

Dynamic Early Warning System for Financial Crashes

Mokha Lerthsuwanroj

School of Computing Science

Sir Alwyn Williams Building

University of Glasgow

G12 8RZ

A dissertation presented in part fulfillment of the requirements of the Degree of Master of Science at the University of Glasgow

<Date of submission placed here>

**Abstract**

<Abstract goes here…>

Education Use Consent

I hereby give my permission for this project to be shown to other University of Glasgow students and to be distributed in an electronic form.

<**Please note that you are under no obligation to sign this declaration, but doing so would help future students.>**

Name: Signature:

Acknowledgements

<Acknowledgements go here>

Contents

<Update the table of contents by right-clicking on it and selecting Update Field… and then select page numbers only.>

Chapter 1 Introduction <This is Heading 1> 1

1.1 A section <This is style Heading 2> 1

1.1.1 A subsection <This is style Heading 3> 1

Chapter 2 Survey 2

Chapter 3 Further Chapters 3

Chapter 4 Conclusion 4

Appendix A <Name of appendix> 1

Appendix B <Another appendix> 2

Bibliography 3

# Introduction

## Motivation

Financial market crashes have historically affected to the economic and social disruptions, leading to the downturn, unemployment, and loss of investor confidence. Early detection and intervention are critical for minimizing the impact of such crises on those affected.

In recent years, sentiment analysis has emerged as a powerful tool in financial risk monitoring, offering insights into investor expectations, fears, and behaviours. When sudden shifts in investor sentiment are combined with traditional volatility indicators, they can provide early signals of impending market instability (Liu et al., 2023) [3]. However, most existing early warning systems rely on static models that struggle to adapt to the fast-changing dynamics of the modern financial markets (Kustina et al., 2023) [1]. By integrating both market and sentiment volatility within a dynamic framework, this research aims to develop more responsive and accurate tool for crash prediction.

## Purpose

The primary purpose of this research is to develop a dynamic early warning system (EWS) that integrates both market-based and sentiment-based volatility indicators to enhance the early detection of financial market crashes. Given the increasing complexity and unpredictability of financial markets, especially during periods of heightened uncertainty, there is a growing need for more adaptive and timely forecasting models.

This study seeks to address by analysing how time-varying patterns in market-based and sentiment-based volatility relate to the occurrence of past financial crises. Specifically, the research will make use of historical financial news data sourced via the Yahoo Finance (yfinance) API to extract relevant sentiment signals. These sentiment indicators will then be combined with traditional market-based volatility measures within a dynamic modelling framework designed to capture the evolving relationships between these variables over time.

# Survey

## Background Survey

### Current Methods for Early Warning Systems (EWS)

Early warning systems (EWS) for financial crashes have evolved from simple statistical models to more advanced machine learning and nonlinear approaches. The traditional statistic models, such as logistic regression, have been widely used to detect early signals for financial crashes using predefined relationships between risk indicators and crash probabilities. Such models often suffer from rigid parameterization and lagging indicators, limiting their ability to capture regime shifts or sudden market changes (Kustina et. al., 2023) [1], in contrast with the more recent research that explored the nonlinear approach to overcome this limitation. Nonlinear algorithms, support vector machines (SVM), and neural networks have shown improved capacity for capturing the complex relationships in the real-world financial markets (Song et al., 2024) [2], allowing more flexibility when modelling market risks and crash probabilities as new data becomes available. Empirical evidence supports that dynamic nonlinear methods outperform static models, providing better crisis prediction under changing market environments (Song et al., 2024) [2].

Beside the market-based indicators, sentiment analysis has also gained attention in financial crash predictions. The rise of social media platforms such as Twitter (now known as X), along with financial news sources, has provided rich datasets for capturing investor mood and behaviours (Liu et al., 2023) [3]. However, extracting signals from this unstructured data often produces noisy which remains a challenge. Liu, Leu, and Holst (2023) [3] proposed a method using FinBERT combined with an ensemble SVM to reduce noise and filter out irrelevant content from social media discussions (Liu et al., 2023) [3].

Huang et al. (2020) [7] showed that FinBERT, which is specifically pre-trained on financial texts including earnings call transcripts, analyst reports, and financial news articles, significantly outperforms general-purpose language models like BERT and traditional approaches in various financial information extraction tasks, including the LM dictionary, NB, SVM, RF, CNN, and LSTM. The model's specialized training on domain-specific vocabulary and financial terminology enables it to better understand the context inherent in financial communications, resulting in improved accuracy for sentiment classification, named entity recognition, and relationship extraction from financial documents (Huang et al., 2020) [7]. However, it is important to note that FinBERT demonstrated superiority applies specifically to financial text analysis tasks, and not directly to modelling financial market volatility.

