The stock market is typically used as a metric to gain an understanding of how healthy a country is in terms of its economy. Research of stock price correlation with economic or global events have been carried out before and prediction of the stock market based on global events is a challenging task as unexpected events also need to be considered. Furthermore, these unexpected events have a greater effect on stock market returns and prices.

Parab (2020) [1] investigated the impact of economic events on stock market returns in India by selecting thirteen (three international) economic events and analysing the return price before, during and after the event date. The data was obtained from the National Stock Exchange (NSE) of India for the NOFTY 50 and other indices within India (Auto, Financial Services, FMCG (fast-moving consumer goods), IT, Media, Metal, Pharma Index, PSU (public sector undertaking) Bank Index, Private Bank, and Realty). Dummy variables were created for each index. He mentions that since the USA trades with a lot of countries, any economic event will have a global affect such as the subprime mortgages in 2008. Through regression analysis, Parab discovered that the responses from investors to instantaneous reactions to economic events were short-term. Furthermore, majority of unexpected events that exerted a dramatic effect allowed investors and trades to reap short-term profits. However, there is a gap in this research model where volatility analysis was not considered.

Jabarin (2019) conducted a different approach compared to Parab by using an event study methodology [2]. The aim of this study was to view the impact that various macroeconomic factors and political events had on the market index returns of the Palestine and Amman stock markets. The study made use of time series monthly data the events that were selected were ones that caught the media’s attention through the years of 2011 - 2017. Through her literature review, Jabarin states that the Consumer Price Index (CPI), GDP and exchange rate had a significant impact on stock market returns. Moreover, the real GDP is the greatest determinant of stock market returns and that there was a positive relationship between balance of trade and stock market returns. However, different countries’ markets behave differently to macroeconomic factors and political events such as Nepal, which saw that new political information affected its market prices 2-3 days from the event date and there were abnormal positive returns before and after the event date. This is contrary to Jabarin’s conclusions where she states that Palestine and Amman’s stock markets are inefficient and do not absorb uncertain information and noisy events. She also highlights the importance of the implementation of economic policies, such as working on reducing the inflation rate, since it has a significant effect on stock returns. Lastly, Jabarin encourages in taking the study on a wider scale by implementing other economic factors, such as unemployment rates and interest rates. Moreover, studies would benefit greatly in understanding by including other markets, both in advanced and developing countries, for comparison.

Building on Jabarin’s work, Ahmed (2023) contradicts this in terms of geopolitics and war. He states that fear of political instability has had a negative effect on the stock market return and the risk profiles of financial assets [3]. Through his review, it was found that there is an inverse relationship between political risk and stock returns and political risk also influences currency carry trade returns. Ahmed’s research objective was to find the affects that the Russia-Ukraine crises had on the European stock market. Firms belonging to the STOXX Europe 600 Index were used from the Refinitiv DataStream. This index represents publicly traded firms with large, medium, and small capital from major European countries. However, thirteen firms had to be dropped from the dataset since daily price or market value data were not available for the entire estimation period. Event-study methodology was utilised over a period from the 25th trading day before the event date of 21 February 2022 to the 25th trading day following the event. However, one must keep in mind that the author’s sample countries are not ‘direct’ participants in the war, although they are providing humanitarian and military support to Ukraine and imposing restrictions on Russia. The key findings through this event-study were that geopolitical crises negatively impact European stock returns (military build ups, threats of war, terrorism) and small and median-cap firms were more heavily affected. Ahmed concluded from this study that the impact of this crisis is considerably broader and deeper than the impact of previous political events with the financial sector having the most severe effect. Moreover, findings demonstrate significantly negative average abnormal returns (AARs) surrounding the short-event windows (-3 to +3 days). Finally, further research was encouraged to comprehensively examine the outcomes and repercussions of the crisis, including its effects on global stock markets and the broader geopolitical landscape.

With regards to machine learning, Vijh (2020) focused on comparing different machine learning techniques to predict the stock closing price [4]. The main objective was to compare prediction accuracy between ANN (Artificial Neural Network) and RF (Random Forest). The stocks observed were Nike, Goldman Sachs, Johnson and Johnson, Pfizer and JPMorgan Chase and Co. In terms of the dataset itself, it was obtained through Yahoo Finance and the information that was used were the high, low, open, close, adjacent close and volume. However, Vijh created new metrics for the training model which comprises of stock high minus low price (h-l), stock close minus open price (o-c), stock price’s seven days’ moving average (7 days ma), stock price’s fourteen days’ moving average (14 days ma), stock price’s twenty-one days’ moving average (21 days ma), stock price’s standard deviation for the past seven days (7 days std dev). After performing the tests for each model, it was concluded that the ANN proved to be a better technique. Vijh came to this conclusion by finding the final minimalized errors in the predicted price. They were subjected to the Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Bias Error (MBE). Lastly, Vijh states that prediction accuracy can be improved using developing deep learning models and by adding external variables to the dataset such as financial news articles and financial parameters (closing price, traded volume, profit, and loss statements).

