Chihlee University of Technology
Department of International Trade

# Predicting raw material Prices through Keyword (SEO) using Sentiment Analysis

Advisor:

Joint first author:Wong,C. K.,Yip,C. M.

(have equal contributions to this article)

Student:Fong,Y. T.,Tsao,J. S.,Xu,W. W.

2023

# Abstract

The first part will explore the development of SEO and e-commerce, while the latter part will delve into artificial intelligence sentiment analysis and the prediction of raw material prices.

In recent years, artificial intelligence (AI) has emerged as a transformative trend in various industries, including business, e-commerce, and search engine optimization (SEO). AI is being used to automate processes, improve efficiency, and reduce costs, making it a valuable tool for businesses looking to stay competitive. In this report, we will explore the ways in which AI is being used in operations and supply chain management, as well as its impact on e-commerce and SEO. Specifically, we will discuss how AI can be used to control SEO and drive e-commerce trends. By analyzing and optimizing search engine rankings using AI algorithms, businesses can increase their online visibility and attract more customers, ultimately driving sales and revenue growth. This report will provide insights and recommendations for businesses looking to harness the power of AI in their SEO and e-commerce strategies.

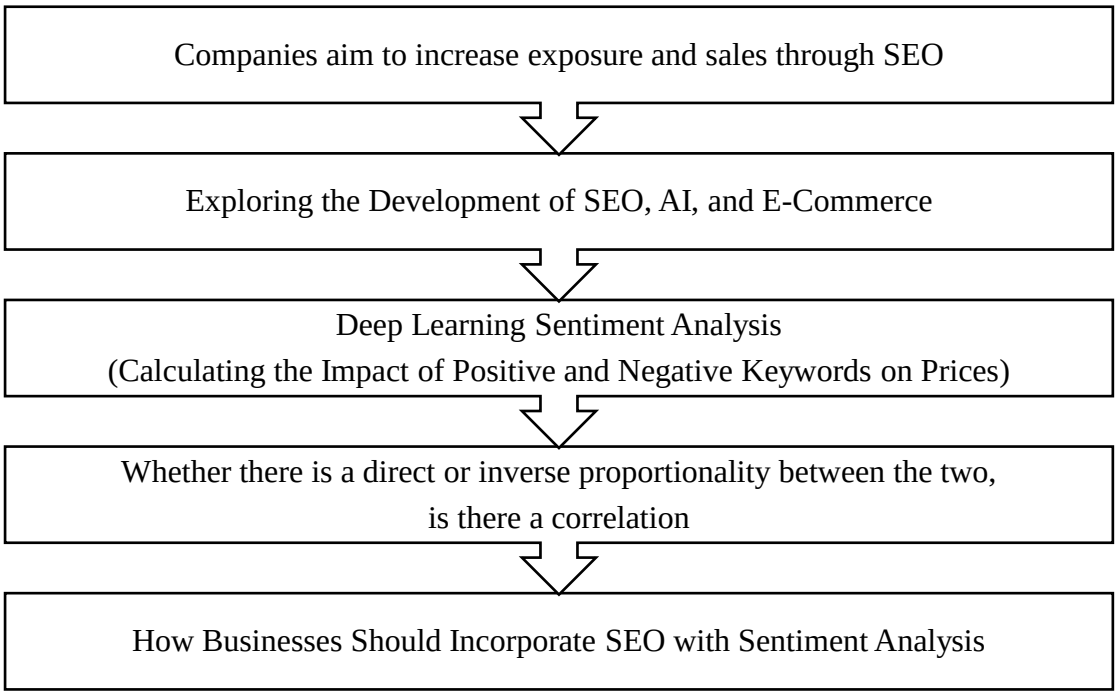


Fig. 1 The main logic flow

# Contents

# List of Figures

# Chapter 1 Introduction

## Section 1 Introduction

It refers to the transaction or service activities in the form of network or electronic transactions. Compared with the traditional sales industry, e-commerce is an electronic and networked sales form. The only consistent factor across all groups was cost – perceived as a barrier. The study concluded that factors are perceived differently by adopters, intended adopters and those not intending to adopt. These results should serve as a basis for more accurate use of these factors in adoption models. (Wymer&Regan,2005,p. 438-453)

This requires a concern of company management settling their product sales strategies to continue going up, mainly online trading.

A Web site that wants to increase its number of visitors can pay for search engine ads or attempt to improve its natural search engine ranking. Nobody really knows, which, if either, of these methods provides a positive return on investment (ROI). A search engine optimization (SEO) project was undertaken at a new e-commerce site. The site's search engine rankings and traffic were measured after each phase in the project. The results indicate that SEO is an effective method for improving search engine rankings and site traffic. In addition, the costs and benefits of the SEO project are compared with a pay-per-click (PPC) search marketing campaign. The SEO project proved more cost effective than the PPC campaign.(Malaga,2007,p. 68-82)

SEO, or search engine optimization, is the process of optimizing a website or online content to rank higher in search engine results pages (SERPs) for specific keywords or phrases. The goal of SEO is to increase visibility and traffic to a website, which can ultimately lead to more sales and revenue for businesses. In the early days of e-commerce, SEO was not a priority for most businesses. However, as the internet grew and more businesses began selling online, competition increased and the importance of SEO became more apparent. In the early 2000s, Google emerged as the dominant search engine and its algorithm became increasingly sophisticated, making it harder for websites to rank without proper optimization.

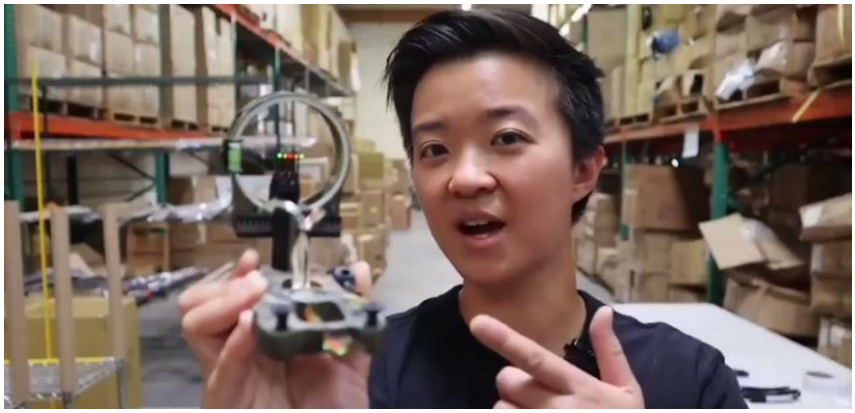

Fig. 2 E-Commerce is the Trend

The reason for using raw material price prediction as an example is primarily due to the opacity and lack of transparency in commodity prices, making it difficult to access and work with raw data. In contrast, raw material prices are more transparent and publicly available, making it relatively easier to perform data prediction. In the future, this approach can be applied to commodity prices as well.

# Research Motivation

As the field of Artificial Intelligence (AI) continues to rapidly grow, there is an increasing demand for professionals with AI expertise across a range of industries. Studying AI offers exciting career opportunities in areas such as healthcare, finance, and technology (Pedro & Valverde, 2019). Moreover, AI is being used to solve complex problems, such as climate change, disease detection, and resource management(Wahl & Schwalbe, 2018). Studying AI can equip individuals with the knowledge and skills to contribute to these efforts, as well as to develop new, groundbreaking technologies such as self-driving cars, intelligent virtual assistants, and personalized healthcare(Drukker & Papageorghiou, 2020, pp. 498-505).

AI is a highly interdisciplinary field that combines computer science, mathematics, statistics, and other fields, studying AI can expose individuals to a diverse range of topics and help them develop a broad set of skills(Lee & Perret, 2022, pp. 12783-12791). Additionally, AI is intellectually challenging, requiring creative thinking, problem-solving skills, and a deep understanding of complex algorithms(Greeno, 1978, pp. 239-270). This makes studying AI a rewarding and enriching experience for individuals with a personal interest in AI.

