

Home › Trends



By [Kezia Grace Jungco](#) April 23, 2024

Datamation content and product recommendations are editorially independent. We may make money when you click on links to our partners. [Learn More.](#)

Sentiment analysis, also known as opinion mining, is the process of using computational techniques to extract subjective information from textual data like emails, customer reviews, or social media feeds. Sentiment analysis tools determine the emotional tone or overall sentiment expressed toward a topic, product, service, brand, or individual. Businesses and organizations can use these tools to monitor online conversations, analyze customer feedback, understand public opinion, manage brand reputation, and predict trends.

While sentiment analysis initially focused on classifying text as positive, negative, or neutral, the field has become more sophisticated, using such methods as natural language processing (NLP), machine learning (ML), and deep learning to identify emotions and emotional undertones behind text.

### Featured Partners: Data Analysis Software



[Learn More »](#)



# Understanding How Sentiment Analysis Works

Sentiment analysis involves a step-by-step process that transforms raw text—from social media content, blog posts, reviews, customer support tickets, case studies, webchats, or community forums, for example—into a quantifiable sentiment score. The process includes preprocessing the data, engineering features to help [the model](#) identify the sentiments it communicates, integrating a lexicon to quantify those sentiments, using [machine learning](#) to process the data, and refining the results.

## Step 1: Preprocessing Data

Data preparation is a foundational step to ensure the quality of the sentiment analysis by cleaning and preparing text before feeding it to a machine learning model. Common data preprocessing techniques include lowercasing, which converts text to lowercase for consistency; tokenization, which breaks down text into words or phrases (tokens); stop word removal, which eliminates unnecessary words; and normalization, which corrects spelling errors and handles slang or abbreviations.

## Step 2: Engineering Features

The next step is to establish features to help the model identify sentiments. This process involves the creation, transformation, extraction, and selection of the features or variables most suitable for creating an accurate machine learning algorithm. Part-of-speech (POS) tagging is used to identify the grammatical functions of each word, as some parts of speech carry more sentiment weight, and n-grams, which analyze sequences of words that hold stronger sentiment value than individual words.

## Step 3: Integrating A Sentiment Lexicon

Lexicon-based sentiment analysis is a popular technique for extracting the emotional polarity of text. It relies on lexicons, or predefined dictionaries of words and phrases that have an associated sentiment value. These lexicons are integrated into the analysis to allow models to assign sentiment scores based on the presence of sentiment-laden words within the text. Examples of sentiment lexicons that can be integrated include [NRCLexicon](#) and [SentiWordNet](#).

## Step 4: Performing Machine Learning Analysis

In this step, machine learning algorithms are used for the actual analysis. These can include supervised or unsupervised learning methods. Supervised learning trains the model on labeled data where the text is paired with its corresponding sentiment, while unsupervised learning allows the model to identify sentiment clusters within the data without pre-labeled examples, which enables exploring emerging sentiment trends.

## Step 5: Evaluating And Refining

The final step involves evaluating the model's performance on unseen data by setting metrics to help assess how well the model identifies the sentiment. Users can refine the model through other methods, such as parameter tuning or exploring a different algorithm based on these evaluations.



## 5 Types Of Sentiment Analysis

Sentiment analysis goes beyond classifying text as positive or negative. It can be categorized in different ways based on the level of granularity and the methods used. Popular methods include polarity based, intent based, aspect-based, fine-grained, and emotion detection.

### Polarity-Based Sentiment Analysis

Polarity-based sentiment analysis determines the overall sentiment behind a text and classifies it as positive, negative, or neutral. Polarity can be expressed with a numerical rating, known as a sentiment score, between -100 and 100, with 0 representing neutral sentiment. This method can be applied for a quick assessment of overall brand sentiment across large datasets, such as social media analysis across multiple platforms.

### Intent-Based Analysis

Intent-based analysis can identify the intended action behind a text—for instance, whether a customer wants to seek information, purchase a product, or file a complaint. This type of sentiment analysis can be applied to developing chatbots for efficient conversation routing or helping marketers identify the right B2B campaign for their target audience.

### Fine-Grained Analysis

Fine-grained analysis delves deeper than classifying text as positive, negative, or neutral, breaking down sentiment indicators into more precise categories. Fine-grained analysis provides a more nuanced understanding of opinions, as it identifies why customers or respondents feel the way they do.

