

Sentiment Analysis: Comprehensive Beginners Guide

Every piece of text carries emotion—whether it's a glowing review, a frustrated tweet, or a neutral survey response. **Sentiment analysis** helps businesses make sense of this emotional tone by automatically determining whether a piece of text is positive, negative, or neutral.

This technology is widely used across industries for analyzing customer feedback, survey responses, and product reviews. It plays a crucial role in social media monitoring, reputation management, and customer experience improvement. For example, a company analyzing thousands of product reviews can quickly uncover insights about pricing, product features, and overall satisfaction—helping them make data-driven decisions faster.

In this guide, we'll break down **how sentiment analysis works**, explore key business use cases, and discuss the challenges and limitations of different approaches. Whether you're new to sentiment analysis or looking to refine your understanding, this guide will give you the foundation you need to get started.



Preferences

What is Sentiment Analysis?

Sentiment analysis is a technique used to determine the emotional tone behind a piece of text—whether it conveys a positive, negative, or neutral sentiment. Also known as **opinion mining** or **emotion AI**, this method is widely used to analyze text data from customer reviews, surveys, and social media conversations.

At its core, **sentiment analysis explained** means a process that uses **text analytics** and AI to interpret human emotions from written language.

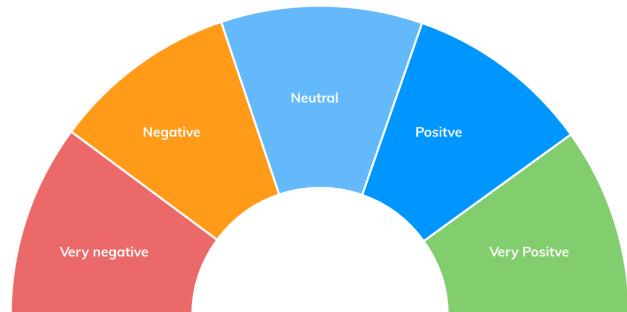
Traditionally, this was done using **natural language processing (NLP)** and **machine learning**, which helped classify text into sentiment categories. More recently, **large language models (LLMs)** or **generative AI** like ChatGPT have revolutionized the field, offering more nuanced and accurate sentiment detection by understanding **context, intent, and tone** at a much deeper level.

Sentiment Scoring

A key aspect of sentiment analysis is polarity classification. Polarity refers to the overall sentiment conveyed by a particular text, phrase or word. This polarity can be expressed as a numerical rating known as a “sentiment score”. For example, this score can be a number between -100 and 100 with 0 representing neutral sentiment. This score could be calculated for an entire text or just for an individual phrase.

Fine-grained Sentiment Analysis

Sentiment scoring can be as fine-grained as required for a specific use case. Categories can expand beyond just “positive”, “neutral” and “negative”. For example, you may choose to use five categories.



As a rule of thumb, things that customer might mention as dislikes in their 1-star reviews are “very negative”. And conversely, positive things in 5-star reviews could be tagged as “very positive”.

You can also refine the sentiment further into specific emotions. For example, positive sentiment can be further refined into happy, excited, impressed, trusting and so on. This is typically done using emotion analysis, which we've covered in one of [our previous articles](#).

Aspect-based Sentiment Analysis (ABSA)

Sentiment analysis is most useful, when it's tied to a specific attribute or a feature described in text. The process of discovery of these attributes or

features and their sentiment is called Aspect-based Sentiment Analysis, or ABSA. Here at Thematic we call these aspects "**themes**". For example, for product reviews of a laptop you might be interested in *processor speed*. An aspect-based algorithm can be used to determine whether a sentence is negative, positive or neutral when it talks about *processor speed*.

Sentiment analysis can be more granular too. For example, Thematic might discover a theme *processor speed is slow* or *processor speed is below competitors*.

Contents

What is Sentiment Analysis?

Benefits Of Sentiment Analysis

Business Applications For Sentiment Analysis

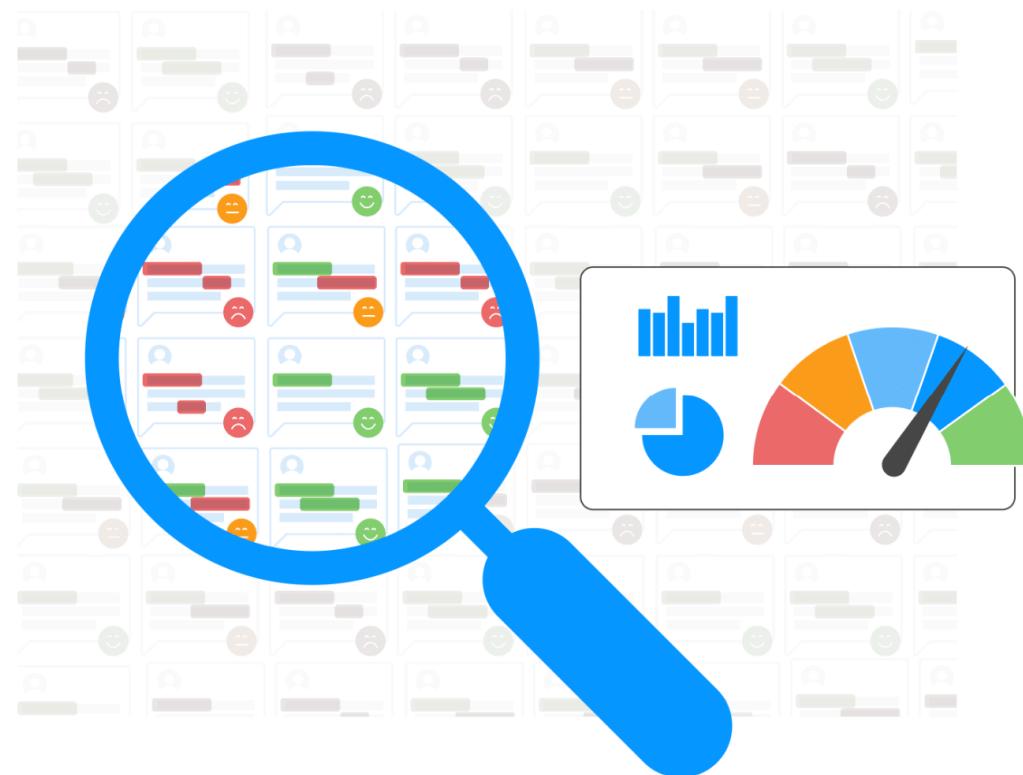
Sentiment Analysis Case Study

How Does Sentiment Analysis Work?

