

# NATURAL LANGUAGE PROCESSING ON CUSTOMER DESCRIPTIVE REVIEW ON E-COMMERCE APPLICATION

**Awantika Singh**

Assistant Professor

Department of Computer Science, Kalinga University, Naya Raipur

## ABSTRACT

The reason for this paper is to acquire bits of knowledge from the web-based item audits of internet business destinations like Flipkart and Amazon and examine its effect on outsider dealers. To pass judgment on the realness of an item, surveys are more valuable than evaluations, since evaluations don't give a total picture. It is continuously liked to consider both the item and merchant surveys to have a consistent conveyance and deformity less item. In this paper, regular handling techniques are utilized to acquire bits of knowledge by considering on the web surveys of an item. Strategies, for example, opinion investigation, sack of words model assistance to comprehend the effect of online item audits on the merchant's evaluations and their exhibition throughout some time. The surveys are ordered into positive, negative, and unbiased utilizing opinion examination. Further, subject demonstrating is done to figure out the point audits are significantly alluding to. The merchant surveys for an explicit item after examination are contrasted with the general dealer surveys with judge the credibility. The aftereffects of this paper would be useful to both the purchasers and venders.

**KEY WORDS:** Bag of Words, Natural Language Processing, Online Product Reviews, Sentiment Analysis, Topic Modeling.

## Introduction

Internet shopping has totally impacted approach towards shopping. Include India alone, online business development is expected towards be \$120 billion, developing at a yearly pace of 51%, by 2020 (IBEF Report, 2020). Online customers stand at 25 million include India. Nowadays, it needs only a single tick and we are only one day from item showing up at doorstep. With innovation changing step by step and that's only tip of iceberg choices coming in, individuals prefer shopping on web than visiting stores (Xiong et al., 2020). Online goliath retailers like Amazon, Flipkart, and Myntra are making it simpler by giving different items include different classifications consistently. Once include a while, items that are not accessible in disconnected stores can be purchased on web. Likewise, value will include general be on lower side when contrasted with store cost. These are a portion of motivations behind why individuals presently really like towards buy on web. Online retailers enjoy a ton of benefits when contrasted with conventional blocks and concrete stores. Online retailers can sell however many items as they need. They have no imperatives with regard towards stock or space executives. These goliath retailers have likewise acquired trust of buyers by furnishing them with one-day conveyance choices, lower costs, confided include venders, item audits, vender surveys, and appraisals. So an ordinary client visits site, picks item, takes a gander at cost and evaluations, peruses surveys, and afterward looks at (Vineet et al., 2020). main factors that he/she typically considers are appraisals and audits of item. These two variables assume a vital part while a customer chooses an item. include any case, on web shopping has its own concerns. There is an opportunity of clients getting a defective or another item rather than one they requested. Likewise, evaluations include some cases don't give a total picture of worth of item. For instance, a buyer buying a cell phone exclusively founded on evaluations could generally dislike battery duration of portable. audits help towards figure out such issues. Online surveys are an incredible wellspring of data and assume a urgent part when a buyer buys an item. Buyer chooses whether towards go on or back down based on surveys (Wang and Chen, 2020). Surveys are composed by shoppers in wake of buying an item and involving it for a short/long span. These surveys are composed for both item what's more, vender. Item surveys significantly characterize various parts of item and experience of client subsequent towards utilizing it. Merchant audits discuss conveyance, administration given, and some of time way of behaving of dealer (Xiaolin et al., 2019). Larger part of past investigations discuss online item surveys, their significance, opinion examination being utilized towards dissect surveys, and effect of these audits on deals (Ali et al., 2020). These studies have considered client surveys at a significant level or were restricted towards just item audits. An examination has been done on those surveys utilizing various methods (feeling investigation, ML calculations) towards give item suggestions and figure out assessment of client with respect to item. Yet, merchant audits are totally disregarded. This study thinks about both item and vender surveys for additional examination towards work on current interaction. Considering above angles, this paper intends towards study and examine web-based audits (both of item and dealer) utilizing Natural Language Processing (NLP) methods. NLP was utilized include this study since it helps include figuring out information, regardless of whether it is unstructured, and then takes choices include view of bits of knowledge accumulated. Additionally, feeling analysis which is one of main fields of