### Key Crash Indicator

Most studies operationalize a financial crash using a binary crash indicator equation, where a crash is identified based on a significant drop in asset prices or index returns over a specified time window. A common method involves calculating the log return of closing prices over a fixed period (e.g., 5-day or 10-day intervals) and labelling an observation as a "crash" if the return falls below a predefined threshold which often set at the 10th percentile of historical returns or a fixed percentage drop, such as −10% (Kaminsky et al., 1998) [8]. However, early warning models also extend to other types of financial crises, such as currency and sovereign debt. Kaminsky and Reinhart (1999) [9], for example, define currency crises based on a sharp depreciation of the exchange rate coupled with reserve losses, using an exchange market pressure index. Bussière and Fratzscher (2006) [11] extend this framework to sovereign debt crises by incorporating a wide range of macroeconomic variables, flagging a crisis when key thresholds are breached.

Focusing on stock prices offers key advantages. Equity data is available at high frequency, enabling timely detection of systemic stress. Stock prices react quickly to macroeconomic shocks and shifts in investor sentiment, providing an early market-driven signal of financial instability. In contrast, currency and sovereign debt indicators often suffer from lower frequency, policy interventions, and reporting delays. Thus, stock price-based crash indicators are well suited for dynamic, real-time early warning systems.

Volatility remains one of the most important indicators in crash prediction research. To flag the potential financial instability, both realized volatility (observed historical price variability) and price-volatility feedback rate have been used (Allaj & Sanfelici, 2023) [4]. Pattern of increased volatility generally precede market downturns, making it useful for EWS frameworks.

Mentioning the traditional risk measures, Value-at-Risk (VaR) and Expected Shortfall (ES) are widely used as quantitative measures to assess market risk and potential losses under various conditions. However, both VaR and ES forecasts often rely on models with specific distributional or structural assumptions (Allaj & Sanfelici, 2023) [4], which may not capture sudden market regime shifts, nonlinear behaviours. This is especially true in emerging markets, where volatility is typically higher and market dynamics are less predictable.

A recent study by Le (2024) [6] examined the effectiveness of combining multiple VaR and ES forecasting models in the context of the Vietnamese stock market. The research found that forecast combination techniques, such as weighted averaging of outputs from different models (e.g., GARCH, CAViaR, and ES-CAViaR), significantly improved the accuracy and reliability of risk forecasts, especially during periods of high market volatility. The combined models showed better back testing performance and greater compliance with regulatory risk thresholds, compared to any single model (Le, 2024) [6].

### Rationale for Focusing on Index-Level Predictions

Index-level models offer several advantages, including aggregation benefits that help reduce noise and unexpected shocks from individual stocks (Park et al., 2024) [5]. Research has shown that top-down index forecasts tend to be more accurate and informative than bottom-up aggregation of individual stock predictions, particularly for systemic risk assessment (Park et al., 2024) [5]. By concentrating on index-level sentiment and market volatility, the model can better capture macro-level signals that reflect broader market conditions.

## Research Objectives

The goal of this project is to assess the following objectives:

1. To analyse the predictive power of both market-based and sentiment-based volatility indicators, examining how their time-varying patterns relate to the occurrence of past financial crashes at the index level.
2. To develop a dynamic modelling framework, incorporating features such as time-varying thresholds and regime-switching mechanisms, to capture the evolving relationship between sentiment-driven and market-based volatility indicators.
3. To evaluate and compare the predictive performance of the proposed dynamic EWS against traditional static statistical models, with a particular focus on assessing the added value of sentiment-based inputs.
4. To validate the robustness and generalizability of the developed model across different market environments and historical crisis periods.

## Research Questions

Based on the stated objectives, this study seeks to answer the following key research questions:

1. How do market-based and sentiment-based volatility indicators individually and jointly relate to the timing and occurrence of past financial crashes at the index level?
2. Can a dynamic early warning system (EWS) outperform traditional static models in predicting financial market crashes?
3. To what extent does sentiment volatility enhance the performance of early warning models compared to using market-based indicators alone?
4. Is the proposed dynamic EWS model robust and generalizable across different market conditions and historical crisis periods?

