IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

BIG DATA AND NEURAL NETWORK APPROACH ON ONLINE PRODUCT SALES PREDICTION USING CUSTOMER REVIEWS AND PROMOTION STRATAGEM - A REVIEW

[1] G. Midhu Bala, [2] Dr. K. Chitra

[1]Assistant Professor ,[2]Assistant Professor,

[1]Department of Computer Science,

[1]Mangayarkarasi College of Arts and Science for Women,, Madurai, Tamil Nadu, India

Abstract: With the advent of Information Technology (IT) and Data Sciences, businesses now have the opportunity to better understand and forecast consumer demands using quantitative methods. The use of Big Data to better understand business processes and results is an evolving IT trend that has piqued the interest of researchers and practitioners. Big Data systems have the potential to assist businesses in better understanding complicated business relationships by delivering knowledge that was previously unavailable. Previous research has shown that providing an effective supply chain gives a producer a competitive advantage over its competitors, how data from online marketplaces or e-commerce allows manufacturers to better understand product demands is a field that has not been thoroughly researched by previous researchers. This paper provides a comprehensive survey about the ongoing researches done in this field by reviewing the articles of the other computer scientists. This paper paves a better way for the upcoming researchers in this field.

***Index Terms* - Web scraping, Neural Network, Locating files in websites, online sale promotion strategies.**

I. INTRODUCTION

Today's businesses operate in a more competitive and diverse world. Manufacturers have traditionally competed by lowering their manufacturing costs and providing higher-quality products. However, as manufacturers shift from providing standardized goods and services to one that focuses on customizations, cost and product quality rivalry is becoming more difficult. Data analytics has long been used to better understand business processes. Wal-Mart and Kohl's, for example, use a variety of sales, pricing, economic, and demographic data to better understand consumer behavior and product demands. Big Data technology and the Internet, on the other hand, offer businesses better access to and analysis of data from various sources, allowing them to uncover previously untapped business data [2].

On the Internet, data is ubiquitous and searching for valuable data and knowledge for analysis has become a regular task. The information on the websites is presented in an unstructured format, such as tables, posts, and comments, which are nested in various HTML tags. Gathering a large volume of data from the internet is a difficult process, but it is a good way to collect data for future research. Can knowledge gleaned from online user-generated data help manufacturers better understand and forecast product demand? User-generated content on the internet is starting to have a greater impact on consumer choices than conventional media like newspapers and television. According to statistics, up to 53% of Twitter users suggest items or brands in their messages, and 48% of those who receive the tweets follow through with the recommendations. The condition of online retail is that using customer feedback as goods are sold online has become standard practice.

Companies selling goods online will provide product details such as price and descriptions, as well as promotional marketing information such as the availability of coupons or "savings" from product discounts. The vast amount of data available on products, marketing campaigns, and online feedback can theoretically be used to forecast product demand and assist businesses in better planning their logistics. The aim of this study is to compare the impact of promotional marketing tactics such as discounts and the availability of free delivery options, as well as online review information such as product ratings and the percentage of positive and negative product reviews. The online sales rank of a product is used to determine its demand in this analysis. According to recent posts, manufacturers can better understand real-time demand and patterns for their goods by collecting social media chatter, and thus help to mitigate the bullwhip effect.

Web scraping is important for rising companies, and collecting big data is considered a necessity for staying competitive. The internet is like an infinite ocean with a tons of unstructured data, and with that data comes new possibilities that have yet to be discovered. Web Scraping, Web Data Extraction, Web Harvesting is a technique employed to extract large amount of unstructured data from websites, saved to a local file or to a database in structured format [3].

1.1 Essential of web scraping

Web scraping tools can be used to collect pricing information from adversaries, which is the most popular web scraping use case listed by most companies in the field. A web crawler can be configured to make requests on various competitor websites' product pages and then gather information such as price, shipping information, and availability. It can be used to fetch product data from e-commerce company web sites. It also may used for company trade name protection. Brands can quickly detect online material that it can damage their brand by using web scraping. Brands will take legal action against those responsible for the material after it has been identified. Lead generation can assist companies in reaching out to new customers. The marketer starts interacting with relevant leads by sending out messages in this process. Scraping contact information from the web, such as email, phone numbers, and social media pages, aids in reaching out to leads. The essentials of web scraping are listed below. General process of web scraping and web crawler is shown in Figure 1.

