International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Study of Sentiment Analysis for Analyzing Human Behavior

Manjiri S Deshmukh[1], Dr. Mukta Dhopeshwarkar[2], Dr. Swapna Kolhatkar[3], Sarika Swami[4]

[1]Assistant Professor, Computer Science Department, MMCC, Pune
[2]Assistant Professor, CS and IT Department, Dr. BAMU, Chh. Sambhajinagar
[3]Assistant Professor, Computer Science Department, MMCC, Pune
[4]Research Scholar, CS and IT Department, Dr. BAMU, Chh. Sambhajinagar

Abstract
Sentiment analysis, which examines feelings and viewpoints represented in text, is an essential method for comprehending human behavior by analyzing text to determine the emotional tone of a message. The analysis is done through data collection, data processing, data analysis and data visualization. Its applicability in media monitoring, employee sentiment etc. helps businesses understand customer feedback and opinions, which can lead to improved customer satisfaction, product development, and brand awareness. On the other hand, social media sentiment analysis automatically determines how people feel about a brand or product on social media. It can help businesses understand their audience, manage crises, and improve their brand reputation.
This paper offers a thorough review of the many sentiment analysis techniques used to assess human behavior in various fields. Deeper understanding of psychological and social factors is made possible by emotional intelligence, which is crucial to this process. The study looks at methods, uses, and the importance of sentiment analysis in fields like social interactions, marketing, and mental health. This paper demonstrates the efficacy of sentiment analysis in deciphering human emotions and behaviors by reviewing the literature.

Keywords: Sentiment, Human behavior, TF-IDF, BERT, Machine Learning, Deep Learning etc.

1. Introduction

Human behavior analysis is a multidisciplinary field that examines the factors influencing human actions, thoughts, and emotions. Albert Bandura's Social Learning Theory posits that individuals acquire behaviors through observation and imitation, emphasizing the significance of modeling in learning processes.[simplypsychology.org] Additionally, research highlights the profound impact of social contexts on individual actions, with social norms and group dynamics playing pivotal roles in shaping behavior.
Sentiment analysis is a valuable tool for analyzing human behavior. By examining textual data from sources such as social media posts, product reviews, and online forums, sentiment analysis enables researchers to assess public opinion, monitor emotional responses, and predict behavioral trends. For instance, during the COVID-19 pandemic, sentiment analysis was utilized to evaluate the emotional and

behavioral states of affected populations, aiding in timely mental health interventions[13].In consumer behavior research, sentiment analysis helps businesses understand customer opinions and satisfaction levels by analyzing feedback and reviews, thereby informing product development and marketing strategies. Overall, sentiment analysis serves as a valuable tool in decoding the complex interplay between language and human behavior, offering quantitative measures of subjective experiences that are essential for research and practical applications across various domains.

1.1. Human Behavior Analysis

Human behavior analysis has garnered significant attention across various fields, including robotics, affective computing, biometrics, and pattern recognition. This multidisciplinary interest is driven by societal needs such as security, natural interfaces, and assisted living. The primary objective is to accurately and non-invasively detect and recognize human behaviors, which remains a substantial challenge in research.

Traditional methods for identifying human behavior often involve continuous data collection from physical sensing devices like cameras, GPS, and RFID. These devices can be worn, attached to objects, or deployed in the environment. Subsequently, recognition algorithms or classification models are applied to this data to identify behavior types, facilitating advanced applications. However, such approaches are often intrusive, requiring specific sensing devices and raising concerns about privacy and deployment costs. Recent advancements in machine learning, particularly deep learning, have significantly enhanced the analysis of human behavior. Deep learning addresses the challenge of extracting high-level and abstract features, such as individual pixels in images, where variations can be erratic. These methods have dramatically improved the state-of-the-art in processing images, video, speech, and audio, leading to breakthroughs in applications like human action recognition. The sophistication in behavior analysis has led to a hierarchical arrangement of different levels of abstraction: action primitives, actions, and activities. Action primitives refer to atomic movements at the limb level (e.g., left leg forward), actions describe whole-body movements composed of action primitives (e.g., running), and activities consist of several subsequent actions (e.g., jumping hurdles). This taxonomy aids in interpreting the complexity of human movements within images.

In summary, human behavior analysis is a rapidly evolving field that integrates various technologies and methodologies to understand and interpret human actions. Ongoing research focuses on developing non-invasive, accurate, and efficient methods to analyze behaviors, with applications spanning multiple domains.

1.2. Sentiment Analysis

Sentiment analysis, a subfield of natural language processing also known as "opinion mining" or "emotion artificial intelligence, involves the computational identification and categorization of opinions expressed in text to determine the writer's attitude toward a particular topic. This technique has become instrumental in analyzing human behavior by providing insights into individuals' emotions, opinions, and attitudes as conveyed through language. It primarily focuses on extracting sentiment from user-generated content, such as online reviews, surveys, and social media discussions.

