}_{\mathrm{HK}} = \sqrt{\frac{1}{T\sum_{i=1}^{N}\mathbb{I}(i \in HK)}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{I}(i \in HK)\left(Y_{it} - \widehat{Y_{it}}\right)^2} \qquad \textbf{(2.4)}$$

where $\mathbb{I}(i \in HK)$ is an indicator variable which returns 1 if it is a local financial institution, and 0 otherwise; Y_{it} is the representative ratio of financial institution i at time t; and $\widehat{Y_{it}}$ is the predicted value of the representative ratio. Eventually, the results of both models will be compared against each other to determine whether the machine learning model can provide genuine benefit over the current econometric model in risk profile prediction. The predictions of these models are also fed into the RVI for further comparison (see Section 2.2).

Econometric model

In line with the conventional econometric approach, this project deployed five dynamic panel data models to predict the risk profile represented by the selected ratios. In relation to the generalized model equation 2.1, the dynamic panel data model is a linear representation of the terms on the right-hand side of the equation, i.e.,

$$Y_{it} = \rho Y_{it-1} + X'_{it-1}\beta + u_{it} \qquad \textbf{(2.5)}$$

where the independent variables are in one lag. Among the lagged macro-factors X_{it-1}, it is noteworthy that logarithmic transformation was performed on the real GDP and stock market index which are not expressed in percentage, so as to gain the economic intuition of semi-elasticity. Moreover, due to the use of the panel data model, it is common to account for the unobserved individual fixed effects in the error term u_{it} to reflect the underlying heterogeneity that is correlated with the independent variables (Cheng, 2007). These individual fixed effects are time-invariant and entity-dependent characteristics such as the business models and financial market infrastructure in the respective jurisdiction. Accordingly, the error term u_{it} can be expressed as

$$u_{it} = \eta_i + \varepsilon_{it}, \qquad \varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2) \qquad \textbf{(2.6)}$$

where η_i are the individual fixed effects and ε_{it} is the white noise. Based on this specification, it is assumed that the fixed effects have linear relationship with the dependent variable Y_{it} and are independently distributed across financial institutions (Cheng, 2007). However, since these unobserved effects are inside the error term, leaving them untreated can cause biased estimations of the model parameters, and thus undermine the predictive power of the model (Bond, 2002). Hence, under the above assumptions, the first-differencing technique was taken to remove the fixed effects. Consequently, the model equation becomes

$$\Delta Y_{it} = \rho \Delta Y_{it-1} + \Delta X'_{it-1}\beta + \Delta\varepsilon_{it} \qquad \textbf{(2.7)}$$

With this revised model equation, the ordinary least squares (OLS) method was implemented to estimate the model parameters ρ and β through the OLS dependency in the statsmodels module (Statsmodels, 2023).

$$\widehat{\Delta Y_{it}} = \hat{\rho}\Delta Y_{it-1} + \Delta X'_{it-1}\hat{\beta} \qquad \textbf{(2.8)}$$

Specifically, data from 2015 to 2020 were used to train the model to obtain the model parameters. These estimated parameters were applied to the data from 2021 to 2022 to produce out-of-sample forecasts in model testing. The predictions on the dependent variable in first difference were then transformed back to the level value by adding its lagged value so that the model performance could be evaluated with the RMSE.

To find the best econometric models for each ratio, variable selection was performed with a subset selection method in view of a few independent variables involved. Under this method, $\sum_{k=1}^{6}\binom{6}{k} = 63$ subsets of variables were first generated by algorithm. The model was fitted on these subsets to evaluate the RMSE of the training and testing data on local banks. In assessing the overall performance of the model under each subset, a single performance score was computed using the weighted average of the RMSE_{HK} of the training and testing data in a 30:70 ratio to give preference to the testing result. Finally, the subset of variables which resulted in the lowest weighted RMSE_{HK} among other subsets was selected. This model alongside the chosen variables was compared against the machine learning model for subsequent analysis.

Machine learning model

In parallel to the econometric models, five machine learning models were trained to predict each of the CAMEL ratios. Initially, neural networks with a single long short-term memory (LSTM) layer were trained, as they are RNNs specialized for time-series data. The input data involved the first six years of data (the CAMEL ratio to be predicted and macroeconomic indicators) for each bank, and the CAMEL ratio of the last two years was used as the label for training.

However, the LSTM models were not adequate (for reasons to be explained in Section 4.1.2). An approach similar to the econometric models was thus taken, using a simpler dense neural network (NN). The input data for the models were 2-lag (i.e. a ratio was predicted based on data from the past two years), including the macroeconomic factors in year $t-2$ and the value of the ratio in years $t-2$ and $t-1$ to predict its value in year t. The architecture of the NNs are as follows:

1. Input NN layer with 128 units, ReLU activation and L1 regularizer
2. Hidden NN layer with 64 units and ReLU activation
3. Output layer with 1 unit, no activation

Two NN layers were used to find nonlinear relationships between the variables. ReLU activation was chosen to prevent the vanishing gradient problem, where the gradients with respect to weights exponentially decrease during the backpropagation process in model training, leading to ineffective weight updates (Grosse, 2017). The L1 regularizer acts as "feature selection," eliminating unnecessary or irrelevant variables during the training process.

Models were trained in batch sizes of 60, and in 40 training epochs. These values were obtained through experimentation, and loss was minimized using these parameters.

2.2 Relative Vulnerability Index

2.2.1 Index Equation

In this project, a Relative Vulnerability Index (RVI) was employed to provide an automatic assessment of the relative vulnerability levels (i.e., the risk) of different financial institutions. A standardized RVI can be expressed as follows:

$$\text{Standardized RVI}_t = Z_{C,t} + Z_{A,t} + Z_{M,t} + Z_{E,t} + Z_{L,t} \tag{2.9}$$

(Wong & Wei, 2023)

In accordance with the methodology proposed by Wong and Wei (2023) for constructing a bank vulnerability index, each CAMEL ratio at time t would be transformed into a z-score under a standard normal distribution. However, it is noteworthy that computing the vulnerability index over all sampled financial institutions can be unjustified because institutions of great significance in the financial system which are under stricter supervisory scrutiny often have better CAMEL ratio, thereby introducing potential bias in the one-size-fits-all z-score. Furthermore, it cannot be guaranteed that each CAMEL ratio of the sampled financial institutions strictly adheres to a normal distribution. Accordingly, the project addresses these challenges by devising appropriate strategies to mitigate the impact of inherent disparities and non-normal distributions.

2.2.2 K-means Clustering

A viable approach to solve the one-size-fits-all issue is to create subgroups for financial institutions. Indeed, this practice is commonly seen in the domestic and global supervisory framework on systemically important financial institutions, under which institutions are subject to regulatory requirements proportionate to their assigned buckets or ranks based on the values of some selected indicators in the previous year (BIS, 2016; Brogi et al., 2021). In this regard, financial institutions were categorized in respect of their level of systemic importance, i.e., the importance in the financial system[3]. Nonetheless, without considerable background knowledge, the procedures for categorizing financial institutions and determining the number of subgroups can be challenging. Hence, an unsupervised machine learning algorithm, namely K-means clustering, was considered (Educational Ecosystem, 2018). Essentially, K-means clustering is well-adapted to group data points based on their shared characteristics and features (Educational Ecosystem, 2018). Moreover, by evaluating the optimal number of K which minimizes the distortion loss, this can determine the suitable number of subgroups to be created. Thus, the algorithm was performed with different numbers of K, i.e., between 1 to 10, to categorize institutions and examine the optimal number each year. After that, each cluster label of a given year t was mapped to a rank label for the subsequent year $t+1$ based on the centroid value ($label_{k,t} \mapsto label_{r,t+1}$), assuming a higher value corresponds to a higher rank of systemic importance. Eventually, these rank labels were used for computing the RVI in the final step.