Further, AI professionals are in high demand, and studying AI can open up opportunities for a lucrative career. With advancements in AI technology expected to continue rapidly, studying AI can enable individuals to be at the forefront of new technological advancements and contribute to developing new and innovative AI applications, as well as better anticipate the impact AI may have on society (Kumar & Lecinski, 2019,p. 135-155).

The demand for AI expertise and the opportunities for innovation and career advancement make studying AI a compelling research motivation for individuals interested in contributing to this rapidly advancing field.

The use of deep learning in sentiment analysis within AI stemmed from an interesting observation. On one occasion, it was noticed that in Taiwan, the severity of penalties for certain offenses could vary based on the specific keywords used. This observation sparked an idea - why not use a similar concept to categorize and rate different keywords as positive or negative, and then predict future price trends based on varying degrees of sentiment.

台灣罵人價目表（本價格已經過法官核准）

| | 用詞 | 罰款金額 | 法官罰款說明 |
|---|---|---|---|
| 1 | 幹 | 判賠 0元 | 「幹」只是台語發語詞 |
| 2 | 更年期到了 | 判賠 2千 | 涉嫌公然污辱 |
| 3 | 賤貨 | 判賠 3千 | 名譽受損 |
| 5 | 米蟲 | 判賠 6千 | 公然侮辱罪 |
| 6 | 王八蛋 | 判賠 1萬 | 法官認為罵人屬實，侵害名譽權 |
| 7 | 不要臉的髒東西 | 判賠 2萬 | 公然侮辱罪 |
| 8 | 你去吃屎啦 | 判賠 2萬5千 | 公然侮辱罪，身心受創，要到醫院就診 |
| 4 | 敗類 | 判賠 3萬 | 法官認為言論確實貶損當事人的人格評價 |
| 9 | 智障 | 判賠 5萬 | 損害大於口語以精神撫慰金5萬元為適當 |
| 10 | 賤人就是矯情 | 判賠 5萬 | 公然侮辱應賠償精神損失5萬元 |
| 11 | 頭殼裝屎 | 判賠 5萬 | 造成身心俱疲，罹患憂鬱症，精神撫慰金 |
| 12 | 白痴 | 判賠 5萬5千 | 不堪受辱、身心疾病復發，公然侮辱罪 |
| 13 | 婊子 | 判賠 6萬 | 名譽受損，且感到愧對母親而精神耗弱 |
| 14 | 下流 | 判賠 6萬 | 是罵人的話有輕蔑之意，構成公然侮辱罪 |
| 15 | 幹X娘 | 判賠 8萬 | 公然侮辱罪 |
| 16 | 死番仔 | 判賠 10萬 | 無禮蔑視踐踏人格損及人性尊嚴侵害名譽權 |
| 17 | 人渣公務員 | 判賠 30萬 | 言論已侵害對方名譽，精神上受到相當痛苦 |
| 18 | 神經病 | 判賠 30萬 | 公然侮辱被判無罪但民事求償30萬元定讞 |
| 19 | 特殊性關係 | 判賠 100萬 | 加重誹謗罪，並在4大報頭版登報道歉1天。 |

Fig. 3 Taiwan Insult Price List

# Section 2 Research Questions

The primary objective of this research is to leverage sentiment analysis techniques to understand how news articles impact market sentiment and influence price movements in raw material markets. The research aims to overcome challenges related to data collection and scarcity by shifting focus from scraping news headlines to analyzing categorized news articles specifically related to raw materials. Ultimately, the goal is to enhance the accuracy of predictions regarding changes in raw material prices and provide valuable insights for future market analysis.

How did the research address challenges in web scraping to ensure the quality of data for sentiment analysis, and what were the outcomes of these efforts?

How did the shift from scraping news headlines to analyzing categorized news articles related to raw materials contribute to improving the relevance and accuracy of the data for sentiment analysis?

How did the decision to switch to pre-existing sentiment analysis models, specifically Microsoft Power Apps, impact the accuracy of price predictions compared to the initial basic program?

How can businesses practically use sentiment analysis as a reference tool for decision-making, particularly in situations where there is a negative sentiment environment, and what insights can be gained from such analyses based on the research findings?

How did the research evaluate the impact of sentiment analysis on market sentiment and price movements, and what were the key findings in this regard?

# Purposes

The primary purpose of applying AI to the search engine of an e-commerce platform is to improve the accuracy and relevance of search results for users. AI-powered search engines can analyze large amounts of data, including user behavior and product information, to provide more personalized and relevant search results. This can help users find the products they are looking for more quickly and easily, which can lead to increased sales and customer satisfaction.

Personalization: ("7 major Benefits of AI in E-commerce" by Pragnya Priyadarsini)AI can analyze user behavior and preferences to provide personalized search results, such as recommending products that the user is more likely to be interested in. This can lead to a better user experience and increased customer satisfaction. Relevance:( "Why Artificial Intelligence is the Future of E-commerce Search" by Tomas Miklas, Prowly Magazine)AI-powered search engines can analyze product data and user behavior to provide more relevant search results, which can help users find what they are looking for more easily and quickly. Efficiency:("The Benefits of AI for E-commerce" by Kenneth Evans, The Next Web) AI can automate certain aspects of the search engine, such as suggesting related products or optimizing search results to prioritize high-margin or low-inventory items. This can help businesses to improve their operational efficiency and increase sales. Insights:("10 Ways AI is Improving the E-commerce Customer Experience" by Daniel Faggella, Forbes) AI can provide valuable insights into customer behavior and preferences by analyzing search queries, click patterns, and purchase history. This information can be used to improve product offerings, marketing strategies, and overall customer experience. Multilingual support:("How AI Can Enhance the Search Experience for Multilingual E-commerce Websites" by Joel House)AI can help support multiple languages, which can be important for e-commerce businesses with a global customer base. Overall, the main purposes of applying AI to e-commerce search engines are to provide a more personalized and efficient customer experience, improve the relevance and accuracy of search results, and gain valuable insights into customer behavior and preferences.

# Section 3 Research Objects

The research subjects are hopefully office workers, because they may not have enough time to buy daily necessities.Therefore, then take office workers as the target customer group.

The text covers the entire landscape of e-commerce. The key message is that faculty who want to teach an introductory class on e-commerce and focus on the “strategy” parts of e-commerce first and technology second, should adopt this book. Faculty who teach marketing, management, strategy and entrepreneurship as the “core” discipline prefer this book over “technology-oriented” e-commerce books. Introduction to e-Commerce gives present and future practitioners of e-Commerce a solid foundation in all aspects of conducting business in the networked economy. The text focuses on what a manager needs to know about Internet infrastructure, strategy formulation and

implementation, technology concepts, public policy issues, and capital infrastructure in order to make effective business decisions (Rayport&Jaworski,2004).

# Limitations

One of the major limitations of this study was the lack of a strong database. Due to the unavailability of a comprehensive and reliable database, we had to rely on data collected from a variety of sources. As a result, some of the data may have been incomplete, inaccurate or outdated. This limitation may have affected the generalizability of the findings, and may limit the scope of the study.

In addition, a significant amount of data was private, and therefore could not be included in the analysis. This may have limited our ability to draw definitive conclusions about certain aspects of the research question. The study was limited by the sample size. The number of participants in the study was relatively small, which may have limited the statistical power of the analysis and the ability to detect significant differences or relationships between variables.While this study provides valuable insights into the research question, the limitations outlined above suggest that further research is necessary to confirm and expand upon these findings.