Figure 1: Process of Sentiment Analysis

Data collection: The initial stage of sentiment analysis is data collecting, which includes compiling user-generated information from news articles, blogs, forums, social networks, online reviews, and surveys. Manual analysis of these data sources is practically hard due to their frequent disarray and diverse expressions using a variety of vocabularies, slang, and contextual meanings. Text analytics and natural language processing (NLP) methods are used to efficiently extract and categorize sentiments in order to manage this complexity. The information gathered may be unstructured, like user comments and tweets, or structured, like survey results.

Data preprocessing is a crucial step in preparing text data for sentiment analysis. It involves cleaning and standardizing text to remove elements that introduce noise and ensure consistency. The process begins with noise removal, where punctuation, special characters, and stopwords (common words like "and," "the," and "is") are eliminated since they do not contribute significantly to sentiment detection. To maintain uniformity, all text is converted to lowercase, ensuring that words like "Happy" and "happy" are treated identically. Next, tokenization breaks sentences or phrases into individual words or phrases, enabling detailed word-level analysis. Following this, lemmatization or stemming reduces words to their base or root form, consolidating variations like "running," "ran," and "runs" into "run." These preprocessing steps ensure that the text is in a structured format, enhancing the accuracy and efficiency of sentiment analysis.

Feature Extraction is a critical step in sentiment analysis that converts raw text into a structured format suitable for computational analysis. Since machine learning models and statistical algorithms cannot process raw text directly, feature extraction transforms words and phrases into numerical representations that retain meaningful information.

One common approach is the Bag of Words (BoW) model, which represents text as word frequency counts without considering word order or context. Another technique, TF-IDF (Term Frequency-Inverse Document Frequency), assigns importance to words by measuring how frequently they appear in a document while reducing the influence of commonly used words across multiple documents. More advanced methods involve word embeddings, such as Word2Vec, GloVe, or BERT, which capture contextual meaning and relationships between words by mapping them into high-dimensional vector spaces. These techniques enable sentiment analysis models to better understand text structure, semantics, and sentiment nuances.

Sentiment Detection is the process of classifying text into predefined sentiment categories, such as positive, negative, or neutral. This step is crucial in sentiment analysis as it determines the emotional tone of the text. Sentiment detection can be performed using different approaches, including lexicon-based methods, which rely on predefined sentiment dictionaries (e.g., SentiWordNet, VADER) to assign sentiment scores to words, and machine learning-based methods, which use algorithms like Naïve Bayes, Support Vector Machines (SVM), and deep learning models (e.g., LSTMs, BERT) trained on labeled datasets. Some models provide sentiment scores, indicating the strength of positive or negative emotions,

rather than just categorical labels. More advanced sentiment detection systems can also handle contextual sentiment, sarcasm, and domain-specific language, improving the accuracy of sentiment classification.
Result : the main objective of sentiment analysis is to convert unstructured data into meaningful information.To facilitate interpretation and decision-making, sentiment analysis data are arranged and shown during the Output Generation stage. In order to allow users to monitor sentiment patterns over time, results are frequently presented in dashboards or reports that incorporate visual components like graphs, pie charts, or time-series plots. Furthermore, the sentiment data may be exported in organised formats such as CSV or JSON files, which are easily incorporated into other systems or utilised for additional analysis. Stakeholders may easily analyse and act upon the insights gleaned from the sentiment analysis process thanks to these outputs.

2. Data Collection:

Data collection is the first step of sentiment analysis consisting of collecting data from user generated content contained in blogs, forums, social networks.following table shows the different platform for data collection and description about the type of data available on that platform. Data can be classified into two main types: conventional and microblogging. Conventional data refers to structured information that is organized in a clear format, often found in databases or spreadsheets. This type includes numerical data, survey results, and transaction records, making it easy to analyze and typically stored in relational databases. In contrast, microblogging data is generated from platforms where users share short messages and updates, such as Twitter and Instagram. This data is usually unstructured, comprising free-form text, hashtags, and multimedia content. It is valuable for applications like sentiment analysis and trend tracking, and is often stored in NoSQL databases.

