DataRobot

BLOG   AI FOR PRACTITIONERS   INTRODUCTION TO SENTIMENT ANALYSIS: WHAT IS...

Introduction to Sentiment Analysis: What is Sentiment Analysis?

MARCH 26 BY DATAROBOT | 10 MIN READ

This article was originally published at Algorithimia's website. The company was underlined acquired by DataRobot in 2021. This article may not be entirely up-to-date or refer to products and offerings no longer in existence.

What is sentiment analysis?

Sentiment analysis is the process of using natural language processing, text analysis, and statistics to analyze customer sentiment. The best businesses understand the sentiment of their customers—what people are saying, how they're saying it, and what they mean. Customer sentiment can be found in tweets, comments, reviews, or other places where people mention your brand. Sentiment Analysis is the domain of understanding these emotions with software, and it's a must-understand for developers and business leaders in a modern workplace.

As with many other fields, advances in deep learning have brought sentiment analysis into the foreground of cutting-edge algorithms. Today we use natural language processing, statistics, and text analysis to extract, and identify the sentiment of words into positive, negative, or neutral categories.

What is sentiment analysis used for?

Sentiment analysis for brand monitoring

One of the most well documented uses of sentiment analysis is to get a full 360 view of how your brand, product, or company is viewed by your customers and stakeholders. Widely available media, like product reviews and social, can reveal key insights about what your business is doing right or wrong. Companies can also use sentiment analysis to measure the impact of a new product, ad campaign, or consumer's response to recent company news on social media. Private companies like Unamo offer this as a service.

Sentiment analysis for customer service

Customer service agents often use sentiment or intent analysis to automatically sort incoming user email into "urgent" or "not urgent" buckets based on the sentiment of the email, proactively identifying frustrated users. The agent then directs their time toward resolving the users with the most urgent needs first. As customer service becomes more and more automated through machine learning, understanding the sentiment and intent of a given case becomes increasingly important.

Sentiment analysis for market research and analysis

Sentiment analysis is used in business intelligence to understand the subjective reasons why consumers are or are not responding to something (e.x. why are consumers buying a product? What do they think of the user experience? Did customer service support meet their expectations?). Sentiment analysis can also be used in the areas of political science, sociology, and psychology to analyze trends, ideological bias, opinions, gauge reactions, etc.

A lot of these applications are already up and running. Bing recently integrated sentiment analysis into its Multi-Perspective Answers product. Hedge funds are almost certainly using the technology to predict price fluctuations based on public sentiment. And companies like CallMiner offer sentiment analysis for customer interactions as a service.

Challenges of sentiment analysis

Sentiment analysis runs into a similar set of problems as emotion recognition does – before deciding what the sentiment of a given sentence is, we need to figure out what "sentiment" is in the first place. Is it categorical, and sentiment can be split into clear buckets like happy, sad, angry, or bored? Or is it dimensional, and sentiment needs to be evaluated on some sort of bi-directional spectrum?

In addition to the definition problem, there are multiple layers of meaning in any human-generated sentence. People express opinions in complex ways; rhetorical devices like sarcasm, irony, and implied meaning can mislead sentiment analysis. The only way to really understand these devices are through context: knowing how a paragraph is started can strongly impact the sentiment of later internal sentences.

Most of the current thinking in sentiment analysis happens in a categorical framework: sentiment is analyzed as belonging to a certain bucket, to a certain degree. For example, a given sentence may be 45% happy, 23% sad, 89% excited, and 55% hopeful. These numbers don't add up to 100 – they're individual indications of how "X" a sentence's sentiment is.

To address the context issue, a lot of research surrounding sentiment analysis has focused on feature engineering. Creating inputs to a model that recognize context, tone, and previous indications of sentiment can help increase accuracy and get a better overall sense of what the author is trying to say. For an interesting example, check out this paper in Knowledge-Based Systems that explores a framework for this kind of contextual focus. Search engines also use a similar technique called semantic search that determines the intent and contextual meaning of users' search terms.

Finally, one more challenge in sentiment analysis is deciding how to train the model you'd like to use. There are a number of pre-trained models available for use in popular Data Science languages. For example, TextBlob offers a simple API for sentiment analysis in Python, while the Syuzhet package in R implements some of research from the NLP Group at Stanford.

These modules can help you get off the ground quickly, but for the best long term results you're going to want to train your own models. Getting access to labeled training data for sentiment analysis can be difficult, but it's key to building models that work for your specific use case. You may execute a workflow where you gather your proprietary data (e.x. customer service conversations) and use a service like CrowdFlower to label and prepare it.

> Try now for free: build models without any code

How is sentiment analysis done?

Sentiment analysis is done using algorithms that use text analysis and natural language processing to classify words as either positive, negative, or neutral. This allows companies to gain an overview of how their customers feel about the brand.

