# Literature Review:

Don Nguyen (Nguyeng, 2020) discusses his experiences with computational text analysis in an article that focuses on complex social and cultural issues. The writers want to share their experiences for three main reasons. They first seek to address and draw attention to difficult problems that aren't typically given enough consideration in conversations regarding computational text analysis techniques. They hope to advance a more thorough comprehension of the challenges associated with computationally assessing social and cultural concepts by illuminating these concerns.

Secondly, Nguyeng (2020) aims to offer a series of fundamental inquiries that might function as a roadmap for scholars conducting study in this field. The writers' varied disciplinary backgrounds and research techniques served as the foundation for these questions, which helped them to uncover commonalities and capture a wide range of ideas that would be of interest to other researchers. Helping academics navigate the complexities of computational text analysis is the goal, especially when working with social and cultural data. The article's third objective is to promote interdisciplinary teamwork.

The author highlights that in order to fully realize the potential of computational text analysis incorporating social and cultural themes, collaborations between academics from other domains and insights from many disciplines are crucial. They think that by overcoming disciplinary barriers, their work can be more productive and result in fresh understandings of culture and society (Nguyeng, 2020). The growing prevalence of insight-driven algorithmic analysis of text is acknowledged by the author. It enables scholars to take on new and controversial topics concerning culture and society, as well as to see things from wider angles and recognize subtle patterns more skillfully. They do, however, also stress how difficult, labor-intensive, and time-consuming the research process in this field is—especially when collaborating across disciplines.

Disagreements and compromises about the operationalization of ideas and the methodology of the study may result from disparate disciplinary backgrounds (Nguyeng, 2020).

The impact of ethnic culture on customer reviews in social commerce was the main topic of this study (Nakayama, 2019), which especially looked at reviews of Japanese restaurants on Yelp.com (in English) and Yelp.co.jp (in Japanese). The researchers found different attitude distribution patterns regarding meal quality, service, ambiance, and price fairness between Japanese and Western clients utilizing bilingual text mining tools. These results shed important information on the differences in review content and ratings between domestic and international users on global social commerce platforms, as well as the impact of culture in this setting (Nakayama, 2019).

Big data analysis and cognitive computing combined to analyze social media tweets for sentiment (SA) (Boyapati, 2022). Big data analytics, or BDA, is becoming more and more important as human data volume increases. An AI-based technology called cognitive computing is presented to overcome problems encountered during BDA. The Term Frequency-Inverse Document Frequency (TF-IDF) technique is used for feature extraction, classification, and pre-processing to clean the data in the suggested model. The Hadoop Map Reduce tool is used to handle large amounts of data. In order to improve classification performance, the research also presents the Binary Brainstorm Optimization (BBSO) feature selection algorithm. Fuzzy Cognitive Maps (FCMs) are classifiers used in sentiment analysis to identify positive and negative sentiments in tweets. The authors provide thorough experimental findings that show how their BBSO-FCM model performs better at classification on a benchmark dataset. The model performs well on a number of sentiment analysis metrics (Boyapati, 2022).

Regarding sentiment analysis of social media data, which comprises user-posted reviews, comments, and critiques (Joo, 2014). The authors contend that social media data offers insightful information that is frequently missed by conventional survey questions. They do, however, point out a number of shortcomings in the analytical techniques now in use, most notably the narrow focus on data only available in English. The study suggests creating efficient analytical techniques for sentiment analysis of non-English language data in order to close this research gap. To gain a deeper understanding of attitudes, especially in multilingual settings, the authors stress the significance of incorporating studies of both English and non-English data (Joo, 2014). A further issue brought up by recent research on sentiment analysis is the necessity for high accuracy techniques that don't necessitate significant training dataset customisation. In response, a novel sentiment analysis technique utilizing languages processors and an adaptive fuzzy inference method (FIM) is presented in this study. This method integrates and develops a multi-source lexicon with the goal of reducing semantic ambiguity. It is claimed that the suggested approach will address the drawbacks of current methods by increasing algorithmic accuracy and permitting non-English language analysis. Several case examples are included in the study to demonstrate the usefulness and implementation of their suggested methodology (Joo, 2014).

The significance of user-generated brand content as an invaluable big data source for deriving relevant insights. Managers must monitor this data in real-time to make successful decisions because a large volume of customer evaluations and ratings are submitted every day (Jil, 2020). The main objective is to extract trustworthy textual customer reviews and translate them into numerical ratings in order to determine which companies or goods are the best. The suggested approach focuses on automatically interpreting reviews to convert user feedback—both favorable and unfavorable—into a score that can be measured. Amazon online reviews are analyzed using sentiment analysis. The paper also presents the false Review Detection Framework (FRDF), which ensures more accurate and dependable outcomes by detecting and eliminating false reviews using Natural Language Processing technology (Jil, 2020). Reviews from high-tech businesses were used to test the FRDF, and consumer sentiment was used to grade brands. The findings indicate that the tool would help customers and brand managers make better judgments when used in conjunction with the conventional 5-Star rating (Jil, 2020).

Sentiment analysis evaluates a representative's and a user's feelings around a problem (Kang, 2019). The tone and style of a writer convey their opinions (Chaki, 2015). Numerous algorithms are available for the analysis, forecasting, and assessment of emotions from textual data, like customer or product reviews. Sentiment analysis substantially aids in polarity recognition. Nevertheless, it encounters difficulties with spam, erroneous information, domain dependence, negation, sophisticated NLP, bipolar terminology, and large vocabulary. Preprocessing is essential to arrange the text and extract data for sentiment analysis (Palomino, 2023). Preprocessing is getting the text ready and cleaned up so that characters can be grouped according to specific letters (Haddi, 2013). Arbitrary text sequences can be handled well by taking punctuation, emoticons, emojis, and white spaces into consideration (Haddi, 2013). Text preparation techniques improve Sentiment analysis accuracy, according to the research. (P, 2017) To increase data mining's efficacy, the problems listed above must be resolved (R., 2020). It has been demonstrated that semantic sentence analysis improves research technique, accuracy, and consistency [27]. Marketing researchers use a variety of techniques to examine the growing number of text-based customer reviews on digital platforms (Kanakaraj, 2015). Prediction accuracy is increased by using ensemble approaches for classification and advanced NLP techniques that include word senses and reference senses into vector features (Kanakaraj, 2015).

NLP analyzes, highlights, and extracts meaning from text and audio using machine learning and algorithms (Castilo, 2023). It enables computers to identify speaker emotions, sometimes even better than humans. Sentiment research can succeed by creatively applying advanced AI tools.

Artificial Intelligence (AI) pertains to computer programs that possess human-like abilities, such as speech recognition, visual perception, decision making, and language translation (Lenard, 2019). It is used in product development, decision support, and other commercial domains. Deep learning and machine learning are key components of AI strategy (Lenard, 2019). Supervised learning helps with classification tasks by connecting data to labels, while deep learning finds applicability in real-world scenarios (Lenard, 2019).

To allow computers to self-train and generate programs, machine learning integrates statistics, mathematics, and artificial intelligence [31]. By using historical data and interactions, this "intelligent" software generates proactive solutions for the future [32]. Businesses employ machine learning to provide executives with actionable projections by utilizing massive volumes of data.