### Emotion Detection Analysis

Emotion detection analysis defines and evaluates specific emotions within a text, such as anger, joy, sadness, or fear. This type of sentiment analysis is ideal for businesses or brands that aim to deliver empathic customer service, as it can help them understand the emotional triggers in advertising or marketing campaigns.

### Aspect-Based Analysis

Aspect-based analysis identifies the sentiment toward a specific aspect of a product, service, or topic. This technique categorizes data by aspect and determines the sentiment attributed to each. It is usually applied for analyzing customer feedback, targeting product improvement, and identifying the strengths and weaknesses of a product or service.

## Sentiment Analysis Methodologies

Sentiment analysis uses computational techniques to determine the emotions and attitudes within textual data. Natural language processing (NLP) and machine learning (ML) are two of the major approaches that are used.



## Natural Language Processing (NLP)

NLP is a branch of artificial intelligence (AI) that combines computational linguistics with statistical and machine learning models, enabling computers to understand human language. In sentiment analysis, NLP techniques play a role in such methods as tokenization, POS tagging, lemmatization or stemming, and sentiment dictionaries.

## Machine Learning (ML)

ML is a branch of AI and computer science that uses algorithms that learn from massive amounts of data to identify patterns and make predictions. It enables AI to imitate how humans learn and has revolutionized the field of sentiment analysis in many ways. With ML, algorithms can be trained on labeled data (supervised learning) or it can identify patterns in unlabeled data (unsupervised learning). It also allows advanced neural networks to extract complex data from text through deep learning.

## Data Collection And Preparation For Sentiment Analysis

The reliability of results depends on the quality and relevance of the data being analyzed—as such, careful consideration must be given to choosing the sources and strategies of data collection. It's also important to address challenges in the data collection process accordingly and follow the best practices in processing data for sentiment analysis.

### Gathering Data For Sentiment Analysis

Data for sentiment analysis can be mined from a variety of sources for both online and offline platforms. Choosing the right data source for sentiment analysis depends on the specific goals, needs, and research questions of a business or organization. Here are the most common sources of data:



- **Social Media:** Social media platforms provide a constant source of public opinions on topics, brands, and current events that can be collected through manual searching, social listening tools, and API data retrieval.
- **Review Sites:** Textual data can be gathered from review sites like Yelp, Amazon, and Google Reviews through web scraping and APIs to provide insights into how customers feel about a company, product, or service.
- **Surveys and Focus Groups:** These examples of structured data gathering are aimed at collecting feedback on specific topics.
- **Articles and Publications:** Online news, websites, blogs, and industry publications are powerful tools for understanding public sentiment and social issues. They can be searched using news aggregators or web scraping methods.
- **Voice of Customer (VoC) Data:** This non-traditional data source includes chat histories, customer support transcripts, customer emails, and more, and can be incorporated using CRM integrations or transcript analysis tools.
- **Public Information:** Public information available on government or company websites, including press releases and financial reports, can be a rich source of useful data.
- **Employee Feedback:** Companies can collect employee feedback through internal surveys, performance reviews, and communication channels using human resources (HR) systems and feedback platforms.
- **Electronic Medical Records:** Patient notes and records within the healthcare system can be accessed through HIPAA-compliant (Health Insurance Portability and Accountability Act of 1996) platforms or CRMs. Medical records require secure access, so it's important to collaborate with healthcare providers and adhere to privacy regulations.
- **Podcasts:** Podcasts are a great source of information, such as commentaries and discussions found within podcast transcripts.
- **Gaming Platforms:** Data from in-game chats, forum posts, and reviews is also a rich source of information for sentiment analysis.

## How To Gather Data For Sentiment Analysis

Before collecting data, define your goals for what you want to learn through sentiment analysis. If you're conducting a study, determine your research questions—be as specific as possible—and identify opinions or emotions you're interested in, such as customer satisfaction, brand perception, or attitude towards a social issue.

Select the type of data suitable for your project or research and determine your data collection strategy. When gathering data online, make sure to comply with the websites' terms and conditions, and if you're interacting directly with respondents or customers, ensure that their privacy is protected and secure informed consent forms—especially if you're using surveys and focus groups.