What Are The Current Challenges For Sentiment Analysis?

Using Thematic For Powerful Sentiment Analysis Insights

Final Thoughts On Sentiment Analysis



[From Ignored to Indispensable: The Proven, Step-by-Step CX Insights Mastery Blueprint in < 6 Days! Get free access>](#)

Platform ▾ Solutions ▾ Case Studies Pricing Resources ▾ Book a demo Login

Traditionally, ASBA was performed using Machine Learning. Today, the term "ASBA" isn't used as often, because Large Language Models (LLMs) have replaced the need for specific ASBA models.

Machine Learning is an area of AI that teaches computers to perform tasks by looking at data. Machine Learning algorithms are programmed to discover patterns in data. Traditionally, this task was done by creating training data with sentiment labels. It was difficult to generate such data, but the more data you gathered, the more accurate was your model. The algorithm itself also played a role.

Today, to perform sentiment analysis using an LLM, you need to design a prompt. The more specific and clear the prompt, the more accurate is the analysis.

Preferences

The most effective **sentiment analysis solutions** are able to handle complex negation, can capture sentiment expressed in many ways and produce consistent results each time. For example, “slow to load” or “speed issues” would both contribute to a negative sentiment for the “processor speed” aspect of the laptop. Another requirement for sentiment analysis solution is speed, to ensure sentiment analysis can be picked up in real time.



Using ABSA in Real Time

Companies use AI-based solutions to apply aspect-based sentiment analysis across their social media, review sites, online communities and internal customer communication channels. The results of the analysis can then be explored in data visualizations to identify areas for improvement. These visualizations could include overall sentiment, sentiment over time, and sentiment by rating for a particular dataset.

Sentiment analysis can be especially useful for real-time monitoring. Businesses can immediately act on issues, before they explode into social media storms and lead to PR disasters.

Benefits Of Sentiment Analysis

Let's dig deeper into the key benefits of sentiment analysis.

More trustworthy

Removes human bias through consistent analysis

Sentiment can be highly subjective. As humans we use tone, context, and language to convey meaning. How we understand that meaning depends on our own experiences and unconscious biases. To explore this further, let's look at a customer review about a new SaaS product:

“Gets the job done, but it's not cheap!”

There is both negative and positive sentiment in this sentence. Negative sentiment is linked to the price. Positive sentiment is linked to the functionality of the product. But what's the overall sentiment of the sentence?

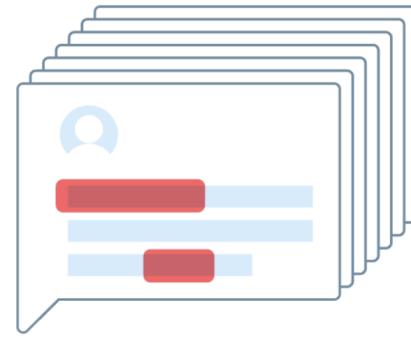
This is where human bias and error can creep in. Human analysts might regard this sentence as positive overall since the reviewer mentions functionality in a positive sentiment. On the other hand, they may focus on the negative comment on price and tag it as negative. This is just one example of how subjectivity can influence sentiment perception.

Sentiment analysis solutions apply consistent criteria to generate more accurate insights. For example, a machine learning model can be trained to recognise that there are two aspects with two different sentiments. It would average the overall sentiment as neutral, but also keep track of the details.

More powerful

Processes data at speed

Sentiment analysis helps businesses make sense of huge quantities of unstructured data. When you work with text, even 50 examples already can feel like Big Data. Especially, when you deal with people's opinions in product reviews or on social media.

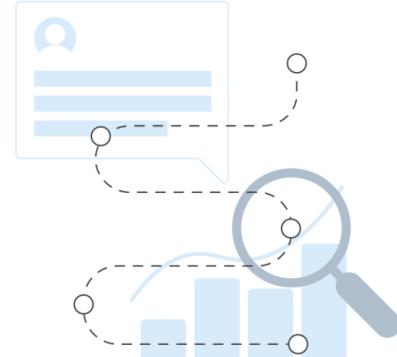


Take the example of a company who has recently launched a new product. Rather than trawling through hundreds of reviews the company can feed the data into a feedback management solution. Its sentiment analysis model will classify incoming feedback according to sentiment. The company can understand what customers think of their new product faster and act accordingly. They can uncover features that customers like as well as areas for improvement.

This type of analysis also gives companies an idea of how many customers feel a certain way about their product. The number of people and the overall polarity of the sentiment about, let's say "online documentation", can inform a company's priorities. For example, they could focus on creating better documentation to avoid customer churn and stay competitive.

Save time with automation!

Sentiment analysis algorithms can analyze hundreds of megabytes of text in minutes. Instead of manually analyzing data in spreadsheets, you can now spend your time on more valuable activities. For example, you can validate the insight: Is this something worth acting on? You can add business context too. If there is an issue, is it seasonal? Have we seen this in other parts of the business? Ultimately, sentiment analysis just provides a signal. But if you get this signal fast and with low effort, you will have time to create the right strategy.

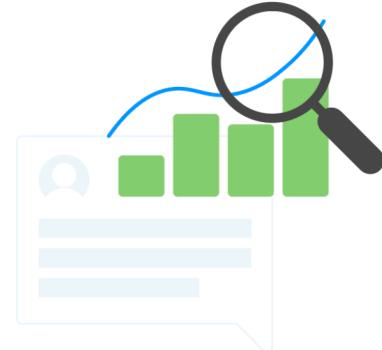


Sentiment analysis algorithms and approaches are continually getting better. They are improved by feeding better quality and more varied training data. Researchers also invent new algorithms that can use this data more effectively. At Thematic, we monitor your results and assess errors. If required, we add more specific training data in areas that need improvement. As a result,

Sentiment analysis is becoming more accurate and delivers more specific insights.

Act faster with real-time analysis and insights

Sentiment analysis is automated using Machine Learning. This means that businesses can get insights in real-time. This can be very helpful when identifying issues that need to be addressed right away. For example, a negative story trending on social media can be picked up in real-time and dealt with quickly. If one customer complains about an account issue, others might have the same problem. By instantly alerting the right teams to fix this issue, companies can prevent bad experiences from happening.



Business Applications For Sentiment Analysis

Sentiment analysis is useful for **making sense of qualitative data** that companies continuously gather through various channels. Let's dig into some of the most common business applications.