2.2.3 Transformation to Normal Distribution

In the context of K-means clustering, it is important to note that the resulting subgroups do not guarantee a normal distribution. Consequently, this study employed the Anderson-Darling test to assess the normality of the identified subgroups. The null hypothesis of the Anderson-Darling test states the sample follows a normal distribution, and this paper set the significance level to $\alpha = .05$. In instances where subgroups were found to deviate from a normal distribution, it was necessary to employ a transformation process to calculate the RVI. Several transformation techniques were considered, including log transformation, exponential transformation, Box-Cox

[3] The indicators include 6 balance sheet items, i.e., total assets, interbank assets, liabilities, customer deposits, loans, and liquid assets, adapted from the HKMA's Systemically Important Authorized Institutions (SIBs) framework and the BIS's Global Systemically Important Banks (G-SIB) framework (BIS, 2011; HKMA, 2021b).

transformation, power transformation, and quantile transformation. Among these options, the quantile transformation, a popular machine learning approach, is considered due to its ability to mitigate the impact of outliers and ensure input independence (StatsTest, 2020). Following the quantile transformation, another round of Anderson-Darling testing was conducted to verify if each transformed subgroup exhibited a normal distribution or demonstrated similarity to a normal distribution.

2.2.4 Compilation of Relative Vulnerability Index

After all subgroups were proven to be normally distributed, the transformed CAMEL ratios would be standardized into z-scores. It is important to note that asset quality and management ratios exhibit a negative relationship with the performance of financial institutions, meaning that a lower value correspond to better performance. Consequently, the additive inverse of these two elements was employed. For example, if the standardized z-score of asset quality for a specific company is 0.2, it signifies that $Z_{A,t} = -0.2$. By utilizing equation 2.9, the RVI for each financial institution could be subsequently calculated.

2.2.5 Examination of Predicted Relative Vulnerability Index

The prediction of the CAMEL ratios, as discussed in Section 2.1, serves as a foundation for predicting the RVI using equation 2.9. Both the econometric and machine learning models were utilized to calculate predicted RVIs and were evaluated based on median values and standard deviations. Given that the Anderson-Darling test solely provides a probability (i.e., p-value) to reject the null hypothesis, the sample does follow a normal distribution, it is not suitable to deploy a paired sample t-test to analyze the differences between the RVI calculated using real CAMEL ratios and the predicted RVI calculated using predicted CAMEL ratios. Instead, the application of the Wilcoxon test, or Wilcoxon signed-rank test, proves more suitable for comparing the two distributions. The null hypothesis of the Wilcoxon test states no significant difference between the two medians of the distributions, and this project set the significance level to $\alpha = .05$. That is, a higher p-value in the Wilcoxon test indicates a greater level of confidence in assuming that the median prediction aligns with the real value. Furthermore, the percentage of predictions falling within one and two standard deviations were calculated to examine deviations between the predicted RVI and the actual RVI.

2.3 Web Application

2.3.1 Frontend Development: Dashboard

The web-based dashboard will show the key statistics of each bank, including historical and predicted risk profiles, RVIs and market sentiment. The dashboard will be written using React, which is a commonly used JavaScript library for developing user interfaces.

2.3.2 Backend Development

The backend is mainly comprised of the server and database. The server will act as a messenger, retrieving and sending data between the database and the frontend. It will also temporarily store images and perform simple computations. The server backend will be written using Express.js, which is a Node.js library that provides high flexibility and simplicity for server backend programming. In addition, the server backend will communicate with a MongoDB database to store and retrieve application data. MongoDB was chosen for its ease of integration with JavaScript, and flexibility in its document structure.

2.3.3 Dashboard User Interface Design

Figma was utilized to design the user interface and visualize ideas. The main features of the dashboard are described below:

1. **Main menu:**

   The dropdown main menu is on the top righthand corner of the screen. There are four main options: “all”, “group 1 – large financial institutions”, “group 2 – medium financial institutions” and “group 3 – small financial institutions”. Clicking on “all” will bring users to the homepage. If groups 1, 2 or 3 are selected, a submenu will appear, showing a list of banks in the peer group for the user to choose from.

2. **Toolbar:**

   The toolbar sits on the left of the screen. It consists of 5 icons. Users can hover their mouse over the icons to reveal the labels: home, dashboard, news, messages and help. The toolbar helps users to navigate between the homepage, dashboard page, news center, notification center and help center. The “Dashboard” and “News” tabs are only available after entering “individual bank mode”.

3. **Individual bank mode:**

   Individual bank mode allows users to view bank-specific information, such as the risk assessment results of the bank and relevant news. Individual bank mode is implemented instead of showing results of all banks together to keep the webpage clean and avoid sensory overload. To enter individual bank mode, users must select a bank first. This can be achieved through selection from the main menu or homepage.

4. **Homepage:**

   The grouping of the banks is displayed on the homepage. There are three containers, one for each peer group: large, medium and small financial institutions. Each container provides a full list of all the banks belonging to the corresponding peer group. Hyperlinks are embedded in which clicking on one of the bank names will bring the user to the individual bank mode of that particular bank. The search bar facilitates the search process when users would like to check which group a bank belongs to.

5. **Dashboard pages:**

   The key financial statistics of each bank are clearly displayed in the “Dashboard” of the respective bank. There are five key components on the page: the predicted CAMEL risk profiles, predicted RVI, historical CAMEL risk profiles, historical RVI and market sentiment. The predicted CAMEL risk profiles and RVI show the forecasted ratings for 2 years. The assigned peer group is shown below the predicted RVI. Historical data of CAMEL risk profiles and RVIs are visualized as broken line graphs. Users can toggle the timeline button to switch between 5-year, 10-year and all-time graphs. The colors of the line graphs are made to match the background color of representative ratio predictions of each CAMEL element for ease of monitoring and enhanced user experience. Market sentiments are represented by a stacked bar chart. The lengths of the green and red bars indicate the percentage of news or social media posts with positive and negative sentiment scores respectively. The headlines of the top 2 positive and negative news articles are also shown.

6. **Notification center:**

   The “Messages” tab keeps users up to date with the latest news and risk assessment results. Users will receive notifications for published news articles mentioning the banks. They will also be alerted when the predicted CAMEL risk profile, predicted RVI, direction of market sentiment or assigned peer group of a bank has been revised.

7. **News center:**

   The “News” tab is available for all banks under individual bank mode. It is a collection of relevant news articles published recently. The page displays the new article headlines. Users can click on the headlines to view the full article. The articles are arranged according to the strength of sentiments. Articles with the most positive and negative sentiments will be listed first.

8. **Help center:**

   The “Help” page provides assistance to new users. It documents questions and answers to frequently asked questions. This includes introduction of basic features, explanations of the terminologies used etc.

3. Project Status

3.1 Current Progress

Thus far, the project progress has been largely on the schedule, with most of the tasks in the backend development of the SupTech application completed. Essentially, both the numerical and textual data required for the CAMEL-based risk profile prediction have been collected. However, only the numerical data was fully processed for model fitting because the team encountered problems with the processing of textual data (see Section 4.2). Notwithstanding this, the econometric and machine learning models were first trained and tested on numerical data to have a grasp of the forecasting performance with these conventional indicators. The results of the model implementation are set out in detail in Section 4. In the meantime, the compilation of RVI from the actual and predicted CAMEL ratios based on the econometric and machine learning models was done. In addition, evaluation on predictions of both models was performed. The user interface design of the web-based dashboard has also been completed using Figma.

3.2 Future Plan

In the following months, the focus of the project will shift from the econometric/machine learning models to the implementation of the web-based dashboard. A Github repository will be created for the dashboard code, and a CS department virtual machine will be used to run the server and database. Additionally, methods proposed in Section 4.2 will be tested out to tackle the current problems of textual data processing. Meanwhile, the existing econometric/machine learning models will continue to be refined through further tuning of hyperparameters for individual models.

After completing the development of the dashboard, the full application will be tested with real users. Users will give feedback based on the presentation and usefulness of the application.