## Best Practices For Processing Data For Sentiment Analysis

Processing raw data before conducting sentiment analysis ensures that the data is clean and ready for algorithms to interpret. While there are several methodical measures that you can take in processing data for



sentiment analysis, it still depends on your goals and the characteristics of the dataset you have.

## Data Cleaning

This process involves removing redundant, incorrect, and irrelevant data not meant for analysis, including HTML tags, hyperlinks, special characters, and other text that shouldn't be in your dataset. Methods of data cleaning and preprocessing include the following:

- **Normalization:** Normalizes the data for analysis by reducing noise and complexity and transforming it into a consistent format using processes like stemming (reducing words to their root form) or lemmatization (grouping words to form a common base).
- **Tokenization:** Breaks down textual data into tokens to make understanding the individual components of the text easier.
- **POS Tagging:** Used for more nuanced sentiment analysis, especially for detecting sarcasm or negation—NLP techniques can identify the grammatical function of each word and other categories such as tense, number, and more.
- **Negation Handling:** Techniques can automatically detect the extent of negation in a text, such as rule-based methods or trained models.
- **Stop Word Removal:** This optional method can remove stop words such as “the,” “a,” or “an” to improve efficiency in analyzing the text.

## Select Your Model

Choose a sentiment analysis model that's aligned with your objectives, size, and quality of training data, your desired level of accuracy, and the resources available to you. The most common models include the rule-based model and a machine learning model.

A rule-based model involves data labeling, which can be done manually or by using a data annotation tool. A machine learning model can be built by training a vast amount of data to analyze text to give more accurate and automated results.

## Analyze And Evaluate

In processing data for sentiment analysis, keep in mind that both rule-based and machine learning models can be improved over time. It's important to assess the results of the analysis and compare data using both models to calibrate them.

## 3 Top Sentiment Analysis Tools And Technologies

Sentiment analysis tools use AI and deep learning techniques to decode the overall sentiment of a text from various data sources. The best tools can use various statistical and knowledge techniques to analyze sentiments behind the text with accuracy and granularity. Three of the top sentiment analysis solutions on the market include IBM Watson, Azure AI Language, and Talkwalker.



## IBM Watson Natural Language Understanding

IBM Watson Natural Language Understanding (NLU) is an AI-powered solution for advanced text analytics. This platform uses deep learning to extract meaning and insights from unstructured data, supporting up to 12 languages. Users can extract metadata from texts, train models using the IBM Watson Knowledge Studio, and generate reports and recommendations in real-time.

IBM Watson NLU stands out as a sentiment analysis tool for its flexibility and customization, especially for users who are working with a massive amount of unstructured data. It's priced based on the NLU item, equivalent to one text unit or up to 10,000 characters. The standard tier starts at \$0.0002 per NLU item per month.

[Visit IBM Watson](#)

## Azure AI Language

Microsoft's Azure AI Language, formerly known as Azure Cognitive Service for Language, is a cloud-based text analytics platform with robust NLP features. This platform offers a wide range of functions, such as a built-in sentiment analysis tool, key phrase extraction, topic moderation, and more.

What sets Azure AI Language apart from other tools on the market is its capacity to support multilingual text, supporting more than 100 languages and dialects. It also offers pre-built models that are designed for multilingual tasks, so users can implement them right away and access accurate results. Azure AI Language offers free 5,000 text records per month and costs \$25 per 1,000 succeeding text records.

[Visit Microsoft Azure](#)

## Talkwalker

Talkwalker is a sentiment analysis tool designed for social media monitoring. As a leading social listening platform, it offers robust tools for analyzing brand sentiment, predicting trends, and interacting with target audiences online.

Users can leverage AI-powered sentiment analysis tools to detect negative comments or sarcasm on social media posts, forums, and images to provide companies and organizations with an in-depth understanding of their online brand perception. Talkwalker offers four pricing tiers, and potential customers can contact sales to request quotes.

[Visit Talkwalker](#)



Learn more about our picks in our [review of the best sentiment analysis tools for 2024](#).

## 5 Practical Application Examples Of Sentiment Analysis

Sentiment analysis has become a valuable tool for organizations in a wide range of industries. Companies can use it for social media monitoring, customer service management, and analysis of customer data to improve operations and drive growth.

