

International Journal of Computers and Informatics

Journal Homepage: https://www.ijci.zu.edu.eg

Int. j. Comp. Info. Vol. 6 (2025) 16–26

### Paper Type: Original Article

# Understanding Big Data: A Comprehensive Guide to Analytics Models and Techniques

Nabil M. Abdel-Aziz <sup>1</sup>, Gehan A. Elwahab <sup>2,\*</sup> and Amira M. Fawzy <sup>1</sup>

<sup>1</sup> Faculty of Computers and Informatics, Zagazig University, Zagazig, Sharqiyah, 44519, Egypt.

Emails: nmabedelaziz@fci.zu.edu.eg; eng\_aamira@zu.edu.eg.

<sup>2</sup> Egyptian Russian University, Badr City, Egypt; gehan-foad@eru.edu.eg.

**Received:** 15 Oct 2024 **Revised:** 12 Dec 2024 Accepted: 31 Dec 2024 Published: 01 Jan 2025

### Abstract

Big data analytics is the process of examining huge and varied sets in order to uncover hidden patterns, unknown correlations, trends or interactions. The importance of this field has increased over the past few years because of a massive explosion in data generated from diverse sources such as social media, sensors and IoT devices. The research and business sectors are increasingly interested in big data analytics as a means for determining trends within large datasets. The deployment of sensor networks, cyber-physical systems and internet of things (IoT), have generated large amounts of data in many areas such as healthcare, social media. It allows creating the smart city, agriculture finance education. It collects data from different sources for example sensors, social media platforms and banking records. Our work focused on Explore Fundamental Analysis Approaches: Investigate various methods used in big data analysis, Address the Challenges of Big Data: Discuss the complexities businesses face due to the exponential growth of data. Provide Competitive Advantage: Emphasize the importance of reliable information for businesses to outperform competitors.

Keywords: Big Data; Digital Transformation; Data Characteristics Analytics Models; Predictive Analytics; Descriptive Analytics; Data Types.

# 1 | Introduction

"Big data" can be defined as massive amounts of high-velocity, complex, and changeable data that necessitate modern techniques and technology to acquire, store, distribute, manage, and analyses the information [1]. For analyzing such huge amounts of data, modern analytical tools for swiftly analyzing and/or anticipating future courses of action with high precision, as well as advanced decision-making techniques, are necessary. As data expands in amount, variety, and speed, so does the uncertainty that comes with charity field of analytics models has seen significant advancements, contributing to a deeper understanding and more efficient processing of complex data structures across various industries. Analytics models are broadly classified into two categories: Predictive and Descriptive. Predictive models are focused on forecasting future events or behaviors based on historical data. Common methods under this category include classification, regression, and time series analysis [2]. Descriptive models, on the other hand, aim to identify patterns and relationships in data. Techniques like clustering, summarization, and association rule mining are used to reveal the underlying structure of datasets [3].



Corresponding Author: gehan-foad@eru.edu.eg

Licensee International Journal of Computers and Informatics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0). The demand for advanced analytics has increased in recent years due to the expanding use of enterprise knowledge discovery and information retrieval methodologies. The exponential growth of data, particularly in domains such as legal informatics and e-discovery, has made privacy management and regulatory compliance crucial issues. In order to meet new data-driven challenges across industries, this presentation will examine the state of analytics models today, their applications, and their prospects. Tell me if there are any sections that need extra explanation or if you would like to concentrate on particular topics.

There are different types of the big data analytics models and each one has its own objectives as following:

- Predictive model A model for determining a given result or variable.
- Descriptive model as the name implies, "describes" the data. After you have collected the data, you will need to convert it into a human-readable format.

