XI SINGEP
Simpósio Internacional de Gestão, Projetos, Inovação e Sustentabilidade
International Symposium on Management, Project, Innovation and Sustainability
ISSN: 2317-8302

CYRUS Institute of Knowledge
MAKE A DIFFERENCE
CIK 11th INTERNATIONAL CONFERENCE

INVESTIGANDO A TESLA ATRAVÉS DA LENTE DA CLUSTERIZAÇÃO: Analisando as tendências da Tesla no mercado de ações ao longo da pandemia de COVID-19.

NVESTIGATING TESLA THROUGH THE LENS OF CLUSTERING: Analyzing Tesla's trends in the stock market over the course of the COVID-19 pandemic.

GUSTAVO HENRIQUE ROMÃO
UFSJ - UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI

HYGOR SANTIAGO LARA
UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP

Agradecimento à orgão de fomento:
Acknowledgment to the funding agency: This work was conducted with the support of the Artificial Intelligence Study Group (GEIA).

XI SINGEP
Simpósio Internacional de Gestão, Projetos, Inovação e Sustentabilidade
International Symposium on Management, Project, Innovation and Sustainability
ISSN: 2317-8302
CYRUS Institute of Knowledge MAKE A DIFFERENCE
CIK 11th INTERNATIONAL CONFERENCE

INVESTIGANDO A TESLA ATRAVÉS DA LENTE DA CLUSTERIZAÇÃO: Analisando as tendências da Tesla no mercado de ações ao longo da pandemia de COVID-19.

Objetivo do estudo
Testar a utilização da aplicação de algoritmos de clusterização em bancos de dados que contêm informações do mercado de ações em diferentes anos concatenadas com as informações das ações da Tesla, como forma de obter mais informações.

Relevância/originalidade
A relevância deste estudo reside principalmente em mostrar o poder dos algoritmos de clusterização como ferramentas para obter mais informações na análise de dados do mercado de ações.

Metodologia/abordagem
Ao aplicar algoritmos de clusterização em vários bancos de dados que contêm dados de ações de mais de 4000 empresas, o presente estudo utilizará os agrupamentos como pontos de partida para análises mais investigativas.

Principais resultados
Nossos resultados sugerem que as tendências do mercado de ações da Tesla podem ser observadas de forma semelhante em outros casos ao longo dos anos anteriores à pandemia de COVID-19 e em empresas fora do setor de fabricação de carros elétricos.

Contribuições teóricas/metodológicas
As principais contribuições deste estudo residem em demonstrar o poder dos algoritmos de clusterização como uma ferramenta na análise do mercado de ações.

Contribuições sociais/para a gestão
Ao chamar a atenção para correlações não intuitivas entre diferentes empresas no mercado de ações, este estudo apresenta uma nova abordagem para descobrir tendências de mercado e pode trazer uma nova dimensão para a análise do mercado de ações.

Palavras-chave: Clustering, COVID-19, Mercado de ações, Tesla, Aprendizado de máquina

NVESTIGATING TESLA THROUGH THE LENS OF CLUSTERING: Analyzing Tesla's trends in the stock market over the course of the COVID-19 pandemic.

Study purpose
To test the use in applying clustering algorithms on databases that contain their stock market information on different years concatenated with Tesla's stock information from the as a way as a way to obtain more information.

Relevance / originality
The relevance of this study lies mainly in showing the power of clustering algorithms as tools in obtaining more information in stock data analysis.

Methodology / approach
By applying clustering algorithms in the many databases containing stock data of over 4000 companies, the present study will use the clusterings as starting points for more investigative analyses.

Main results
Our findings would suggest that Tesla's stock market trends can be similarly observed in other instances across the years before the COVID-19 pandemic and in companies outside the field of electric car manufacturing.

Theoretical / methodological contributions
This study's main contributions lie in showing the power of clustering algorithms as a tool in stock market analysis.

Social / management contributions
By bringing attention to non intuitive correlations between different companies in the stock market, this study presents a new approach to discovering market trends and can bring a new dimension to stock market analysis.

Keywords: Clustering, COVID-19, Stock Market, Tesla, Machine Learning

1 Introduction

On the turn of the year leading into 2020, the world faced an unprecedented crisis due to the COVID-19 pandemic, having lived through its immensely disruptive impact in a wide range of industries in the stock market, known for its high sensibility to external factors. The tech industry in particular experienced an acute bullish trend in stock prices during this time, in large part due to the boosted dependency on communication technology during the long periods of social isolation (GEORGESCU, 2021).

However, another notable growth in stock value was observed in Tesla, an electric car manufacturer, during this period. This raises an interesting discussion on whether labelings such as "tech companies" offer enough explanation for the phenomena behind the sudden hike in stock value in companies like Apple or Amazon in the year of 2020.

And so we're led to question if there are any clearer and more objective ways to group companies as it pertains to their behavior in the stock market. So that certain attributes can be observed in a more pronounced form or if certain non intuitive correlations could be made more apparent once compared to each other in a smaller scaled, more homogenized grouping. If Tesla did indeed show a similar trend to that of tech companies in the stock market during the COVID-19 pandemic, then it would be interesting to look further into if any other companies (not necessarily tech companies themselves) also showed a similar behavior, so as to better understand this trend as a whole.

Could this growth spike during the beginning of the COVID-19 pandemic really be something unique to tech companies? Or could this be another case of industries non intuitively correlated influencing one another (Zhenhua Liu, Hui-Kuan Tseng, Jy S. Wu, and Zhihua Ding., 2020)? Are there better and more concrete ways we could be grouping companies together so as to better understand their behavior in the stock market? And how would we get to that grouping?