Table1: API and Data description

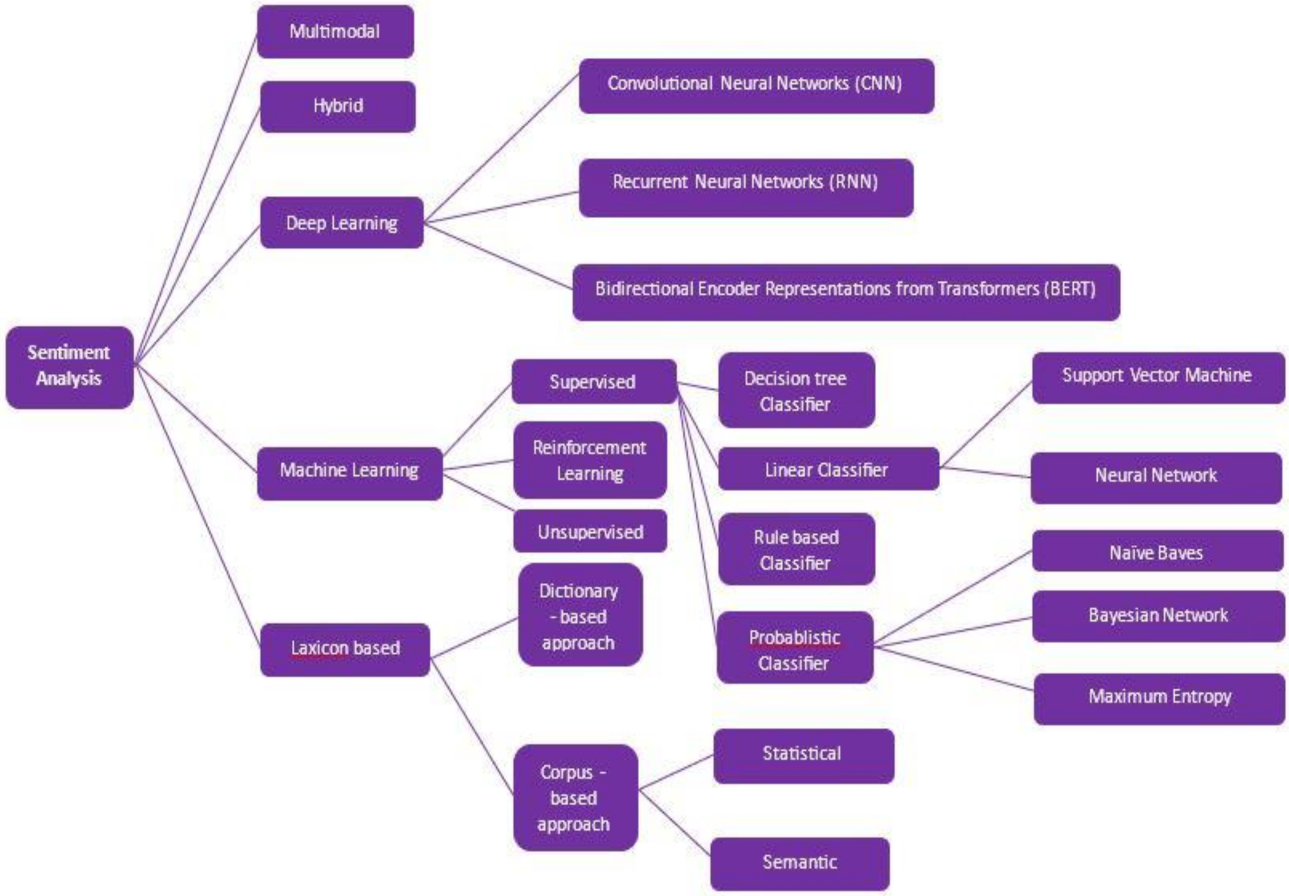
Sr.no	API	Description of Data
1	off-the-shelf tools and APIs	Various customer experience software (e.g. InMoment, Clarabridge) collect feedback from numerous sources.
2	InMoment	provides five products that together make a customer experience optimization platform.
3	Clarabridge	It pulls and analyzes text from chats, survey platforms, blogs, forums, and review sites. The system can convert them into text. They provide social media listening as well.
4	IBM Watson Natural Language Understanding	can extract such metadata as concepts, entities, keywords, as well as categories and relationships.
5	Microsoft Text Analytics API	extract key phrases, entities (e.g. people, companies, or locations), sentiment, as well as define in which among 120 supported languages their text is written.

Sr.no	API	Description of Data
6	The Sentiment Analysis API	returns results using a sentiment score from 0 (negative) to 1 (positive).
7	Google Cloud Natural Language API	sentiment from emails, text documents, news articles, social media, and blog posts.

3. Methodology

Sentiment analysis applies numerous strategies to interpret and classify emotions represented in textual data, each with distinct approaches and applications. There are multiple approaches to implement Sentiment Analysis, the most popular being[14]

Figure 2: Different Techniques of Sentiment Analysis



- Rule based/Lexicon based approach: There are a set of predefined rules which are used to determine the sentiment portrayed by a given text.
- Machine learning based systems: These are mostly classifiers that use existing data to train a model using various Machine Learning algorithms. This is used to classify text into having binary sentiments (positive, negative) or multiple classes. Example for multi-class sentiment analysis include rating the product on a scale of 1 to 5.