This work intended to study and classify the big data analytics techniques for selecting the suitable technique

## 1.1 | Big Data Characteristics

### 1.1.1 | Volume

Refers to the size of the data set. These are measured in terabytes and petabytes. In a poll released by IBM in mid-2012, more than half of the 1144 respondents judged datasets larger than one terabyte to be big data [4]. A terabyte of data might fill 1500 CDS or 220 DVDS, or around 16 million Facebook photographs. Facebook, they claim, processes over one million photos per second. 1024 terabytes make to a petabyte [5].

Big data definitions volumes are relative and change depending on factors such as time and data type. Depending on the nature of data, two datasets of the same size may require different data management solutions.

Examples of unstructured data are the tabular data found in spreadsheets or relational databases. Text, images, audio, and video extensible markup language (xml), a textual language for exchanging data on the web, is a typical example of semi-structured data A high level of variety, a defining characteristic of big data, is not necessarily new Data management technologies and analytics is the innovative aspect. This invaluable information is leveraged in decisions related to product promotions, placement, and staffing.

#### 1.1.2 |Velocity

Relates to the rate at which data is generated, as well as the rate at which it should be examined and acted on. High-frequency data is generated by even traditional retailers. For example, Wal-Mart processes over one million transactions each hour. The data emanating from mobile devices and flowing through mobile apps produces torrents of information that can be used to generate real-time. They enable firms to create real-time intelligence from high volumes of 'perish-able' data [6].

#### 1.1.3 |Veracity

The fourth V, Veracity, was coined by IBM to reflect the unreliability of some data sources. Customer feelings in social media, for example, are inherently unreliable because they involve human judgement[4]. It emphasizes how crucial it is to assess the precision and dependability of data. Not all data can be trusted, even while some may have biases, mistakes, or inconsistencies that compromise their quality. Customer opinions posted on social media platforms, for example, are less trustworthy by nature because they are frequently subjective and impacted by situational, emotional, or cultural variables. This demonstrates how difficult it is for companies to filter, validate, and understand data in order to get insightful information. Addressing veracity calls for strong data governance procedures, sophisticated analytics, and context-aware algorithms to guarantee that judgements are founded on reliable facts [7].

### 1.1.4 | Variability (and complexity).

SAS introduced Variability and Complexity as two additional dimensions of big data. The term "variability" refers to the fluctuation in data flow rates. Big data velocity is frequently irregular, with repeated peaks and troughs. The fact that big data is generated from a variety of sources is referred to as its complexity.

### 1.1.5 | Value

Was presented by Oracle as a defining characteristic of big data. Big data, according to Oracle's definition, is generally characterized by "poor value density." The defining boundaries are determined by the firm's size, sector, and location, and these restrictions change over time. It is also worth noting that these dimensions are not mutually exclusive. As one dimension changes, the likelihood increases that another dimension will also change as a result.

# 2 | Classifying Big Data Analytics Models

Big data analytics enhances a company's understanding of a customer's wants and preferences, allowing it to grow its customer base while also maintaining current consumers through personalized and relevant product or service offers[8]. After organizations have collected large amounts of data, the next critical step is to begin using analytics. Many businesses are unsure where to start, what categories of analytics will help them expand, or what these various types of analytics represent.

There two types of big data analytics models:

## 2.1 | Predictive Model

Is a statistical or machine learning methods employed for the prediction of future events based on the patterns observed in the past. Such models predict certain events or behaviors by studying data trends and data interrelations [9].

## 2.1.1 | Data-Driven

Statistical models, such as remapping for crime forecasting, are built on the foundation of historical data that recognizes significant patterns to project future occurrences.

### 2.1.2 | Feature Importance

They use many predictors; therefore, there are always tries to choose fundamental predictors, which would benefit the model's efficiency.

### 2.1.3 | Interpretability

Some models are relatively more interpretable than other models (predictive deep learning models for instance) and these models make the interpretation dependent on the vocabulary and complexity degree.

### 2.1.4 | Scalability

Usually, the size of the data models and their predictive power can be maintained or scaled even in the cloud if necessary or deployable on-prem.