It is of great interest to investors and financial analysts to have access to more concrete information about joint market trends through the use of more in-depth data analysis techniques. As they can potentially reveal more information that can prove relevant to decision making processes.

This study investigates Tesla's behavior in the stock market in the years of 2020, 2021, and 2022 and analyzes if other companies have shown similar trends in previous or concurrent years. The central idea is to explore whether the trend observed in Tesla, from the beginning to the end of the pandemic, can be generalized to any specific set of companies or if it was indeed unique.

2 Objectives

With the aim of analyzing Tesla's behavior patterns and finding companies with similar trends, historical stock price data from different companies/entities has been collected and analyzed from 2014 and up. Clustering algorithms are then used to identify which companies are grouped together based on their stock market information. The comparative analysis allows us to observe if there were other periods in which certain companies exhibited a behavior similar to Tesla's growth in 2020, for example.

The results of this research are relevant not only to investors and financial analysts but also to demonstrate the power of applying clustering algorithms in identifying similarities between trends in the stock market.

The clusters formed can provide valuable insights into non-intuitive correlations. It is also one of the main interests of this study to learn which companies followed Tesla's trends throughout the entire pandemic period (2020, 2021, 2022) in order to identify if there are any other companies that reacted to this atypical period in a similar manner.

3 Theoretical foundations

3.1 Measures, coefficients, and indices used in the study

- Mean: The mean is obtained by summing all the values and dividing the result by the total number of values.
- Median: The median represents the middle value of a set of numbers when they are arranged in ascending or descending order. If there is an odd number of observations, the median is the middle value. If there is an even number of observations, the median is the average of the two middle values.
- Variance is a statistical measure that quantifies the dispersion or spread of a set of data points around their mean.
- Standard deviation: The standard is calculated as the square root of the variance. The higher the standard deviation, the greater the dispersion of the data.
- Maximum: The maximum is the highest value in a data set.
- Minimum: The minimum is the lowest value in a data set.
- Range: The range is the difference between the maximum value and the minimum value in a data set.
- Volatility: Volatility is often used in finance to assess the risk associated with an investment. The higher the volatility, the greater the variation in prices or returns.
- Average return: Average return is a statistical measure that calculates the average returns of an investment over a specific period.
- Skewness: Skewness indicates whether the distribution is symmetric (no skew) or asymmetric (skewed to one side). Skewness can be positive (right-skewed), negative (left-skewed), or zero (symmetric distribution).
- Correlation coefficient: The correlation coefficient ranges from -1 to +1, with +1 indicating a perfect positive correlation, -1 indicating a perfect negative correlation, and 0 indicating weak or no linear correlation. A value close to +1 or -1 indicates a

strong linear relationship, while a value close to 0 indicates a weak or no linear relationship.

- Silhouette Score: The Silhouette Score is a technique that provides metrics that help identify the internal structure of the data and assist in choosing the optimal number of clusters.

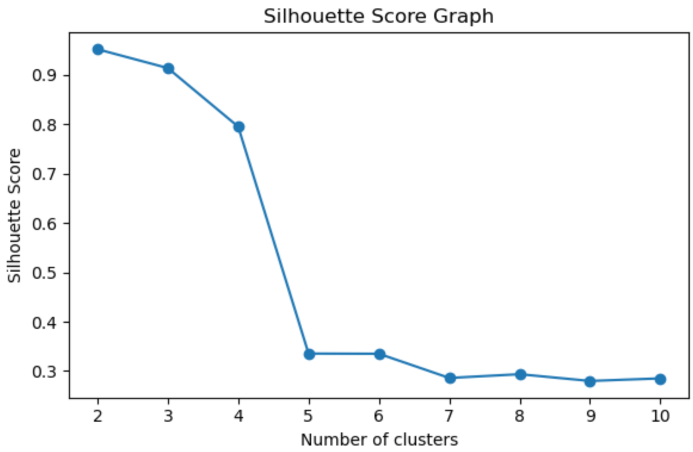

Figure 1. Example of a Silhouette Score graph.
Source: Research data.

The Silhouette Score provides an evaluation of the clustering quality, taking into account intra-cluster compactness and inter-cluster separation. The Silhouette Score ranges from -1 to 1, where values close to 1 indicate that the object is well-fitted to its cluster, values close to 0 indicate overlap or intermingling between clusters, and values close to -1 indicate that the object may be closer to points from other clusters than its own cluster.

- For each object in the dataset, we calculate two metrics: the average intra-cluster dissimilarity (a) and the average minimum dissimilarity to all other clusters (b).
- The average intra-cluster dissimilarity (a) is calculated as the average distance between the object and all other objects within the same cluster.
- The average minimum dissimilarity to all other clusters (b) is calculated as the average distance between the object and all objects from the other clusters, selecting the nearest cluster. The Silhouette Score (s) for each object is then calculated as: s = (b - a) / max(a, b).
- The global Silhouette Score is obtained by calculating the average of the Silhouette Score of all objects.
- Interpreting the Silhouette Score: A Silhouette Score close to 1 indicates that the object is well-fitted to its cluster, with good separation from other clusters. A Silhouette Score close to 0 indicates that the object may be in an overlapping or intermingling region between clusters, making its assignment less clear. A negative

Silhouette Score indicates that the object may be closer to points from other clusters than its own cluster, suggesting a possible incorrect assignment.

3.2 Clustering algorithms to be applied to the databases

3.2.1 K-means

The K-means algorithm is a widely used clustering algorithm that groups a dataset into K distinct clusters. The objective is to find the centroids (centers) of the clusters, so that the sum of distances between each point and the centroid of its cluster is minimized.