- **Cost analysis**: Gather information from online shopping sites and use it to compare product prices.
- **Email address collection**: Businesses collect email addresses and use them to submit mass emails for marketing purposes.
- **Online Media Scraping**: To learn about the latest trends, data is gathered from social media platforms such as Twitter, Face book, and others for sentiment analysis.
- **Research & Development**: Web scraping from websites is used to gather vast amounts of data for analysis, surveys, and R&D.
- **Job postings:** Information about job openings and interviews is gathered from various websites and summarized in one location for users to easily access.
- **Movie Review**: Scraping data to compare movies, drugs, and other products.
- **Digital Designs Classification**: Using images scraped from various websites to train an image
- **Ecopreneurs:** scraping user reviews and feedback from sites like Flipkart and Amazon to create a new product.

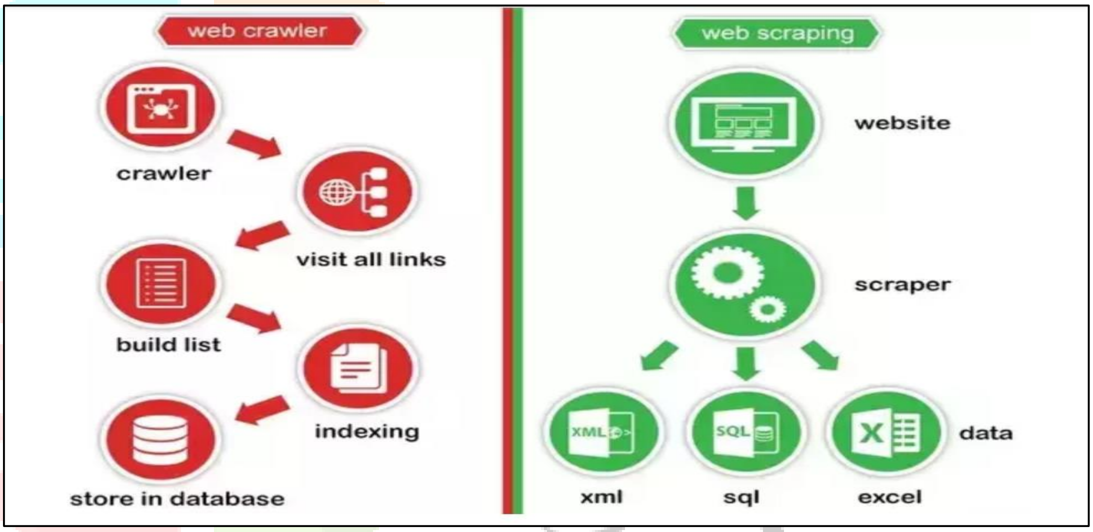

Figure 1.General Process of Web Scrapping and Crawling

II. LITERATURE REVIEW

People are gradually turning to online shopping for their shopping needs, as it is very easy to select items based on customer feedback. Thousands of reviews are regularly posted on e-commerce sites, even for moderately popular items. The constant generation of such a large amount of data can be considered a big data problem for both online companies and customers. This makes it impossible for consumers to read all of the reviews before making a purchase.

According to Internet Live Stats which collects data from different international sources, around 40% of the world population has an internet connection today, while 20 years ago it was less than 1%. The number of internet users has increased tenfold from 1999 to 2013, the third billion being reached in 2014. Of it, today only one billion people are using social networks, but by 2020 this number will double, according to analysts.[6]

Web scraping is an ideal method for performing predictive analytics. Web scraping tools are simple to use and use, regardless of your current level of performance. Scraping Robot has a number of modules for scraping famous websites like Amazon, Google, and eBay. Scraping, on the other hand, has a wide range of applications. To perform predictive research, you can use web scraping to collect data on costs, industry dynamics, and customers [4].

2.1 Promotional advertisement on the internet

Because of today's shorter product life cycles, particularly in the electronic industry, manufacturers are under more pressure to sell their products in a shorter period of time. Consumers now have more choices than in the past due to the increased availability of product details.

Consumers can now easily compare product prices and features online, for example. Companies are investing a considerable amount of their money to market and sell their goods digitally as a result of these business pressures. Price reductions are a popular promotional marketing technique used by businesses.

2.1.1 Discount value

Price reductions are common because they can boost a product's sales in a short period of time .According to transaction utility theory, as there are more discounts, customers' appetite for a commodity will increase because they will think they have gotten a good deal. Price reductions have measurable consequences, and since they improve customer traffic, they help to support partnerships between suppliers and retailers, as well as ensuring that a specific brand is well-stocked and has ample shelf room in retail stores [7].

