**Smarthinking Tutor Response Form**

Your tutor has written overview comments about your essay in the form below. Your tutor has also embedded comments **[in bold and in brackets]** within your essay. Thank you for choosing Smarthinking to help you improve your writing!

Hello, Damian! I am Melvin M., and I look forward to working with you on this **Grammar and Documentation Review**to improve your writing today. Let's get started!

**\*Writing Strength:**

You use transitional devices in some sentences in your analysis. For example, you write:

In contrast to their research, we explore the development of an information driven wave pre-news release and test the confirmation of the tendency on post-news release.

Here, you write “In contrast,” and this phrase helps readers know that you state a contrasting idea in your sentence. Well done, Damian!

**Sentence Structure:**

You write pronouns that do not have clear antecedents in some sentences in your background analysis. Here is an example from the second paragraph of Section 2:

To do that, **they** implement *support vector machines (SVM),* to classify the tendency,using *text Mining* on online news, their respective comments, and stock market data.

Here, you write the pronoun “they,” but this pronoun does not have a clear antecedent. As a result, readers will not know what the pronoun “they” tries to replace. When you write a pronoun, you must make sure that it has an antecedent (a noun or another pronoun). To improve your sentence, you can change the pronoun “they” to the antecedent being replaced. Compare these examples:

Unclear antecedent: To do that, they used different methods so that their data would be valid.

Revised: To do that, the researchers used different methods so that their data would be valid.

In the first example, I have an unclear pronoun reference because the pronoun “they” does not have an antecedent. In the other example, I change the pronoun “they” to the antecedent being replaced, which is “the researchers,” making my sentence clear. What antecedent is being replaced by the pronoun “they” in your sentence? Make sure to clarify vague pronoun references in other sentences in your analysis.

**Grammar & Mechanics:**

In your background analysis, you have sentences with subject-verb agreement errors. Here is an example from Section 1:

This last term, informed traders, refer to the presence of traders that posses fundamental or alternative company’s information and generate trades upon understanding its respective future market impact.

In your sentence, you use the plural verb “refer” for the singular subject “last term,” making the sentence unclear to readers. Keep in mind that when you write sentences, your subjects and verbs should agree in number. Therefore, when you have a singular subject, like “last term,” you use a singular verb. You can get the singular form of a verb by adding –s or –es to the verb. When you have a plural subject, you use a plural verb (like “refer”), which is also a verb’s base form. Here is an example:

The first term, *covert participation*, means participating without letting anyone know.

Here, I use the singular verb “means” because the subject “first term” is singular, making my sentence clear. What is the singular form of the verb “refer”? In your analysis, make sure that subjects and verbs in other sentences agree in number. For more help, consider checking the lesson on [Subject-Verb Agreement](https://services.smarthinking.com/static/document_library/docs/writeman/5_02_06.cfm).

**Documentation:**

You do not write your reference-page entries properly. Here is an example:

[8] López de Prado, Marcos and Rebonato, Riccardo, Kinetic Component Analysis (June 5, 2016). Journal of Investing, Vol. 25, No. 3, 2016. Available at SSRN: <https://ssrn.com/abstract=2422183> or [http://dx.doi.org/10.2139/ssrn.2422183](https://dx.doi.org/10.2139/ssrn.2422183)

Here, you create an entry for an online journal article with two authors, but you do not write the entry properly because of the way you format the elements and the lines. As a result, you will not adhere to IEEE style. When you create an entry, you write the first name initials and the last names of the authors. Specifically, when you create an entry for an online journal article, after the authors, you write the article title with quotation marks, the journal title in italics, the volume number (“vol.”), the issue number (“no.”), the page numbers (“pp.”), the publication month and year, the descriptor “Online” in brackets, the database (in this format: Available: Database Title, URL), and the access date in brackets. In addition, you use hanging indent for the entries, with the bracketed numbers flushed with the left side. Here is an example:

[1] K. Coffman and L. Lewis, “The Effectiveness of Individualized Treatment Plans,” *Journal of Mental Health*, vol. 3, no. 4, pp. 20-30, June 2018. [Online] Available: ProQuest, https://proquest.com. [Accessed May 23, 2020].

In this example, I write my entry for another similar source properly because I use the appropriate formatting for the elements and the lines. Make sure to write reference-page entries properly to help you adhere to IEEE style.

**Summary of Next Steps:**

* Clarify vague pronoun references.
* Make subjects and verbs agree in number.
* Write reference-page entries properly.

Thank you for submitting your essay for a review, Damian. I enjoyed helping you with this step in the revision process. Have a good day! ~Melvin M.

You can find more information about writing, grammar, and usage in the [Smarthinking Writer's Handbook](https://services.smarthinking.com/static/Document_Library/docs/writeman/contents.cfm).

Please look for comments **[in bold and in brackets]** in your essay below.  
Thank you for submitting your work to Smarthinking! We hope to see you again soon.

Big Data Analytics Capstone

Background Analysis

Damian Etchevest

St. Thomas University

CIS-627-170

Table of Contents

[Introduction 3](#_Toc41082763)

[Section 1 – Information Driven Bars & Entropy 3](#_Toc41082764)

[Section 2 – Sentiment Analysis & Momentum 3](#_Toc41082765)

[Section 3 – Network Analysis & Instrument Profiling 4](#_Toc41082766)

[References 5](#_Toc41082767)

# Introduction

These article serves as an extension of the work performed by Maximilian Frankie and Sean Mondesire [1]. On it, they explore the relationship between published news and the stock market reaction. Their research results achieve corresponding *Mean Square Error (MSE)* of 0.039 and an accuracy of 0.652. In contrast to their research, we explore the development of an information driven wave pre-news release and test the confirmation of the tendency on post-news release.