NLP, was utilized include this review. Feeling examination not just aides include figuring out current and notable setting of text yet include addition guides include anticipating what's towards come (Shobana and Murali, 2020). Furthermore, data put away in text can be utilized as a marker like positive, negative and nonpartisan (Al-Sharuee et al., 2020). This part of feeling examination is extremely useful include this review, which can be utilized as a sign for decision makers. Relating something similar towards this review, NLP assists with understanding effect surveys make on dealer and furthermore contrast them and vender surveys overall towards pass judgment on their genuineness. At last, shoppers can settle on buy ignite from a specific merchant based on above examination and furthermore dealer can work on their working include light of client input (Caitlin et al., 2019).further areas of this paper, for example, Literature Review segment examine different studies where comparative examination has been done and hole this study will fill. Different areas include particular Research Methodology and Analysis and Results; examine approach utilized towards get bits of knowledge, examination done, and outcomes. Conversation area talks about outcomes include correlation with different examinations include this field. At last, certain suggestions and restrictions have additionally been incorporated.

## **RESEARCH METHODOLOGY AND DESCRIPTION**

This study was done include two sections. towards start with, essential exploration was directed towards grasp client assessments about utilization of online surveys include buy choices. Also, optional examination of gathering and dissecting client audits was directed. A positivist methodology of assessing mathematical evaluations can't give subtleties of surveys, similar towards purposes behind clients' preferences loathes. It has been seen that a rating of 4 out of 5 might be generally excellent however survey remarks may incorporate grievances about item or administration. Such complicated surveys get dismissed include a positivist methodology. Review expects that client is a complicated person who is unaffected by outside special endeavors and needs top towards bottom examination for better comprehension of their way of behaving. Clients will generally express their genuine thoughts whenever gave a chance of a nitty gritty survey as opposed towards simply a mathematical rating (Gallagher et al., 2019). Subsequently this approach helps include grasping reason why clients feel manner include which they feel. Information assortment segment depicts technique followed towards assemble information. Information extraction segment portrays technique for scratching surveys from online business site. Point depiction segment of procedure gives a short portrayal of multitude of subjects that were examined also, utilized include this review. At long last, information demonstrating segment portrays start towards finish cycle of building a model.

### **Information Collection**

Inessential examination, a poll was drifted and almost 150 reactions were gathered. Questions were connected with clients' inclinations while buying an item on web and their focuses of view about item and merchant audits. Respondents had a place with different age bunches going from 20 towards 60 years. Every last one of them had past experience of buying/shopping on web. of them had option towards separate between item audits and vender surveys. Outcomes from this research framed central part of auxiliary exploration. Include optional examination, online monster retailer Amazon.in was picked. An item with a high number of evaluations and surveys (Sennheiser CX 180) was thought of. Surveys from India were taken for examination. Just a single item and that too specifically above item was picked due to accessibility of impressive number of audits for item as well as for venders. include different items, there was an irregularity in quantity of audits. Surveys of just those clients who had a Verified Purchase tag were thought of. Tag as well as historical backdrop of client, like number of buys, his use, and number of audits given towards date were some other factors thought about.

### **Information Extraction**

Mining overall or information mining is tied include with dissecting information and acquiring experiences from given information. Include this review, mining of online audits was done towards grass preference. Whole course of investigation was done utilizing Python and Excel. Libraries/apparatuses utilized include Python include: Pandas utilized for information control and Analysis, NLTK (Normal Language Toolkit), a device that has various libraries connected with measurable and emblematic NLP, Seaborn for alluring and useful representations, Gensim for unaided subject displaying, furthermore, pyLDAvis, which is an intelligent LDA representation bundle. Web scratching strategy was utilized towards scratch surveys, and an example of 5,000 item audits was thought of. Web scratching is a computerized technique that is utilized towards remove information from sites what's more, save it include a calculation sheet on nearby PC (Lusiana et al., 2019).separated result include this study comprised of various qualities, for example, request number, URL, creator, title, date, content, rating, and hyperlink. This was information that was utilized for opinion examination.