# Section 4 Research Structure

The purpose of this study is to explore the potential benefits of using artificial intelligence (AI) to control search engine optimization (SEO) in e-commerce. SEO is a critical component of e-commerce, as it helps businesses to improve their visibility and ranking on search engines, thereby increasing their website traffic and sales. However, SEO can be a complex and time-consuming process, requiring ongoing optimization and analysis. AI has the potential to streamline this process by automating tasks such as keyword research, content optimization, and link building. By using AI to control SEO, businesses can improve their efficiency and effectiveness in optimizing their websites for search engines, ultimately leading to increased traffic and sales.

## Process

To research e-commerce trends, start by looking for industry reports, consumer behavior studies, and trend analysis from reputable sources. To collect data on SEO, use tools like Google Analytics, Google Search Console, and keyword research tools like Ahrefs and Moz to track website traffic, search engine rankings, and keyword performance. Finally, to use AI to train a model with Python, choose a machine learning framework like TensorFlow or Keras, identify the data you want to use to train your model, and use Python to preprocess, clean, and analyze the data, and train and test your model using machine learning algorithms. With these steps, you can gather valuable insights on e-commerce trends and SEO performance and develop models to inform your business strategy.

# Chapter II Industry Introduction

E-commerce,("E-commerce Industry: Statistics, Trends & Data for 2021-2022" by WebsiteSetup Editorial Team on WebsiteSetup.org) short for electronic commerce, refers to the buying and selling of goods or services over the internet. This can include transactions between businesses, between businesses and consumers, or between consumers. E-commerce has become increasingly popular in recent years, with the rise of online marketplaces like Amazon and eBay, as well as the growth of social media platforms that enable businesses to sell products directly to consumers.

E-business, on the other hand, refers to a broader concept of using digital technology to conduct business operations. This can include not only e-commerce transactions but also other functions like customer relationship management, supply chain management, and online marketing. E-business can be used by both traditional brick-and-mortar businesses and online-only businesses.

An online shop, also known as an e-shop or web store, is a website that allows businesses to sell products or services directly to consumers over the internet. Online shops can be standalone websites or can be integrated into larger e-commerce marketplaces like Amazon or Etsy. They typically include features like product listings, shopping carts, and secure payment processing.

Overall, e-commerce, e-business, and online shops all involve using digital technology to conduct business operations and facilitate transactions. E-commerce specifically refers to the buying and selling of goods or services over the internet, while e-business encompasses a broader range of digital business functions. Online shops are a specific type of e-commerce platform that allows businesses to sell products or services directly to consumers through a website.

# Chapter III Reference Literature Discussion

## The E-commerce

E-commerce:("Ecommerce vs. E-Business vs. Online Business: What's the Difference?" by Lucidpress on Lucidpress.com) "E-commerce refers to the buying and selling of goods or services using the internet, and the transfer of money and data to complete the sales. It encompasses a wide variety of online activities, including online shopping, electronic payments, online auctions, and internet banking." (Source: Investopedia) E-business:("What is E-commerce? How does it work?" by Avinash Prajapati on Medium.com) "E-business refers to the use of electronic platforms and technologies to conduct a company's business. It involves not only buying and selling goods and services online, but also encompasses activities such as online marketing, customer service, supply chain management, and electronic data exchange." (Source: Techopedia) Online shop: "An online shop, also known as an e-shop, is a website that allows businesses to sell products or services to customers over the internet. Online shops typically feature a catalog of products, a shopping cart, and a checkout process that enables customers to purchase items and have them delivered to their doorstep." (Source: BigCommerce)

## Search Engine Optimization

First things of Regex, short for Regular Expression, is a sequence of characters used to define a search pattern. It is commonly used in SEO for tasks such as redirecting URLs,

removing duplicate content, and filtering out unwanted pages. Regex can be a powerful tool for SEO professionals, but it requires some technical knowledge to use effectively.To be more specific.They describe as "a pattern language that specifies a set of strings", and explains how they are commonly used in text processing and search applications. It goes on to describe the decision-making process involved in creating regular expressions, including the need to balance accuracy with efficiency, and the potential trade-offs between specificity and generality.(Michael et al., 2019)Regex can be used in combination with social media developer APIs to scrape and analyze data from social media platforms.(Dewi et al., 2019, p.444-447).Regex as a notation for describing patterns in text or strings, used in a variety of programming languages and applications,They discuss the difficulties and challenges of using regular expressions, including the lack of standardization and portability across different programming languages and applications. They also highlight the importance of testing and validating regular expressions to ensure correct functionality and avoid errors.(Davis et al., 2019, p.1-12)

Secondly,Reverse link: They define reverse link as the communication link from the mobile handset to the base station.The reverse link is used to transmit voice and data from the mobile handset to the base station. They also discuss how soft handoff, a technique used in CDMA systems, can extend cell coverage and increase the capacity of the reverse link.(Viterbi et al., 1994, p.541-551)Reverse link is also defined as the link from the mobile station (MS) to the base station (BS) and discusses the importance of setting the reverse link target SIR to optimize system performance. They propose a new method for setting the reverse link target SIR based on the analysis of the error rate and SIR distribution(Sampath et al., 1997, p.929-932).The idea of following incoming links to a website, rather than following outgoing links from a website.(Yesilada et al. , 2007, p.3-10).

Build a link:.Link building is the process of acquiring hyperlinks from other websites to one's own site with the goal of improving search engine rankings. They explain that links act as a vote of confidence for a website, and the more high-quality links a site has pointing to, the more likely it is to rank higher in search engine results.(Zhang & Cabbage, 2016, p.148-159).However,some of the experts held different opinions with different perspectives. (Shenoy & Prabhu, 2016, p.9-20) The authors provide a more comprehensive overview of link building strategies and techniques, including internal linking, external linking, and backlinking. They also discuss the importance of quality over quantity in link building and the potential risks of engaging in unethical or manipulative link building practices.PageRank is a link analysis algorithm that was developed by Google's co-founders, Larry Page and Sergey Brin(Swati & Ajay, 2013, p.10-13).In short PageRank is a “vote”, by all the other pages on the Web, about how

important a page is. A link to a page counts as a vote of support,is also calculated using a system of equations that takes into account the incoming and outgoing links of each page, and how the algorithm can be used to rank search results in Google.(Rogers, n.d. ,p.3-5)In terms of the relationship with SEO, the authors note that PageRank is one of many factors that can affect a website's search engine ranking.(- et al., 2023, p.3-7).Canonical URL is a concept in SEO that refers to the preferred URL for a specific web page when there are multiple versions of that page available on the web. A canonical URL is "the URL that you want visitors to see. Pretty simple, right?" Having multiple versions of the same content on different URLs can lead to duplicate content issues, which can negatively impact search engine rankings.(*Canonicalization - Moz*, n.d.) By specifying a canonical URL, webmasters can tell search engines which version of the content is the preferred or "canonical" version. A canonical URL is the preferred version of a webpage that search engines should index and rank, and that having multiple versions of the same content on different URLs can lead to duplicate content issues, which can harm a website's SEO performance.(Edgar, 2023, p.137-140).Canonical tags can be used to prevent duplicate content issues, consolidate link equity and improve crawl efficiency. The article also provides examples of how canonical tags can be implemented on different types of web pages.(Canonical Tag - Definition + Best Practices - Seobility Wiki, n.d.)

