### **Chapter 2: Literature Review**

#### **2.0 Introduction**

Sentiment analysis, also known as opinion mining, is an essential aspect of natural language processing (NLP) used to determine the emotional tone behind textual data. Its applications range from gauging customer sentiment in businesses to analyzing public opinion on social or political issues. In retail trading, sentiment analysis provides insights into customer preferences, enabling traders to understand trends and adapt their offerings effectively.

This chapter explores the theoretical framework underpinning sentiment analysis, focusing on its concepts and methodologies. Additionally, it discusses the technologies commonly employed for sentiment analysis, emphasizing their relevance and applicability to retail trading in a campus environment.

#### **2.1 Theoretical Background**

##### **2.1.1 Concept**

Sentiment analysis involves identifying and categorizing opinions expressed in text, often determining whether the sentiment is positive, negative, or neutral. This field emerged from advancements in computational linguistics and data science, driven by the need to automate the understanding of public sentiment across various platforms.

In retail trading, sentiment analysis plays a pivotal role by helping vendors comprehend the preferences and behaviors of their customers. On a campus, where students, staff, and faculty form a dynamic trading ecosystem, sentiment analysis enables traders to offer tailored services and products. For example, analyzing social media posts or feedback forms can help identify trending items or dissatisfaction with certain products.

Key tasks in sentiment analysis include:

1. **Text Preprocessing:** Tokenizing, removing stop words, stemming, and lemmatizing textual data to prepare it for analysis.
2. **Feature Extraction:** Identifying relevant features such as frequency of positive and negative words, linguistic patterns, and contextual usage.
3. **Classification:** Using algorithms to categorize sentiment based on extracted features.

Sentiment analysis can be carried out at various levels, such as:

* **Document-level:** Analyzing the sentiment of an entire document.
* **Sentence-level:** Determining sentiment in individual sentences.
* **Aspect-level:** Focusing on specific aspects mentioned in the text.

##### **2.1.2 Technologies Used**

Modern sentiment analysis leverages various technologies and frameworks, many of which are rooted in machine learning and NLP. The following subsections highlight key technologies relevant to sentiment analysis for retail trading.

**a. Natural Language Toolkit (NLTK):**

NLTK is a leading Python library for NLP, providing tools for text processing tasks such as tokenization, stemming, lemmatization, and sentiment analysis.

* **Advantages for Campus Retail:** NLTK’s simplicity and robust tools allow for quick deployment of sentiment analysis models tailored to the campus setting. For example, NLTK can analyze student reviews to identify popular items or dissatisfaction trends.
* **Key Features:**
  + Prebuilt sentiment analysis models like VADER (Valence Aware Dictionary and sEntiment Reasoner).
  + Support for lexicon-based and machine-learning-based approaches.

**b. Machine Learning in Sentiment Analysis:**

Machine learning algorithms play a vital role in sentiment analysis by learning patterns from labeled data to predict sentiments in new, unseen data. Common machine learning methods include:

* **Logistic Regression and Naïve Bayes:** Suitable for baseline sentiment classification tasks, often yielding high accuracy in well-structured datasets.
* **Support Vector Machines (SVMs):** Effective for handling sparse data, such as reviews and social media posts.
* **Deep Learning Models:** Techniques like Long Short-Term Memory (LSTM) networks and transformers (e.g., BERT) provide superior performance by capturing the nuances and dependencies within text data.

**c. Sentiment Lexicons:**

Lexicons like AFINN, SentiWordNet, and Word2Vec provide precompiled sentiment scores for words and phrases. These lexicons facilitate rapid sentiment analysis without requiring extensive training data.

* **Campus Example:** Lexicons can quickly process customer feedback forms to identify overall satisfaction trends.

**d. APIs and Frameworks for Sentiment Analysis:**

Several APIs and frameworks, such as Google Cloud Natural Language API, TextBlob, and SpaCy, provide prebuilt sentiment analysis functionalities. These tools enable campus traders to integrate sentiment analysis into their systems with minimal development effort.