- Deep Learning-Based Approaches: Leveraging neural networks, particularly architectures like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), these methods automatically learn feature representations from large datasets. They are adept at capturing complex patterns and contextual nuances in text, leading to improved performance in sentiment classification tasks. However, they require substantial computational resources and large amounts of labeled data for effective training
- Hybrid Approach: Combining elements from both lexicon-based and machine learning methods, hybrid approaches aim to leverage the strengths of each. For instance, a system might use a lexicon-based method to handle general sentiment detection while employing machine learning models to fine-tune the analysis based on context-specific nuances. This combination can enhance accuracy and robustness in sentiment analysis tasks
- Multimodal Approache**s:** Beyond textual data, sentiment analysis can benefit from incorporating additional data types such as images, audio, or video. Deep learning models can be designed to process and integrate these diverse data sources, providing a more comprehensive understanding of sentiment, especially in social media contexts where posts often include multimedia elements.

4. CONCLUSION

Sentiment analysis has emerged as a potent instrument for comprehending human feelings and viewpoints, especially with the development of deep learning and machine learning. It may be used in a wide range of fields, from companies looking to improve their marketing tactics to people keeping an eye on how others perceive them. Social media mining's growth improves sentiment analysis even more by revealing insightful information about user interactions and patterns of activity. Decision-making processes, industry formation, and social dynamics will all be significantly impacted by the capacity to evaluate sentiments more accurately and contextually as technology advances.

References

1. Chinemela Queen Adougo, Ovute A.O., Obochi Charles I., “ The influence of social media on the Nigerian youths: Aba residents experience”, Quest Journals Journal of Research in Humanities and Social Service, 2015. Vol. 3, no. 3, pp.12-20.
2. Reza Zafarani, Mohammad Ali Abbasi and Huan Liu, Social Media Mining-An Introduction, 2014, A Book
3. B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, “Short text classification in twitter to improve information filtering,” In Proceeding of the 33rd internationalACM SIGIR conference on Research and development in information retrieval, 2010, pp. 841–842.
4. J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, and J. Sperling, “Twitterstand: news in tweets,” in Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2009, pp. 42–51.
5. Manisha Rani, JyotiArora, Twitter Data Predicting Geolocation Using Data Mining Techniques, International Journal of Innovative Research in Computer and Communication Engineering, June 2016, vol. 4, no. 6, pp. 10446-10453
6. Shruti Wakade, Chandra Shekar, Kathy J. Liszka and Chien-Chung Chan. Text Mining for Sentiment Analysis of Twitter Data.

7. Zhunchen, L., Miles, O., Ting, W., "An effective approach to tweets opinion retrieval", Springer Journal on World Wide Web, May 2015, vol. 18, no. 3, pp 545–566.
8. Sameesksha Shrivastava, Dr. Pramod S. Nair, "Mood Prediction on Tweets Using Classification Algorithm", International Journal of Science and Research, 2015, vol. 4, no. 11, pp. 295-299
9. Luciano Barbosa and Junlan Feng, " Robust sentiment detection on twitter from biased and noisy data", .Proceedings of the 23rd International Conference on Computational Linguistics: Posters , 2010, pp. 36–44.
10. Haoqi Li1,Brian Baucom and Panayiotis Georgiou, "Linking emotions to behaviors through deep transfer learning " ,arXiv:1910.03641v1 [cs.LG] 8 Oct 2019.
11. Muhammad Usman Tariq1 , Muhammad Babar2 *, Marc Poulin1 , Akmal Saeed Khattak3 , Mohammad Dahman Alshehri4 and Sarah Kaleem5, "Human" in Human-Artificial Intelligence Interaction", Human-Media Interaction, a section of the journal Frontiers in Psychology,: 06 July 2021
12. Oliver Sturman1,2, Lukas von Ziegler1,2, Christa Schläppi1,2, Furkan Akyol1,2, Mattia Privitera1,2, Daria Slominski1,2, Christina Grimm2,3, Laetitia Thieren2,4, Valerio Zerbi2,3, Benjamin Grewe2,5,6 and Johannes Bohacek,"Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions", Neuropsychopharmacology (2020) 45:1942 – 1952,25 July 2020
13. Bruno Degardin and Hugo Proença," Human Behavior Analysis: A Survey on Action Recognition",Appl. Sci. 2021, 11, 8324. https://doi.org/10.3390/app11188324
14. Sweta Saraff , Roman Taraban, Rishipal, Ramakrishna Biswal , Shweta Kedasand Shakuntala Gupta,"Application of Sentiment Analysis in Understanding Human Emotions and Behaviour",EAI Endorsed Transactions on Smart Cities 08 2020 - 01 2021 | Volume 5 | Issue 13 | e4