# Search Engine Marketing

Search Engine Optimization (SEO) is the practice of optimizing a website to improve its ranking in organic search engine results pages (SERPs).Search Engine Marketing (SEM) is the practice of using paid advertising on search engines to increase visibility and drive traffic to a website.(Panchal et al., n.d. p.1)

As search engine marketing becomes more popular, the competition for ad space on search engines increases, leading to higher costs per click.(Aswani et al., 2018, p.107-116).One of the main advantages of SEO over SEM is that it can be more cost-effective. While SEM requires businesses to pay for each click on their ads, SEO is a more organic and sustainable approach that can lead to long-term traffic growth without ongoing advertising costs. SEM can be an effective way to increase website traffic and online visibility, it can also be expensive and require ongoing investment. In contrast, SEO is generally considered a more cost-effective approach to improving search engine rankings and driving traffic to a website over the long term.(Xing & Lin, 2006 ,p.519-529) SEM can be expensive, particularly for competitive keywords, as you have to pay every time someone clicks on your ad compared to SEO.(Chaters, 2011, p.95-119)SEM

campaigns can be expensive, especially for competitive keywords(Ramlall, I. 2016, p.19-20).Another point is trust and credibility: Many users view organic search results as more trustworthy and credible than paid ads. A study by Econsultancy found that only 6% of users click on paid search results, while 94% click on organic search results. By prioritizing SEO, companies can build trust with potential customers and establish themselves as credible authorities in their industry.Many internet users may not trust advertisements, which can negatively impact the effectiveness of SEM campaigns(Ramlall, I. 2016, p.19-20)

SEO involves optimizing a website's content, structure, and other factors to improve its ranking in organic search results, with the goal of generating more traffic to the website over time. This process often involves keyword research, on-page optimization, link building, and other techniques. On the other hand, SEM involves paying for advertising space on search engine results pages (SERPs) to drive traffic to a website. This can include paid search ads, display ads, or other types of paid advertising. SEM can provide more immediate results than SEO, but can also be more expensive as advertisers must pay each time a user clicks on their ad. In summary, SEO is focused on optimizing a website to rank higher in organic search results over time, while SEM involves paying for advertising space to drive traffic to a website in the short term.(Panchal et al., n.d. p.3-5)

Another point of websites that prefer to use SEO is because of long-sustainability,as mentioned before,SEO first is the sustained, long-term value that comes with SEO. While SEO comes with a “price” of its hard work and a long haul, the value generated from good SEO can drive sustainable long-term benefits for a business. Investing in technical SEO, building backlinks, and creating relevant content will stay with your site forever and hopefully continue to land you at the top of the search results,on the other hand,if the website owner wants a quick result instead of sustainability,they will more prefer to using the SEM.For instance, Samples can be analyzed in as little as 2 minutes versus the 5 to 15 minutes that’s typical with competitive options.

# Artificial Intelligence in business

Artificial Intelligence (AI) is increasingly being used in business to improve efficiency, enhancecustomerexperiences,andgaincompetitive advantages.(Daqar&Smoudy,2019,p. 22)AI technologies can automate routine tasks, analyze large amounts of data, and provide valuable insights and predictions to inform decision-making. (Jarrahi,2018,p. 577-586)One of the most common applications of AI in business is in customer service. Chatbots and virtual assistants can provide 24/7 support and help customers with basic

inquiries, freeing up human agents to focus on more complex issues. AI-powered personalization can also provide tailored recommendations to customers based on their preferences and past behavior, improving the overall shopping experience. (Rathore,2023,p. 15)AI is also being used in marketing to improve targeting and conversion rates. (Gkikas&Theodoridis,2018,p. 1251-1259)By analyzing customer data, AI can help businesses identify the most effective channels, messages, and offers to engage with customers and increase sales. AI can also help optimize advertising campaigns by adjusting bids, targeting, and messaging in real-time to maximize ROI. In addition, AI is being used in operations and supply chain management to improve efficiency and reduce costs.(Reyes&Jaska,2020,p. 156-162). AI-powered predictive analytics can help businesses optimize inventory levels, reduce waste, and streamline logistics. AI can also be used to monitor equipment and detect issues before they become major problems, reducing downtime and maintenance costs.(Lee&Sutherland,2019,p. 506-511)AI has the potential to transform the way businesses operate and interact with customers. By leveraging AI technologies, businesses can improve efficiency, enhance customer experiences, and gain competitive advantages in today's digital landscape.

# Keywords

First, keyword analysis must be done for the website on which the optimization operation will be performed. Products or services must be defined in both their short and long usage structures. At this point, definitions are made with some words rather than using short and single words. According to statistical research studies, Internet users prefer using 2 or 3 words to perform a search on a search engine because competition with a single word is denser than with other preferences. Some different approaches to using keywords can also be employed to prevent usual typing errors. (Nursel Yalçın, Utku Köse, Procedia - Social and Behavioral Sciences, (2010), 487-493, 9).In recent years, internet news has become one of the most important channels for acquiring information, as more and more people read news through internet-connected computers, tablets, and smartphones, etc. Due to the constantly reproduced news, the number of online media outlets has increased dramatically, and the volume of news has also expanded rapidly. (Zihuan Wang, Kyusup Hahn, Youngsam Kim et al., Multimedia Tools and Applications, (2017), 4339-4353, 77).At first, search engines (Google, Bing, Baidu, Yandex, and others) are essential platforms for marketers that facilitate an exchange of value between three types of users: consumers seeking information, advertisers trying to reach those consumers, and content providers (Rangaswamy et al.,

2020). Recent studies have shown the importance of organic search queries and search engine optimization (SEO) for consumer attraction (e.g., Baye et al., 2015, Méndez-Suárez and Monfort, 2020, An and Jung, 2021).
It elaborates on keyword research. Keyword analysis is the most important part of optimization. Similar to keyword research for online marketing by SEOs of websites, one has to target the audience of video SEO and conduct research on what the audience will enter as keywords in the search engine. This keyword analysis is provided by some keyword analysis tools like Google Adwords, YouTube keyword suggestions, query log analysis, etc. (Krishna Choudhari Vinod K. Bhalla (2015), 691-697, 58). Keywords will display the monitoring results intuitively to users by using an SEO framework. (Wu Di, Luan Tian, Bai Yan et al.) The understanding of the role of the brand in the communication process with the customer through search engines (Sheffield, 2020) can influence the assessment of the brand value (Singh et al., 2020).

# Deep Learning

First, deep learning is a recent and fast-growing field of machine learning. It attempts to model abstraction from large-scale data by employing multi-layered deep neural networks (DNNs), thereby making sense of data such as images, sounds, and texts. (Chensi Cao, Feng Liu, Hai Tan et al. (2018), 17-32, 16)

Deep learning allows computational models that consist of multiple processing layers to learn representations of data with multiple levels of abstraction. The development of deep learning's intricate structure in large data sets involves using the backpropagation algorithm to indicate how a machine should change its internal parameters, which are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional networks have brought about breakthroughs in processing images, video, speech, and audio, while recurrent networks have illuminated sequential data such as text and speech. (Yann LeCun, Yoshua Bengio, Geoffrey Hinton, Nature, (2015), 436-444). Deep learning-based algorithmic frameworks shed light on these challenging problems. (Chensi Cao, Feng Liu, Hai Tan et al. (2018), 17-32, 16).

# Sentiment Analysis

It is a computational study of opinion, sentiment, appraisal, evaluation, and emotion.

The main usage is usually for benchmarking products and services, as well as marketing intelligence. (Bing Liu, Jun Zhao, Kang Liu et al.) To explain the difference between Opinion Mining and Sentiment Analysis: Opinion Mining extracts and analyzes people's opinions about an entity, while Sentiment Analysis identifies the sentiment expressed in a text and then analyzes it. Therefore, the target of SA is to find opinions, identify the sentiments they express, and classify their polarity.(Walaa Medhat, Ahmed Hassan, Hoda Korashy (2014), 1093-1113, 5(4))

Sentiment analysis has been used in several applications including analysis of the repercussions of events in social networks, analysis of opinions about products and services, and simply to better understand aspects of social communication in Online Social Networks (OSNs). There are multiple methods for measuring sentiments, including lexical-based approaches and supervised machine learning methods. (Pollyanna Gonçalves, Matheus Araújo, Fabrício Benevenuto et al.)