### 2.1.5 | Real-time Prediction

And this can be relevant even in predictive maintenance for example. There are many models to make predictions faster which are real-time models.

## 2.2 | Descriptive Model

As the name implies, "describes" the data. After you have collected the data, you will need to convert it into a human-readable format. It is one of the model types utilized in data analysis that captures and interprets the past. Whereas predictive models relate to future scenarios, descriptive ones deal with the present realities of the data pattern, trends, and relationships [6].

Characteristics of Descriptive Models

- Data Summarization: They condense the complex dataset to its essential elements, making it easier to understand the core concept contained in the dataset.
- Pattern Recognition: The descriptive models assist in spotting patterns and relationships in the data that will assist in illuminating the interdependencies of different variables.
- Visualization: Uses visual means such as charts, and graphs to state the conclusion in a more understandable fashion.
- Exploratory Analysis: For assessing data to make predictions or additional analytics procedures.

Association	Clustering	Summarization Analysis	Sequence Discovery Analysis
To identify patterns, association establishes a connection between two or more items. For instance, a supermarket might notice that customers often purchase whipped cream alongside strawberries, and the reverse is also true. Point- of-sale systems commonly use association to uncover prevalent product relationships. Association Rule Mining An association rule is represented as $X \Rightarrow YX$ \Rightarrow $YX \Rightarrow Y$ , where XXX and $YYY$ are sets of items that do not share any common items [10].	Clustering is the process of grouping similar records together. Clustering is the process of dividing a dataset into groups. Items in a dataset are grouped or clustered according to the idea that objects in one cluster are extremely similar to one another but quite different from objects in other clusters. Data objects in clustering do not have a class label. That is, we have no idea what the final clusters will be or by which characteristic the data will be grouped before we begin clustering. As a result, clustering is often referred to as unsupervised learning[11].	Is a method for emphasizing important findings and trends while condensing and presenting data in a more palatable manner. This method is frequently used to promote comprehension and decision-making in a variety of domains, such as business, healthcare, and research [11].	Is a data mining technique that concentrates on identifying and examining sequential patterns within datasets. Its goal is to reveal trends and behaviors that unfold in a particular order over time, commonly used in areas such as marketing, healthcare, and web usage analysis. For example, before purchasing a razor in a store, a user may regularly purchase shaving foam. It all comes down to the user's purchase order, which the business owner may then organize [12].

#### Table 1. Descriptive technique methods.



Figure 1. Techniques of big data analysis based on each model.

Table 2. Predictive technique methods.				
Classification	Regression analysis	Time serious analysis	Prediction analysis	
Is a supervised machine learning technique used to categorize a certain class of items, many qualities can be employed. To properly identify what will occur within a class, classification allocates objects to target groups or classes. Classification, also known as supervised learning, is the process of identifying a collection of models or functions that characterize and differentiate data classes or ideas based on a set of training data (i.e. data objects whose class label is known) [13].	Is a statistical technique for analyzing how variables relate to one another. In order to comprehend how changes in the predictors affect the outcome, it specifically attempts to model the link between a dependent variable (the outcome) and one or more independent variables (the predictors or characteristics) [14].	Is a statistical method that allows for the study of underlying patterns, trends, and seasonal fluctuations throughout time by analyzing time-ordered data points. When predicting future values from previous data, this technique is essential [15].	Is a procedure that forecasts future events by utilizing statistical methods and past data. This analytical method, which identifies patterns and makes informed estimates about potential future events, is extensively utilized in many different disciplines to guide strategic planning and decision-making [16].	

# 3 | Classifying Big Data Analytics Techniques based on Type of Data

Organizations can select the best approach for their unique requirements by grouping big data analytics strategies according to the type of data they handle. Below is a summary of popular analytics methods broken down by kind of data?