4. Results and Discussion

4.1 Findings

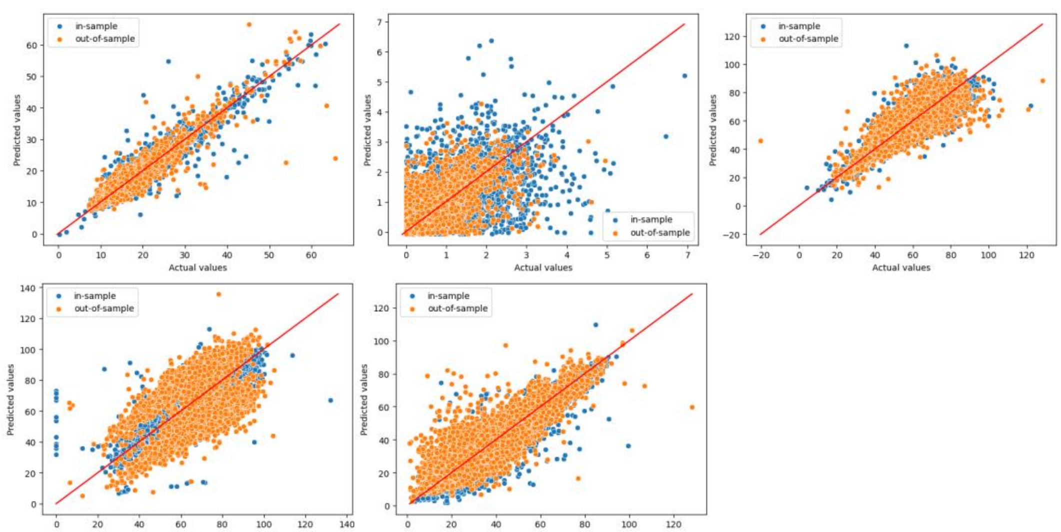
4.1.1 Econometric Model

In the econometric model, $\sum_{k=1}^{6}\binom{6}{k} = 63$ subsets of independent variables were performed. After examination of predicted results, the best-weighted RMSE_{HK} and its corresponding subsets of independent variables are shown in Table 1:

	CAMEL	Inflation rate	3-month interbank rate	Unemploy-ment rate	Real GDP	Stock market index	Training RMSE_{HK}	Testing RMSE_{HK}	Weighted RMSE_{HK}
C	V	V			V		1.4679	1.1433	1.2407
A	V		V			V	0.6505	0.5944	0.6112
M	V			V	V		3.5536	5.0957	4.6331
E	V						3.4672	3.0943	3.2062
L	V						7.1377	4.8449	5.5327

Table 1: The selected variables for each CAMEL ratio and corresponding RMSE_{HK}

It should be noted that the independent variables of Table 1 were represented as the first-differencing outcome with 1 lag as derived from equation 2.7. Figures 1-5 then exhibit the graphical representation of the predicted values alongside the actual values for each CAMEL ratio.



Figures 1-5: From left to right and top to bottom, the comparison plots display the actual and predicted values of C, A, M, E, and L ratios, respectively.

In addition to the general predicted outcomes for CAMEL ratios displayed in Figures 1-5, individual predictions for each local financial institution are also presented (as exemplified in Figure 6). The comprehensive results encompassing all CAMEL ratios for the 22 financial institutions in Hong Kong are anticipated to be included in the front-end design.

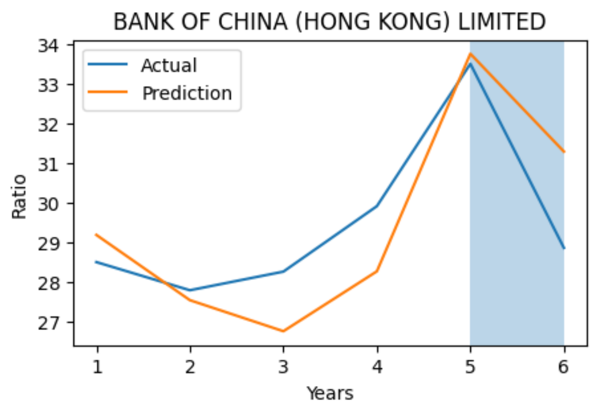


Figure 6: The example of tailored prediction results specifically focusing on the management (M) ratio of the Bank of China (Hong Kong) Limited. The figure showcases a visual comparison between the actual and predicted M ratio, with the testing set being depicted by the blue area.

4.1.2 Machine Learning Model

The decision to move away from LSTM models and to NN models was primarily based on two reasons:

1. The dataset has only 8 points in time total (6 in training), making it insufficient for an LSTM to make predictions on trends, causing it to make inconsistent predictions.
2. The dataset was small, covering only 2,401 banks, making an LSTM model too complicated for the problem.

Figure 7 shows some inconsistent predictions made by the LSTM models:

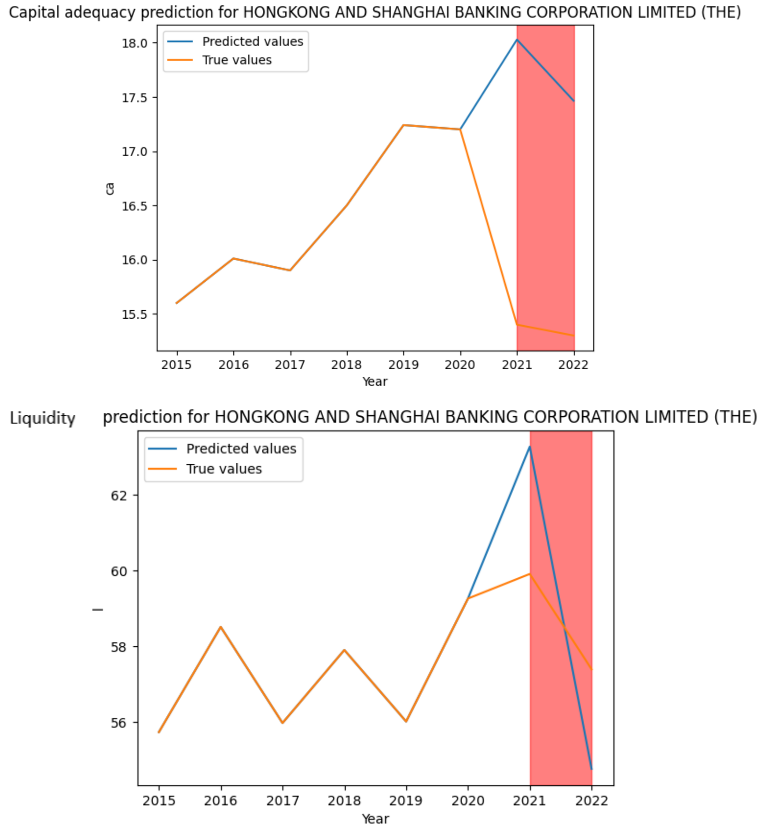


Figure 7: Prediction of the capital adequacy (CA, top) and liquidity (L, bottom) ratio of The Hong Kong and Shanghai Banking Corporation Limited made by the LSTM models. The figure showcases a visual comparison between the actual and predicted CA/L ratio, with the prediction area being depicted in red.

As shown in the graphs above, even with optimized RMSE values (2.3027 and 6.7920 RMSE for the CA and L models respectively), the predictions themselves are relatively far away from the actual values.

On the other hand, preliminary NN models have been completed, and they can somewhat closely predict the values of the ratios. Figure 8 (overleaf) is a visualization of the predictions of the capital adequacy (CA) and liquidity (L) ratios for The Hong Kong and Shanghai Banking Corporation Limited:

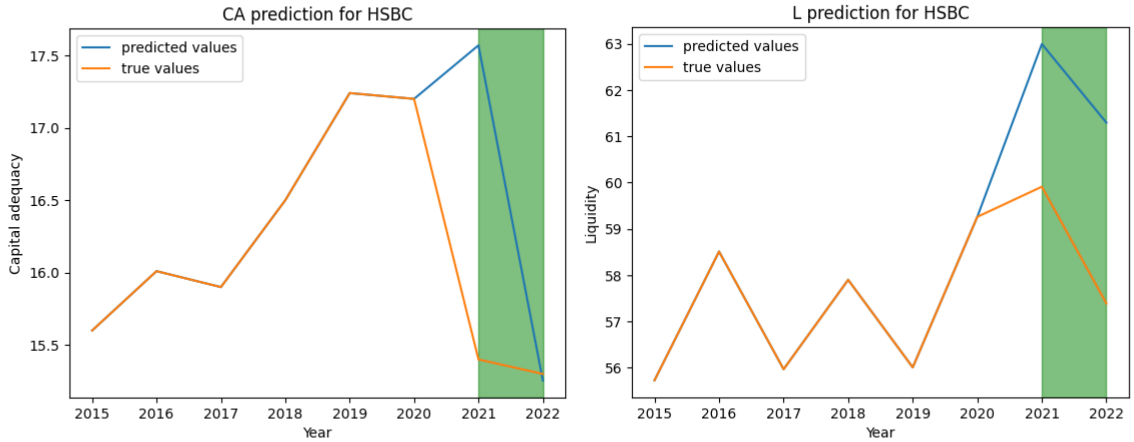


Figure 8: Prediction results of the capital adequacy (CA, left) and liquidity (L, right) ratio of The Hong Kong and Shanghai Banking Corporation Limited. The figure showcases a visual comparison between the actual and predicted CA/L ratio, with the prediction area being depicted in green.