### **Theme Description**

Normal Language Processing Regular Language Processing or NLP is connected with software engineering, phonetics, and counterfeit knowledge. It manages associations among PCs and people utilizing normal language. Specifically, investigating a lot of normal language information by machines towards acquire experiences is what NLP is about. NLP empowers machines towards peruse, comprehend, and get significance from human dialects (Vera et al., 2020). Some utilization cases incorporate catboats, recognizing counterfeit news, halting spam, expectation of illnesses from past wellbeing records, and opinion examination. Additionally, NLP and text examination were utilized include breaking down UK yearly report accounts. However work was restricted exclusively to English language, further examination was done towards execute include other dialects (Mahmoud et al., 2019).

## Feeling Analysis

Feeling investigation is a subfield of NLP that is utilized towards recognize feelings, disposition, and assessments include light of computational treatment of subjectivity include a text (Sameh and Ozgur, 2020). Feeling examination can give a ton of data and bits of knowledge about decisions of clients and their choice drivers (Shobana and Murali, 2020). Include this review, VADER (Valence Aware Dictionary and sEntiment Reasoner) feeling examination is utilized. This is a dictionary and a standard based feeling investigation device that is utilized towards order a text as positive, negative, or impartial (Anton and Martin, 2020). VADER investigation likewise gives a compound score, which is an amount of all vocabulary evaluations standardized between - 1 towards +1. Towards order a text as a positive, negative, or impartial one, Table 1 can be utilized as a source of perspective.

Sentiment	Compound Score
Positive	$\geq 0.05$
Neutral	$> -0.05 \ \& \ < 0.05$
Negative	$\leq -0.05$

Table 1. Compound Score Metric

## Information Modeling

subsequent stage is towards separate significant and valuable subjects include every classification audit. Idea of Latent Dirichlet Allocation (LDA) was utilized towards separate something very similar (Fan et al., 2016). LDA was utilized include this review since it gives preferable exactness over other unaided calculations like K-implies. Accompanying advances are engaged with cycle.

## Information Preprocessing

This is a significant stage wherein accentuations and stop words are taken out and audits are standardized however much possible. The most successive words in information are likewise separated. Since regular words comprise of many articles that are really not applicable, tokenization and lemmatization is finished. This aides include separating important and helpful successive words from audits (Guang et al., 2017).

## Stop Words

Stop words are most generally utilized words, for example, "a," "an," "the," "in," which are normally taken out while handling regular language information. There is no generally followed list for stop words. Rundown can be worked without any preparation. Foremost part, while handling text utilizing NLP procedures, familiar words, which add very little or no esteem, are eliminated. Some of time while eliminating stop words, pertinent data helpful for our examination may likewise be taken out. include such cases, a extremely negligible measure of stop words rundown ought towards be ready, contingent upon objective client needs towards accomplish (Jatinderkumar and Rajnish, 2016).

## Tokenization

Tokenization is interaction wherein text/string is changed over into a rundown of tokens/words. This is helpful include separating significant words from audits. During time spent tokenization, message is sectioned into words or sentences relying upon order of client. These isolated pieces are called tokens. During time spent division of text, many characters, for example, accentuations are taken out. It is confounded towards manage texts wherein there are a ton of brackets, dashes, and other accentuation marks.