# Section 1 – Information Driven Bars & Entropy

Section 1 deep dives into the development of an automated signal generator which performs with live feed stock data to determine when exists the presence of informed traders. **[Your sentence is not clear because you write the verb “exists” before the subject “the presence of informed traders” in the relative clause. Consider switching the places of the subject and the verb to make your sentence clear.]** This last term, informed traders, refer to the presence of traders that posses fundamental or alternative company’s information and generate trades upon understanding its respective future market impact. **[The word “posses” is misspelled because of a missing letter *s* at the last letter *s*. Adding another letter *s* at the end will help you spell the word properly.]**

# Section 2 – Sentiment Analysis & Momentum

Lots of research has been done exploring the sentiment of news and their market reactions. This section utilizes time series predictions to measure the inference of alternative data sources (offers the opportunity to work with truly unique, hard-to-process datasets) on the stock prices as a means to determine the characteristics of the information-driven price waves.

An interesting approach aims to predict Chinese stock market movements, tendency [5]. To do that, they implement *support vector machines (SVM),* to classify the tendency,using *text Mining* on online news, their respective comments, and stock market data. Also, their findings show that source’s news quality and audience number can serve to calculate their influence by predicting results’ difference from the normal prediction results.

A similar study examines the predictability of market reactions to bad news [6]. Their approach is innovative as they first use *time series clustering*, to generate news clusters based on subsequent stock returns, and then apply *SVM*  to classify the features extracted from each cluster, by *natural language processing (NLP)* which is a method that interprets human language. Their research identifies four types of market reactions after the news becomes public: downward drift, short-term reversal, medium-term reversal, upward drift.

In their article [7], the authors explore the impact of *Seeking Alpha* (crowd-sourced content service for financial markets) on the stock market’s movements after SA research articles are published. They use classification and regression methods and find incremental in order imbalance, significantly related to the sentiment of research articles and comments, which begins within half-hour after SA publications.

Lopez de Prado and Rebonato [8] examines the use of *Kinetic Component Analysis* *(KCA -* a state-space application that extracts the signal from a series of noisy measurements by applying a Kalman Filter on a Taylor expansion of a stochastic process - on tick data to determine what they refer as “financial inertia”. **[Your use of the singular verb “examines” is inappropriate because the subject “Lopez de Prado and Rebonato” is plural. What is the plural form of the verb “examines”?]** They compare the use of KCA to other signal processing tools as *Fast Fourier Transform (FFT)* or *Locally Weighted Scatterplot Smoothing (LOWESS),* and find that their method: provides confidence intervals estimates of the signal’s position, define the wave’s characteristics (velocity and acceleration of the series), does not exhibit Gibbs phenomenon, and it can be updated online to be forward-looking and resilient to structural changes. ## Predict Wave

# Section 3 – Network Analysis & Instrument Profiling

Finally, we develop clusters of instruments and their respective networks to understand the impact of neighbor’s news releases on their market prices. The aim is to profile the different stocks to develop behavioral classification models to predict the price wave.

Stephen Taylor [9] fundaments the use of Fisher Information Metric, which is relevant in the fields of information geometry and computing geodesics, to generate clusters of stocks based on their distribution. They contribute valuable theory examples: nearest neighbor comparison using the generalized Pareto distribution, and the maximum daily loss over and annual period distribution hierarchical clustering techniques using the generalized extreme value distribution. Even though his findings do not consider instruments networks, it provides behavioral features that contribute to the stocks profiling.

# References

[6] Yu, Xiaowen and Xin, Xin and Chen, Liangliang and Kim, Hang Sun, Predicting Market Reactions to Bad News (March 19, 2018). Available at SSRN: <https://ssrn.com/abstract=3144041> or [http://dx.doi.org/10.2139/ssrn.3144041](https://dx.doi.org/10.2139/ssrn.3144041)

[5] Xie, Yancong. “Stock Market Forecasting Based on Text Mining Technology: A Support Vector Machine Method.” Journal of Computers (2017): 500–510. Crossref. Web.

[7] Farrell, Michael and Green, T. Clifton and Jame, Russell and Markov, Stanimir, The Democratization of Investment Research and the Informativeness of Retail Investor Trading (October 15, 2019). Available at SSRN: <https://ssrn.com/abstract=3222841> or [http://dx.doi.org/10.2139/ssrn.3222841](https://dx.doi.org/10.2139/ssrn.3222841)

[8] López de Prado, Marcos and Rebonato, Riccardo, Kinetic Component Analysis (June 5, 2016). Journal of Investing, Vol. 25, No. 3, 2016. Available at SSRN: <https://ssrn.com/abstract=2422183> or [http://dx.doi.org/10.2139/ssrn.2422183](https://dx.doi.org/10.2139/ssrn.2422183)

[9] Taylor, Stephen Michael, Clustering Financial Return Distributions Using the Fisher Information Metric (May 22, 2018). Available at SSRN: <https://ssrn.com/abstract=3182914> or [http://dx.doi.org/10.2139/ssrn.3182914](https://dx.doi.org/10.2139/ssrn.3182914)