Sentiment analysis algorithms

Algorithmia provides several powerful sentiment analysis algorithms to developers. Implementing sentiment analysis in your apps is as simple as calling our REST API. There are no servers to setup, or settings to configure. Sentiment analysis can be used to quickly analyze the text of research papers, news articles, social media posts like tweets and more.

Social Sentiment Analysis is an algorithm that is tuned to analyze the sentiment of social media content, like tweets and status updates. The algorithm takes a string, and returns the sentiment rating for the "positive," "negative," and "neutral." In addition, this algorithm provides a compound result, which is the general overall sentiment of the string.

Input Example:

```
{"sentenceList": [
"I like double cheese pizza",
"I love black coffee and donuts",
"I don't want to have diabetes"]}
```

Output Example:

```
[{
"positive": 0.455,
"negative": 0,
"sentence": "I like double cheese pizza",
"neutral": 0.545,
"compound": 0.3612
},
{
"positive": 0.512,
"negative": 0,
"sentence": "I love black coffee and donuts",
"neutral": 0.488,
"compound": 0.6369
},
{
"positive": 0,
"negative": 0.234,
"sentence": "I don't want to have diabetes",
"neutral": 0.766,
"compound": -0.0572
}]
```

Algorithmia also features a flexible, multi-use Sentiment Analysis algorithm, which is great for more general texts, like books, articles, or transcripts. This algorithm is based on the Stanford CoreNLP toolkit. **To get started, you can get 10K credits on us with the invite code** *sentimentanalysis*.

The algorithm takes an input string and returns a rating from 0 to 4, which corresponds to the sentiment being very negative, negative, neutral, positive, or very positive.

Input Example:

```
"Algorithmia loves sentiment analysis!"
```

Output Example:

```
{
"result": 3
}
```

In addition, Algorithmia provides a Sentiment By Term algorithm, which analyzes a document, and tries to find the sentiment for the given set of terms. The algorithm works by taking in a string, a list of terms, and then splits the document into sentences, and computes the average sentiment of each term. This algorithm becomes powerful when combined with an auto-tagging algorithms, such as LDA, Auto-Tag URL, or Named Entity Recognition algorithms.

Input Example:

```
[
"John Brown (Johnny to his friends) is amazing! Johnny is by far the best mechanic in the tri-state area. Bob Bozo is the
worst.",
["john brown","bob"],
{"john brown":["johnny"]}
]
```

Output Example:

```
{
"john brown": 3.5, "bob": 1
}
```

Additional Sentiment Analysis Resources

Reading

An Introduction to Sentiment Analysis (MeaningCloud) – "*In the last decade, sentiment analysis (SA), also known as opinion mining, has attracted an increasing interest. It is a hard challenge for language technologies, and achieving good results is much more difficult than some people think. The task of automatically classifying a text written in a natural language into a positive or negative feeling, opinion or subjectivity (Pang and Lee, 2008), is sometimes so complicated that even different human annotators disagree on the classification to be assigned to a given text.*"

Sentiment Analysis Slides (EMP LCT) – "*Humans are subjective creatures and opinions are important. Being able to interact with people on that level has many advantages for information systems.*"

Thumbs up?: sentiment classification using machine learning techniques (Paper) – "*We consider the problem of classifying documents not by topic, but by overall sentiment, e.g., determining whether a review is positive or negative. Using movie reviews as data, we find that standard machine learning techniques definitively outperform human-produced baselines. However, the three machine learning methods we employed (Naive Bayes, maximum entropy classification, and support vector machines) do not perform as well on sentiment classification as on traditional topic-based categorization.*"

Opinion Mining and Sentiment Analysis (Paper) – "*An important part of our information-gathering behavior has always been to find out what other people think. With the growing availability and popularity of opinion-rich resources such as online review sites and personal blogs, new opportunities and challenges arise as people can, and do, actively use information technologies to seek out and understand the opinions of others. The sudden eruption of activity in the area of opinion mining and sentiment analysis, which deals with the computational treatment of opinion, sentiment, and subjectivity in text, has thus occurred at least in part as a direct response to the surge of interest in new systems that deal directly with opinions as a first-class object.*"

Demos and Tutorials

Data Science 101: Sentiment Analysis in R Tutorial (Kaggle) – "*Welcome back to Data Science 101! Do you have text data? Do you want to figure out whether the opinions expressed in it are positive or negative? Then you've come to the right place! Today, we're going to get you up to speed on sentiment analysis.*"

Deep Learning for Sentiment Analysis (Stanford) – "*This website provides a live demo for predicting the sentiment of movie reviews. In contrast, our new deep learning model actually builds up a representation of whole sentences based on the sentence structure. You can also browse the Stanford Sentiment Treebank, the dataset on which this model was trained.*"

Creating a Sentiment Analysis Model (Google) – "*This document explains how to create a basic sentiment analysis model using the Google Prediction API. A sentiment analysis model is used to analyze a text string and classify it with one of the labels that you provide; for example, you could analyze a tweet to determine whether it is positive or negative, or analyze an email to determine whether it is happy, frustrated, or sad.*"