## Lemmatization

Lemmatization works by distinguishing grammatical features of a given word and afterward applying more complex standards towards change word into its actual root, known as lemma. include general, target of lemmatization is towards diminish word into its base structure and afterward consolidate/assemble various types of a similar word. For instance, swimming, swims, and swam are types of "swim," so "swim" is their lemma. Lemmatization, depending on specific situation, can likewise separate between comparable words that have various implications. Along these lines it might actually take care of issues like disambiguation (Khaltar and Atsushi, 2009). There is one more cycle known as stemming, which is likewise utilized by quite a few people. Stemming is additionally comparable towards lemmatization yet it is less compelling than last option. Methodology followed by two strategies is likewise unique. Most common way of stemming is quick and requires less computational power as analyzed towards lemmatization. Be that as it may, it isn't ideal technique. Lemmatization, then again, is a resource intensive task and requires higher computational power. Include this review, lemmatization was done on tokenized surveys, which gives most applicable continuous terms in audits and furthermore promotes eliminate commotion.

## Subject Modeling

Towards find theoretical subjects that happen include an assortment of records, a factual displaying technique called point displaying was utilized. Subject displaying, as name recommends, is utilized towards recognize points present include a message and furthermore towards grasp deeper significance from that text corpus (Ike and Sathish, 2020). Subject displaying is utilized for different purposes like grouping of archives, include determination, unstructured information data recovery, and arranging literary information into huge lumps. It likewise has a few other certifiable applications like dimensionality decrease, proposal motor, and text synopsis.

## Building a LDA Model

Idle Dirichlet Allocation or LDA is an illustration of a subject model that is utilized towards characterize text include a record towards a specific point. It is perhaps most generally utilized model. LDA is a solo learning strategy that forms subject per report model and words per theme model, demonstrated as Dirichlet dispersions (Qing et al., 2020). towards be more exact, LDA assembles words per point model by appointing words towards irregular subjects. Subjects are not characterized by client before. Just quantity of subjects he/she wishes towards know is given as an info. Then, at that point, calculation will plan every one of pertinent words towards a specific subject. Numerous emphases are finished by calculation include planning words towards various subjects by taking various cases into thought. Probabilities that a word has a place with a specific subject and report will be created by a point are determined on numerous occasions until combination of calculation. When LDA model is fabricated, it is utilized towards print out subjects it has learned. Include this review, LDA model was constructed utilizing archive term framework as one of boundaries. Pack of-words can likewise be utilized as a boundary rather than equivalent.

## Pack of-Words

Pack of-words model is most usually involved model that gives include of words include a piece of text. Count of each word is given as an event network. It doesn't consider language structure and request of words (Hong et al., 2016). Then subjects are imagined include a 2-D space. Perception is intelligent and it shows points it has learned, alongside most important words include every subject. This cycle was followed for every one of three classifications — positive, negative, and unbiased surveys, include request towards figure outpoints include which item succeeds and ones wherein it slacks. This would help customer while settling on item. In subsequent stage, audits comprising of vender watchword, all more especially name of dealer are scratched similarly. A similar cycle as above is continued towards order into positive, negative, and nonpartisan. Then, at that point, subject displaying is done towards distinguish regions where help/conveyance was great or turned out badly. Then, at that point, these surveys and furthermore positive, negative, and nonpartisan rates of particular item picked are contrasted and specific dealer surveys that are given by customers for different items overall. This approves whether dealer is bona fide include their conveyance or not, contingent upon new and by and large execution. This correlation of merchant surveys on a granular level assists customers with picking merchant for particular item depending on above factors. Assuming item audits referencing dealer coordinate with general merchant surveys on a presentation premise, then shopper can go ahead with vender. In other case, they can pick a dealer with better evaluations and better execution match rate. This empowers purchaser towards not pick dealer exclusively founded on evaluations and cost (lower cost than different venders). This additionally assists merchant with breaking down what turned out badly with conveyance and administration of item and furthermore work on their presentation if slacking.

Age	Percentage
Below 18	2.70%
18 - 30	66.90%
31 - 50	21.40%
50+	9%

Table 2. Age distribution of respondents.