- Text analytics •
- Audio analytics
- Video analytics

- Social media analytics
- Predictive analytics

	Text Analytics	Audio analytics	Video analytics	Social media analytics	Predictive analytics
SYNONYM	Text mining	Speech analytics	Video content analysis (VCA)	No	No
TYPE OF DATA	Extract structured data from unstructured text	Unstructured audio data	Unstructured data	Structured and unstructured data	Structured and unstructured data
FIELDS USED	Emails, blogs, online forums, survey replies, business papers, news, and contact center logs are all examples of data sources.	Medical Feedback Area	Social media Facebook and other Fields	Social networks, blogs), microblogs, social news, social bookmarking, media sharing In biology,	Many fields
METHODS	Information extraction (IE) Text summarization techniques Question answering (QA) Sentiment analysis (opinion mining)	TRANSCRIP T BASED APPROACH( LVCSR) Phonetic- based approach	Namely server- based and edge- based	Content-based analytics Structure- based analytics	Two techniques regression techniques and machine learning techniques
ADVANTA GES	<ol> <li>Statistical analysis, computational linguistics, and machine learning should all be used.</li> <li>Enable organizations to turn enormous amounts of human-generated material into relevant summaries [21] which can be used to make evidence-based decisions.</li> <li>Can be used to forecast the stock market using data gleaned from financial news [22].</li> </ol>	<ol> <li>Analyze hundreds, if not millions, of hours of recorded calls using audio analytics.</li> <li>Can be programmed to evaluate a live conversation and provide cross-selling and up-selling suggestions based on the customer's previous and current contacts.</li> <li>Provide real-time feedback to agents</li> </ol>	<ol> <li>Compared to other methods of big data analysis , video analytics is still in its infancy.</li> <li>Has been used in automated security and surveillance systems.</li> <li>Can execute surveillance activities efficiently and effectively, such as detecting violations of prohibited zones and identifying things that have been removed or left unattended [20].</li> </ol>	<ol> <li>Social media may be classified as follows: social networks, blogs), microblogs, social news, social bookmarking, media sharing, question- and-answer sites, and review sites [19].</li> <li>Social media is an umbrella word for a number of internet platforms that enable users to produce and share information.</li> </ol>	Can be used in practically any field, from anticipating jet engine failure based on a stream of data from thousands of sensors to predicting consumers' future actions based on what they purchase, when they buy, and even what they post on social media [17] .Seek to uncover patterns and capture relationships in data

Table 3. A com	parison betweer	n big data anal	vtics techniques	based on type of data.
- 4010 01 11 00111	parroon been eer	i olg aata alla	, and we we will all all all all all all all all all	buood on type of data.

		4. Assist with the identification and treatment of medical disorders that alter a patient's			The majority of predictive analytics strategies are based on statistical methodologies
		communicatio n style [20] .			utilized two methods Machine learning techniques (e.g. Neural networks) and regression techniques (e.g. Multinomial legit models)[18]
DIS- ADVANTA GES	<ol> <li>It has nothing to do with technology and everything to do with copyright laws.</li> <li>The ability to copy, scan, and then text mine is severely limited by the limitations imposed on many texts in order to protect copyright holders' rights.</li> </ol>	<ol> <li>It's simpler to lose track of time.</li> <li>It necessitates a well-crafted presentation or material.</li> <li>It's possible that participants will pay more attention to the images than to the audio.</li> </ol>	the expense of additional equipment Additional considerations about data management A higher total cost Aggregation can be difficult[23] .	Results from real- time processing might be skewed. One of the most appealing aspects of social data is how much of it is created fast and regularly. Many analytics companies claim to be able to examine this stream of information in real time.	It's possible that the information isn't comprehensive. If you're using survey data, bear in mind that individuals don't always give you correct answers. The quality and presentation of data obtained from various sources might differ.

## 3.1 | Text Analytics

Test analytics is the technique of using unstructured text data to extract insightful information. It analyses, interprets, and extracts valuable information from text by using methods from machine learning, computational linguistics, and natural language processing (NLP) [24].