Voice of Customer (VoC) Programs

Understanding how your customers feel about your brand or your products is essential. This information can **help you improve the customer experience or identify and fix problems** with your products or services. To do this, as a business, you need to collect data from customers about their experiences with and expectations for your products or services. This feedback is known as **Voice of the Customer** (VoC).

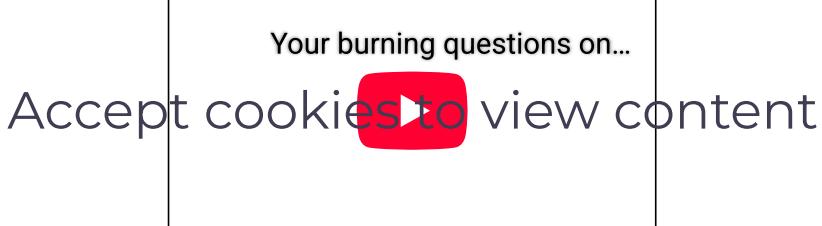
Net Promoter Score (NPS) surveys are a common way to assess how customers feel. Customers are usually asked, "How likely are you to recommend us to a friend?" The feedback is usually expressed as a number on a scale of 1 to 10. Customers who respond with a score of 10 are known as "promoters". They're the most likely to recommend the business to a friend or family member. High NPS means better customer retention. More promoters also means better word-of-mouth advertising. This means that you need to spend less on paid customer acquisition.

A drawback of NPS surveys is they don't give you much information about why your customers really feel a certain way. Open-ended questions supplement the NPS rating questions. They capture why customers are likely or unlikely to recommend products and services. Sentiment analysis turns this text into the **drivers of NPS**.

NPS is just one of the VoC survey types. The same idea applies to any metric that you might care about: Customer Effort Score, Customer Satisfaction etc. It really doesn't matter that much what metric is used. What's driving the ups and downs of the metric is more important.

A great VOC program includes listening to **customer feedback** across all channels. You can imagine how it can quickly explode to hundreds and thousands of pieces of feedback even for a mid-size B2B company. Sentiment analysis is critical to make sense of this data.

Finally, companies can also quickly identify customers reporting strongly negative experiences and rectify urgent issues. Tracking your customers' sentiment over time can help you identify and address emerging issues before they become bigger problems.



Customer Service Experience

A great customer service experience can make or break a company. Customers want to know that their query will be dealt with quickly, efficiently, and professionally. Sentiment analysis can help companies streamline and enhance their customer service experience.

Sentiment analysis and text analysis can both be applied to customer support conversations. Machine Learning algorithms can automatically rank conversations by urgency and topic. For example, let's say you have a community where people report technical issues. A sentiment analysis algorithm can find those posts where people are particularly frustrated. These queries can be prioritized for an in-house specialist. Regular questions can be answered by other community members.

As you can see, sentiment analysis can reduce processing times and increase efficiency by directing queries to the right people. Ultimately, customers get a better support experience and you can reduce churn rates.

Product Experience

Sentiment analysis can identify how your customers feel about the features and benefits of your products. This can help **uncover areas for improvement** that you may not have been aware of.

For example, you could mine online product reviews for feedback on a specific product category across all competitors in this market. You can then apply sentiment analysis to reveal topics that your customers feel negatively about. This could reveal opportunities or common issues.

For example, when we analyzed sentiment of US banking app reviews we found that the most important feature was mobile check deposit. Interestingly, most apps had issues with this feature. Companies that have the least complaints for this feature could use such an insight in their marketing messaging.

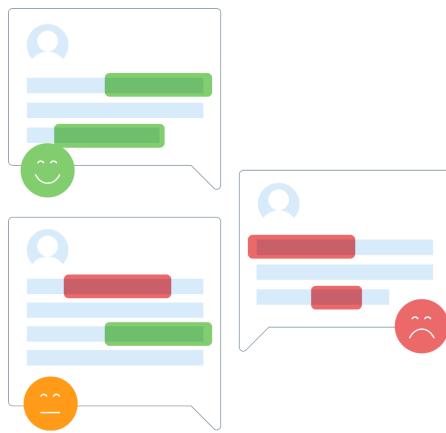
Product managers can iterate on improving the feature. They can then use sentiment analysis to monitor if customers are seeing improvements in functionality and reliability of the check deposit.

Brand Sentiment Analysis

How customers feel about a brand can impact sales, churn rates, and how likely they are to recommend this brand to others. In 2004 the “Super Size” documentary was released documenting a 30-day period when filmmaker Morgan Spurlock only ate McDonald’s food. The ensuing media storm combined with other negative publicity caused the company’s profits in the UK to fall to the **lowest levels in 30 years**. The company responded by launching a PR campaign to improve their public image.

Sentiment analysis can help brands monitor how their customers feel about them. They can analyze communities, forums and social media platforms to keep an eye on their brand reputation. Or they can conduct surveys to understand what issues their customers feel strongly about.

Companies also track their brand, product names and competitor mentions to build up an understanding of brand image over time. This helps companies assess how a PR campaign or a new product launch have impacted overall brand sentiment.



Social media sentiment analysis

Social media is a powerful way to reach new customers and engage with existing ones. Good customer reviews and posts on social media encourage other customers to buy from your company. But the reverse is also true. Negative social media posts or reviews can be very costly to your business.

Research by Convergys Corp. showed that a negative review on YouTube, Twitter or

Facebook can cost a company about 30 customers. Negative social media posts about a company can also cause big financial losses. One memorable example is **Elon Musk's 2020 tweet** which claimed the Tesla stock price was too high.

The viral tweet wiped \$14 billion off Tesla's valuation in a matter of hours. Sentiment analysis can **help identify these types of issues in real-time** before they escalate. Businesses can then respond quickly to mitigate any damage to their brand reputation and limit financial cost.

Market research

Sentiment analysis can help companies identify emerging trends, analyze competitors, and probe new markets. Companies may want to analyze reviews on competitors' products or services. Applying sentiment analysis to this data can identify what customers like or dislike about their competitors' products. These insights might reveal how to gain a competitive edge. For example, sentiment analysis could reveal that competitors' customers are unhappy about the poor battery life of their laptop. The company could then highlight their superior battery life in their marketing messaging.