Twitter Sentiment Analysis Using Python (GeeksForGeeks) – "*Sentiment Analysis is the process of 'computationally' determining whether a piece of writing is positive, negative or neutral. It's also known as opinion mining, deriving the opinion or attitude of a speaker.*"

Courses and Lectures

Text Mining, Scraping and Sentiment Analysis with R (Udemy) – "*This course will teach you anything you need to know about how to handle social media data in R. We will use Twitter data as our example dataset. During this course we will take a walk through the whole text analysis process of Twitter data.*"

Sentiment Analysis in R: The Tidy Way (Datacamp) – "*Text datasets are diverse and ubiquitous, and sentiment analysis provides an approach to understand the attitudes and opinions expressed in these texts. In this course, you will develop your text mining skills using tidy data principles. You will apply these skills by performing sentiment analysis in several case studies, on text data from Twitter to TV news to Shakespeare.*"

Text Mining and Analytics (Coursera, University of Illinois) – "*Detailed analysis of text data requires understanding of natural language text, which is known to be a difficult task for computers. However, a number of statistical approaches have been shown to work well for the "shallow" but robust analysis of text data for pattern finding and knowledge discovery. You will learn the basic concepts, principles, and major algorithms in text mining and their potential applications.*"

Books

Sentiment Analysis: Mining Opinions, Sentiments, and Emotions (Bing Liu) – "*Sentiment analysis is the computational study of people's opinions, sentiments, emotions, and attitudes. This fascinating problem is increasingly important in business and society. It offers numerous research challenges but promises insight useful to anyone interested in opinion analysis and social media analysis. This book gives a comprehensive introduction to the topic from a primarily natural-language-processing point of view to help readers understand the underlying structure of the problem and the language constructs that are commonly used to express opinions and sentiments.*"

Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More (O'Reilly) – "*How can you tap into the wealth of social web data to discover who's making connections with whom, what they're talking about, and where they're located? With this expanded and thoroughly revised edition, you'll learn how to acquire, analyze, and summarize data from all corners of the social web, including Facebook, Twitter, LinkedIn, Google+, GitHub, email, websites, and blogs.*"

Sentiment Analysis in Social Networks (Morgan Kaufmann) – "*Sentiment Analysis in Social Networks begins with an overview of the latest research trends in the field. It then discusses the sociological and psychological processes underlying social network interactions. The book explores both semantic and machine learning models and methods that address context-dependent and dynamic text in online social networks, showing how social network streams pose numerous challenges due to their large-scale, short, noisy, context- dependent and dynamic nature.*"

Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence (Marketing Science) – "*Bridging the gap between the marketer who must put text analytics to use and data analysis experts, Practical Text Analytics is an accessible guide to the many advances in text analytics. It explains the different approaches and methods, their uses, strengths, and weaknesses, in a way that is relevant to marketing professionals. Each chapter includes illustrations and charts, hints and tips, pointers on the tools and techniques, definitions, and case studies/examples.*"

TRIAL

Try now for free: build models without any code

Try for Free

ABOUT THE AUTHOR

DataRobot
Value-Driven AI

≣ DataRobot

DataRobot is the leader in Value-Driven AI – a unique and collaborative approach to AI that combines our open AI platform, deep AI expertise and broad use-case implementation to improve how customers grow and optimize their business. The DataRobot AI Platform is the only complete AI lifecycle platform that interoperates with your existing investments in data, applications and business processes, and can be deployed on-prem or in any cloud environment. DataRobot and our partners have a decade of world-class AI expertise collaborating with AI teams (data scientists, business and IT), removing common blockers and developing best practices to successfully navigate projects that result in faster time to value, increased revenue and reduced costs. DataRobot customers include 40% of the Fortune 50, 8 of top 10 US banks, 7 of the top 10 pharmaceutical companies, 7 of the top 10 telcos, 5 of top 10 global manufacturers.

Related Posts

AI FOR PRACTITIONERS

DataRobot with NVIDIA: The fastest path to production-ready AI apps and agents

MARCH 18 | 5 MIN READ

AI FOR PRACTITIONERS

How to use DeepSeek-R1 for enterprise

FEBRUARY 18 | 7 MIN READ

See other posts in AI for Practitioners ›

Why DataRobot

DISCOVER

AI Leaders

AI Practitioners

VALIDATION

Awards and Recognition

Customers

Solutions

INDUSTRY

Energy

Financial Services

Healthcare

Manufacturing

Product

OVERVIEW

AI Apps

AI Platform

AI PLATFORM

Generative AI

Predictive AI

AI Governance

AI Observability

AI Foundation

OPEN SOURCE

Covalent

LEARN MORE

Integrations

Services

Resources

WATCH

Product Tours

Webinars

READ