Text summarization			Sentiment analysis (opinion
Information extraction (IE)	techniques	Question answering (QA)	mining)
Text analytics (sometimes known as text mining) refers to methods for extracting information from textual data. Textual data is maintained by organizations in the form of social media feeds, emails, blogs, online forums, survey responses, corporate papers, news, and call center records. There are Two sub-tasks in IE are Entity Recognition (ER) and Relation Extraction (RE). ER finds names in text and classifies them into predefined categories such as per-son, date, location, and organization. RE finds and extracts semantic relationships between entities (e.g., persons, organizations, drugs, genes, etc.) in the text [25].	This method generates a summary of a single or more documents automatically. Summarization used two approaches: the extractive approach and the abstractive approach . Extractive approach created a summary from the original text units (sentences as usual). The summary is sub classed from the original document. Creating a summary involves identifying and connecting the most important text units. Its importance is evaluated by analyzing their location and frequency in the text. In summarization, it is not important to understand the text. Abstractive summarization It entails extracting semantic data from a text. Text units that were not included in the original text are included in the summaries. This method interprets the original text and creates the summary using contemporary Natural Language Processing (NLP) techniques .	It responds to inquiries that are posed in normal language. These systems have been implemented by healthcare, finance, marketing, and education. QA systems, like abstractive summarization, rely on complicated NLP approaches. There are three approach of this technique : Information Retrieval (IR) knowledge-based hybrid approach. Information Retrieval (IR) Often have three sub- components :(question processing, document processing answer processing). The Knowledge systems are particularly useful for restricted domains, such as tourism, medicine, and transportation, where large volumes of pre- written documents do not exist. In hybrid QA systems, like IBM's Watson, while the question is semantically analyzed, candidate answers are generated using the IR methods [26].	Techniques for sentiment analysis (opinion mining) examine opinionated material, which comprises people's feelings about things like products, organizations, people, and events Businesses are increasingly capturing more data about their customers 'sentiments that has led to the proliferation of sentiment analysis[27]. Document-level, sentence- level, and aspect-based sentiment analysis approaches are the three sub-groups of sentiment analysis techniques [28].

### Table 4. Text analytics methods.

## 3.2 | Audio Analytics

Is the procedure for deriving insights and important information from audio data. This discipline interprets and analyses voice, noises, and other audio signals by combining methods from machine learning, signal processing, and natural language processing [28].

Table 5. Audio analytics methods Speech analytics follows two common technological approaches.

vocabulary continuous speech recognition (LVCSR)	phonetic-based approach
Systems follow a two-phase process: indexing and searching	Systems work with sounds or phonemes. Phonemes are the distinct units of sound in a specified language that distinguish one word from another.

## 3.3 |Video Analytics

Video analytics, also known as video content analysis (VCA), includes a number of techniques for monitoring, analyzing, and extracting useful data from video streams.

	server-based architecture	Edge-based architecture
Feature set	<ul> <li>In this configuration, the video captured through each camera is routed back to a centralized and dedicated server that performs the video analytics[29].</li> <li>They offer limited number of features.</li> </ul>	<ul> <li>Analytics are applied at the 'edge' of the system.</li> <li>They offer a very small number of features.</li> </ul>
Detection performance	It provides medium to high performance due to limited quality of the videos available and limited processing resources.	Low to high due to limited processing resources, sometime for high performance algorithms are run in isolation.

## 3.4 | Social Media Analytics

The study of structured and unstructured data from social media platforms is referred to as social media analytics. The phrase "social media" refers to a multitude of internet platforms that allow users to produce and share information. The two sources of information in social media are user-generated material (e.g., sentiments, photos, videos, and bookmarks) and the relationships and interactions between network entities (e.g., individuals, companies, and goods) [30].