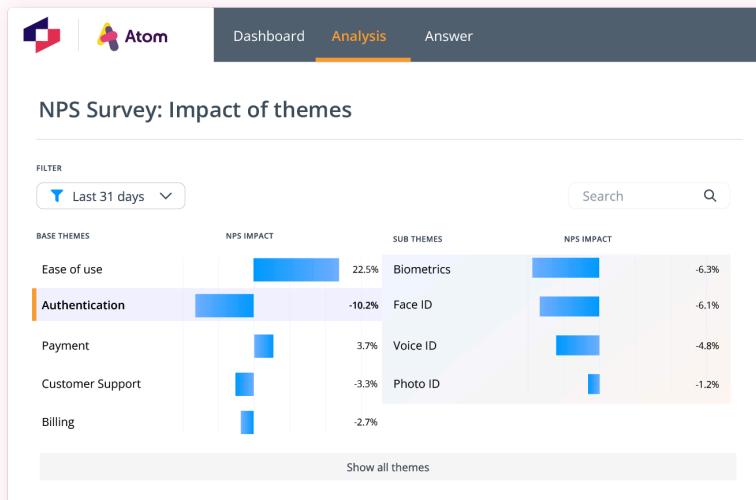
Sentiment analysis could also be applied to market reports and business journals to pinpoint new opportunities. For example, analyzing industry data on the real estate market could reveal a particular area is increasingly being mentioned in a positive light. This information might suggest that industry insiders see this area as a good investment opportunity. These insights could then be used to gain an early advantage by investing ahead of the rest of the market.

Sentiment Analysis Case Study

Atom Bank Customer Feedback

Atom bank is a newcomer to the banking scene that set out to disrupt the industry. They take customer feedback seriously. These insights are used to continuously improve their digital customer experiences.

Atom bank's VoC programme includes a diverse range of feedback channels. They ran regular surveys, focus groups and engaged in online communities. This gave them A LOT of unstructured and structured data.



Working with Thematic, Atom bank transformed their banking experience. As you can see above, combining thematic and sentiment analysis identified what mattered most to their customers. Some themes

such as “authentication” were associated with negative sentiment in Atom bank customer feedback. Other themes like “ease of use” were associated with positive sentiment.

Sentiment analysis also helped to identify specific issues like “face recognition not working”. Atom bank then used these insights to rectify these issues.

With all these **customer sentiment insights**, the team could prioritize the app features they knew would have the most impact. These improvements made Atom bank the highest rated bank according to Trustpilot. They also now have an App Store Rating of 4.7/5. And contact centre failure demand reduced by 30%!

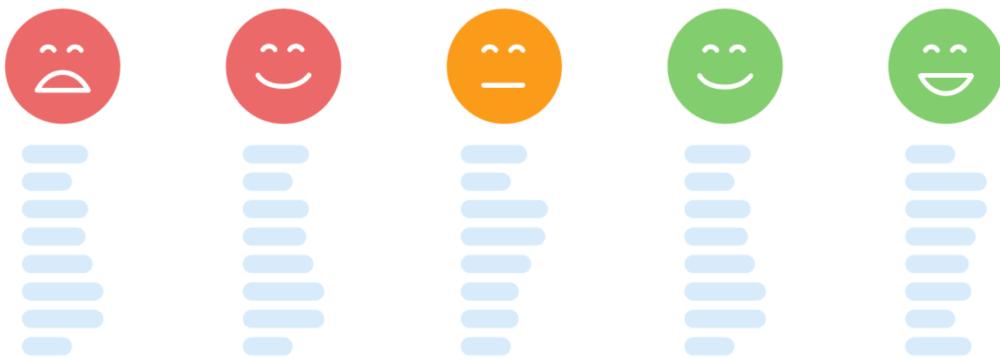


How Does Sentiment Analysis Work?

Sentiment analysis can be implemented using a variety of methods. Below, we'll review rule-based, a traditional (supervised) Machine Learning method and performing sentiment analysis using a Large Language Model like ChatGPT.

Rule-based Sentiment Analysis

This is the traditional way to do sentiment analysis based on a set of manually-created rules. This approach includes NLP techniques like lexicons (lists of words), stemming, tokenization and parsing.



Rule-based sentiment analysis works like this:

1. "Lexicons" or lists of positive and negative words are created. These are words that are used to describe sentiment. For example, positive lexicons might include "fast", "affordable", and "user-friendly". Negative lexicons could include "slow", "pricey", and "complicated".
2. Before text can be analyzed it needs to be prepared. Several processes are used to format the text in a way that a machine can understand. Tokenization breaks up text into small chunks called tokens. Sentence tokenization splits up text into sentences. Word tokenization separates words in a sentence. For example, "the best customer service" would be split into "the", "best", and "customer service". Lemmatization can be used to transforms words back to their root form. A lemma is the root form of a word. For example, the root form of "is, are, am, were, and been" is "be". We also want to exclude things which are known but are not useful for sentiment analysis. So another important process is stopword removal which takes out common words like "for, at, a, to". These words have little or no semantic value in the sentence. Applying these processes makes it easier for computers to understand the text.
3. A computer counts the number of positive or negative words in a particular text. A special rule can make sure that negated words, e.g. "not easy", are counted as opposites.
4. The final step is to calculate the overall sentiment score for the text. As mentioned previously, this could be based on a scale of -100 to 100. In this case a score of 100 would be the highest score possible for positive sentiment. A score of 0 would indicate neutral sentiment. The score can also be expressed as a percentage, ranging from 0% as negative and 100% as positive.

Disadvantages of Rule-based Sentiment Analysis

	Guessed sentiment based on "great"	Actual sentiment
My coffee was great	Positive	Positive
My coffee was awful	Negative	Negative
My coffee was not great	Positive	Negative
My coffee was not that great	Positive	Neutral?
I did not think my coffee was great	Positive	Negative
I did not expect my coffee to be this great	Positive	Positive
I was disappointed with the quality of the coffee	Negative	Negative
I wasn't disappointed with the coffee quality	Negative	Positive

Example of rule-based sentiment analysis failures

Rule-based approaches are limited because they don't consider the sentence as whole. The complexity of human language means that it's easy to miss complex negation and metaphors. Rule-based systems also tend to require regular updates to optimize their performance.

Machine Learning-based Sentiment Analysis

Traditional machine learning (ML) techniques require training a model on labeled data. This is also called supervised learning. The success of this approach depends on the quality of the training data set and the algorithm.

Step 1: Feature Extraction

Before the model can classify text, the text needs to be prepared so it can be read by a computer. Tokenization, lemmatization and stopword removal can be part of this process, similarly to rule-based approaches. In addition, text is transformed into numbers using a process called vectorization. These numeric representations are known as "features". A common way to do this is to use the bag of words or bag-of-ngrams methods. These vectorize text according to the number of times words appear.