## Information ANALYSIS AND RESULTS

Of145 reactions got from essential examination, 71 were from females and 74 were from guys. Of them, 97 had a place with age bunch 18-30, 31 to age bunch 31-50, 13 to age bunch of 50+, and 4 individuals were under 18. This is displayed include Table 2. Around 58% of respondents shopped online once include a while and 35 percent shopped frequently. Just 1 individual never shopped on web. This assists with figuring out ascent of internet business. This is shown include Table 3. However there is a little piece of individuals who lean toward just internet shopping, it very well may be expressed that individuals will quite often favor choosing item practically than going to store and attempting it. On account of online item audits, around 91% denoted that they read surveys previously checkout. They would include general peruse around 12-15 surveys on a normal while buying an item on web. This can be seen include Table 4. Out of over 91%, almost 75% likewise went through various dealer choices accessible furthermore, their surveys. Most of individuals consider item audits towards be a higher priority than dealer audits, as

found include Table 5. However it is by all accounts genuine considered from a customer point of view, vender appraisals and surveys ought to towards likewise be given equivalent significance.

Shopping Frequency	Percentage
Always	5.90%
Often	35.20%
Occasionally	57.90%
Never	1%

Table 3. Frequency of shopping online.

Read Reviews	Percentage
Yes	91%
No	3%
Maybe	6%

Table 4. Respondents views about Product Review.

Category	Low Importance	High Importance
Price	74	71
Product Reviews	5	140
Product Ratings	15	130
Seller Ratings	52	93
Seller Reviews	50	95

Table 5. Importance of different attributes according to respondents.

In Table 5, it is plainly apparent that dealer appraisals and surveys are additionally given low significance; also, it is more noteworthy when contrasted with item audits and evaluations. Taking into account every one of above-expressed realities, optional examination was finished by scratching item surveys for a specific item — Sennheiser CX 180Street II in-ear earphone (dark), without mic — from Amazon India. Chose item had around 58,000 appraisals and 48,000 audits. A test comprising of 5,000 surveys was utilized for examination. VADER opinion investigation was utilized towards order audits as sure, negative, and nonpartisan. Compound opinion score metric include Table 6 was thought about and afterward audits were grouped. Out of 5,000 audits, 81 percent (4,054) were positive, 15% (736) were negative and 210 (4%) were nonpartisan. Positive surveys were considered for additional examination. Audits were then tokenized and then lemmatization was done towards get regular words utilized include sure surveys. Table 7 shows main 30 continuous words before stop words are eliminated. Main 30 continuous words after stop words are taken out are displayed include Table 8. Tokenization and lemmatization are done towards get more important regular words in separated surveys. A portion of successive words incorporate "quality," "magnificent," "commotion dropping," "guarantee," "lucidity," "high pitch," and "volume," as displayed include Table 9. LDA model was underlying request towards get main 5 important points include view of above sure audits. This subject model can be envisioned include a 2-D space and every point likewise has its own pertinent words. Words displayed include Figure 1 are more disposed toward working subject of earphone



## Conversation

The significance of online surveys include purchaser navigation has proactively been deeply grounded (Huang and Pape, 2020). Since most recent couple of years, as web based business has been on ascent, it's worth is expanding step by step. Greater part of purchasers like towards peruse others' perspective prior towards going forward with item they have chosen. More item audits, almost certainly, buyer will invest more energy settling on that item as opposed towards any of others, and will probably wind up buying it (Kunlin et al., 2020). Many examinations have previously discussed effect of online surveys include client choices, feeling examination on web-based item audits towards figure out assessment of client with respect to item, suggestion frameworks, foreseeing what client would arrange later on utilizing AI and profound learning calculations (Dennis et al., 2020). component that separates