Cui (2023) shows an implementation of the K-means algorithm. The author highlights that the K-means algorithm is widely used in finance and other areas, being considered a simple and easy-to-apply algorithm. However, it also presents some limitations, such as the difficulty in determining the value of K and finding the initial cluster centers across the dataset. It should be noted that it also makes the assumption that the clusters to be formed in the dataset take on a circular shape.

3.2.2 DBSCAN

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a clustering algorithm that groups data points based on local density. It identifies dense regions of nearby points and classifies the points as core, boundary, or noise. DBSCAN does not require prior specification of the number of clusters and is capable of finding clusters of arbitrary shapes and sizes. It uses two main parameters: the maximum distance between points in the dataset to consider them neighbors and the minimum number of points in a neighborhood to form a cluster.

3.2.3 Spectral Clustering

Spectral clustering organizes the data into a graph structure, where each data point represents a node and the edges represent the similarity between points. This graph structure is built using an affinity matrix, which measures the proximity between points. From this graph representation, the algorithm performs spectral decomposition to obtain the eigenvectors associated with the largest eigenvalues. These eigenvectors are used to allocate the points into clusters using partitioning techniques like K-means. Thus, spectral clustering utilizes spectral information and the graph structure to perform data clustering. It is known for its high computational cost, despite being a valuable tool in cluster analysis due to there not being much need to make assumptions on the shape of the dataset.

3.2.4 Gaussian Mixture

The Gaussian Mixture Model (GMM) is a machine learning technique that aims to model the probabilistic distribution of a dataset by combining multiple Gaussian distributions. It assumes that the data is generated from a mixture of different Gaussian components, each representing a cluster or group in the dataset. The algorithm uses the Expectation-Maximization (EM) process to estimate the parameters of the Gaussian distributions and determine the data assignment to each component. GMM is widely used in clustering, pattern recognition, and probability density modeling.

4 Methodology

In order to explore stock market trends, information on Tesla's stock market in 2020, 2021, and 2022 is concatenated with different databases containing information on stock market values of other companies from 2014 onwards until 2022. These datasets are subjected to clustering algorithms, allowing the identification of groups of companies with similar behaviors.

This approach enables us to observe which companies were grouped together with Tesla, providing insights into which companies exhibited similar trends in the stock market. This comprehensive analysis allows us to identify not only if there were companies other than technology companies in 2020, for example, that had similar fluctuations but also in which years these trends occurred.

This offers a broader and more objective view, allowing for the identification of patterns and correlations between different companies and sectors over time, contributing to a deeper understanding of the stock market and the factors influencing its behavior.

4.1 Defining the parameters

The dimensions of the analysis are defined based on the chosen parameters for the database. It is from these parameters that the unsupervised clustering algorithms will detect behavior similarities.

This study defines as relevant dimensions for the analysis the ones which include the closing price of stocks and traded volume. Therefore, there are individual columns for each of the following metrics for each company/entity: mean, median, standard deviation, maximum, minimum, range, volatility, average return, and skewness.

4.2 Establishing the number of clusters (K)

Based on the analysis of the algorithms used in this study, it is important to highlight that some of them require the predefinition of the number of clusters (K) to be formed. Although there is no definitive method to determine the ideal number of clusters, this study adopts an approach based on the Silhouette Score and the quality of the formed clusters (e.g., clusters that are too large are considered unsatisfactory for the analysis).

4.3 Choosing the most appropriate algorithm for the analysis and applying it to each period

Among the four algorithms mentioned, the ones that display the best performance in the tests on the database are chosen. Given the large number of tests to be applied to these databases, the algorithm's execution time is taken into account. An algorithm that takes a long time to execute is considered unsuitable for the purposes of this study. Another criterion to be considered in the choice of algorithms is the quality of the formed clusters. An algorithm that groups Tesla with too many other companies is considered to have unsatisfactory performance for the analysis.

4.4 Defining the comparisons to be made

Firstly, the chosen algorithms are applied to seven different databases. They consist of all the companies in the database from 2014 onwards until 2022 concatenated with Tesla's information in 2020, 2021, and 2022.

With the clusters having been formed, the correlations between companies within the same cluster are then analyzed. The goal is to examine the efficiency of the constructed model and the selected clustering algorithms to determine if the grouped companies/entities are indeed correlated.

A second analysis is then performed on three databases from 2020, 2021, and 2022, where all companies (including Tesla) have information for their respective years. The clusters formed with Tesla throughout these three years are observed and the companies that exhibited similar trends to Tesla during the entire pandemic period can be identified and further looked into.

5 Results and Discussions

5.1. On the choice of algorithms to be applied to the database

During the development of the study, the algorithm's execution time also proved to be an important consideration, as several tests needed to be performed. In this regard, the spectral clustering algorithm was found to be ineffective due to its longer execution time for such a large database. The K-means, Gaussian Mixture, and DBSCAN algorithms demonstrated to be more efficient in terms of execution time.

Table 1. Table containing the execution times of the selected algorithms when applied to one of the study's databases.

Model	K-Means	Gaussian Mixture	DBSCAN	Spectral
Tempo de execução	0.3 seconds	0.5 seconds	0.5 seconds	3 hours and 24 minutes

Due to the high dimensionality of the dataset, the DBSCAN algorithm did not produce satisfactory clusters (Tang, J., Chen, Z., Fu, A. W., Cheung, D. W., & Liu, J. 2018). In several instances, all companies were grouped into a single cluster, which is not useful for the intended analyses.

Considering the need to perform tests within a viable time frame and the high dimensionality of the database, the DBSCAN and Spectral algorithms were scrapped from further use in this study.

5.2. Notable results from the tests

The comparison of Tesla's stock trends in 2020, in particular, with companies in other years is of great interest to the study, given that it was a year of significant stock market value increase (the initial year of the pandemic). Identifying similarly bullish trends in companies in previous years is one of the main focuses of this study.