In Deep Learning, a predecessor of LLMs, text vectorization is performed using a neural network model. The neural network can be taught to learn word associations from large quantities of text. One example is a word embeddings model which represents each distinct word as a vector. The advantage of this approach is that words with similar meanings are given similar numeric representations. This can help to improve the accuracy of sentiment analysis.

Step 2: Training & Prediction

In the next stage, the algorithm is fed a sentiment-labelled training set. The model then learns to associate input data with the most appropriate corresponding label. For example, this input data would include pairs of features (or numeric representations of text) and their corresponding positive, negative or neutral label. The training data can be either created manually or generated from reviews themselves.

Step 3: Predictions

The final stage is where ML sentiment analysis has the greatest advantage over rule-based approaches. New text is fed into the model. The model then predicts labels (also called classes or tags) for this unseen data using the model learned from the training data. The data can thus be labelled as positive, negative or neutral in sentiment. This eliminates the need for a pre-defined lexicon used in rule-based sentiment analysis.

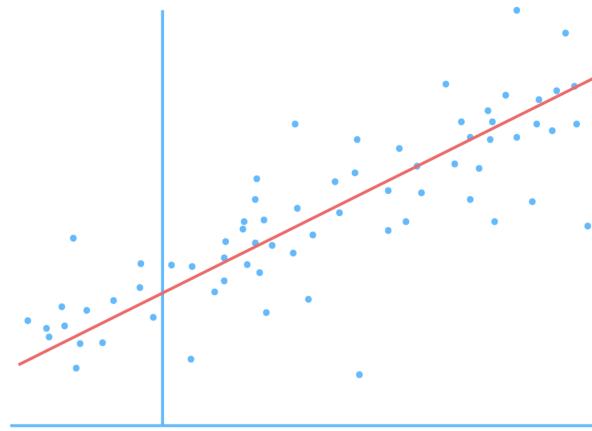
Classification algorithms

Classification algorithms are used to predict the sentiment of a particular text. As detailed in the steps above, they are trained using pre-labelled training data. Classification models commonly use Naive Bayes, Logistic Regression, Support Vector Machines, Linear Regression, and Deep Learning. Let's explore these algorithms in a bit more detail.

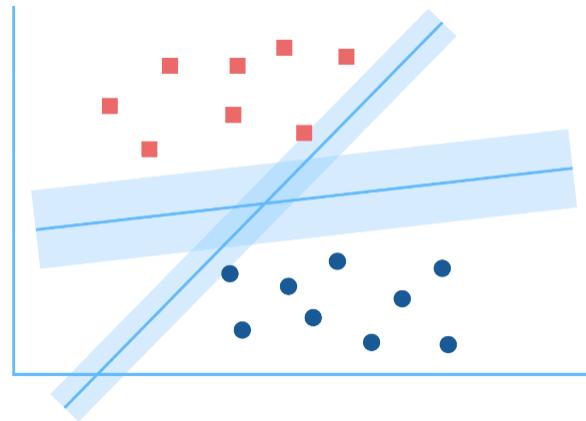
Naive Bayes: this type of classification is based on Bayes' Theorem. These are probabilistic algorithms meaning they calculate the probability of a label for a particular text. The text is then labelled with the highest probability label. "Naive" refers to the fundamental assumption that each feature is independent. Individual words make an independent and equal contribution to the overall outcome. This assumption can help this algorithm work well even where there is limited or mislabelled data.

Logistic Regression: a classification algorithm that predicts a binary outcome based on independent variables. It uses the sigmoid function which outputs a probability between 0 and 1. Words and phrases can be either classified as positive or negative. For example, "super slow processing speed" would be classified as 0 or negative.

Linear Regression: algorithm that predicts polarity (Y output) based on words and phrases (X input). The objective is to learn a linear model or line which can be used to predict sentiment (Y). Accuracy of the model can be improved by reducing the error.

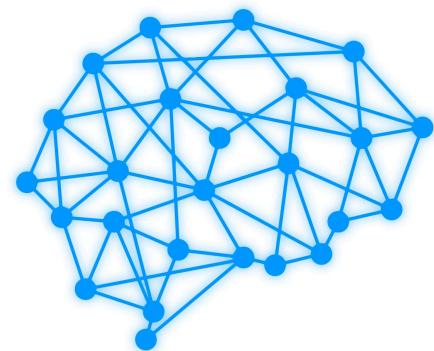


Example of simple linear regression.



Support Vector Machines: a model that plots labelled data as points in a multi-dimensional space. The hyperplane or decision boundary is a line which divides the data points. In the example below, anything to the left of the hyperplane would be classified as negative. And everything to the right would be classified as positive. The best hyperplane is one where the distance to the nearest data point of each tag is the largest. Support vectors are those data points which are closer to the hyperplane. They influence its position and orientation. These are the points which help to build the support vector machine.

Deep Learning: here, an artificial neural network performs multiple layers of processing. Deep learning is a diverse set of algorithms that imitate human brain learning through associations and abstractions. Deep learning has significant advantages over traditional classification algorithms. These neural networks can understand context, and even the mood of the writer.



Deep Learning & Sentiment Analysis

Traditional ML techniques require manual work to define classification features. They also often fail to consider the impact of word order. Deep learning and artificial neural networks have transformed this approach.

Deep learning algorithms were inspired by the structure and function of the human brain. This approach led to an increase in the accuracy and efficiency of sentiment analysis. In deep learning the neural network can learn to correct itself when it makes an error. With traditional machine learning errors need to be fixed via human intervention.

Long Short-Term Memory

One important Deep Learning approach is the Long Short-Term Memory or LSTM. This approach reads text sequentially and stores information relevant to the task.

The LSTM consists of three parts which are known as “gates”:

- Forget Gate: This first part decides whether previous data is to be remembered. If it is irrelevant to the task, it can be forgotten.
- Input Gate: In the second part the cell tries to learn new information from the new data.
- Output Gate: The final part is where the cell passes updated information to the next timestamp.

For sentiment analysis it's useful that there are cells within the LSTM which control what data is remembered or forgotten. Negation is crucial in accurate sentiment analysis. For example, it's obvious to any human that there's a big difference between “great” and “not great”. An LSTM is capable of learning that this distinction is important and can predict which words should be negated. The LSTM can also infer grammar rules by reading large amounts of text.

Transformer models

LSTMs have their limitations especially when it comes to long sentences. The model can often forget the content of distant words. And the sentence has to be processed word by word.