2.1.2 Discount rate

The discount ratio as compared to the actual price of a commodity is another indicator used in this study. The power of a consumer's market perception in absolute or relative terms to affect their price discount impressions. According to the 'psychophysics-of-price-heuristics' theory, customers. The psychological utility derived from saving a fixed sum of money is inversely related to the price of an object. As a result, a business may either have an absolute price discount or a relative price discount. Price reduction expressed as a percentage.

2.1.3 Free shipping

Consumers are more likely to buy a product online if free shipping is available, in addition to using price discounts. In their study of users' expectations for website attributes, [8] discovered that, in addition to special offers; free delivery is a function that will attract online users. However, not all websites that offer free shipping as a perk are popular. With the maturation of e-commerce, customers will expect free delivery, so those who do not provide it will lose out to their competitors.

2.2 Online credentials

With the rise of online media, today's consumers frequently and actively share their thoughts on goods and services with others on a variety of online forums such as product reviews, blogs, Twitter, and wiki. [9] looked at a three-year panel data set from an online restaurant review website and found evidence of a connection between valence and volume and product sales.

2.2.1 Online review valence (average rating)

The evaluation score of a particular product or service is known as the online review valence. While previous research has suggested that online review valences have compelling effects on consumers' purchase decisions [11], the results have been inconsistent. Researchers including [9][10], on the other hand, found evidence for a connection between online review valence and product sales. In an WOM (electronic word of mouth) world, product and service ratings are becoming more relevant, as customers are more likely to make decisions based on the wisdom of the crowds. [12] also found that valence is one of the strongest predictor of sales among all the other word of mouth attributes.

2.2.2 Online review volume

The amount of online reviews for a product or service is known as the online review volume [9]. More discussions about a product or service in e-WOM can lead to increased awareness among customers, resulting in changes in pricing, according to studies on why online review volume increases product sales. . In this study, online review volume refers to the number of comments from reviewers about the specific electronic product. [13]on the other hand, claim that the amount of online feedback is more relevant in experience goods than in search items like electronics. Users of experience products prefer to rely on extrinsic cues such as product popularity, which is reflected in the number of e-WOM reviews, since they are unable to experience product attributes. [13] Conclude that the valence of e-WOM is more significant than its volume in the case of electronic goods. Given the importance of online review volume in predicting product sales, we discovered that it is critical to use volume as one of the predictors in our model.

2.2.3 Negative online review percentage

Negative e-WOM has been shown to affect a customer's buying decision more than positive e-WOM in previous studies [11]. Negative e-WOM is said to spread much more quickly than positive e-WOM [13]. A positive e-WOM comment indicates a product's quality and credibility, while a negative comment reflects users' lack of trust in a product, which could lead to lower product sales [13]. Thus, in addition to product scores, the proportion of positive and negative feedback will affect customer buying decisions. When opposed to severe positive information, psychologists have discovered that negative information has a much stronger effect on evaluations. As compared to positive information, negative information is perceived to be more valuable for decision-making purposes and is given greater weight.

Customers today face a dilemma in terms of the amount of online information accessible to them. When it comes to making purchases, the abundance of information will make it difficult to make decisions. When it comes to advertising their goods online, many businesses are less likely to use only one marketing technique. Firms today are more likely to provide multiple knowledge channels from suppliers to consumers and from previous customers to users to reach out to them, according to a report by [13]. Will a company's promotional marketing campaigns, on the other hand, be a better indicator of product sales, or would online reviews? This study extends on [13]'s study, and examine if online reviews such as its volume, valence and percentage of positive comments' interactions with price discounts can improves the prediction on product sales and customer demand of products using our proposed methodology.

III. PROPOSED METHODOLOGY

3.1 Background and data for the study

We have proposed our study that aims to show how manufacturers can use online stores to forecast consumer demands. Amazon.com, a retailer-hosted e-WOM website, was used as the source of our data. Since Amazon.com is a popular, well-known, and well-respected company, it is a good source of data for our research on the impact of e-WOM and online marketing strategies. To reduce the possibility of biased data due to low reputation and sales, we avoided pages with lower rankings.