Figure 2. Graph of Tesla's stock closing prices from 2019 to 2023.
Source: Research data.

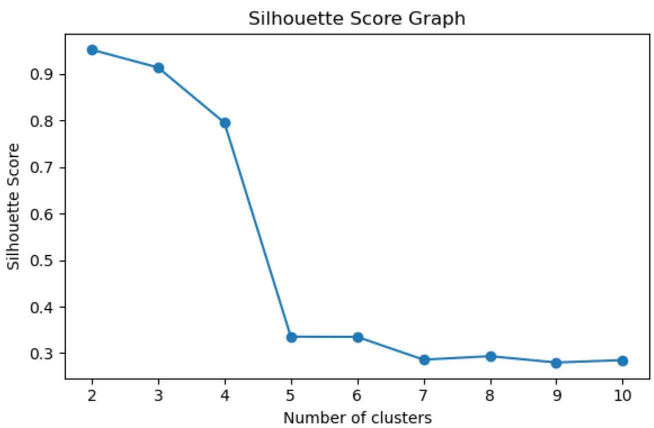

Figure 3. Silhouette Score Graph for the database containing information about Tesla's stocks in 2020 and the rest in 2014.
Source: Research data.

By setting the value of K as 3 (the ideal number of clusters according to the silhouette score), the K-means algorithm grouped Tesla with only Apple. However, the GMM did not produce a satisfactory clustering for Tesla, resulting in a cluster of a very large size. Therefore, the analysis proceeds with K equal to 4. And so, both algorithms (K-means and GMM) produced identical results when applied to the dataset, which contains information about Tesla's stocks in 2020 (the initial period of the pandemic) and the other companies in 2014. In this case, Tesla in 2020 was exclusively grouped with Apple. The fact that they alone were grouped together suggests a similar and particular trend in the growth of Tesla's stock closing price and traded volume in 2020 with the growth of Apple’s in 2014.

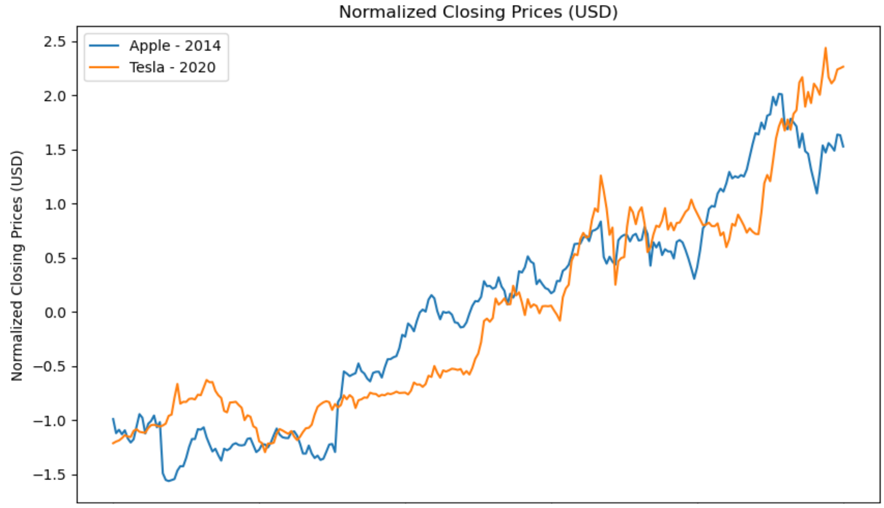

Figure 4. Graph of Tesla's normalized closing prices in 2020 and Apple's normalized closing prices in 2014.
Source: Research data.

It should be noted that it was in the year of 2014 that Apple brought the Iphone into the world’s largest mobile services provider (serving over 760 million customers as of 2014) in Mobile China after signing a multi-year deal at the end of 2013. By opening itself to a much broader market, Apple could have inspired significant confidence in its investors at the time.

By observing the strongly correlated trend between Tesla in 2020 and Apple in 2014, we now look into Tesla’s 2021 trend and Apple’s 2015 trend, looking to find out whether the correlation persisted in the subsequent years for each.

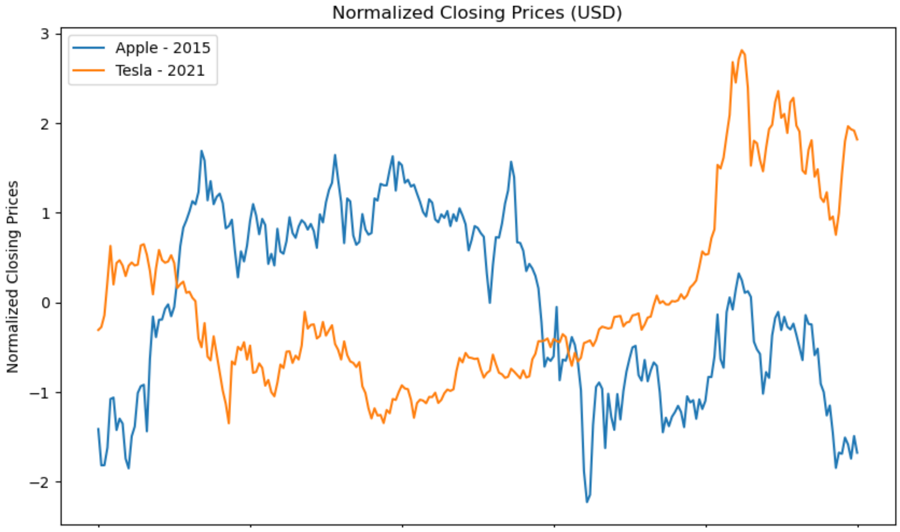

Figure 5. Graph of Tesla's normalized closing prices in 2021 and Apple's normalized closing prices in 2015.
Source: Research data.