An alternative solution is to use a transformer. This model differentially weights the significance of each part of the data. Unlike a LTSM, the transformer does not need to process the beginning of the sentence before the end. Instead it identifies the context that confers meaning to each word. This is known as an attention mechanism. Transformers have now largely replaced LTSMs as they're better at analyzing longer sentences.

Large Language Models and Sentiment Analysis

Machine Learning approaches described above were focused on creating a single pre-trained model for the sentiment analysis task. However, with the rise of Large Language Models, the way we analyze language data has changed.

LLMs superseded all previous research by creating a universal model that can be applied to many language understanding and language generation tasks. Now, in order to perform sentiment analysis, you can use an off-the-shelf model and write a prompt describing the task you want it to do.

For example, you could ask "What's the theme and sentiment of the sentence below explained in simple terms" or "What's the sentiment of each sentence below? Say Positive, Neutral or Negative".

Open AI's GPT-4 - Testing sentiment

What's the sentiment of each sentence below? Say Positive, Neutral or Negative

My coffee was great - Positive
My coffee was awful - Negative
My coffee was not great - Negative
My coffee was not that great - Negative
I did not think my coffee was great - Negative
I did not expect my coffee to be this great - Positive
I was disappointed with the quality of the coffee - Negative
I wasn't disappointed with the quality of the coffee - Positive

Rule-based accuracy 30%

LLM accuracy 100%

Testing ChatGPT on tricky sentiment examples.

The good news is that sentiment analysis is now 100% accurate. The bad news, it requires careful prompt engineering, it's slower and it's costly.

If your prompt is too vague or if you provide too much data in one go, you are likely to get inconsistent results. But it provides a quick solution to analyze sentiment and themes. AI and text analytics solutions specializing in feedback analysis, like Thematic, come in with pre-defined prompts that are optimized for accuracy, speed and consistency. But if you use a generic LLM, via Gemini, ChatGPT or CoPilot, make sure you review the data. We wrote a [guide on how to perform sentiment analysis using ChatGPT](#).

What Are The Current Challenges For Sentiment Analysis?

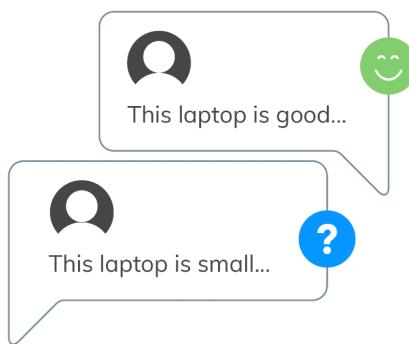
Sentiment analysis uses AI methods to identify whether a text is negative, positive, or neutral. Depending on the approach you use, there are various challenges listed below.

Subjectivity

Preferences

Texts can be objective or subjective.

Consider the following sentences as an example:



The first sentence (*This laptop is good*) is clearly subjective and most people would say that the sentiment is positive. The second sentence (*This laptop is small*) is objective and would be classified as neutral. In this “good” is considered more subjective than “small”.

The challenge here is that machines often struggle with subjectivity. Let's take the example of a product review which says “the software works great, but no way that justifies the massive price-tag”. In this case the first half of the sentence is positive. But it's negated by the second half which says it's too expensive. The overall sentiment of the sentence is negative.

Large training datasets that include lots of examples of subjectivity can help algorithms to classify sentiment correctly. Alternatively, you can craft a prompt that explains your preferred way of handling these cases.

Context

Context is crucial when it comes to understanding sentiment. Opinion words can change their polarity depending on the context. Machines need to learn about context in order to correctly classify a text.

For example, the question “what did you like about our product” could produce the following answers:

“Versatility”

“Features”

The first answer would be classified as positive. The second answer is also positive, but on its own it is ambiguous. If we changed the question to “what did you not like”, the polarity would be completely reversed. Sometimes, it's not the question but the rating that provides the context.

The solution to this is to preprocess or post-process the data to capture the necessary context. This can be a complex and lengthy process.

May I say how considerate it is of you to enable your passengers such an in-depth and thorough tour of the Caribbean. Most other airlines I have travelled on would simply wish to take me from point A to B in rather a hurry. I was intrigued that we were allowed to stop at not a lowly one or two but a magnificent six airports yesterday.

Irony & Sarcasm

Humor and sarcasm can present big challenges for machine learning techniques! Take the real life example of a complaint letter sent to LIAT Caribbean Airlines by passenger Arthur Hicks:

With irony and sarcasm people use positive words to describe negative experiences. It can be tough for AI to understand the sentiment here without knowledge of what people expect from airlines. In the example above words like 'considerate' and "magnificent" would be classified as positive in sentiment. But for a human it's obvious that the overall sentiment is negative.

Luckily, in a business context only a very **small percentage of reviews use sarcasm**.

Comparisons

Comparison is another potential stumbling block to correct sentiment classification. Consider these example online reviews:

In the first case ("*the best product on the market*") it's obvious sentiment is positive. The second one ("*much better than alternatives*") is trickier since they rely on comparisons. Without knowing what the product is being compared to, it's hard to know if these are positive, negative or neutral. In the second sentence it depends on the "*alternatives*". If the person considers the other products they've used to be very poor, this sentence could be less positive than it seems at face value.



The best product on the market



Much better than alternatives."

Speaking about Competitors:

If you are company X and your competitor is company Y, it is impossible to have one sentiment model that captures positive sentiment about Y as negative sentiment about X. Let's say you get these comments:

- I love the service that I get from company X
- I love the service that I get from company Y

A general model can only say both are positive. If you want to say that a comment speaking highly of your competitor is negative, then you need to train a custom model.

Emojis

Emojis can require extensive preprocessing especially when using data sources like social media platforms. There are two key types of emojis, Western emojis and Eastern emojis. Western emojis use only a couple of characters, such as :).



Eastern emojis use more characters in a vertical combination, such as `_(^\\)_/` which means something like “smiley sideways shrug” in Japan.

Idioms

Machine Learning algorithms struggle with idioms and phrases. An example is “not my cup of tea”. This would potentially confuse the algorithm. If a reviewer uses an idiom in **product feedback** it could be ignored or incorrectly classified by the algorithm. The solution is to include idioms in the training data so the algorithm is familiar with them.

Neutrality

For accurate sentiment analysis defining the neutral label appropriately is important. The criteria need to be consistent to generate good quality and reliable analysis. Examples of texts that should be classified as neutral include objective statements like the example we looked at above: “This laptop is black”. There are no obvious sentiments expressed in this sentence.