Sentiment	Lifetime Percentage
Positive	89%
Neutral	4%
Negative	7%

Table 6. Compound Score Metri

This review from every one of previously mentioned ones is that even merchant surveys have been considered for further examination. Merchant surveys ordinarily are available for every dealer on a significant level, for every one of orders that vender has conveyed. This can be utilized towards pass judgment on presentation of merchant on a general level. Be that as it may, when buyer needs towards choose a specific merchant while requesting a specific item, it is important towards try and present vender's exhibition include conveying that specific item. Past conveyance experience and audits that are given by purchasers for conveyance of that item ought towards likewise be thought of. However this is on an additional granular level, this denotes a significant perspective for merchant. This part of considering merchant surveys, which was not done include past examinations, helps us towards not just find out about advantages and disadvantages of merchant yet include addition towards dissect their execution and settle on dealer. As per Huang and Pape (2020), writing proposes that two impacts of customer audits on deals are mindfulness impact and convincing impact. Creators found impact of both, however convincing impact is somewhat bigger than mindfulness impact. Thus we dissected impact of both purchaser audits and dealer surveys. Thus, momentarily, this concentrate primarily mirrors significance of mining on web item surveys and furthermore dissecting effect these surveys make on outsider dealers. This study is advantageous towards both shopper and merchant. Through this review, it is clarified that vender surveys likewise convey equivalent significance as item surveys. Dealer surveys don't just mean audits include general yet include addition surveys of dealer given by clients in item audit itself. Dealer ought towards likewise think of them as towards take further choices on how and include what regions towards get to next level. Procedures utilized and technique followed can be used for different sorts of items excessively towards investigate effect surveys make on outsider venders. These outcomes might be valid for more youthful grown-up customers yet not really for more established grown-up shoppers (von Helversen et al., 2018). Additionally, review done by von Helversen et al. (2018) proposes that more established grown-up shoppers don't think about certain audits zeroing include on sure encounters however are effectively convinced by audits revealing negative encounters.

## CONCLUSION

However a portion of internet based monster retailers have understood reality and are gaining ground toward dissecting these surveys and acquiring bits of knowledge, time has come for others towards likewise understand something similar. Separated from surveys, individuals likewise take a gander at how they are introduced. Route ought towards be without any problem seen so guest can track down them and ought towards get a kick out of chance towards go through them. Vender surveys ought towards likewise be plainly apparent to shopper, towards keep away from those situations where there is an issue with vender conveyance. However retailers consequently suggest merchants with high evaluations and ubiquity, when various merchants additionally grab eye This concentrate essentially mirrors significance of mining on web item audits and furthermore examining influence these audits make on outsider venders. This study is useful towards both buyer and dealer. Through this review, it is clarified that merchant surveys likewise convey equivalent significance as item audits. Vender audits don't just mean surveys all include all yet include addition audits given by clients to merchant in item survey itself. Merchant ought towards likewise think about them towards take further choices on how and include what regions towards get to next level. Taking into account examination done include this review, experiences noticed/accumulated help both purchaser also, and dealer. purchaser benefits from way that as opposed towards going through a ton of surveys include request towards be aware of upsides and downsides of item, this investigation helps him/her towards straightforwardly see level of positive, negative, and impartial surveys and pertinent successive words include every classification, and furthermore drives home way that there are sure subjects that end user can see straightforwardly with relating words include every point. This would assist him with finding out about item significantly quicker. Dealer examination includes general and on a specific item would likewise help buyer towards choose whether towards choose that specific vender or pick another.

According to dealer's perspective, towards enhance item deals execution; merchant audit investigation would be of extraordinary assistance. This review, by taking into account merchant surveys, which was not done include past examinations, helps us towards not just find out about advantages and disadvantages of dealer yet include addition towards examine exhibition and settle on vender. Procedures utilized and technique followed can be used for different sorts of item excessively towards investigate effect audits make on outsider dealers.