With the correlation between the two now being -0.49, we can observe that Apple doesn't follow the same bullish trend into 2015 that Tesla did entering 2021.

Upon applying the clustering algorithms on a database containing companies with information from 2016 and Tesla with that of 2020's, both K-Means and GMM grouped Tesla together with Bank of America Corporation, among others.

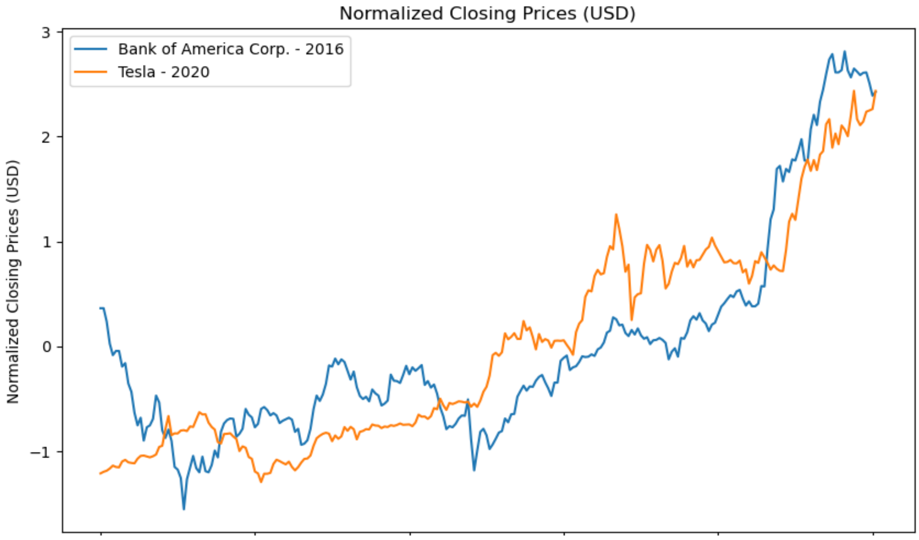

Figure 6. Graph containing normalized closing prices of Tesla (year 2020) and Bank of America Corporation (year 2016).
Source: Research data.

Tesla's stock behavior in 2020 and that of Bank of America's in 2016 showed a correlation of 0.85, indicating a strong positive correlation in their market trends in each respective year.

Among the most notable, game changing world events of 2016 one could cite Donald Trump's win over the U.S. presidency or even Britain voting to leave the European Union. With both events signaling a shift in the social political landscape on a worldwide scale.

It is in the interest of this study to observe whether Bank of America's market trend in the years following 2016 is similar to the one observed in Tesla from the year of 2020 onwards.

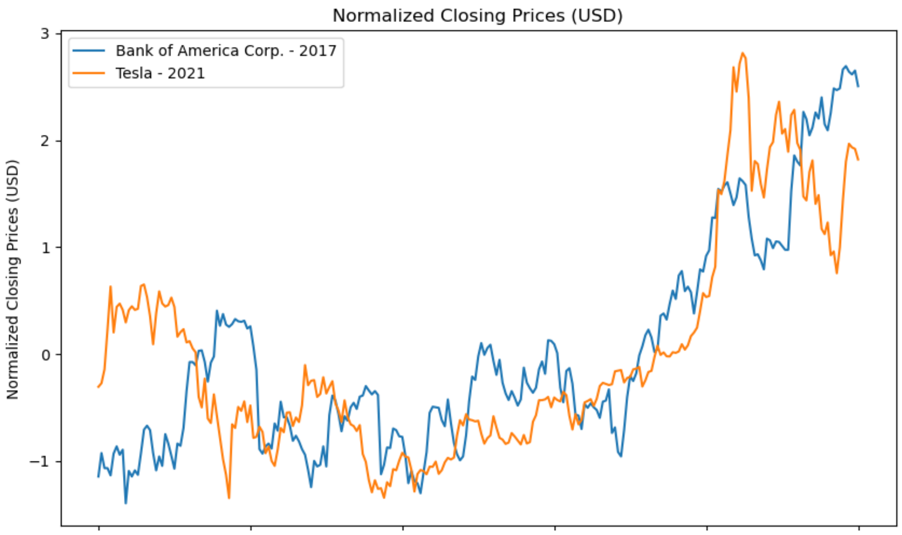

Figure 7. Graph containing normalized closing prices of Tesla (year 2021) and Bank of America Corporation (year 2017).
Source: Research data.

Tesla's stock market behavior in 2021 and Bank of America Corporation's in 2017 exhibited a correlation coefficient of 0.73, also indicating a strong positive correlation between their closing stock prices in each respective year.

By 2017, Bank of America's stock had surged over 600% since its post-crisis lows, in large part due to its successful embrace of digital banking. By reducing physical branches and increasing its active mobile-banking user base, the bank had significantly improved efficiency and lowered expenses. This strategic shift towards digital banking had not only attracted a larger customer base but also allowed for convenient and cost-effective transactions. With its strong growth, cost reductions, and commitment to digital innovation, Bank of America had positioned itself as a well-run and forward-thinking banking operation, likely driving investor confidence and contributing to its stock price rise.