Irrelevant data can be classified as neutral. Another approach is to filter out any irrelevant details in the preprocessing stage.

Use of the word “wish” may indicate neutral sentiment. Consider the example, “I wish I had discovered this sooner.” However, you’ll need to be careful with this one as it can also be used to express a deficiency or problem. For example, a customer might say, “I wish the platform would update faster!” This word can express a variety of sentiments.

Negation

Negation can also create problems for traditional sentiment analysis models. For example, if a product reviewer writes “I can’t not buy another Apple Mac” they are stating a positive intention.

In the past, it was difficult to train models to recognize such nuances. Today, LLMs recognize them with ease.

User

What's the sentiment of this sentence: I can't not buy another Apple Mac

Assistant

The sentiment of the sentence "I can't not buy another Apple Mac" appears to be positive. The use of the double negative "can't not" suggests a strong inclination or desire to purchase another Apple Mac, indicating a positive sentiment towards the product.

GPT-4o acing the tricky sentiment sentence.

If using an LLM is not an option, a Long Short-Term Memory model is one of the better solutions for dealing with negation efficiently and accurately. This is because there are cells within the LSTM which control what data is remembered or forgotten. A LSTM is capable of learning to predict which words should be negated. The LSTM can “learn” these types of grammar rules by reading large amounts of text.

Audiovisual Content

Video and audio are a very different type of data to text. Audio on its own or as part of videos will need to be transcribed before the text can be analyzed using Speech-to-text algorithm. Sentiment analysis can then analyze transcribed text similarly to any other text. There are also approaches that determine sentiment from the voice intonation itself, detecting angry voices or sounds people make when they are frustrated. These techniques can also be applied to podcasts and other audio recordings.

Limitations Of Human Annotator Accuracy

As we mentioned above, even humans struggle to identify sentiment correctly. This can be measured using an inter-annotator agreement, also called **consistency**, to assess how well two or more human annotators make the same annotation decision. Since machines learn from training data, these potential errors can impact on the performance of a ML model for sentiment analysis.

Based on a recent test, Thematic's sentiment analysis correctly predicts sentiment in text data 96% of the time. But we also talked extensively about **the meaning of accuracy** and how one should take any reports of accuracy with a grain of salt.

That said, when it comes to aspect based sentiment analysis (ABSA), as defined earlier, we did run a study where we compared aspects discovered by 4 people

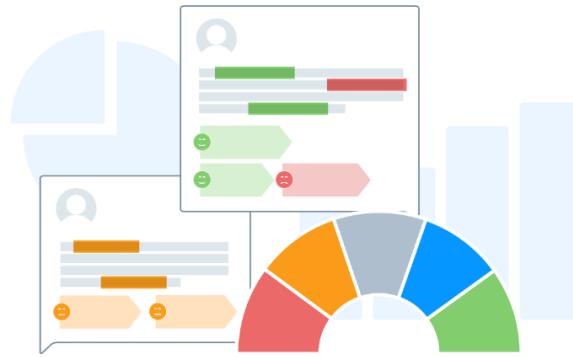
Using Thematic For Powerful Sentiment Analysis Insights

For many businesses the most efficient option is to purchase a SaaS solution that has sentiment analysis built in. Thematic is a great option that makes it easy to perform sentiment analysis on your customer feedback or other types of text.

Thematic uses **sentiment analysis algorithms** that are trained on large volumes of data using machine learning. A unique feature of Thematic is that it combines sentiment with themes discovered during the thematic analysis process.

Thematic Analysis Vs. Sentiment Analysis

Before we dig into the benefits of combining **sentiment analysis and thematic analysis**, let's quickly review these two types of analysis.



Thematic Analysis

Thematic analysis is the process of discovering repeating themes in text. A theme captures what this text is about regardless of which words and phrases express it. For example, one person could say "the food was yummy", another could say "the dishes were delicious". In both cases, it's the same theme. We could call it "tasty food".

AI researchers came up with Natural Language Understanding algorithms to automate this task. Thematic software is powered by these algorithms. You can learn more about **how it works in our blog post**.

Where does the Sentiment Analysis come in?

Preferences

We talked earlier about Aspect Based Sentiment Analysis, ABSA. Themes capture either the aspect itself, or the aspect and the sentiment of that aspect. In addition, for every theme mentioned in text, Thematic finds the relevant sentiment.

How To Use Sentiment Analysis And Thematic Analysis Together

Let's walk through how you can use sentiment analysis and thematic analysis in Thematic to get more out of your textual data.

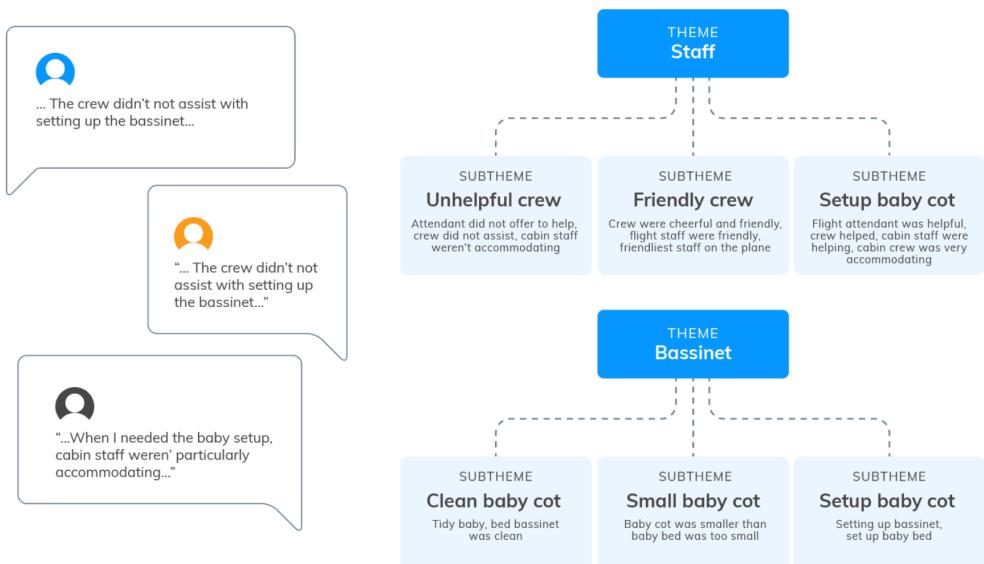
Step 1: Upload Your Data

The first step is to upload your unstructured data to a feedback analytics tool like Thematic. This could include online survey feedback, chat conversations, or social media mentions. Thematic has a wide range of one-click integrations that make it really easy to connect all your channels. These include Qualtrics, Trustpilot, Amazon, Facebook, Intercom, Twitter, Tripadvisor, and many more. Thematic then automatically cleans and prepares your data so it's ready to be analyzed.