Sentiment analysis involves five steps to analyze feelings and opinions in data: Data Collection: This first step gathers user-generated content from various sources like blogs, forums, and social networks. This data is often messy, written in different ways, and uses various vocabularies and slang. Manual analysis is impractical, so we use text analytics and natural language processing to organize and categorize it. Text Preparation: This step involves cleaning the collected data by removing non-textual and irrelevant content. Sentiment Detection: We examine the sentences in the data to find those with subjective expressions like opinions, beliefs, and views. Objective information like facts is discarded. Sentiment Classification: Subjective sentences are classified as positive, negative, good, bad, liked, disliked, or with multiple points. Presentation of Results: The goal of sentiment analysis is to turn unstructured text into meaningful information. After analysis, we present the results using graphs like pie charts, bar charts, and line graphs. We can also create a timeline showing sentiment changes over time, using chosen values like frequency, percentages, and averages.( Pollyanna GonçalvesMatheus AraújoFabrício Benevenuto et al.)

# Chapter IV Research Methods

This study will employ a variety of tools and methods for research. Firstly, we will use different tools to conduct preliminary analysis of the required search keywords to gather useful information about these keywords. Next, we plan to use web scraping techniques

to capture keywords from the headlines of news websites. By analyzing the keywords in these headlines, we will be able to gain insights into current events and topics.

One of the main objectives is to perform sentiment analysis on the headline keywords, evaluating the emotional tone of each headline and assigning a corresponding sentiment score. This will help us understand the emotional trends of news events or topics and provide a foundation for subsequent analysis.

Ultimately, our research will be based on this data to conduct predictive analysis of commodity prices. By combining keyword analysis, sentiment analysis, and market data, we aim to provide deeper insights into the dynamic changes in commodity prices and offer valuable information to investors and decision-makers to support future decision-making and strategy development.

# Chapter V Data Analysis and Results

This research project aims to utilize web scraping techniques to automatically capture news headlines related to the top ten keywords associated with raw materials on the ETtoday news website within a single day. The goal is to improve efficiency and ensure comprehensive data collection. To accomplish this, we have chosen to use the Python programming language along with relevant web scraping libraries like BeautifulSoup.

The objective of this tool is to search and retrieve all news headlines related to the top ten raw material keywords on the ETtoday news website. We will develop a program that initiates an HTTP request to access the ETtoday website, followed by parsing the webpage content to locate the news headlines of interest. To ensure that the news articles collected cover a one-day timeframe, we will configure parameters for specifying the date range, restricting the crawler to search for news published within the specified date.

The captured news headlines will be stored in a data structure for subsequent analysis. The primary purpose of this tool is to assist researchers in saving a significant amount of time while ensuring access to the latest and most comprehensive information regarding raw material prices.


Fig. 4 ETtoday News List

This research aims to develop a web scraper using the Python 3.11.4 environment to capture the headlines of news articles related to the top ten commodity prices listed on the StockQ.org website. This task requires the utilization of key Python libraries, including requests and BeautifulSoup, for acquiring and parsing web data.

Fig. 5 The top ten commodity prices listed on StockQ.org.

respectively 黃金(Gold)、銀(Silver)、白金(Platinum)、鈀(Palladium)、銠(Rhodium)、銅(Copper)、鎳(Nickel)、鋁(Aluminum)、鋅(Zinc)、鉛(Lead)

## Search for suitable keywords

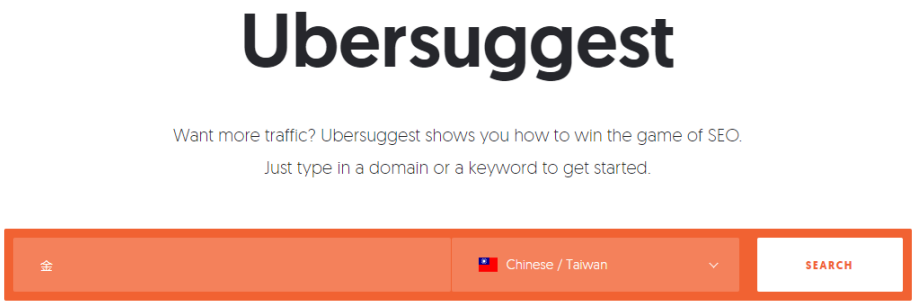


Fig. 6 Ubersuggest

Ubersuggest is a digital marketing tool developed by internet marketing expert Neil Patel, primarily designed to assist website owners and digital marketing professionals in improving their performance on search engines. Its features include keyword research, which helps users find potential keywords related to their business or website niche to optimize content and enhance search engine rankings. Additionally, Ubersuggest provides competitive intelligence, allowing users to access data about competitors' rankings, traffic, and used keywords, aiding in better marketing strategy formulation. This tool can also examine website health, identify issues, and offer improvement suggestions to enhance website performance and SEO.

Content suggestions are another significant feature, offering recommendations for creating high-quality content to attract more visitors and increase user engagement. Finally, Ubersuggest enables users to monitor changes in keyword rankings and website traffic while viewing past data trends. In summary, Ubersuggest is a powerful digital marketing tool that helps website owners and marketing professionals optimize their online presence, improve search engine visibility, and attract more visitors. It offers both free and paid versions, allowing users to choose based on their needs. It's important

to note that the tool's features and interface may change over time, so it's advisable to check the official website for the latest information.

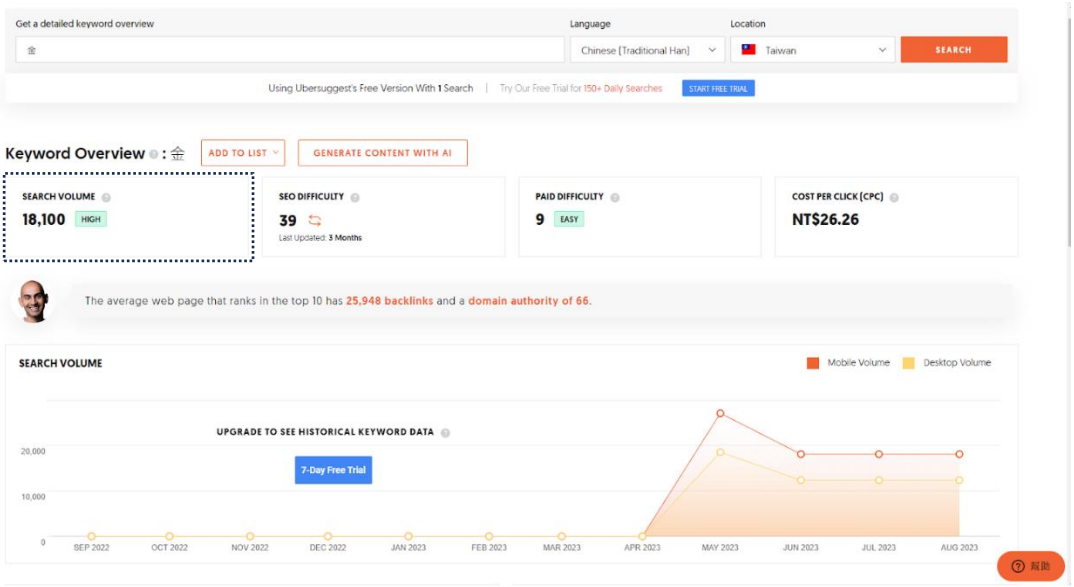


Fig. 7 Data on Ubersuggest for the keyword "金"

According to the above illustration, let's take the keyword "金" (gold) as an example for further exploration. The search volume for this keyword is shown as high, indicating that it is a popular search term and likely attracts a significant amount of user attention. However, due to the character limit imposed on titles, it may be challenging to provide a wealth of relevant elements in the title. Because of this character limit, we need to cleverly incorporate the keyword "金" (gold) into the title while also providing enough context to ensure that readers can understand the content. This may require the use of other related words or phrases in the title to supplement and expand upon the topic. Additionally, we can use subtitles or summaries to further clarify the content to meet reader needs. However, subtitles or summaries are less common in news headlines, making this approach more difficult.