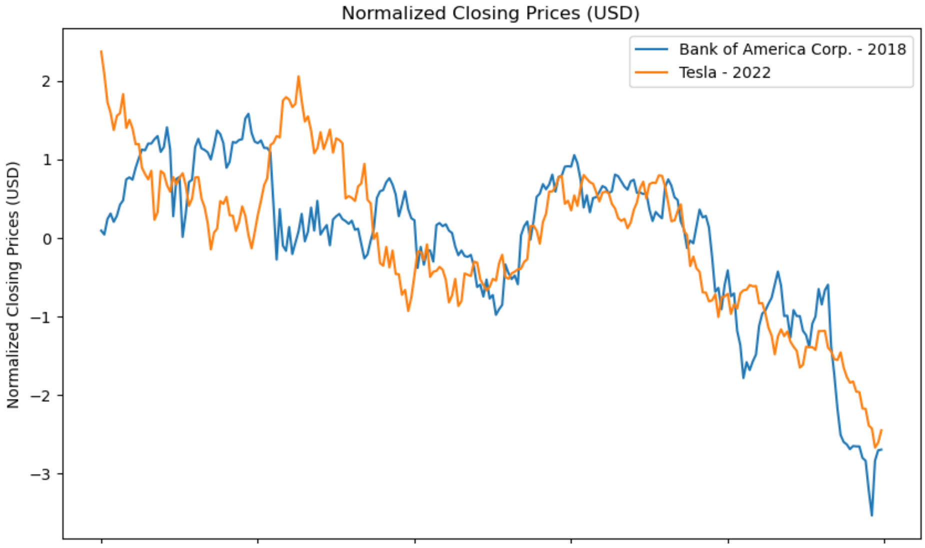

Figure 8. Graph containing normalized closing prices of Tesla (year 2022) and Bank of America Corporation (year 2018).
Source: Research data.

Tesla in 2022 and Bank of America in 2018's historical stock information exhibited a correlation coefficient of 0.71, indicating a strong positive correlation. So as Tesla's stocks turned bearish over the course of 2022, a similar trend can be observed in Bank of America's in the year of 2018.

In 2018, the banking industry, including Bank of America, experienced a weak performance despite positive expectations. Bank of America's business had actually been experiencing increased earnings, loan and deposit growth, and improved efficiency ratios. However, several industry-wide factors contributed to the subdued stock performance. Rising interest rates did not lead to significant margin expansion as anticipated, deregulation measures did not benefit big banks like Bank of America, and the benefits of tax reform were already priced into the stock before 2018. Overall, the stock's lackluster performance could be attributed to already high expectations and broader industry dynamics.

This can be seen as a parallel to the year of 2022 in which many companies, mostly within the tech field, experienced mass layoffs due to stock earnings reports not being able to keep up the momentum built over the previous two years in which the following set of events took place: a massive shakeup in the social-political landscape across the globe and subsequent embrace of new communications technologies that, for one reason or another, aimed to reduce the physical footprint of its users as they went on about their daily lives. Through the lens of clustering, a non intuitive historic parallel to Tesla's stock behavior along the years marked by the effects of the COVID-19 pandemic (2020 to 2022) could be found in that of Bank of America's from 2016 to 2018.

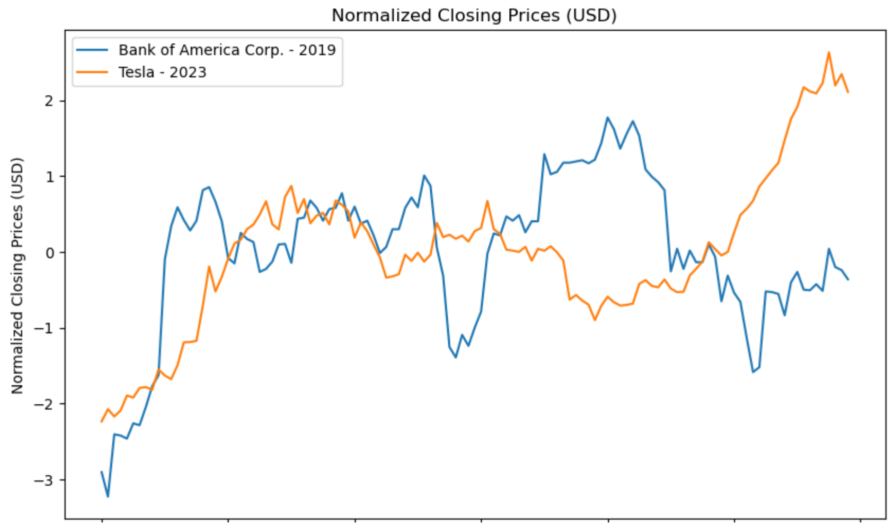

Figure 9. Graph containing normalized closing prices of Tesla (year 2023) and Bank of America Corporation (year 2019).
Source: Research data.

We now look into Tesla in 2023 (up to June 25th) and Bank of America Corporation (up to June 25th) in 2019 and we can see that they showed a correlation coefficient of -0.17, suggesting that the previously observed strong positive correlation between Bank of America from 2016 to 2018 and Tesla from 2020 to 2022 did not carry into the following year for each company.

In the database containing Tesla's 2020 stock historical data and other companies also in the year of 2020 (the year where the impact of the COVID-19 pandemic was truly first made present), both the K-means algorithm and the GMM algorithm grouped the following companies together with Tesla: American Airlines Group Inc. (Airline), Apple Inc. (Technology), Amazon.com Inc. (E-commerce), Bank of America Corporation (Banking), Bionano Genomics Inc. (Genomics), Ford Motor Company (Automotive Manufacturer), Gevo Inc. (Renewable Energy), Nio Inc. (Electric Vehicle Manufacturer), Ocugen Inc. (Biotechnology), and Palantir Technologies Inc. (Data Technology). This finding suggests a similarity in their stock market behavior to that of Tesla throughout the year of 2020, when the impact of the COVID-19 pandemic truly became widespread.