Step 2: Analysis

Thematic Analysis

Thematic analysis can then be applied to discover themes in your unstructured data. Thematic's AI groups themes into a 2-level taxonomy. For a given text there will be core themes and related sub-themes. For example, a core theme could be "staff behavior". A sub-theme could be "friendly crew". This helps you easily identify what your customers are talking about, for example, in their reviews or survey feedback.



Preferences

Sentiment Analysis

Sentiment analysis builds on thematic analysis to help you understand the emotion behind a theme. Sentiment analysis scores each piece of text or theme and assigns positive, neutral or negative sentiment.



In the example above the theme “print boarding passes” has been selected within the Thematic dashboard. Here you can get an overview of the sentiment associated with this theme across your textual data. Overall this theme has negative sentiment with 61.2% of theme appearances classified as negative. You can also see that this theme appears in 0.4% of customer reviews.

Another option is to filter your themes by sentiment. This allows you to quickly identify the areas of your business where customers are not satisfied. You can then use these insights to drive your business strategy and make improvements.

Combining these two types of analysis can be very powerful. It allows you to understand how your customers feel about particular aspects of your products, services, or your company.

Step 3: Sentiment Analysis + Metrics

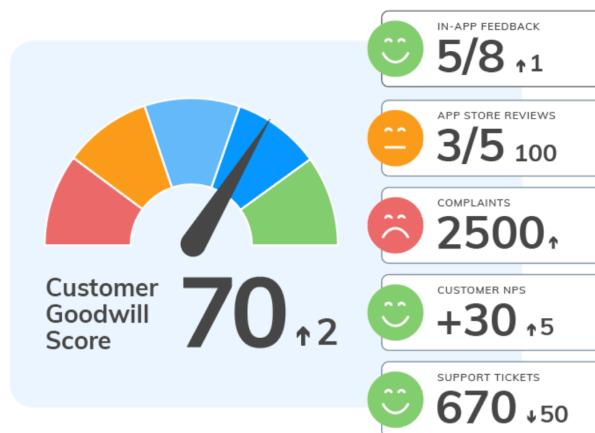
Combining Thematic and Sentiment analysis can also help you understand metrics like NPS or customer churn.

This example from the Thematic dashboard tracks customer sentiment by theme over time. You can see that the biggest negative contributor over the quarter was “bad update”. This makes it really easy for stakeholders to understand at a glance what is influencing key business metrics.



With Thematic you also have the option to use our Customer Goodwill metric. This score summarizes customer sentiment across all your uploaded data. It allows you to get an overall measure of how your customers are feeling about your company at any given time.

In the example below you can see the overall sentiment across several different channels. These channels all contribute to the Customer Goodwill score of 70.



Step 4: AI + Human

Thematic's platform also allows you to go in and make manual tweaks to the analysis. Combining the power of AI and a human analyst helps ensure greater accuracy and relevance.

For example, you may want to scan through the themes and delete any which are not useful. You also have the option to merge themes together, create new themes, and switch between themes and sub-themes.

The screenshot shows a list of customer feedback items in Thematic:

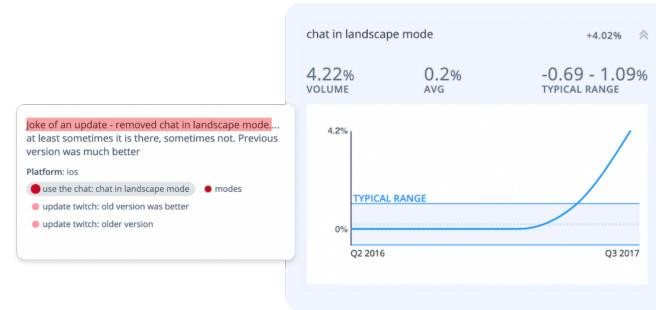
- Hard to find MERGED
- Hard to get around the store
- Not enough choice MERGED
- Not in stock
- Hard to find correct sizes
- Hard to find my size
- Range of sizes MERGED
- Correct sizes
- Larger sizes
- Size of trousers Rename

Actions available for each item include Merge & Unmerge, Delete, and Move.

Step 5: Real-Time Monitoring

The final step in the process is continual real-time monitoring. This can help you stay on top of emerging trends and rapidly identify any PR crises or product issues before they escalate.

In the example above you can see sentiment over time for the theme “chat in landscape mode”. The visualization clearly shows that more customers have been mentioning this theme in a negative sentiment over time. Looking at the customer feedback on the right indicates that this is an emerging issue related to a recent update. Using this information the business can move quickly to rectify the problem and limit possible customer churn.



Final Thoughts On Sentiment Analysis

Sentiment analysis uses AI methods to identify whether a text is negative, positive, or neutral. There are many approaches which vary in accuracy, cost, ease of setting up, speed and convenience. The rule-based approaches are now obsolete, but traditional Machine Learning solutions still have a place, when speed and cost are an issue. Large Language Models deliver the best accuracy but require careful prompt engineering, are too slow for some applications, and can get costly.

We hope this guide has given you a good overview of sentiment analysis and how you can use it in your business. Sentiment analysis can be applied to everything from brand monitoring to market research and HR. It's helping companies to glean deeper insights, become more competitive, and better understand their customers.

Sentiment analysis is also a fast-moving field that's constantly evolving and developing. That's why it's important to stay on top of the latest trends. Another option is to work with a platform like Thematic that's continually being upgraded and improved. For more information about how Thematic works you can request a personalized guided trial [right here](#).



Preferences

Product	Resources	Guides	Company
Overview	Blog	Sentiment Analysis Guide	About Us
Pricing	Community	How Thematic Analysis Software works	Science Behind Thematic
Accuracy	Webinars & eBooks	How to Analyze Surveys	Contact Us
Security & Privacy	Case Studies		Privacy Terms

Security & Compliance	Tools	How to Analyze Customer and Product Reviews	Terms of service
	NPS Calculator		
Support	Thematic vs Clarabridge	How to Code Qualitative Data	
Knowledge Base	Thematic vs TextIQ		
Integrations	Thematic vs Medallia	How to Code and Analyze Open-Ended Questions	
Thematic API	Thematic vs Chattermill		
Login	Thematic vs Kapiche		

© Copyright Thematic Analysis Inc.

Preferences