On a marketplace like Amazon.com, our risk is small. This research will look at e-WOM using Amazon.com, and the items will include cameras, televisions, hi-fi systems, notebooks, and other electronic devices. The focus of this study will be on the effects of valence, volume, and percentages of positive and negative online feedback on electronic product sales can analyzed and promote

3.2 Big data technology

Our technology choice lays the groundwork for future research while also serving its role in this study. On a typical desktop with traditional network connections, processing I/O for hundreds of pages of web items is possible. Scalable technology is required, however, to manage tens to hundreds of thousands of web pages, including data cleaning during runtime. In the twenty-first century, Big Data technology has become a requirement for science, and we lay the groundwork for extracting data to guide our research here. Before deploying our Web crawlers, we will find the product connections. Once our crawlers had gathered enough data, we deploy them. We used our asynchronous Web scraper agents to populate product links by parsing the incoming HTML info and use regular expressions to remove specific elements in real time. Incoming data is automatically saved in the database. After the scraper agents finished their work, a CSV file will be created on the MongoDB server.

3.3 Neural networks

The neural network is a machine learning technique based on the human brain [14], in which the networks are viewed as structures of interconnected neurons that can compute values from input data. From sample data, a neural network may learn the intrinsic nature of patterns or processes [15]. A neural network is made up of a number of nodes that are arranged in a hierarchy of layers [14]. In most neural networks, there will be three layers: input, secret, and output [16]. The input node is the receiver that receives data files, as the name implies. The final information produced is stored in the output layer. The hidden layers will obtain feedback from the input layer's neurons, and the interneuron relation strengths (i.e. synaptic weights) will be used to store information. The neural network will analyse the data set using a supervised learning algorithm, and the synaptic weights of the neural weight will be modified to achieve the desired design objective. They are then used to store information and make it accessible in the future [15].However, there is an increase in the need to apply predictive analytics into information systems. One of the key advantages that predictive model such as neural network can offer is that they are able to create useful and practical model which can help researchers to develop new theory.

3.4 Limitations of neural networks

- Neural network is non-parametric model
- Compared to parametric regression models, neural network does not assume about probability distributions of the variables
- Neural network's 'black-box' nature is its main disadvantage.

3.5 Research questions

We assess our proposed approach using the following two research questions:

1. **Research Question 1**: Can deep neural networks be used to reliably identify product innovator firms solely based on their website texts?
2. **Research Question 2:** When such a prediction model is applied to a broad out-of-sample dataset of firm website texts, are the resulting firm-level, regional, and sectoral trends similar to the patterns observed from existing innovation indicators?
3. **Research Question 3:** can increase the promotion strategy features for band purchase probability?
4. **Research Question 4:** can apply any probability model to optimize the promotion strategy for product sale in terms of profit maximization or quality maximization or customer maximization?
5. **Research Question 5:** can proposed model in this research recommend a system for non-discount product?
6. **Research Question 6:** can analysis the inter relationship among promotion strategies.

Hence we have proposed an architecture using Deep Neural networks to be carried out in future. We hope that this proposed methodology would outperform well. The Figure 2 shows a big data architecture for customer demand product prediction via promotion strategy using big data and eagle strategy techniques.

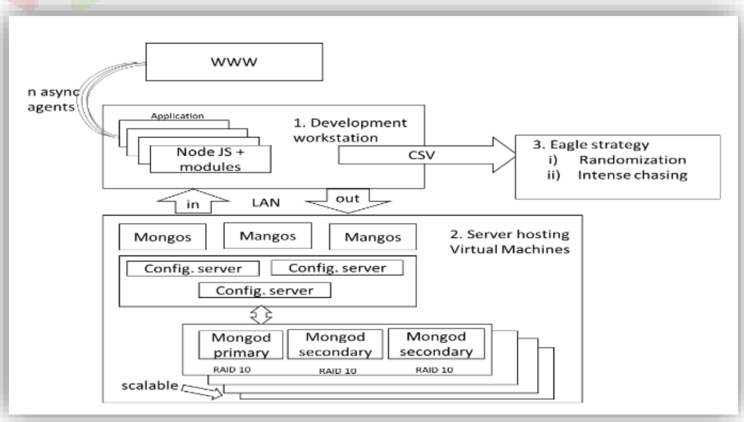

Figure 2. Architecture for promoting product sale via promoting strategy in Big Data

IV. CONCLUSION

This paper investigates the commodity demands of electronic devices using Big Data technology and neural network modelling on the internet. The findings indicate that all of the variables studied are useful predictors of product sales. In addition, the outcomes show that variables from previous studies' online reviews and online promotional marketing variables can be used to create a model that forecasts online sales of electronic goods. We've also suggested a methodology for the potential work we'll be doing. The proposal is made in light of neural network limitations. The procedure will be carried out in stages in the future.