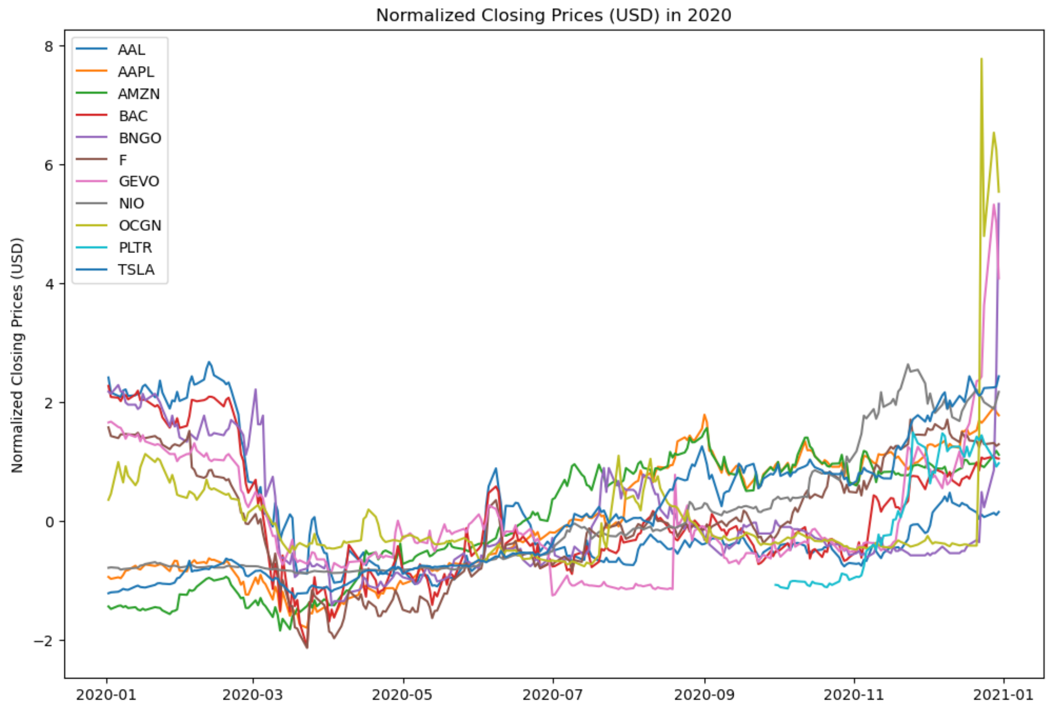

Figure 10. Graph containing normalized closing prices of the cluster formed with Tesla from 2020 and the rest of the companies with information from 2020.
Source: Research data.

Among the companies in this cluster, the ones that show the highest correlation coefficient with Tesla are Apple and Advanced Micro Devices. And, out of the companies in the cluster, Apple showed the highest sum of correlations with all the others. This implies that Apple's stock prices or financial performance have a relatively stronger relationship with the rest of the companies in the cluster compared to the other companies' relationships with each other.

Finally, it was in the year 2022 that many technology companies carried out mass layoffs due to the decline in the stock market in the first year of widespread break-off from the social isolation restrictions imposed by the pandemic. With larger clusters in both algorithms being observed even when applying the K values with the best silhouette scores. This suggests that Tesla followed a broader trend when slipping into a bear market in 2022.

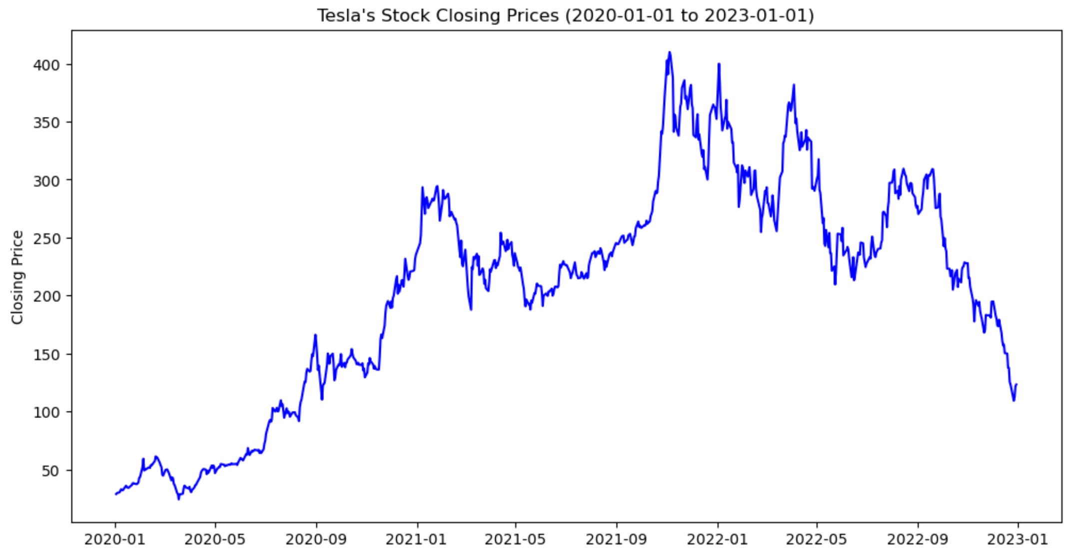

Figure 11. Closing prices of Tesla's stocks entering and exiting the COVID-19 pandemic.
Source: Research data.

During the joint analysis of Tesla's information over the years 2020, 2021, and 2022 with other companies/entities (in the same years), a result of note was that Nio Inc, a Chinese electric car manufacturer, was consistently grouped in the same clusters as Tesla for each year and by all selected algorithms. Nio Inc. is commonly referred to as the "Chinese Tesla" and has the highest market capitalization in the US exchange market among its Chinese-based electric vehicle competitors (Yi Xuan Lim and Consilz Tan, 2022). This finding suggests that the model detected similarities in the closing price data and trading volume between Nio and Tesla throughout the entire period.