### Customer Experience (CX)

Sentiment analysis is a valuable tool for improving customer satisfaction through brand monitoring, product evaluation, and customer support enhancement. Sentiment analysis tools can help businesses and organizations monitor what customers say about their brands, products, or services, and sentiment analysis data can help brands resolve issues, retain customers, and identify the best communication platforms to engage with leads and customers.

### Marketing And Sales

Sentiment analysis tools enable sales teams and marketers to identify a problem or opportunity and adapt strategies to meet the needs of their customer base. They can help companies follow conversations about their business and competitors on social media platforms through social listening tools. Organizations can use these tools to understand audience sentiment toward a specific topic or product and tailor marketing campaigns based on this data.

### Social Issues And Politics

Sentiment analysis tools are valuable in understanding today's social and political landscape. For instance, users can understand public opinion by tracking sentiments on social issues, political candidates, or policies and initiatives. It can also help in identifying crises in public relations and provide insights that are crucial for the decision-making process of policymakers.

### Finance

Sentiment analysis is essential for financial markets, as it helps professionals analyze news articles, social media channels, and financial reports to predict stock trends. Sentiment analysis tools can help calibrate investor sentiments toward companies that could affect stock prices.

### Healthcare

Healthcare practitioners can leverage patient sentiment data to understand their needs and support them, which is a helpful tool in advancing mental health research. Sentiment analysis also enables service providers to analyze patient feedback to improve their satisfaction and overall experience.





# Advanced Topics In Sentiment Analysis

Sentiment analysis is a complex field and has played a pivotal role in the realm of [data analytics](#). Ongoing advancements in sentiment analysis are designed for understanding and interpreting nuanced languages that are usually found in multiple languages, sarcasm, ironies, and modern communication found in multimedia data.

## Multilingual Sentiment Analysis

Analyzing sentiments across multiple languages and dialects increases the complexity of [data analysis](#). Different languages and dialects have unique vocabularies, cultural contexts, and grammatical structures that could affect how a sentiment is expressed. To understand the sentiments behind multiple languages, you can make use of AI-driven solutions or platforms that include language-specific resources and sentiment-aware models.

## Sarcasm, Irony, And Other Complexities

Despite the advancements in text analytics, algorithms still struggle to detect sarcasm and irony. Rule-based models, machine learning, and deep learning techniques can incorporate strategies for detecting sentiment inconsistencies and using real-world context for a more accurate interpretation.

## Sentiment Analysis And Multimedia Data

Multimodal sentiment analysis extracts information from multiple media sources, including images, videos, and audio. Analyzing multimodal data requires advanced techniques such as facial expression recognition, emotional tone detection, and understanding the impact between modalities.

## Challenges And Limitations Of Sentiment Analysis

Sentiment analysis is a powerful tool for businesses that want to understand their customer base, enhance sales marketing efforts, optimize social media strategies, and improve overall performance. However, sentiment analysis also has challenges and limitations.

It requires accuracy and reliability, but even the most advanced algorithms can still misinterpret sentiments. Accuracy in understanding sentiments is influenced by several factors, including subjective language, informal writing, cultural references, and industry-specific jargon. Continuous evaluation and fine-tuning of models are necessary to achieve reliable results.

Sentiment analysis should also adhere to ethical considerations, as the process involves personal opinions and private data. In conducting sentiment analysis, prioritize the respondents' privacy and observe responsible data collection processes. Identify and address potential biases in datasets by using diverse and representative data that covers different demographics, cultures, and viewpoints, or by employing re-sampling and specialized algorithms.



## Bottom Line: The Continuing Relevance Of Sentiment Analysis

Sentiment analysis can help organizations understand the emotions, attitudes, and opinions behind an ever-increasing amount of textual data. While certain challenges and limitations exist in this field, sentiment analysis is widely used for enhancing customer experience, understanding public opinion, predicting stock trends, and improving patient care.

In the future, sentiment analysis systems might employ more advanced techniques for recognizing nuanced languages and capturing sentiments more accurately. Ultimately, sentiment analysis will remain an essential tool for businesses and researchers alike to better understand their audience and stay on top of the latest trends.

**Learn more about other things you can discover through different types of analysis in our articles on key benefits of [big data analytics](#) and [statistical analysis](#).**