Compared to Vijh’s work, a different approach was conducted by Purnama (2023) as he compared Support Vector Machine (SVM) and Linear Regression (LR) for stock price prediction [5]. He also obtained the data from Yahoo Finance and the stock used was PT. Vale Indonesia. Moreover, the data was taken over a nine-year period between June 2013 and June 2023. The variables that made the dataset were the date, open, high, low, and close. Prior to creating and running the mode two models, Purnama had done data cleaning manually by deleting missing values and min-max normalisation was done. For the methodology, the data was split for 3 different scenarios and RMSE (root mean square deviation) evaluation was carried out for comparison results. The three scenario data splits comprised of 70:30, 80:20 and 90:10. After running these tests, Purnama concluded that the LR algorithm was more accurate at predicting the stock prices than that of SVM. He also states that when more training data is used, the more accurate the prediction results were. Lastly, he suggests conducting research using deep learning algorithms or combining machine learning and deep learning as a comparison to obtain better prediction results.

Koch utilised event-study methodology for gathering news sentiment towards the equity markets in the UK over a 5-year timeframe during the Brexit period. After using the time varying connectedness of Dilmoid and Yilmaz and analysing over 34000 news articles, NAME’S findings suggest that investors saw Brexit as a smaller threat to EU-based companies and therefore, less threat to EU companies. Moreover, the spillover increased when articles were on topics related to investors and through this, determined that financial market sentiment in news articles were better indicators rather than general ones.

Alongside financial news articles, social media is another spectacle to consider when performing sentiment analysis, as Khan demonstrates. This was performed by predicting the accuracy of STOCK price for 10 subsequent days and comparing the results between social media and financial news articles. Khan’s research displays the importance of reducing the spam that is present in social media comments to obtain workable data. After comparison, it was found that the highest prediction score was achieved using social media and financial news with an RF model providing consistency and achieving the highest accuracy score. It also worked well since dataset was mixed with numerical and categorical features. Social media was found to have more influence on stock price prediction. Moreover, a positive effect on prediction performance when applying feature selection and spam tweet reduction. NAME concluded that for more in depth analysis, other social media data could be used to compare effects such as Facebook.

Straying away from regional and company specific news, Jin provides insight on the effect of international news articles on the global stock market. The study demonstrates that after analyzing 35 countries, overall, international news sentiment positively influences stock market returns globally. It is crucial to note that Jin highlights that the effect of news sentiment is weaker in countries with higher financial development and stronger financial institutions. Overall access to international media was another factor that demonstrates the effect on stock markets as an increased access amplifies its effect. Additionally, NAME states that the impact is more pronounced in countries with high levels of economic openness. Lastly, news sentiment’s influence on markets grew stronger after the 2008 financial crisis, especially in emerging markets relative to developed ones.

Pashkov’s approach was at company level however, through applying sentiment analysis on company reports to then be used for a prediction model for the returns of the next quarter. Data was consisted of US based companies’ filings over the past 20 years. Based on the sections of the reports, a Bag of Words (BOW) corpus was conducted. After that, TF-IDF was used to weigh the words used in the BOW which were then scored into different categories which is comprised of 'Positive', 'Negative', 'Weak' and 'Understatement'. Pashkov opted for a CatBoost model for fitting rather than Linear Regression as it relied significantly on pre-processing. Based on the results obtained, Pashkov concludes that specific categories (e.g., “Weak & Underst”) are more informative than general positive/negative scores. He also found that weak emotional category tops the ranking, outweighing the importance of all previous quarters' returns. As for the data itself, adding sector and month features was explored to enhance model performance to which sector feature notably improved the model, pushing the AUC above 0.56. This cannot be said the same for the month feature as it has a minimal impact, adding just 0.4 percentage points over the sentiment-augmented model. As for improvements, Pashkov challenges were that MD&A extraction needs improvement due to varying keywords and section labels and HTML was unavailable in some filings which in turn complicates parsing.