# Design and Implementation

## Data Collection and Preparation

* Market from …, Sentiment from …
* Kaggle, yfinance, mediastack, etc.

## Model Architecture Design

* Python, framework, library, etc.
* Crashes factor
* Allaj and Sanfelici (2023) [4] explore the use of varying estimation window lengths, denoted as T\_w, to assess the predictive performance of different forecasting models. Specifically, they consider values of T\_w = 22, 66, and 132, which correspond to approximately one month, one quarter, and half a year of trading days, respectively.

## Data Preprocessing

* Split, label, calculate score, augmentation?

## Model Training

* Generate sentiment score and compute sentiment volatility
  + FinBERT tokenizer
  + Calculate \*mean\* (group by ‘date’)
* Align with market crashes
* Modelling

## Model Evaluation

## Out of Sample Analysis

* Test the goodness of the previous model

The content of these chapters depends on the project and should be agreed with your supervisor (e.g. description of the solution, evaluation results, etc).

<Figure below is in style “figure” which continues to style “figure caption” when you press Enter and then back to “Normal” when you press Enter again.>

Figure 1: Some important shapes.

<If you wanted to show any code fragments, you could use the following style called code, which could then be followed by figure caption..>

*# This is a little bit of Python*

**for** i in range( 10 ):

**for** j in range( 10 ):

**print** i\*j,

**print**

Figure 2: A crucial algorithm for the project.

# Conclusion

Main conclusions of your project. Here you should also include suggestions for future work.

###### <Name of appendix>

<Use Heading 6 for the Appendix heading>

###### <Another appendix>

# Bibliography

[1] Lisa Kustina, Rachmat Sudarsono and Nury Effendi (2023). Market crash factors and developing an early warning system: Evidence from Asia, <https://doi.org/10.21511/imfi.20(3).2023.10>

[2] Song and Li (2024), Early warning signals for stock market crashes: empirical and analytical insights utilizing nonlinear methods, <https://doi.org/10.1140/epjds/s13688-024-00457-2>

[3] Liu, J., Leu, J., & Holst, S. (2023). Stock price movement prediction based on Stocktwits investor sentiment using FinBERT and ensemble SVM. PeerJ Computer Science, 9. <https://doi.org/10.7717/peerj-cs.1403>

[4] Erindi Allaj, Simona Sanfelici, Early Warning Systems for identifying financial instability, International Journal of Forecasting, Volume 39, Issue 4, 2023, 1777-1803, ISSN 0169-2070, <https://doi.org/10.1016/j.ijforecast.2022.08.004>

[5] Park, M., Peterson, M., & Weisbrod, E.H. (2024). Top-Down vs. Bottom-Up Index Forecasts: Are Market Strategists Strategically Pessimistic? <http://dx.doi.org/10.2139/ssrn.4695279>

[6] Le, T.H. (2024), "Forecasting value-at-risk and expected shortfall in emerging market: does forecast combination help?", Journal of Risk Finance, Vol. 25 No. 1, 160-177. <https://doi.org/10.1108/JRF-06-2023-0137>

[7] Huang, Allen H. and Wang, Hui and Yang, Yi, FinBERT - A Large Language Model for Extracting Information from Financial Text (July 28, 2020). <http://dx.doi.org/10.2139/ssrn.3910214>

[8] Kaminsky, Graciela, Lizondo, Saul and Reinhart, Carmen, (1998), Leading Indicators of Currency Crises, IMF Staff Papers, 45, issue 1, 1-48, <https://EconPapers.repec.org/RePEc:pal:imfstp:v:45:y:1998:i:1:p:1-48>

[9] Kaminsky, G., Reinhart, C., 1999. The twin crises: the causes of banking and balance-of-payments problems. American Economic Review 89 (3), 473-500. <https://doi.org/10.1257/aer.89.3.473>

[10] Jemovic, M., Marinkovic, S., 2019. Determinants of financial crises—an early warning system based on panel, logit regression. International Journal of Finance and Economics, 1–15. <https://doi.org/10.1002/ijfe.1779>

[11] Matthieu Bussiere, Marcel Fratzscher, Towards a new early warning system of financial crises, Journal of International Money and Finance, Volume 25, Issue 6, 2006, 953-973, ISSN 0261-5606, <https://doi.org/10.1016/j.jimonfin.2006.07.007>.