## REFERENCES

- A. R. (2020). The importance of online reviews depends on when they are presented. *Decision Support Systems*, 133, 113307. doi:10.1016/j.dss.2020.113307
- Ali, M. M., Doumbouya, M. B., Louge, T., Rai, R., & Karray, M. H. (2020). Ontology-based approach to extract product's design features from online customers' reviews. *Computers in Industry*, 116, 103175. doi:10.1016/j.compind.2019.103175
- Al-Obeidat, F., Spencer, B., & Al Taei, M. (2017). Identifying major tasks and minor tasks within online reviews. *Future Generation Computer Systems*.
- AL-Sharuee, M. T., Liu, F., & Pratama, M. (2020). Sentiment analysis: Dynamic and temporal clustering of product reviews. *Applied Intelligence*, 1–20.
- Anton, B., & Martin, B. (2020). Using VADER Sentiment and SVM for Predicting Customer Response Sentiment. *Expert Systems with Applications*, 113746.
- Caitlin, D., Theresa, A., Philip, E., & Suzzane, B. (2019). A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *International Journal of Medical Informatics*, 125, 37–46. doi:10.1016/j.ijmedinf.2019.02.008 PMID:30914179 Camilleri,
- Chang, Y. C., Ku, C. H., & Chen, C. H. (2019). Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *International Journal of Information Management*, 48, 263–279. doi:10.1016/j.ijinfomgt.2017.11.001
- Chauhan, U. A., Afzal, M. T., Shahid, A., Abdar, M., Basiri, M. E., & Zhou, X. (2020). A comprehensive analysis of adverb types for mining user sentiments on Amazon product reviews. *World Wide Web (Bussum)*, 1–19. doi:10.1007/s11280-020-00785-z
- Dennis, K., Stefan, L., & Markus, S. (2020). Predicting online shopping behaviour from clickstream data using deep learning. *Expert Systems with Applications*, 113342.
- Doi, N., & Hayakawa, H. (2020). Electronic word-of-mouth: A survey from an economics perspective. *International Journal of the Economics of Business*, 1–18.
- Fan, C., Tianlin, Z., Yuexin, W., & Nan, G. (2016). Research on Classifying Web Information with Latent Semantic Based on Modified LDA Model. *International Conference on Networking and Network Applications*. doi:10.1109/NaNA.2016.25
- Gallagher, C., Furey, E., & Curran, K. (2019). The application of sentiment analysis and text analytics to customer experience reviews to understand what customers are really saying. *International Journal of Data Warehousing and Mining*, 15(4), 21–47. doi:10.4018/IJDWM.2019100102
- Guang, W., Jainfang, J., & Shen, Y. (2017). Quality-related Fault Detection Approaches Based on Data Preprocessing. *IFAC-PapersOnLine*, 50(1), 15740–15747. doi:10.1016/j.ifacol.2017.08.2305
- Hong, L., Hao, T., Wei, X., & Ziyi, G. (2016). Sequential Bag-of-Words model for human action classification. *CAAI Transactions on Intelligence Technology*.
- Huang, M., & Pape, A. D. (2020). The Impact of Online Consumer Reviews on Online Sales: The Case-Based Decision Theory Approach. *Journal of Consumer Policy*, 43(3), 463–490. doi:10.1007/s10603-020-09464-y
- Ifie, K. (2020). Excellent product... but too early to say: Consumer reactions to tentative product reviews. *Journal of Interactive Marketing*, 52, 35–51. doi:10.1016/j.intmar.2020.03.002 Ike, V., &
- Jatinderkumar, R., & Rajnish, M. (2016). On continent and script-wise divisions-based statistical measures for stop-words lists of international languages. *Procedia Computer Science*, 89, 313–319. doi:10.1016/j.procs.2016.06.076.
- Jiaxiu, H., Xin, W., Mark, B., & Barrie, R. (2020). Revealed preference in online reviews: Purchase verification in the tablet market. *Decision Support Systems*, 113281.
- Khaltar, B. O., & Atsushi, F. (2009). A lemmatization method for Mongolian and its application to indexing for information retrieval. *Information Processing & Management*, 45(4), 438–451. doi:10.1016/j.ipm.2009.01.008
- Kordrostami, E., & Rahmani, V. (2020). Investigating conflicting online review information: Evidence from Amazon. com. *Journal of Retailing and Consumer Services*, 55, 102125. doi:10.1016/j.jretconser.2020.102125
- Kunlin, L., Yuhan, C., & Liyi, Z. (2020). Exploring the influence of online reviews and motivating factors on sales: A meta-analytic study and the moderating role of product category. *Journal of Retailing and Consumer Services*, 55, 102107. doi:10.1016/j.jretconser.2020.102107
- Li, T., Wang, X., & Wu, Y. (2020). Pricing strategies in presence of online consumer ratings-from the product customization perspective. *Journal of Theoretical and Applied Electronic Commerce Research*, 15(3), 84–100. doi:10.4067/S0718-18762020000300107
- Lusiana, C., & Meilianaand Alvin, C. (2019). Social Media Web Scrapping using Social Media Developers API and Regex. *Procedia Computer Science*, 157, 444–449. doi:10.1016/j.procs.2019.08.237