While the NN models correctly predict the trends of the ratios, they have yet to be able to closely predict the true values of the ratios in 2021 and 2022.

Comparison of RMSE with Econometric Models

Below is a table (Table 2) summarizing the RMSE values (on the testing set) of the machine learning models:

Type of ML model	Capital adequacy	Asset quality	Management	Earnings	Liquidity
LSTM	2.3026	0.4298	6.9524	6.7920	8.7925
NN	1.9827	0.4565	6.7175	9.3528	8.9175
NN (HK only)	1.2660	0.7713	6.0473	6.2215	6.2551

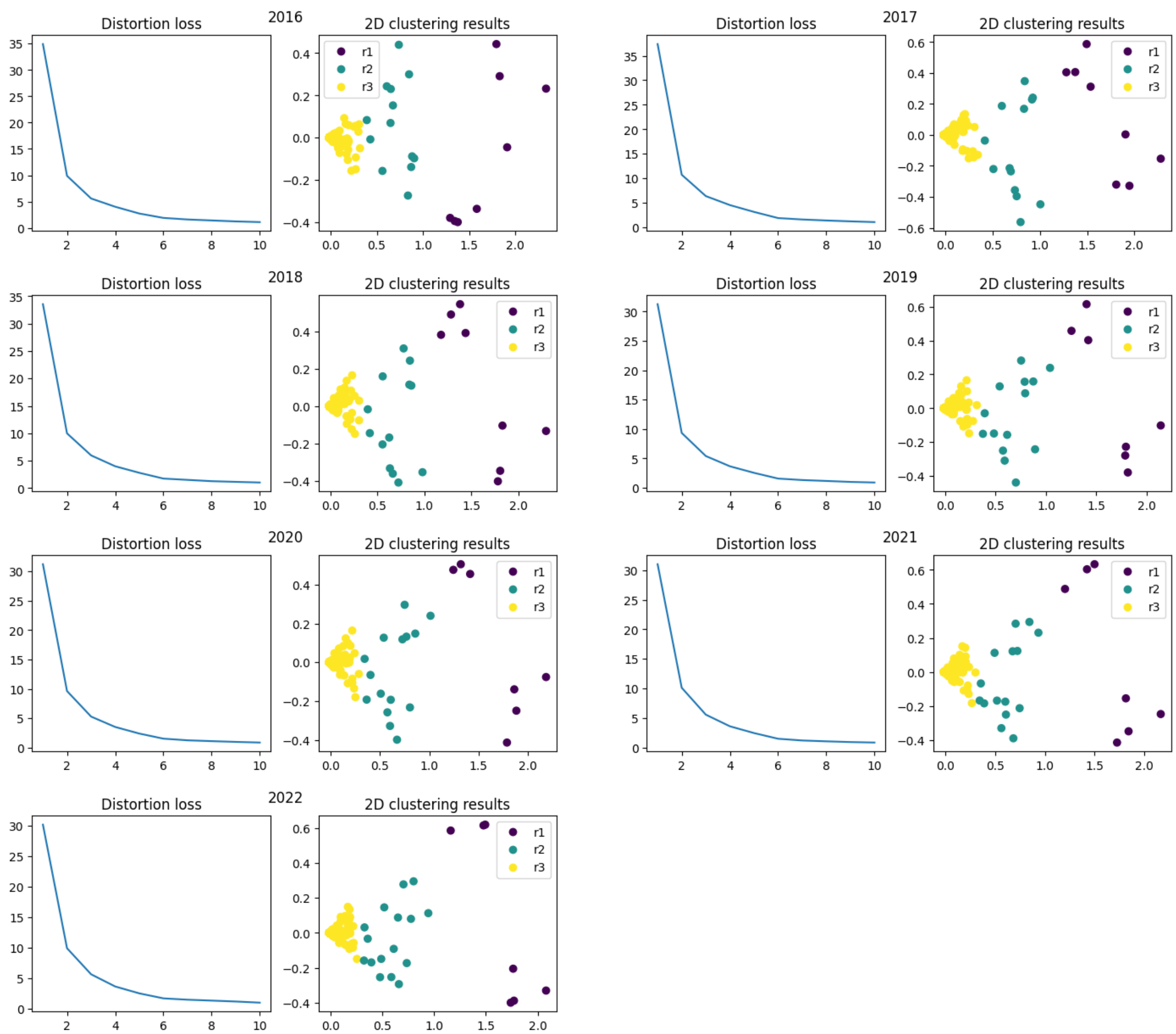
Table 2: The testing RMSE of every CAMEL ratio under the machine learning models

Currently, the machine learning models have higher RMSE values than the econometric models (Table 1), but these values are expected to be lower as the models are individually further tuned and optimized, bringing their accuracy closer to that of the econometric models.

4.1.3 Relative Vulnerability Index

Clustering

As mentioned in Section 2.2, the progress of compiling RVI can be divided into three steps: clustering, transformation to the normal distribution, and the final compilation. K-means clustering for 7 years (2016-22) was conducted. Using the elbow method through graphical representation of the distortion loss, it was found that the best clustering size was $K = 3$ for every year held, as shown in Figures 9-15:



Figures 9-15: From left to right and top to bottom (2016-22). Each figure contains two subplots: the distortion loss and the 2D clustering results.

Thus, the total subgroups created are $105 = 7(\text{years}) \times 5(\text{ratios}) \times 3(\text{clustering sizes})$.

Transformation to Normal Distribution

After preliminary Anderson-Darling tests, none of the 105 subgroups followed normal distributions (i.e., $p < .05$), as exemplified in Figure 15. Therefore, quantile transformation was utilized to transform the 105 samples into normal distributions.

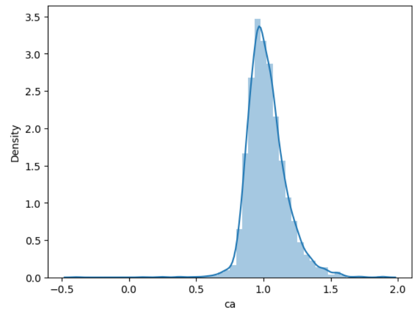
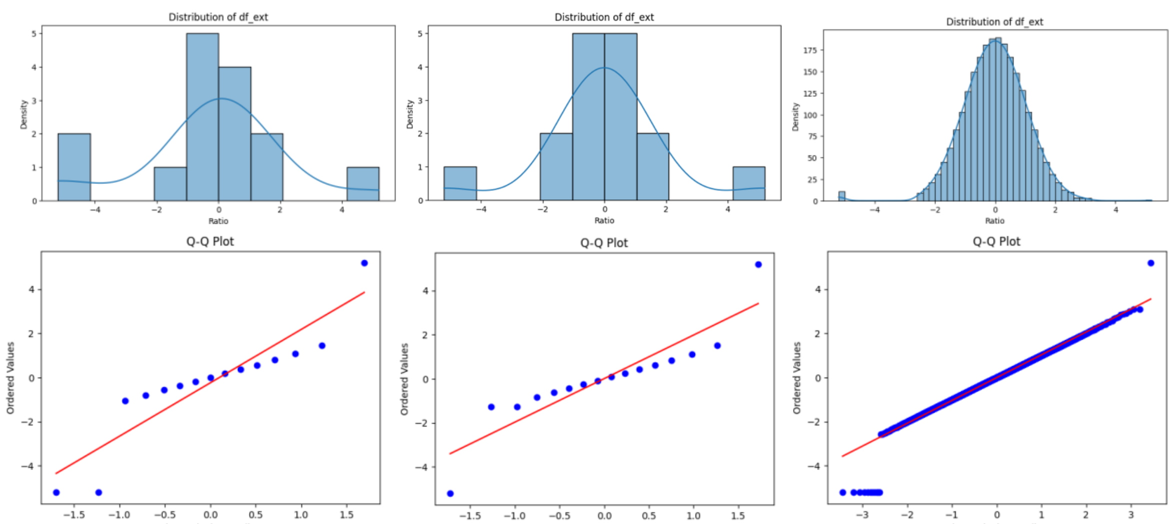


Figure 16: The example of Anderson-Darling test. This test was performed on capital adequacy (C) ratio for Rank 1 ($p = 3.02 \times 10^{-18}$), and the null hypothesis of the sample being normally distributed was rejected, suggesting the necessity of transformation.