Another approach is to delve into the keyword "金" (gold) in more detail within the body of the article, which can provide more relevant information and depth to the content. This approach helps offer a more comprehensive understanding while still including the topic keyword. In conclusion, when dealing with titles constrained by character limits, it's essential to strike a balance between providing keywords and offering sufficient context to ensure that the article can engage readers and pique their interest. This may require using different strategies, such as elaborating on the topic within the body of the article, to ensure valuable information is provided. This operation

requires more powerful computing resources and programming skills, and it may involve secondary web scraping to analyze both the keywords within the article's body and the keywords in the title. Unfortunately, in this experiment, we were unable to achieve such an in-depth analysis. We apologize for any inconvenience, and we aim to improve this in future experiments.

# Chapter VI Research Conclusions and Suggestions

## Section 1 Conclusions and Suggestions

```python
from selenium import webdriver
from bs4 import BeautifulSoup
from datetime import datetime, timedelta
import time

#初始化Selenium WebDriver並打開新聞網站
browser = webdriver.Chrome()
browser.get("https://www.ettoday.net/news/news-list.htm")

# 計算一天前的日期時間
one_day_ago = datetime.now() - timedelta(days=1)

#設置要搜索的關鍵字列表
keywords = ["黃金", "銀", "白金", "鈀", "銠", "銅", "鎳", "鋁", "鋅", "鉛"]

# 循環控制變量
go = True

while go:
    # 模擬滾動操作以加載更多新聞
    browser.execute_script("window.scrollTo(0, document.body.scrollHeight);")
    time.sleep(2)
    html_source = browser.page_source
    soup = BeautifulSoup(html_source, "lxml")

    # 獲取頁面上的新聞時間和標題
    for d in soup.find(class_="part_list_2").find_all('h3'):
        news_time = datetime.strptime(d.find(class_="date").text, '%Y/%m/%d %H

        # 如果新聞時間早於一天前，停止循環
        if news_time < one_day_ago:
            go = False
            break

        # 檢查標題是否包含關鍵字
        title = d.find_all('a')[-1].text
        for keyword in keywords:
            if keyword in title:
                print("日期時間:", news_time, "標題:", title)

# 關閉瀏覽器
browser.close()
```

Fig. 8 The code for web scraping ETtoday

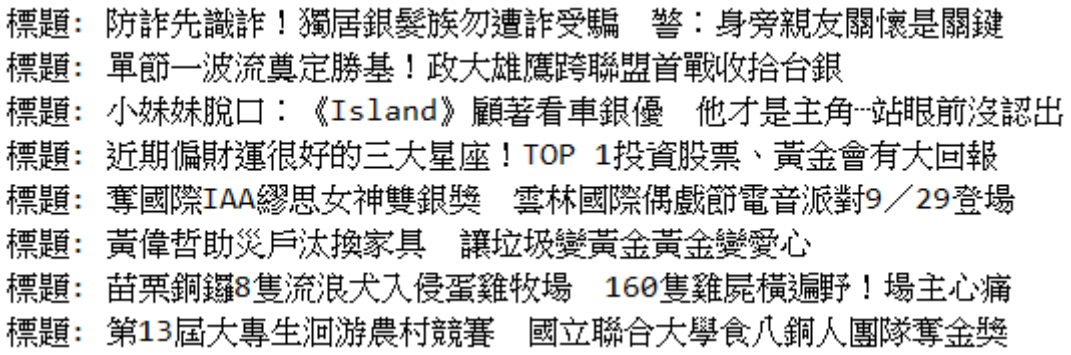


Fig. 9 The web scraping results

Through practical web scraping operations, we encountered some challenges and limitations. Firstly, we found that using web scraping directly on the news list could not precisely match the content we needed, potentially resulting in a large number of unrelated news articles in the results. This might be due to the relatively short nature of news headlines, making it difficult to achieve precise matching, leading to content that does not meet our requirements.

We observed that news related to raw materials is relatively scarce, and it may take a week or even a month to find a few relevant articles. The limited quantity of such samples could pose challenges to our analysis and predictions because we lack sufficient data to build reliable models. Furthermore, the non-reproducible nature of news makes it difficult to extrapolate future price movements, as news content may change over time and be influenced by various factors.

Our research faced difficulties and constraints in web scraping operations, including the challenges of precise matching, limited news quantity, and variability in news content. These issues may require further research and methods to overcome, ensuring that we can obtain reliable data to support predictive analysis of raw material prices.

Therefore, in the upcoming research, we plan to shift towards using news articles originally labeled as the raw materials section for sentiment analysis. The primary objective of this strategy is to assess the impact of past news content on relevant price movements. Compared to directly scraping news headlines, this approach may have the potential to provide deeper insights into the raw material market.

This approach allows us to access news articles that have already been categorized as related to raw materials, saving time and enhancing the relevance of the data. Through sentiment analysis, we can evaluate the emotional tone of these news articles and

understand their potential influence on market sentiment and prices. This strategy will enable us to better leverage past news to predict changes in raw material prices and provide more valuable data for future market analysis. We expect this approach to enhance the accuracy and reliability of our predictions.

# Sentiment Analysis

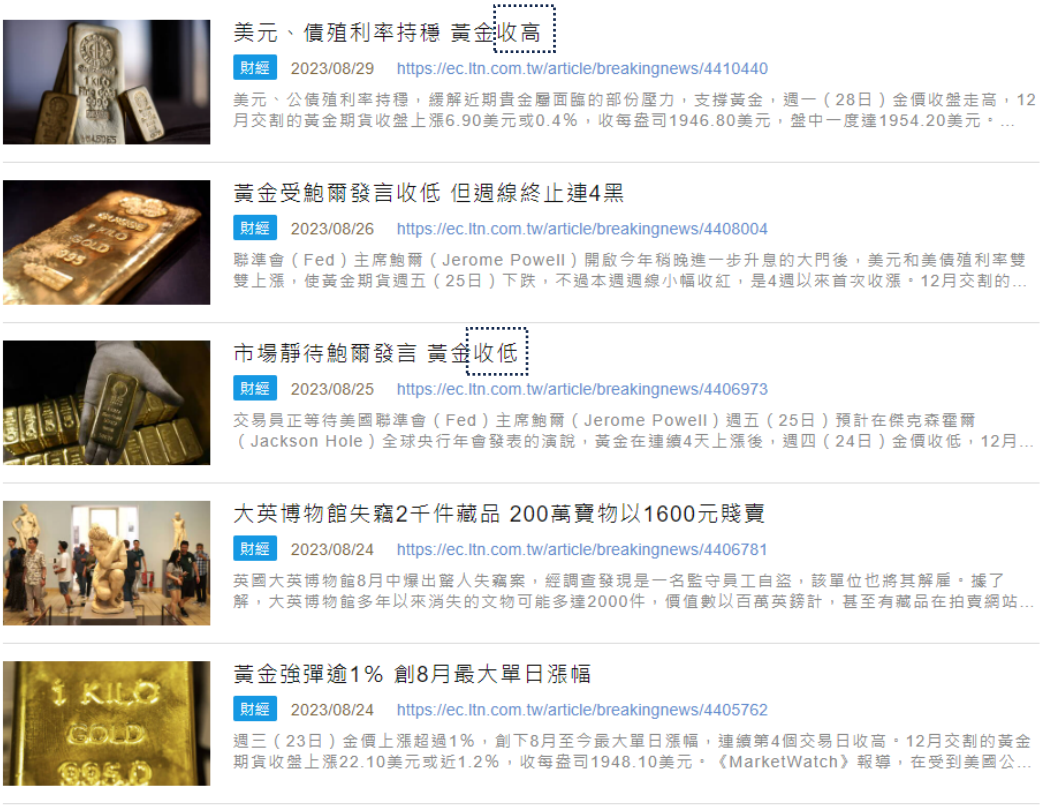


Fig. 10 August Gold-Related News in Liberty Times

As evident from the above titles, within a span of just five days, there is a stark contrast between 'Closing Higher' and 'Closing Lower,' as inferred solely from the headlines. In general investment practices, the preference naturally leans toward assets that show potential for appreciation, leading to a willingness to hold such assets. Of course, there are also strategies that involve shorting.

