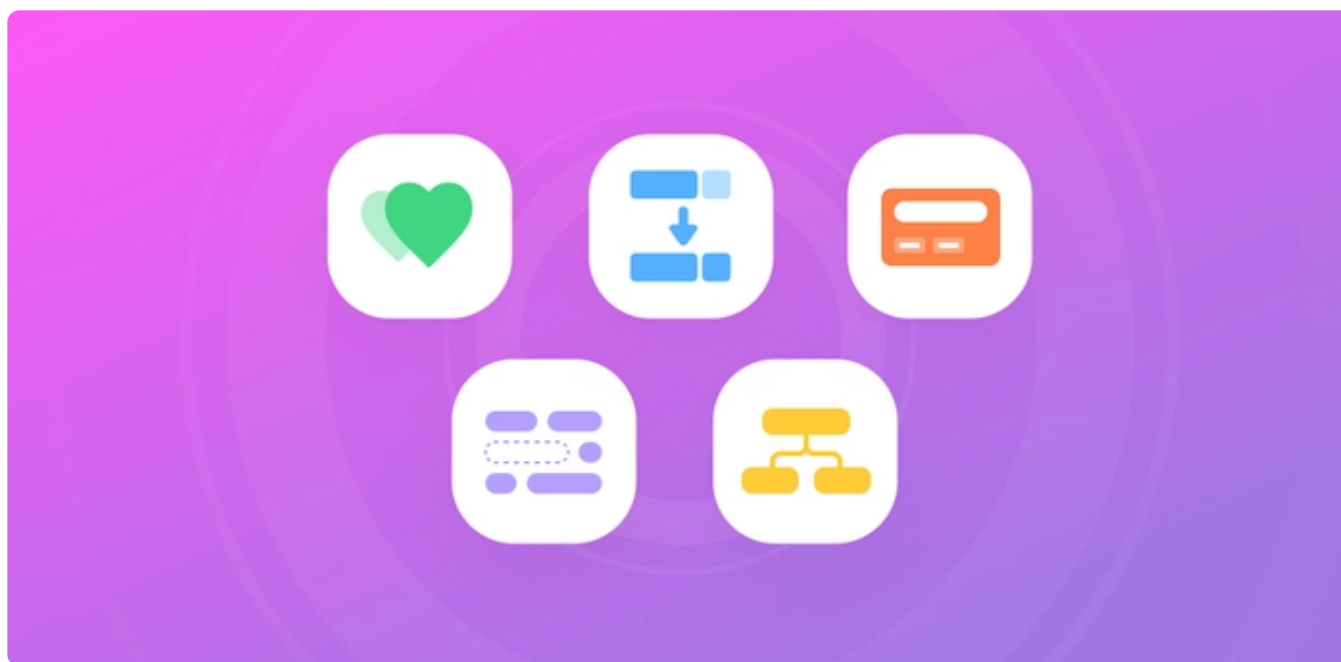**Blog**

Natural Language Processing (NLP): 7 Key Techniques

[Natural language processing](#), the deciphering of text and data by machines, has revolutionized [data analytics](#) across all industries.

Examples of NLP in action are everywhere you look. But how you use natural language processing can dictate the success or failure for your business in the demanding modern market.

By dissecting your NLP practices in the ways we'll cover in this article, you can stay on top of your practices and streamline your business.

Feel free to click through at your leisure, or jump straight to natural language processing techniques.

What is Natural Language Processing?

Natural language processing is the artificial intelligence-driven process of making human input language decipherable to software.

Why is Natural Language Processing Important?

Imagine your business software speaks a foreign language that you're not fluent in – natural language processing, or NLP, is your translator. It takes your human input, reorganizes it, and explains what you've said in a way that your software can parse.

Why should you care? Well, because communication is important and NLP software can improve how businesses operate and, as a result, customer experiences.

Let's take a look at common NLP techniques, and how you can leverage them.

Natural Language Processing Techniques

The top 7 techniques Natural Language Processing (NLP) uses to extract data from text are:

1. Sentiment Analysis

2. Named Entity Recognition

3. Summarization

4. Topic Modeling

5. Text Classification

6. Keyword Extraction

7. Lemmatization and stemming

Let's go over each, exploring how they could help your business.