Have recently emerged to extract information from the structure of social networks, Community detection, also referred to as community discovery, extracts implicit communities within a network. Social influence analysis refers to techniques that are concerned with modeling and . Assessing the impact of social network participants and linkages in a given time interval, link prediction algorithms forecast the occurrence of interaction, collaboration, or affect between network components.

<b>Table</b> 7. Social media analytics methods.					
Community detection	Social influence analysis	Link prediction			
CD using DBSCAN algorithm, CD with edge connection in social networks, CD using new- man Girvan algorithm, CD method using Bayesian network and expectation maximization technique, graph mining technique etc [31].	Computational models for social influence analysis :learning influence probability, influence maximization model, influence action and dynamics, influence and interaction [32].	Similarity-based algorithms(local and global similarity index ), maximum likelihood methods(hierarchical structure model, stochastic block model) and probabilistic methods(probabilistic relational model, stochastic relational models) [33].			

## Table 7. Social media analytics methods

## 3.5 | Predictive Analytics

The majority of predictive analytics approaches employed two methodologies: regression techniques (e.g. Multinomial legit models) and machine learning techniques to predict future events based on historical and present data (e.g. neural networks). Because big data is so large and has properties like heterogeneity, noise building, misleading correlation, and unintended indigeneity, new statistical methods are needed. Due to the difficulties of computational efficiency and the unique qualities of big data, new approaches to manage massive data and get considerable insight should be created [33].

## 4 | Conclusion

Our work focuses on the fundamentals of big data analytics and emphasizes how important it is for negotiating the challenges presented by the exponential development of data in contemporary industries. This

study offers a structured understanding of the descriptive and predictive approaches by classifying big data analytics models and methodologies. It also compares analytics methods in detail across various data kinds, including social media, text, audio, and video. These findings highlight the need for organizations to implement customized analytics strategies in order to maximize decision-making, generate competitive advantages, and exploit actionable intelligence. In order to further assist data-driven innovations across a variety of sectors, future research should concentrate on developing these methodologies to handle new problems in computational efficiency, scalability, and accuracy.

### Acknowledgments

The authors are grateful to the editorial and reviewers, as well as the correspondent author, who offered assistance in the form of advice, assessment, and checking during the study period.

## Author Contribution

All authors contributed equally to this work.

### Funding

This research has no funding source.

### Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available due to the privacy-preserving nature of the data but are available from the corresponding author upon reasonable request.

## **Conflicts of Interest**

The authors declare that there is no conflict of interest in the research.

## Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors

## References

- [1] Lele, A. and A. Lele, Big data. Disruptive Technologies for the Militaries and Security, 2019: p. 155-165.
- [2] Roy, D., et al., A Complete Overview of Analytics Techniques: Descriptive, Predictive, and Prescriptive, in Decision Intelligence Analytics and the Implementation of Strategic Business Management, P.M. Jeyanthi, et al., Editors. 2022, Springer International Publishing: Cham. p. 15-30.
- [3] Gorokhovatskyi, V., et al., Search for visual objects by request in the form of a cluster representation for the structural image description. Advances in Electrical and Electronic Engineering, 2023. 21(1): p. 19-27.
- [4] Gandomi, A. and M. Haider, Beyond the hype: Big data concepts, methods, and analytics. International journal of information management, 2015. 35(2): p. 137-144.
- [5] AmirGandomiMurtazaHaider, Big data concepts, methods, and analytics. International Journal of Information Management, 2014.
- [6] Vendor Independent Big Data Qualifications.
- [7] Lugmayr, A., et al., Cognitive big data: survey and review on big data research and its implications. What is really "new" in big data? Journal of Knowledge Management, 2017. 21(1): p. 197-212.
- [8] Hariri, R.H., E.M. Fredericks, and K.M. Bowers, Uncertainty in big data analytics: survey, opportunities, and challenges. Journal of Big Data, 2019. 6(1): p. 44.
- [9] THE FOUR V'S OF BIG DATA.
- [10] Verhoef, P.C., E. Kooge, and N. Walk, Creating value with big data analytics: Making smarter marketing decisions. 2016: Routledge.
- [11] Huang, D., S. Wang, and Z. Liu, A systematic review of prediction methods for emergency management. International Journal of Disaster Risk Reduction, 2021. 62: p. 102412.

