Empirical Review of Models used for Predicting Financial Market Crashes Using Market Data

Project: Literature Review

By Ping-Chieh Tu, Adrien Bélanger, and Inigo Torres

November 22nd 2024

Background and Introduction Adrien (Goes with Background)

We are not the first to approach this subject. Multiple papers have approached prediction of stock market crashes using Machine Learning. Some have used Support Vector Machines, random forests, and Neural Networks [Okpeke, Predicting Stock Market]. In a time series analysis approach, models used in Time Series Analysis have traditionnally included RNNs and Arima, and more recently Transformers. [Sabeen Ahmed, Transformers ...] [Arunkumar, Comparative Analysis]. These models have been used in Market Crash prediction. Others models have been used, to varied success, such as LTSM models. Others have tried using diverse databases such as Social Media interactions [Chhajer, The Applications..] using Natural Language Processing. As this is a very economically valuable foresight, a lot of research has been done on the subject.

Reviews and comparisons of these models, such as this project aims to do, have been made such as [Okpeke, Predicting], but a comprehensive empirical review of Time-Series Analysis models on equal footing is lacking in litterature. This project aims to address this lack, by providing an empirical comparison of three commonly used models in Time-Series Analysis (Arima models, RNNs and Transformers) to predict Market Crashes. [Sabeen Ahmed, Transformers ...] [Arunkumar, Comparative Analysis]. We shall use data freely available on the Yahoo finance database, and tag historically factual Market Crashes by hand, as there are few. This methodology (or closely related) is common procedure, and has been used for these kind of projects [][]. Specific Criterias for comparison and analysis shall be discussed in a further section.

Methodology

- Justify why we chose those three models by finding similar work Oscar ? This might be good: https://oarjst.com/sites/default/files/OARJST-2024-0095.pdf

Arima Oscar

- Overview of ARIMA
  - Review its application in time series analysis in our context

Reccurent Neural Networks Adrien

Recurrent Neural Networks are a class of neural network architectures designed to detect patterns in sequential data, such as handwriting, genomes, text, or numerical time series. [Schmidt, Recurrent Neural Networks...]. They have been used in multiple projects accounting to market Crashes, such as [] and [].

Transformers Inigo

- Overview of Transformers
  - Review its application in time series analysis in our context

Criterias and Analysis Oscar?

- Summarize criterias and metrics in literature for comparing models.
  - Justify the selection of those comparison methods based on sources

Adapt proposal: Inigo