1. Sentiment Analysis

Our specialty here at MonkeyLearn is sentiment analysis. This is the dissection of data (text, voice, etc) in order to determine whether it's positive, neutral, or negative.

As you can see in our classic set of examples above, it tags each statement with 'sentiment' then aggregates the sum of all the statements in a given dataset.

In this manner, sentiment analysis can transform large archives of customer feedback, reviews, or social media reactions into actionable, quantified results. These results can then be analyzed for customer insight and further strategic results.

Try out our sentiment analyzer to see how NLP works on your data.

To complement this process, MonkeyLearn's AI is programmed to link its API to existing business software and trawl through and perform sentiment analysis on data in a vast array of formats.

2. Named Entity Recognition

Named Entity Recognition, or NER (because we in the tech world are huge fans of our acronyms) is a Natural Language Processing technique that tags 'named identities' within text and extracts them for further analysis.

As you can see in the example below, NER is similar to sentiment analysis. NER, however, simply tags the identities, whether they are organization names, people, proper nouns, locations, etc., and keeps a running tally of how many times they occur within a dataset.

Ousted **WeWork** founder **Adam Neumann** lists his **Manhattan** penthouse for **$37.5 million**

[organization]   [person]   [location]   [monetary value]

How many times an identity (meaning a specific thing) crops up in customer feedback can indicate the need to fix a certain pain point. Within reviews and searches it can indicate a preference for specific kinds of products, allowing you to custom tailor each customer journey to fit the individual user, thus improving their customer experience.

The limits to NER's application are only bounded by your feedback and content teams' imaginations.

3. Text Summary

This is a fun one. Text summarization is the breakdown of jargon, whether scientific, medical, technical or other, into its most basic terms using natural language processing in order to make it more understandable.

This might seem daunting – our languages are complicated. But by applying basic noun-verb linking algorithms, text summary software can quickly synthesize complicated language to generate a concise output.

Try out text summarization by adding your own text to the model below:

Test with your own text

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus. If the symptoms of COVID-19 are mild to moderate people can recover without special treatment, but if strong symptoms are found, special treatment is needed to recover. The most common symptoms include fever, dry cough and tiredness. The virus that causes COVID-19 is usually transmitted through droplets generated when an infected person

Extract Text

Results

TAG	VALUE
	The virus that causes COVID-19 is usually transmitted through droplets generated when an infected person
SUMMARY	coughs, sneezes, or exhales.

4. Topic Modeling

Topic Modeling is an unsupervised Natural Language Processing technique that utilizes artificial intelligence programs to tag and group text clusters that share common topics.

You can think of this a similar exercise to keyword tagging, the extraction and tabulation of important words from text, except applied to topic keywords and the clusters of information associated with them

5. Text Classification

Again, text classification is the organizing of large amounts of unstructured text (meaning the raw text data you are receiving from your customers). Topic modeling, sentiment analysis, and keyword extraction (which we'll go through next) are subsets of text classification.

Text classification takes your text dataset then structures it for further analysis. It is often used to mine helpful data from customer reviews as well as customer service slogs.

6. Keyword Extraction

The final key to the text analysis puzzle, keyword extraction, is a broader form of the techniques we have already covered. By definition, keyword extraction is the automated process of extracting the most relevant information from text using AI and machine learning algorithms.

You can mold your software to search for the keywords relevant to your needs – try it out with our sample keyword extractor.

7. Lemmatization and Stemming

More technical than our other topics, lemmatization and stemming refers to the breakdown, tagging, and restructuring of text data based on either root stem or definition.

That might seem like saying the same thing twice, but both sorting processes can lend different valuable data. Discover how to make the best of both techniques in our guide to Text Cleaning for NLP.

That's a lot to tackle at once, but by understanding each process and combing through the linked tutorials, you should be well on your way to a smooth and successful NLP application.

Takeaways

Natural language processing bridges a crucial gap for all businesses between software and humans. Ensuring and investing in a sound NLP approach is a constant process, but the results will show across all of your teams, and in your bottom line.

MonkeyLearn can make that process easier with its powerful machine learning algorithm to parse your data, its easy integration, and its customizability. Sign up to MonkeyLearn to try out all the NLP techniques we mentioned above. .

Rachel Wolff
October 19th, 2021

Posts you might like...