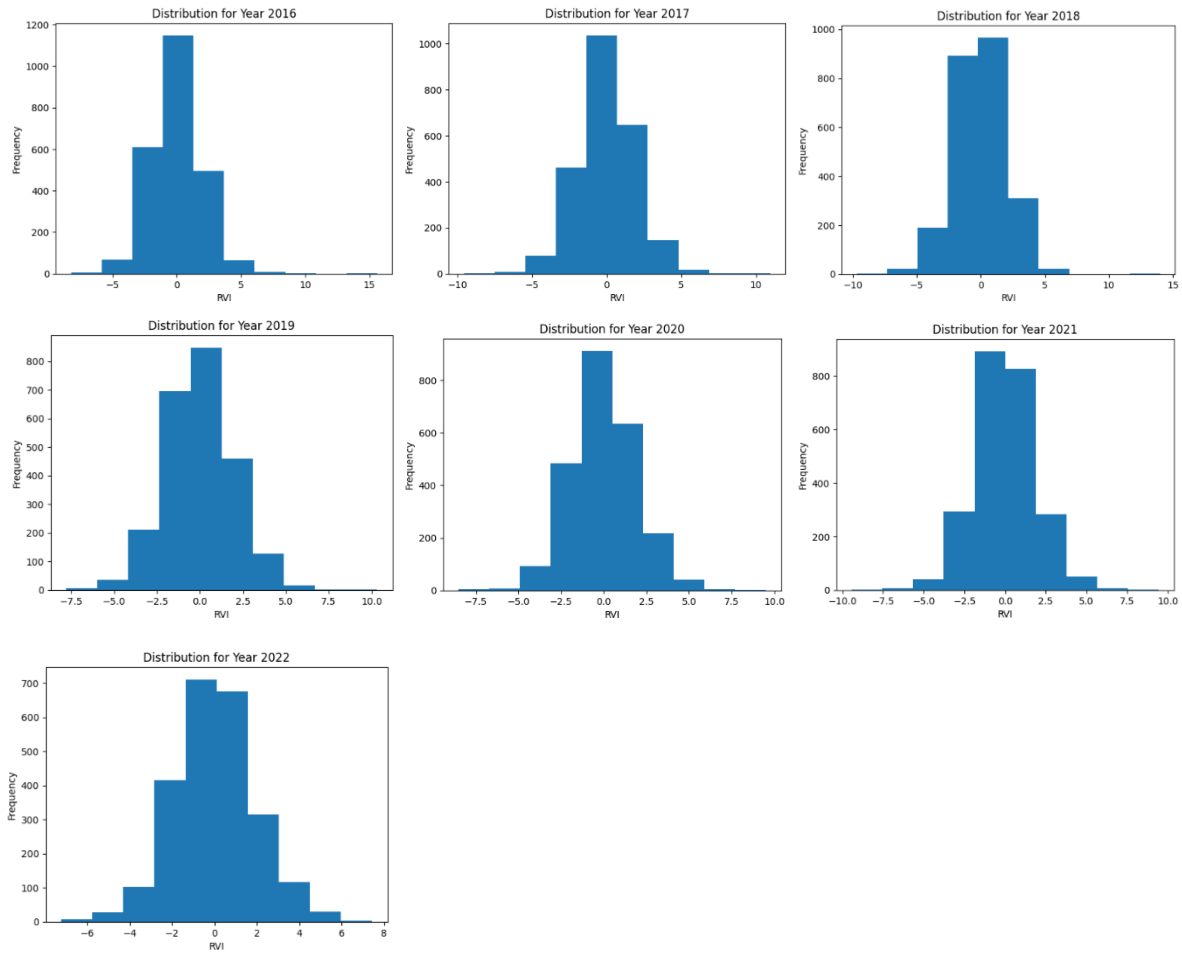
After quantile transformation and standardization, the 105 subgroups were tested with the Anderson-Darling test in the second round. 3 subgroups of exception were observed, which were the asset quality ratios of Rank 2 in 2019 and 2021, as well as the earnings ratio of Rank 3 in 2017 (as depicted in Figures 17-19). However, the Q-Q plots indicated that the null hypothesis was rejected because of a few extreme values, and the data points were overall a straight line (as depicted in Figures 20-22). Therefore, the exceptions are acceptable for calculations of RVI.



Figures 17-22: From left to right and top to bottom. Figures 17 and 18 show the asset quality ratios in 2019 and 2021 after the quantile transformation of Rank 2, while Figure 19 shows the earnings ratio of Rank 3 in 2017. Figures 20-22 are the Q-Q plots for Figures 16-18, respectively.

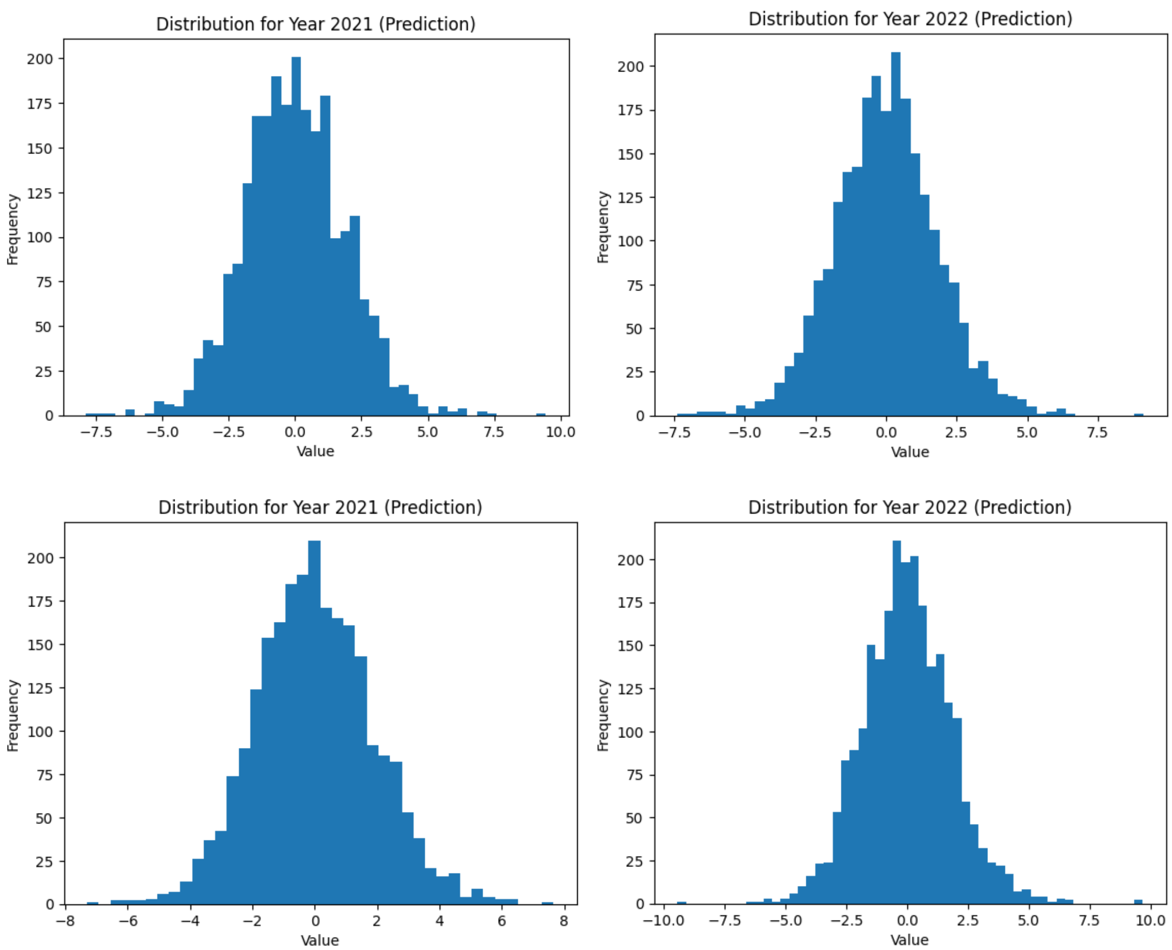
Compilation

With all the standardized $Z_{C,t}$, $Z_{A,t}$, $Z_{M,t}$, $Z_{E,t}$, and $Z_{L,t}$ being collected, the standardized RVI could be computed for every financial institution each year. The distributions of the RVI are demonstrated in Figures 23-29:



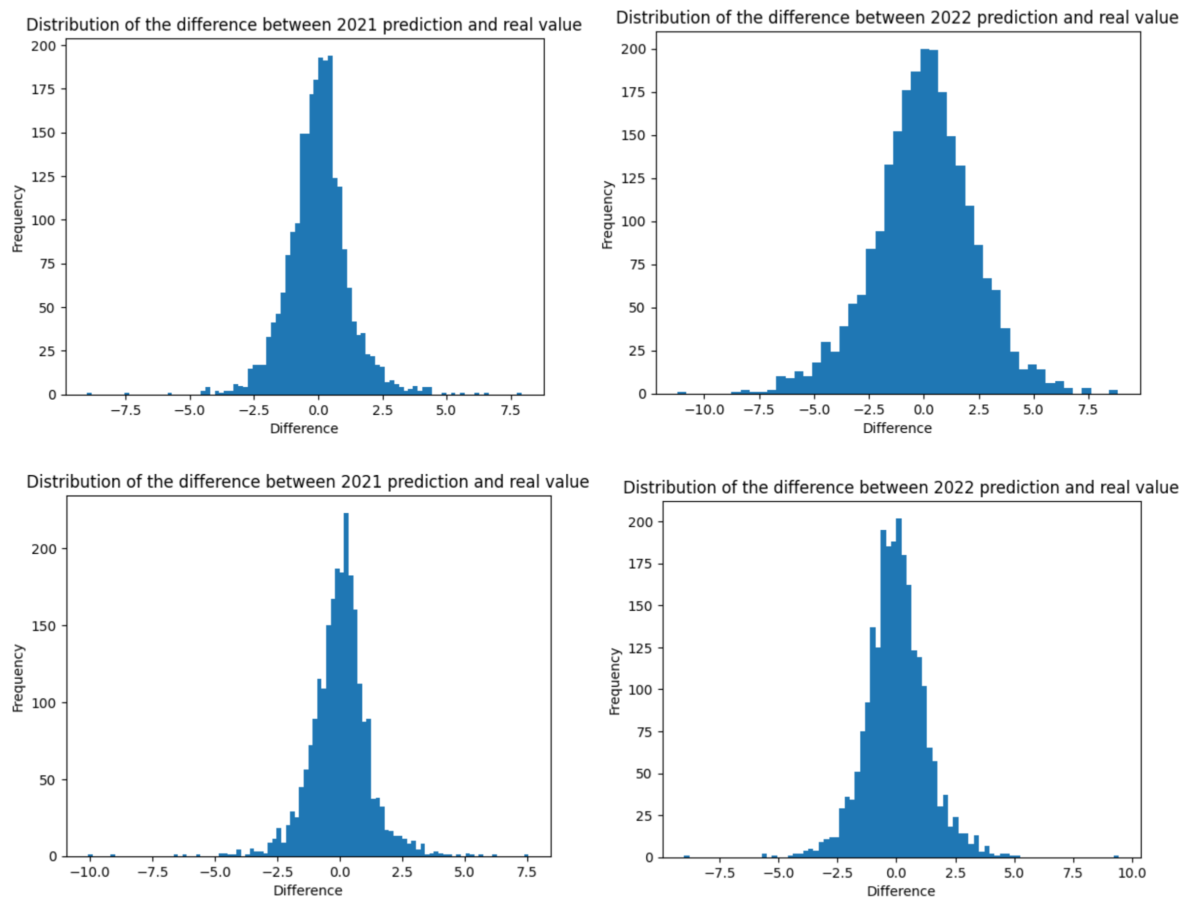
Figures 23-29: From left to right and top to bottom provide the distributions of RVI from 2016 to 2022.

Simultaneously, the predictions of CAMEL risk profiles for 2021 and 2022 stemming from the econometric model and the machine learning model were compiled into predicted RVI as well, as depicted in Figures 30-33:



Figures 30-33: From left to right and top to bottom. Figures 30 and 31 provide the distribution of predicted RVI given by the econometric model in 2021 and 2022. Figures 32 and 33, on the other hand, provide the distribution of predicted RVI given by the machine learning model in 2021 and 2022.

Despite the observed tendency for both the real RVI (as depicted in Figures 28 and 29) and the predicted RVI (as depicted in Figures 30-33) to cluster around zero, the Wilcoxon test was still conducted to ascertain any potential deviation of the median of the predicted RVI from the real RVI. To illustrate this, Figures 34-37 present the distributions of the discrepancies between the real RVI and the predicted RVI, considering both the econometric model and the machine learning model. The statistics of the Wilcoxon test and the percentage of predictions falling within one and two standard deviations are then displayed in Table 3.



Figures 34-37: From left to right and top to bottom. Figures 34 and 35 provide the distribution of the difference between the real RVI and predicted RVI based on the econometric model in 2021 and 2022, while Figures 36 and 37 provide the distribution of the difference between the real RVI and predicted RVI based on the machine learning model in 2021 and 2022.

	Econometric Model		Machine Learning Model	
	2021	2022	2021	2022
Wilcoxon P-value	0.8786	0.6406	0.9741	0.4922
1 standard deviation	0.7601	0.7497	0.7659	0.7530
2 standard deviations	0.9471	0.9454	0.9492	0.9517

Table 3: The Wilcoxon p-value and the percentage of predictions falling within one and two standard deviations. The higher values of p-value and percentage interpret a better performance in terms of predictions.

In Table 3 above, it is noted that the Wilcoxon p-value for both models exceeded the significance level, suggesting that the real RVI and the predicted RVI shared the same median in both predictive models for both years. While the econometric model scored a higher p-value in year 2022, the machine learning model got a higher p-value in year 2021. Both models generally possessed higher p-values and prediction accuracy within one and two standard deviations in 2021 compared to 2022, except for the metric of two standard deviations in the machine learning model. It is reasonable since the long-term prediction is more susceptible to deviating from the median. In terms of the prediction falling within one and two standard deviations, the machine learning model outperformed the econometric model in both metrices. As a result, overall, it was proven that the predicted RVI made by the machine learning model had a more consistent predictions than the traditional econometric model.

4.2 Problems Encountered

Natural language processing

While the textual data has been collected and processed, the results from the data processing were unsatisfactory; most of the sentiment scores obtained have been very close to zero (shown in Figure 8, overleaf).

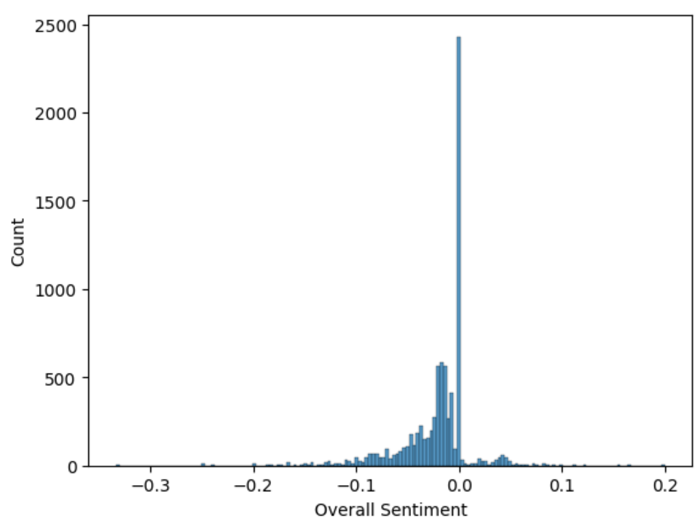


Figure 38: A histogram showing the frequency of banks with overall sentiments of different values. Note that most banks have an overall sentiment score of close to 0.

Two reasons are proposed for the convergence of the sentiment scores to zero:

1. News articles from reliable sources, such as the Guardian and New York Times, maintain an objective and informative tone, leading to a lack of language that exhibits strong sentiment.
2. The news articles, which were intended for a general audience, are incompatible with the Dictionary, which contains a lot of financial jargon.