V. ACKNOWLEDGMENT

G. Midhu Bala, Assistant Professor, Department of Computer Science, Mangayarkarasi College of Arts and Science for Women, Madurai, Tamil Nadu, India, *Email ID*: midhug.research@gmail.com.

Dr. K. Chitra, Assistant Professor, Department of Computer Science, Government Arts College, Melur, Madurai, Tamil Nadu, India.

REFERENCES

[1]. Dobrzykow ski, David, Vafa Saboori Deilami, Paul Hong, and Seung-Chul Kim. 2014. "A Structured Analysis of Operation and Supply Chain Management Research in Healthcare (1982–2011)." International Journal of Production Economics 147, Part B: 514–530. doi:10.1016/j.ijpe.2013.04.055

[2]. Alain Yee LoongChonga, Eugene Ch'ng , Martin J. Liu and Boying Li. 2015. "Predicting consumer product demands via Big Data: the roles of online promotional marketing and online reviews". International Journal of Production Research, http://dx.doi.org/10.1080/00207543.2015.1066519

[3]. G. Midhu Bala.2021." Data Extraction and Scratching Information Using R". International Journal of Arts, Science and Humanities, Volume: 8 Issue: 3 , January, P-ISSN: 2321-788X E-ISSN: 2582-0397

[4]. Predictive Analytics Tools and Web Scraping for the Future (scrapingrobot.com)

[5]. JyotiPrakash Singh, SedaIrani, Nripendra P. Rana, Yogesh K. Dwivedi, Sunil Saumya, Pradeep Kumar Roy,Predicting the "helpfulness" of online consumer reviews,Journal of Business Research,Volume 70,2017,Pages 346-355,ISSN 0148-2963,https://doi.org/10.1016/j.jbusres.2016.08.008.

[6]. Internet Live statistics, article retrieved October 15, 2015, from http://www.internetlivestats.com/intern et-users/

[7]. Liu, Martin J., Natalia Yannopoulou, XuemeiBian, and Richard Elliott. 2015. "Authenticity Perceptions in the Chinese Marketplace." Journal of Business Research 68 (1): 27–33

[8]. Yip, Lily, and Rob Law. 2002. "User Preferences for Web Site Attributes: A Study of the Hong Kong Disneyland." Asia Pacific Journal of Tourism Research 7 (1): 33–44. doi:10.1080/10941660208722108.

[9]. Lu, X., S. Ba, L. Huang, and Y. Feng. 2013. "Promotional Marketing or Word-of-Mouth? Evidence from Online Restaurant Reviews." Information Systems Research 24 (3): 596–612. doi:10.1287/isre.1120.0454.

[10]. Zhu, Feng, and Xiaoquan Zhang. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics." Journal of Marketing 74 (2): 133–148. doi:10.1509/jmkg.74.2.133.

[11]. Cheung, Christy M. K., and Dimple R. Thadani. 2012. "The Impact of Electronic Word-of-Mouth Communication: A Literature Analysis and Integrative Model." Decision Support Systems 54 (1): 461–470. doi:10.1016/j.dss.2012.06.008.

[12]. Dellarocas, Chrysanthos, NeveenAwad, and Xiaoquan Zhang. 2004. "Exploring the Value of Online Reviews to Organizations: Implications for Revenue Forecasting and Planning." ICIS 2004 Proceedings 30: 379–386.

[13]. Cui, Geng, Hon-KwongLui, and XiaoningGuo. 2012. "The Effect of Online Consumer Reviews on New Product Sales." International Journal of Electronic Commerce 17 (1): 39–58

[14]. Chiang, Wei-yu Kevin, Dongsong Zhang, and Lina Zhou. 2006. "Predicting and Explaining Patronage Behavior toward Web and Traditional Stores Using Neural Networks: A Comparative Analysis with Logistic Regression." Decision Support Systems 41 (2): 514–531. doi:10.1016/j.dss.2004.08.016

[15]. Sim, Jia-Jia, Garry Wei-Han Tan, Jessica C. J. Wong, Keng-Boon Ooi, and Teck-Soon Hew. 2014. "Understanding and Predicting the Motivators of Mobile Music Acceptance – A Multi-Stage MRA-Artificial Neural Network Approach." Telematics and Informatics 31 (4): 569–584. doi:10.1016/j.tele.2013.11.005.

[16]. Chong, Alain Yee-Loong. 2013. "Predicting M-Commerce Adoption Determinants: A Neural Network Approach." Expert Systems with Applications 40 (2): 523–530. doi:10.1016/j.eswa.2012.07.068.