- [12] Cai, Q., Cause analysis of traffic accidents on urban roads based on an improved association rule mining algorithm. IEEE access, 2020. 8: p. 75607-75615.
- [13] Oyewole, G.J. and G.A. Thopil, Data clustering: application and trends. Artificial Intelligence Review, 2023. 56(7): p. 6439-6475.
- [14] Thawkar, O., et al., Xraygpt: Chest radiographs summarization using medical vision-language models. arXiv preprint arXiv:2306.07971, 2023.
- [15] Bailey, T.L., STREME: accurate and versatile sequence motif discovery. Bioinformatics, 2021. 37(18): p. 2834-2840.
- [16] Mekuye, B. and B. Abera, Nanomaterials: An overview of synthesis, classification, characterization, and applications. Nano Select, 2023. 4(8): p. 486-501.
- [17] Arkes, J., Regression analysis: a practical introduction. 2023: Routledge.
- [18] Yi, K., et al., A survey on deep learning based time series analysis with frequency transformation. arXiv preprint arXiv:2302.02173, 2023.
- [19] Mistry, J., Impact of Model Selection on Pulmonary EffusionDiagnosis Using Prediction Analysis Algorithms. Journal of Xidian University, 2024. 18(1): p. 611-618.
- [20] Kumar, V. and M. Garg, Predictive analytics: a review of trends and techniques. International Journal of Computer Applications, 2018. 182(1): p. 31-37.
- [21] Kapoor, K.K., et al., Advances in social media research: Past, present and future. Information Systems Frontiers, 2018. 20: p. 531-558.
- [22] Tlili, F., S. Ayed, and L.C. Fourati, Exhaustive distributed intrusion detection system for UAVs attacks detection and security enforcement (E-DIDS). Computers & Security, 2024. 142: p. 103878.
- [23] Stewart, M., et al., Patient-centered medicine: transforming the clinical method. 2024: CRC press.
- [24] Hakam, H.T., et al., Human-Written vs AI-Generated Texts in Orthopedic Academic Literature: Comparative Qualitative Analysis. JMIR Formative Research, 2024. 8: p. e52164.
- [25] Moreno-Pérez, C. and M. Minozzo, Natural language processing and financial markets: semi-supervised modelling of coronavirus and economic news. Advances in Data Analysis and Classification, 2024: p. 1-25.
- [26] Kang, D., P. Bailis, and M. Zaharia. Challenges and Opportunities in DNN-Based Video Analytics: A Demonstration of the BlazeIt Video Query Engine. in CIDR. 2019.
- [27] Ittoo, A. and A. van den Bosch, Text analytics in industry: Challenges, desiderata and trends. Computers in Industry, 2016. 78: p. 96-107.
- [28] Small, S.G. and L. Medsker, Review of information extraction technologies and applications. Neural computing and applications, 2014. 25: p. 533-548.
- [29] Aggarwal, C.C., Opinion Mining and Sentiment Analysis. March 2018.
- [30] Balaji, P., D. Haritha, and O. Nagaraju, An overview on opinion mining techniques and sentiment analysis. International Journal of Pure and Applied Mathematics, 2018. 118(19): p. 61-69.
- [31] Shoumy, N.J., et al., Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals. Journal of Network and Computer Applications, 2020. 149: p. 102447.
- [32] Olatunji, I.E., Video Analytics for Visual Surveillance and Applications: An Overview and Survey. researchgate, July 2019.
- [33] Zachlod, C., et al., Analytics of social media data–State of characteristics and application. Journal of Business Research, 2022. 144: p. 1064-1076.