Nevertheless, since the sentiment score could still provide useful insights into risk assessment, potential measures are proposed to address this problem. First, a better set of keywords might be used for better sentiment detection. If the refinement of the NLP techniques is not effective, the score may only focus on banks that are more well-known and have non-trivial sentiment scores. Accordingly, a composite sentiment score for the banking sector of a given jurisdiction will be computed. This score could serve as the market sentiment of the overall banking sector, which may signal potential systemic risk in the sector (Svetlana et al., 2017). In addition, the approach of quantile transformation in RVI computation may help mitigate the problem as well, given that quantile transformation can deal with over-concentrated samples and force them to transform into easily analyzable distributions, such as uniform distributions or normal distributions.

5. Conclusion

To conclude, this project aims to develop a SupTech application to assist the HKMA in its supervisory process by providing early detection and forecast of financial institution failures using NLP and ML techniques. Data processing and model training have been mostly completed and the next step would be front-end development and further refining the models. The current problem encountered in this project is the unsatisfactory results in the processing of textual data which serve to construct the bank confidence index in model training. Nonetheless, potential measures are proposed to address this problem so that the bank confidence index can be computed. Ultimately, this application will establish a blueprint for how regulators can harness advanced technologies in enhancing the current supervisory process. At the practical level, future works can build on this application by including other risk models and stress-testing. This would help make the application more versatile for real-world applications.

References

Almarzoqi, R., Naceur, S. B., & Scopelliti, A. D. (2015). *How does bank competition affect solvency, liquidity and credit risk? Evidence from the MENA countries* (IMF Working Paper No. 15/210). https://www.imf.org/external/pubs/ft/wp/2015/wp15210.pdf

Arathi, M., & Govardhan, A. (2015). Effect of Mahalanobis distance on time series classification using shapelets. In Satapathy, S., Govardhan, A., Raju, K., & Mandal, J. (Eds.), *Emerging ICT for bridging the future – Proceedings of the 49th annual convention of the computer society of India CSI volume 2* (pp. 525-535). Springer.

Bank of England [BoE]. (2022). *Stress testing the UK banking system: Key elements of the 2022/23 annual cyclical scenario*. https://www.bankofengland.co.uk/stress-testing/2022/key-elements-of-the-2022-stress-test

Bank for International Settlements [BIS]. (2010). *Calibrating regulatory minimum capital requirements and capital buffers: A topdown approach*. https://www.bis.org/publ/bcbs180.pdf

Bank for International Settlements [BIS]. (2011, November). *Global systemically important banks: Assessment methodology and the additional loss absorbency requirement*. https://www.bis.org/publ/bcbs207.pdf

Bank for International Settlements [BIS]. (2016, June). *Regulatory consistency assessment programme (RCAP): Assessment of Basel III G-SIB framework and review of D-SIB frameworks – Japan*. https://www.bis.org/bcbs/publ/d371.pdf

Bank for International Settlements [BIS]. (2023, July 12). *RCAP jurisdictional assessments: Regulatory implementation consistency*. https://www.bis.org/bcbs/implementation/rcap_jurisdictional.htm?m=3060

Bond, S. R. (2002). *Dynamic panel data models: A guide to micro data methods and practice* (CeMMAP Working Paper No. CWP09/02). https://doi.org/10.1920/wp.cem.2002.0902

Broeders, D., & Prenio, J. (2018, July). *FSI insights on policy implementation no.9: Innovative technology in financial supervision (suptech) – the experience of early users*. https://www.bis.org/fsi/publ/insights9.pdf

Brogi, M., Lagasio, V., & Riccetti, L. (2021). Systemic risk measurement: Bucketing global systemically important banks. *Annals of Finance, 17*, 319-351. https://doi.org/10.1007/s10436-021-00391-7

Buckley, R. P., Arner, D. W., Zetzsche, D. A., & Selga, E. (2019). The dark side of digital financial transformation: The new risks of FinTech and the rise of TechRisk. *EBI Working Paper Series*, *2019*(54), 7–9. https://doi.org/10.2139/ssrn.3478640

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing, 300*. 70-79. https://doi.org/10.1016/j.neucom.2017.11.077

Cheng, H. (2007). Panel data analysis—advantages and challenges. *TEST, 16*, 1-22. https://doi.org/10.1007/s11749-007-0046-x

Chevalier, G. (2018). *LSTM Cell. Wikimedia Commons*. Retrieved September 23, 2023, from https://upload.wikimedia.org/wikipedia/commons/9/93/LSTM_Cell.svg.

D'Andrea, A., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications, 125*(3), 26–33. https://doi.org/10.5120/ijca2015905866

Dordevic, L., Ferreira, C., Kitonga, M., & Seal, K. (2021). Strengthening bank regulation and supervision: National progress and gaps. *Departmental paper series, 21*(5). 5-10. https://doi.org/10.5089/9781513566658.087

Educational Ecosystem. (2018, September 12). *Understanding K-means clustering in machine learning*. Medium. https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1

European Central Bank. (2010, September). *Beyond ROE – how to measure bank performance*. https://www.ecb.europa.eu/pub/pdf/other/beyondroehowtomeasurebankperformance201009en.pdf

European Central Bank. (2023, July 12). *ECB publishes supervisory banking statistics on significant institutions for the first quarter of 2023*. https://www.bankingsupervision.europa.eu/press/pr/date/2023/html/ssm.pr23712~c5bce797c5.en.html#:~:text=The%20non%2Dperforming%20loans%20(NPL,the%20first%20quarter%20of%202023.

European Securities and Markets Authority [ESMA]. (2019). *ESMA report on trends, risks and vulnerabilities*. https://www.esma.europa.eu/sites/default/files/library/esma50-report_on_trends_risks_and_vulnerabilities_no1_2019.pdf

Federal Deposit Insurance Corporation [FDIC]. (2018, July). *Earnings*. https://www.fdic.gov/resources/supervision-and-examinations/examination-policies-manual/section5-1.pdf

Federal Deposit Insurance Corporation [FDIC]. (2022a, August). *Capital*. https://www.fdic.gov/resources/supervision-and-examinations/examination-policies-manual/section2-1.pdf

Federal Deposit Insurance Corporation [FDIC]. (2022b, March). *Management*. https://www.fdic.gov/resources/supervision-and-examinations/examination-policies-manual/section4-1.pdf

Federal Deposit Insurance Corporation [FDIC]. (2024, January). *Asset quality*. https://www.fdic.gov/resources/supervision-and-examinations/examination-policies-manual/section3-1.pdf

Federal Reserve Board [Fed]. (2022). *2022 Stress test scenarios*. https://www.federalreserve.gov/publications/2022-Stress-Test-Scenarios.htm

Financial Stability Board [FSB]. (2020, October 9). *The use of supervisory and regulatory technology by authorities and regulated institutions: Market developments and financial stability implications*. https://www.fsb.org/wp-content/uploads/P091020.pdf

Garcia, M. T. M., & de Mendonça, F. C. (2023). Compliance with the Basel core principles and supervisory structure: A cross-country analysis. *Central Bank Review, 23*(2). 1-15. https://doi.org/10.1016/j.cbrev.2023.100117

Grosse, R. (2017). Exploding and Vanishing Gradients. Department of Computer Science, University of Toronto. Retrieved January 17, 2024, from https://www.cs.toronto.edu/~rgrosse/courses/csc321_2017/readings/L15%20Exploding%20and%20Vanishing%20Gradients.pdf.