This discovery reveals a correlation between the two companies, indicating a possible relationship in market fluctuations and investor behavior during this challenging period. The graph below shows the normalized closing prices of Tesla and Nio throughout the years 2020, 2021, and 2022.

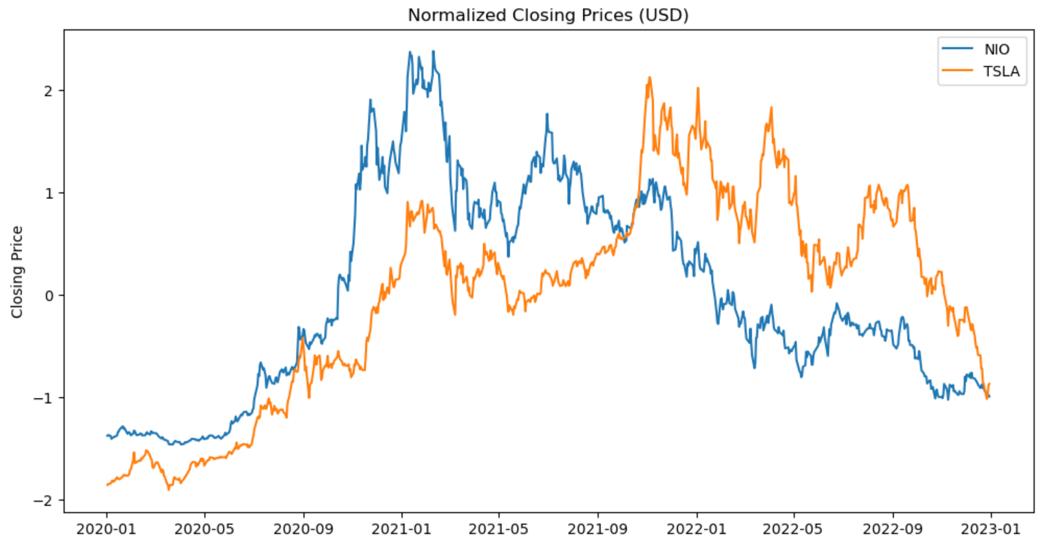

Figure 12. Normalized closing prices of Tesla and Nio Inc. throughout the pandemic period.
Source: Research data.

The correlation coefficient between Tesla and Nio (over the course of these 3 years) is 0.508, suggesting a moderate positive correlation between Tesla and Nio. This indicates that there is a tendency for the stock prices of Tesla and Nio to move in the same direction, although not necessarily in a linear manner.

5.3 Discussions

This study reveals the power in using clustering algorithms to help identify market trends, as it has identified clusters among correlated companies that were not initially obvious. Nio, for example, was grouped with Tesla in every stage of the pandemic, a period known in large part for its disruptive effect on tech companies in particular.

Furthermore, based on the obtained results, it can be observed that the impact of the initial year of the pandemic (2020) on Tesla was indeed relatively noteworthy, as the clusters with companies from previous years were small (e.g., in a database containing Tesla with 2020 information and the rest with their 2014 historical stock information, only Apple, out of all of the companies in a dataset containing over 2000 observations, was grouped with Tesla).

Also, by grouping Tesla (from 2020 onwards) with different companies over different years, this study was able to detect a similar trend to Tesla's in the COVID-19 pandemic onwards in that of Bank of America Corporation's in 2016 up to 2019. Which shows that Tesla's spike in stock value and subsequent trend was something that could be similarly observed in another time in history by another company and in a seemingly completely different context.

It should be noted that as Tesla slipped into a bear market, so too did a lot of other companies in the year of 2022 (the year of widespread disruption in social distancing measures). Considering the selected algorithms grouped Tesla into larger clusters when applied to databases containing Tesla's and the rest of the companies' historical stock data of the year of 2022, through the lens of clustering, that can be seen as an indication of a larger trend having taken place.

All in all, it is expected that this study has demonstrated the power in making use of clustering algorithms as tools for investors and financial analysts to help identify joint behaviors in the stock market. And so we're led to believe that it may prove useful as a starting point for new approaches to analyses of trends in the stock market.

6 References

- CUI, Mengyao. Introduction to the K-Means Clustering Algorithm Based on the Elbow Method. Accounting, Auditing and Finance, v. 1, p. 5-8, 2020. DOI: 10.23977/accaf.2020.010102. ISSN 2616-3136.
- Gao, M., & Shi, G. (2020). Ship-handling behavior pattern recognition using AIS sub-trajectory clustering analysis based on the T-SNE and spectral clustering algorithms. Ocean Engineering, 205, 106919.
- GEORGESCU, Tiberiu-Marian. A Study on how the Pandemic Changed the Cybersecurity Landscape. Informatică Economică, vol. 25, no. 1/2021, p. [42-43]. DOI: 10.24818/issn14531305/25.1.2021.04.
- Liu, Z., Tseng, H. K., Wu, J. S., & Ding, Z. (2020). Implied volatility relationships between crude oil and the U.S. stock markets: Dynamic correlation and spillover effects. Resources Policy, volume 66, June 2020, 101637.
- Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2018). A comprehensive survey of cluster analysis techniques. Annals of Data Science, 2(2), 165-209.
- Tang, J., Chen, Z., Fu, A. W., Cheung, D. W., & Liu, J. (2018). On density-based clustering in high-dimensional space: A summary and perspective. Data Science and Engineering, 3(1), 1-17.
- Yi Xuan Lim and Consilz Tan. "Do negative events really have deteriorating effects on stock performance? A comparative study on Tesla (US) and Nio (China)." Journal of Asian Business and Economic Studies, vol. 1, no. 1, 2022, pp. 105-123.