The next step involves analyzing the content of the articles and assessing whether there is any significant impact on prices one week later.

In the initial phase, a basic program was designed to facilitate sentiment scoring and

rating functionality.

```python
def 文章情感分析(文章):
    # 分割文章為單詞
    單詞列表 = 文章.split()

    # 初始化情感分數為0
    情感分數 = 0

    # 定義正面和負面詞彙列表（目前資料庫不夠完整日後將會擴充 此為示範所用 ）
    正面詞彙 = ["收高", "持穩", "上漲", "正面", "攀升", "光明"]
    負面詞彙 = ["壓力", "暴跌", "通膨", "威脅", "下跌", "下降", "承壓"]

    # 遍歷文章中的每個單詞，計算情感分數
    for 單詞 in 單詞列表:
        if 單詞 in 正面詞彙:
            情感分數 += 1
        elif 單詞 in 負面詞彙:
            情感分數 -= 1

    return 情感分數

# 測試文章情感分析函數
文章 = input("請輸入要分析的文章：")
總分 = 文章情感分析(文章)
print(f"文章情感分數：{總分}")
```

Fig. 11 Simple Sentiment Analysis

However, due to its simplicity, it resulted in less accurate price predictions. Given the limited resources in terms of both capability and time, the decision was made to switch to using pre-existing sentiment analysis models, specifically Microsoft Power Apps.

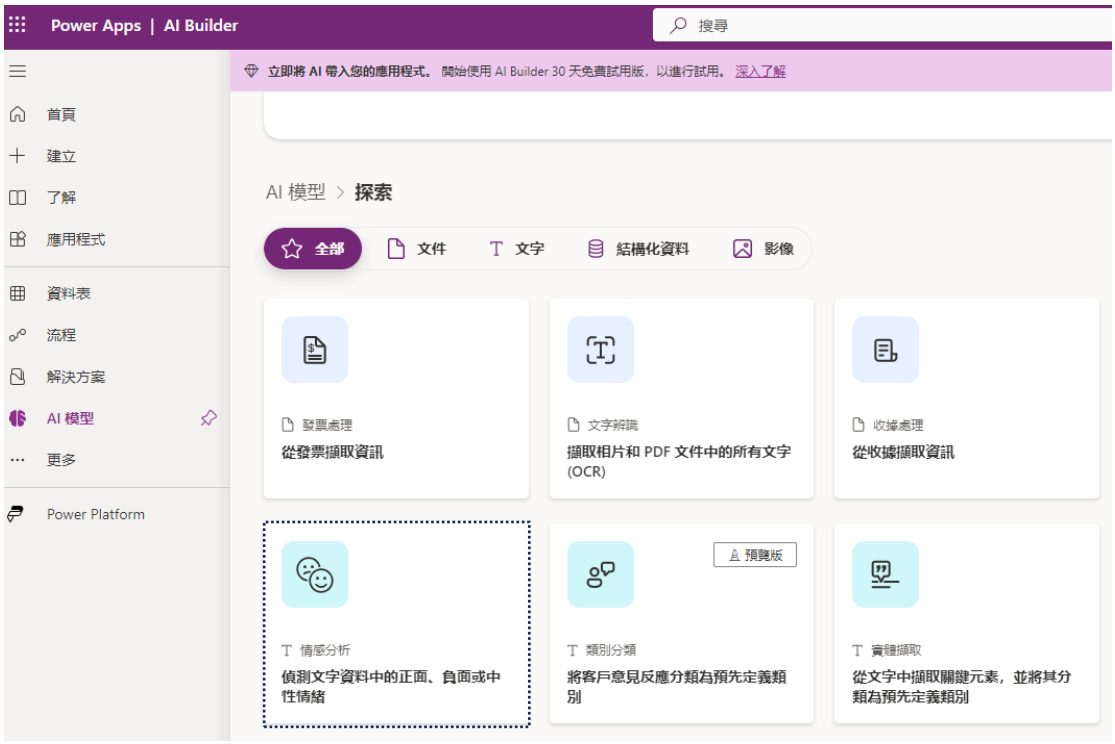


Fig. 12 Power Apps

Since the current system does not support Traditional Chinese, the decision has been made to translate the articles into Simplified Chinese before analysis.


Fig. 13.1 Sentiment Analysis

According to sentiment analysis conducted on August 25, 2023” 市場靜待鮑爾發言黃金收低”, which resulted in a negative sentiment rating of 78%, it is anticipated that prices will predominantly decrease in the future.


Fig. 13.2 Sentiment Analysis

According to sentiment analysis conducted on August 26, 2023” 黃金受鮑爾發言收低 但週線終止連 4 黑”, which resulted in a negative sentiment rating of 86%, it is anticipated that prices will predominantly decrease in the future.


Fig. 13.3 Sentiment Analysis

According to sentiment analysis conducted on August 29, 2023” 美元、債殖利率持穩 黃金收高”, which resulted in a negative sentiment rating of 86%, it is anticipated that prices will predominantly decrease in the future.

Considering the overall sentiment analysis for this week, which yielded a negative sentiment rating of 83%, it is hypothesized that the upcoming week will be influenced predominantly by negative factors.

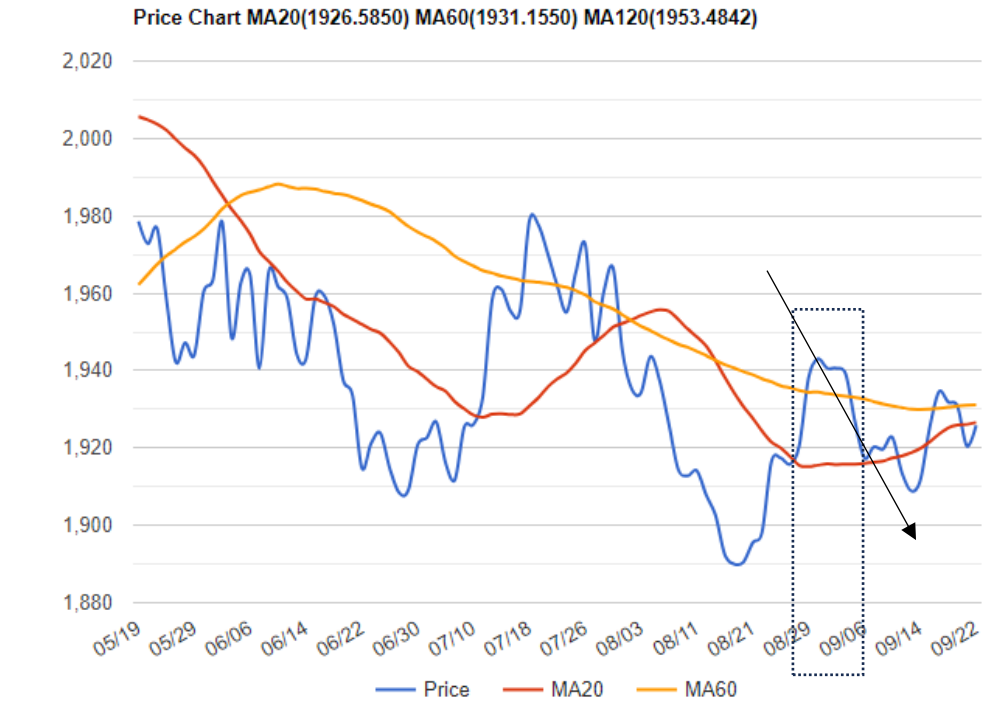


Fig. 14 Gold Price Trend Chart

Indeed, it can be observed from the conclusion that prices have been influenced and continue to decline.