Hong Kong Monetary Authority [HKMA]. (2015, February 27). *Supervisory policy manual: Overview of capital adequacy regime for locally incorporated authorized institutions*. https://www.hkma.gov.hk/media/eng/img/key-functions/banking-stability/supervisory-policy-manual/CA-G-1.pdf

Hong Kong Monetary Authority [HKMA]. (2019, August 26). *Liquidity*. https://www.hkma.gov.hk/eng/key-functions/banking/banking-legislation-policies-and-standards-implementation/liquidity/

Hong Kong Monetary Authority [HKMA]. (2020, January 24). *Supervisory review process*. https://www.hkma.gov.hk/media/eng/doc/key-functions/banking-stability/supervisory-policy-manual/CA-G-5.pdf

Hong Kong Monetary Authority [HKMA]. (2021a, June). *HKMA's adoption of supervisory technology*. https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210617e1a1.pdf

Hong Kong Monetary Authority [HKMA]. (2021b, April 23). *Supervisory policy manual: Systemically important banks*. https://www.hkma.gov.hk/media/eng/doc/key-functions/banking-stability/supervisory-policy-manual/CA-B-2.pdf

Hong Kong Monetary Authority [HKMA]. (2022, December 23). *Supervisory policy manual: Risk-based supervisory approach*. https://www.hkma.gov.hk/media/eng/doc/key-functions/banking-stability/supervisory-policy-manual/SA-1.pdf

Hong Kong Monetary Authority [HKMA]. (2023, September). *Half-yearly monetary and financial stability report*. https://www.hkma.gov.hk/media/eng/publication-and-research/quarterly-bulletin/qb202309/E_Half-yearly_202309.pdf

Holley, E. (2018, November 2). *BankFocus reveals how asset quality is holding Eurozone banks back*. https://www.bvdinfo.com/en-gb/blog/bank-credit/bankfocus-reveals-how-asset-quality-is-holding-eurozone-banks-back

Hutto, C. J., & Gilbert, E. E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media, 8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Khatun, N. (2021). Applications of normality test in statistical analysis. *Open Journal of Statistics*, *11*(1). 113-122. https://doi.org/10.4236/ojs.2021.111006

Liu, L., Moon, H. R., & Schorfheide, F. (2018, September). *Forecasting with dynamic panel data models* (NBER Working Paper No. 25102). http://www.nber.org/papers/w25102

Loughran, T., & McDonald, B. (2010). When is a liability not a liability? textual analysis, dictionaries, and 10-KS. *The Journal of Finance, 66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Moody's Analytics. (2023). *BankFocus*. https://www.moodysanalytics.com/product-list/bankfocus

Muharemi, F., Logofătu, D., & Leon, F. (2018). Review on general techniques and packages for data imputation in R on a real world dataset. In Nguyen, N., Pimenidis, E., Khan, Z., & Trawiński, B. (Eds.), *Computational collective intelligence: 10th international conference, ICCCI 2018 Bristol, UK, September 5-7, 2018 proceedings, part II* (pp. 386-395). Springer.

News API. (2023). *News API – search news and blog articles on the web. News API - Search News and Blog Articles on the Web*. https://newsapi.org/

O'Keefe, J. P. (2022, January 24). *The use of machine learning for bank financial stress tests: Prepared for the 3rd annual Marcus Evans conference - best practices for stress testing in financial institutions, New York USA*. http://dx.doi.org/10.2139/ssrn.4019533

Ong, L. L., Jeasakul, P., & Kwoh, S. (2013, August). *HEAT! A bank health assessment tool*. https://www.imf.org/external/pubs/ft/wp/2013/wp13177.pdf

Park, J., Shin, M., & Heo, W. (2021). Estimating the BIS capital adequacy ratio for Korean banks using machine learning: Predicting by variable selection using random forest algorithms. *Risks, 9*(2). 32. https://doi.org/10.3390/risks9020032

Petropoulos, A., Siakoulis, V., Vlachogiannakis, N., & Stavroulakis, E. (2019, March). *Deep-Stress: A deep learning approach for dynamic balance sheet stress testing*. https://www.eba.europa.eu/sites/default/documents/files/document_library/Calendar/Conference-Workshop/2019/8th%20annual%20workshop%20documents/2%20DeepStress-A%20deep%20learning%20approach%20for%20dynamic%20balance%20sheet%20stress%20testing.pdf

Statsmodels. (2023, December 14). *Linear regression: Module reference*. https://www.statsmodels.org/stable/generated/statsmodels.regression.linear_model.OLS.html#statsmodels.regression.linear_model.OLS

StatsTest. (2020, November 3). Wilcoxon signed-rank test. StatsTest.com. https://www.statstest.com/wilcoxon-signed-rank-test/

Staudemeyer, R. C., & Morris, E. R. (2019). *Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks*. arXiv. https://arxiv.org/abs/1909.09586

Svetlana, B., Garmaev, E., Lammers, P., & Rustige, J. (2017, April). *SenSR: A sentiment-based systemic risk indicator* (DNB Working Paper No. 553). https://www.dnb.nl/media/4qhpfjyd/working-paper-no-553_tcm47-356707.pdf

The World Bank. (2024a). *Metadata glossary: Bank nonperforming loans to total gross loans (%)*. https://databank.worldbank.org/metadataglossary/world-development-indicators/series/FB.AST.NPER.ZS

The World Bank. (2024b). *Metadata Glossary: Liquid assets to deposits and short term funding (%)*. https://databank.worldbank.org/metadataglossary/global-financial-development/series/GFDD.SI.06

Van Greuning, H., & Bratanovic, S. B. (2009). *Analyzing banking risk: A framework for assessing corporate governance and risk management* (3rd ed.). The World Bank.

Wong, J., Choi, K. F., & Fong, T. (2005). *Determinants of the capital level of banks in Hong Kong*. https://www.hkma.gov.hk/media/eng/publication-and-research/quarterly-bulletin/qb200509/fa2.pdf

Wong, S. L., & Wei, M. (2023, June). *The bank vulnerability index: A health check on ASEAN+3 banks and banking sectors*. https://amro-asia.org/wp-content/uploads/2023/08/AMRO-Working-Paper-23-02_The-Bank-Vulnerability-Index_-Aug-2023.pdf

Wooldridge, J. M., Wadud, M., & Lye, J. (2017). *Introductory econometrics*. Cengage.

<u>Appendices</u>

Appendix 1. Glossary of terms

Term	**Definition**
Asset quality	It evaluates the quality of assets in relation to the amount of credit risk exposed by financial institutions (Federal Deposit Insurance Corporation [FDIC], 2024).
Capital adequacy	It assesses whether a financial institution has adequate capital to absorb potential losses incurred from assets so that it can operate viably (FDIC, 2022a).
Credit risk	It is the risk that the debtor may default on its borrowed money (Almarzoqi et al., 2015).
Earnings	It indicates the profitability of financial institutions (FDIC, 2018).
Liquidity	It determines the ability of financial institutions to settle short-term obligations on the due date (HKMA, 2019).
Management	It reflects the management capability of financial institutions, namely in terms of operational efficiency (FDIC, 2022b).
Semi-elasticity	It is an economic concept which refers to the effect of a percentage change in x on y (Wooldridge et al., 2017).
Solvency risk	It is the risk that an entity cannot fulfill its long-term obligations (Almarzoqi et al., 2015).

Appendix 2. Formulae of the selected CAMEL ratios

$$\text{Common Equity Tier 1 Ratio} = \frac{\text{Common equity tier 1 capital}}{\text{Risk weighted assets}}$$

(HKMA, 2015)

$$\text{Non performing Loan Ratio} = \frac{\text{Non performing (impaired) loan}}{\text{Total gross loan}}$$

(Holley, 2018; The World Bank, 2024a)

$$\text{Cost to Income Ratio} = \frac{\text{Operating expense}}{\text{Operating revenue}}$$

(European Central Bank, 2010)

$$\text{Return on Risk weighted Assets Ratio} = \frac{\text{Net income}}{\text{Risk weighted assets}}$$

(BIS, 2010)

$$\text{Liquid Assets to Short term Funding Ratio} = \frac{\text{Liquid assets}}{\text{Short term funding}}$$

(The World Bank, 2024b)

Appendix 3. Project schedule (table)

Schedule		**Milestone**		**Status**
Phase 1: Project inception				
2023	Sep	• Literature review • Feasibility assessment • Project plan formation	**Deliverables**: Detailed project plan + website	✓
Phase 2: Project elaboration				
	Oct	• Data collection and processing (numerical and textual data) • Econometric model preparation • Machine learning model preparation		✓
	Nov	• Data collection and processing (textual data) • Econometric model preparation • Machine learning model preparation		✓
	Dec	• Econometric model training and testing • Machine learning model training and testing • RVI construction (subgroup creation + transformation to normal distribution) • Dashboard design • Interim report drafting		✓
2024	Jan	• Econometric model evaluation • Machine learning model evaluation • RVI compilation and evaluation	**First presentation** **Deliverables**: Prototype + interim report	✓

		• Dashboard design • Interim report drafting		
Phase 3: Project construction				
	Feb	• Model fine-tuning • Troubleshooting textual data processing • Frontend and backend development		
	Mar	• Model fine-tuning • Frontend and backend development • Debugging and code documentation • Final report drafting		
	Apr	**Final presentation** **Deliverables**: Finalized product + final report **Project exhibition**		

Appendix 4. Project schedule (Gantt chart)