- Mahmoud, E., Paulo, A., Paul, R., Martin, W., & Steven, Y. (2019). Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. *Accounting and Business Research*, 50(1), 6–34.
- Obeidat, Z. M. I., Xiao, S. H., Iyer, G. R., & Nicholson, M. (2017). Consumer revenge using the internet and social media: An examination of the role of service failure types and cognitive appraisal processes. *Psychology and Marketing*, 34(4), 496–515. doi:10.1002/mar.21002
- Qing, X., Xinyuan, Z., Ying, D., & Min, S. (2020). Monolingual and multilingual topic analysis using LDA and BERT embeddings. *Journal of Informetrics*, 14(3), 101055. doi:10.1016/j.joi.2020.101055
- Ramaswamy, S., & DeClerck, N. (2018). Customer perception analysis using deep learning and NLP. *Procedia Computer Science*, 140, 170–178. doi:10.1016/j.procs.2018.10.326
- Sameh, A., & Ozgur, T. (2020). A comparative assessment of sentiment analysis and star ratings for consumer reviews. *International Journal of Information Management*, 54, 102132. doi:10.1016/j.ijinfomgt.2020.102132
- Sann, R., & Lai, P. C. (2020). Understanding homophily of service failure within the hotel guest cycle: Applying NLP-aspect-based sentiment analysis to the hospitality industry. *International Journal of Hospitality Management*, 91, 102678. doi:10.1016/j.ijhm.2020.102678
- Sasikala, P., & Sheela, L. M. I. (2020). Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS. *Journal of Big Data*, 7(1), 1–20. doi:10.1186/s40537-020-00308-7
- Sathish, A. (2020). A review of topic modeling methods. *Information Systems*, 101582. India Brand Equity Foundation (IBEF). (2020). Indian Ecommerce Industry Report. Accessed at <https://www.ibef.org/industry/ecommerce-presentation>
- Shobana, J., & Murali, M. (2020). A survey on sentiment analysis algorithms and techniques. *Journal of Advanced Research in Dynamic and Control Systems*, 1893-1906.
- Vera, S., Yiftach, B., Eli, K., & Eyal, K. (2020). Deep Learning for Natural Language Processing in Radiology—Fundamentals and a Systematic Review. *Journal of the American College of Radiology*. PMID:32004480
- Viji, D., Asawa, N., & Burreja, T. (2020). Fake reviews of customer detection using machine learning models. *International Journal of Advanced Science and Technology*, 86-94.
- Vineet, K., Arpita, K., Rosy, B., & Marta, B. C. (2020). Why do online retailers succeed? The identification and prioritization of success factors for Indian fashion retailers. *Electronic Commerce Research and Applications*, 39, 100906. doi:10.1016/j.elerap.2019.100906
- von Helversen, B., Abramczuk, K., Kopeć, W., & Nielek, R. (2018). Influence of consumer reviews on online purchasing decisions in older and younger adults. *Decision Support Systems*, 113, 1–10. doi:10.1016/j.dss.2018.05.006
- Wang, Z., & Chen, Q. (2020). Monitoring online reviews for reputation fraud campaigns. *Knowledge-Based Systems*, 195, 105685. doi:10.1016/j.knosys.2020.105685