# Section 2 Implications of Management Practice

Price movements are influenced by a multitude of factors, and sentiment analysis is just one of the assessment methods available. It can serve as a valuable reference tool for businesses to gauge market sentiment and make informed decisions. However, it is crucial to recognize that sentiment analysis is not the sole determinant of price fluctuations.

Various other economic, geopolitical, and industry-specific factors play their roles in shaping market dynamics. Thus, businesses should consider sentiment analysis as part of their broader management practices when assessing market conditions and potential investment strategies. In this context, sentiment analysis offers valuable insights, but it should be integrated into a comprehensive decision-making framework that takes into account the multifaceted nature of financial markets.

# Section 3 Suggestions for follow-up research

In future research, there is potential for a more in-depth exploration of the intertextual relationships and the contextual analysis of preceding and subsequent text segments. This approach can lead to more precise and comprehensive results. By delving deeper into the connections between textual elements, researchers may uncover nuanced patterns and correlations that contribute to enhanced accuracy in sentiment analysis and predictive modeling. Therefore, we suggest that future investigations place a greater emphasis on examining the interplay between textual components, as this may hold the key to unlocking a deeper understanding of market sentiment and price dynamics. Such an approach could pave the way for more advanced sentiment analysis techniques and predictive models.

# References


Section 1 Introduction

*Wymer, S. A., & Regan, E. A. (2005). Factors influencing e-commerce adoption and use by small and medium businesses. Electronic markets, 15(4), 438-453.*

*Malaga, R. A. (2007). The value of search engine optimization: An action research project at a new e-commerce site. Journal of Electronic Commerce in Organizations (JECO), 5(3), 68-82.*

Research Motivation

*Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development.*

*Wahl, B., Cossy-Gantner, A., Germann, S., & Schwalbe, N. R. (2018). Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings?. BMJ global health, 3(4), e000798.*

*Drukker, L., Noble, J. A., & Papageorghiou, A. T. (2020). Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. Ultrasound in Obstetrics & Gynecology, 56(4), 498-505.*

*Lee, I., & Perret, B. (2022, June). Preparing High School Teachers to Integrate AI Methods into STEM Classrooms. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 11, pp. 12783-12791).*

*Greeno, J. G. (1978). Natures of problem-solving abilities. Handbook of learning and cognitive processes, 5, 239-270.*

*Kumar, V., Rajan, B., Venkatesan, R., & Lecinski, J. (2019). Understanding the role of artificial intelligence in personalized engagement marketing. California Management Review, 61(4), 135-155.*

Artificial Intelligence in business

*Daqar, M. A. A., & Smoudy, A. K. (2019). The role of artificial intelligence on enhancing customer experience. International Review of Management and Marketing, 9(4), 22.*

*Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business horizons, 61(4), 577-586.*

*Rathore, B. (2023). Integration of Artificial Intelligence& It's Practices in Apparel Industry. International Journal of New Media Studies (IJNMS), 10(1), 25-37.*

*Gkikas, D. C., & Theodoridis, P. K. (2019). Artificial intelligence (AI) impact on digital marketing research. In Strategic Innovative Marketing and Tourism: 7th ICSIMAT, Athenian Riviera, Greece, 2018 (pp. 1251-1259). Springer International Publishing.*

*Reyes, P. M., Visich, J. K., & Jaska, P. (2020). Managing the dynamics of new technologies in the global supply chain. IEEE Engineering Management Review, 48(1), 156-162.*

*Lee, W. J., Wu, H., Yun, H., Kim, H., Jun, M. B., & Sutherland, J. W. (2019). Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. Procedia Cirp, 80, 506-511.*

Section 3 Research Objects (felix)

*Rayport, J. F., & Jaworski, B. J. (2004). Introduction to e-commerce. McGraw-Hill Irwin MarketspaceU.*

Search Engine Optimization

*Michael, L. G. (2019). Regexes are Hard: Decision-Making, Difficulties, and Risks in Programming Regular Expressions. 415-426.*

*Citra, L. D., Chandra, A., & Meiliana. (2019). Social Media Web Scraping using Social Media Developers API and Regex. Procedia Computer Science, 157, 444-449.*

*Himanshu, M., Ruchika, G., & Sumit, A. (2020). Search Engine optimization in Digital Marketing: Present Scenario and Future Scope. Proceedings of International Conference on Intelligent Engineering and Management, ICIEM 2020, 530-534.*

*Luz, M. N., Kam's, K. K., Francis, M. L., & Camile, L. B. (2023). Pagerank SEO Algorithm : Issues, Complexity and Implementation. IJFMR - International Journal For Multidisciplinary Research, 5(2).*

*P, S. P., & S, A. P. (2013). Search Engine Optimization: A Study. Research Journal of Computer and Information Technology Sciences, 1(1), 10-13.*

*Moz. (n.d.). Canonicalization - Moz*

*Edgar, M. (2023). Conclusion: Tech SEO Audit. Tech SEO Guide, 137-140.*

*Ian, R. (n.d.). The Google Pagerank Algorithm and How It Works. IPR Computing Ltd., 1-18.*

Search Engine Marketing

*Ramlall,, I. (2016). Drawbacks of SEM. Applied Structural Equation Modelling for Researchers and Practitioners, 19-20.*

*Aswani, R. (2018). Search engine marketing is not all gold: Insights from Twitter and SEOClerks. 38(1), 107-116.*

*Digital Marketing - Search Engine Optimization (SEO) and Search Engine Marketing (SEM) - ProQuest. (n.d.).*

*SEO vs SEM: Complete Guide to Finding the Right Search Strategy. (n.d.). Noble Desktop. Retrieved April 27, 2023*

*Home. (n.d.). YouTube. Retrieved April 27, 2023, from https://www.thermofisher.com/blog/materials/desktop-sem-microstructure-analysis-for-breakthrough-research/*

*D'Andrea, A., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, Tools, and*

*Applications for Sentiment Analysis Implementation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 125, pp. 1-8).*

*Gonçalves, P., Araújo, M., Benevenuto, F., & Cha, M. (2013). Comparing and combining sentiment analysis methods. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 123, pp. 27-38).*

*Medhat, W., Hassan, A., & Korashy, H. (2014). Ain Shams Engineering Journal, 5(4), 1093-1113.*

*Liu, B., Zhao, J., Liu, K., & Xu, L. (2016). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions (Vol. 42).*

*Wang, Z., Hahn, K., Kim, Y., Song, S., & Seo, J-M. (2017). A news-topic recommender system based on keywords extraction. Multimedia Tools and Applications, 77, 4339-4353.*

*Yalçın, N., & Köse, U. (2010). What is search engine optimization: SEO? In Procedia - Social and Behavioral Sciences (Vol. 9, pp. 487-493).*

*Cao, C., Liu, F., Tan, H., Song, D., Shu, W., Li, W., Zhou, Y., Bo, X., & Xie, Z. (2018). Genomics, Proteomics & Bioinformatics. Genomics, Proteomics & Bioinformatics, 16, 17-32, Volume 16.*

*LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521, 436-444.*

*Di, W., Tian, L., Yan, B., Liyuan, W., & Yanhui, L. Study on SEO monitoring system based on keywords & links.*

*Choudhari, K., & Bhalla, V. K. (2015). Procedia Computer Science, 58, 691-697.*

Chapter IV Research Methods

*Lohtia, R., Donthu, N., & Hershberger, E. K. (2003). The impact of content and design elements on banner advertising click-through rates. Journal of advertising Research, 43(4), 410-418.*

*Jansen, B. J., & Clarke, T. B. (2017). Conversion potential: a metric for evaluating search engine advertising performance. Journal of Research in Interactive Marketing, 11(2), 142-159.*