9 Natural Language Processing (NLP) Trends in 2021

9 Natural Language Processing (NLP) Trends in 2021

Natural Language Processing (NLP) is one of the most exciting fields in AI and has already given rise to technologies like chatbots, voice assistants, translators, and many other tools we use every day.

The study of natural language processing began in the 1950s, with the first attempts of automated translation from Russian to English laying the groundwork for research in natural language processing. Around the same time, the Turing Test – originally known as the imitation game – was developed to determine if a machine was capable of behaving in the same way as a human being.

Since then, NLP has come a long way. And, as unstructured data grows, NLP technology continues to evolve to better understand the nuances, context, and ambiguities of human language. Curious to know what's ahead for natural language processing? In this article, we'll explore some of the main NLP trends for 2021.

9 Trends in Natural Language Processing for 2021

It seems that more and more companies are beginning to see the benefits of NLP to draw out insights from large amounts of data, and automate tedious and repetitive tasks like question answering and ticket routing. Even though budgets were hit hard by the COVID-19 pandemic, 53% of leaders said their NLP budget was at least 10% higher compared to 2019.

So what's next for NLP, and how will businesses benefit? Well, the latest trends in NLP point towards applications that require limited labeled data and simplified processes that make NLP accessible to everyone.

One thing is certain: NLP is only going to grow in 2021.

Transfer Learning

Transfer learning is a machine learning technique where a model is trained for one task and repurposed for a second task that's related to the main task. So, instead of building and training a model from scratch, which is expensive, time-consuming, and requires huge amounts of data, you'll just need to fine-tune a pre-trained model. This means that businesses can complete NLP tasks faster using smaller amounts of labeled data.

Popularized in the field of computer vision, transfer learning is now being used in NLP tasks like intent classification, sentiment analysis, and named entity recognition.

Transformers Will Lead the Way: BERT & ELMO

One of the biggest breakthroughs in NLP this year has been the creation of machine learning models that create articles from scratch, with GPT-3 (Generative Pre-trained Transformer 3) leading the way. What's unique about transformers is that they're able to understand the context of words in a way that wasn't possible before.

Come 2021, the NLP community is going to be focusing more on BERT (Bidirectional Encoder Representations from Transformers) and ELMo (Embeddings from Language Models). These models have been trained on colossal amounts of data and are able to drastically improve the performance of a wide range of NLP problems.

Low-Code Tools Going Mainstream

Before, if you wanted to build an NLP model you needed a solid background in the field, coding skills to use open-source libraries, and machine learning knowledge. Not any more.

Although low-code / no-code tools have been around for a while now, they're set to become commonplace in 2021. SaaS companies like MonkeyLearn aim to democratize NLP and machine learning technology, allowing non-technical users to perform NLP tasks

that were once only accessible to data scientists and developers.

MonkeyLearn's point-and-click model builder makes it easy to build, train, and integrate text classification or sentiment analysis models in just a few steps, which means we can expect to see more and more businesses implementing NLP tools in 2021.

Multilingual NLP Will Grow

Most NLP advances to date have been focused on English. But companies like Google and Facebook are now publishing pre-trained multilingual models, which perform just as well or better than monolingual models.

Before 2019 multilingual models were unheard of, then Facebook introduced XLM-R and more recently M2M-100, the first multilingual machine translation model that can translate 100 languages without relying on English data.

With recent advances in language-agnostic sentence embeddings, zero-shot learning, and the availability of multilingual embeddings, open-source libraries are also following in the footsteps of Google and Facebook, so we can expect to see a growing trend in multilingual NLP models this year.

Combining Supervised & Unsupervised Machine Learning Methods

When training a model for NLP, combining both supervised and unsupervised methods seems to provide more accurate results.

Supervised learning, commonly used for tasks such as topic classification, requires a large amount of tagged data and many iterations until a model can make accurate predictions. In unsupervised learning, on the other hand, there's no labeled data: the model learns from input data and is able to detect patterns and make inferences on unseen data, on its own. An example of this is clustering, where similar objects are grouped together.

Combining supervised and unsupervised has shown to boost a machine learning model's performance, especially for text analysis.

Training Models with Reinforcement Learning

Reinforcement learning is an area of machine learning that is predicted to grow in 2021. Basically, reinforcement algorithms learn by doing, through a process of trial and error using feedback from previous actions and experiences.

In NLP, reinforcement learning can be used to speed up tasks like question answering, machine translation, and summarization. Currently, NLP models are trained first with supervised algorithms, and then fine-tuned using reinforcement learning.

Automating Customer Service: Tagging Tickets & New Era of Chatbots

This has been a tough year for customer service. As a result of the Covid-19 pandemic, there has been a tremendous increase in support tickets across all industries, from travel to finance. For businesses, it has been a challenge to deal with increasing ticket volume and provide fast responses to urgent queries.

More than ever, companies are realizing they need to automate simple customer service tasks to help them handle customer queries in a faster and more effective way.

Integrating NLP tools with help desk software, for example, could automate tedious tasks like tagging and routing customer support tickets, freeing agents from agents time-consuming and tedious tasks, and allowing them to focus on higher-value tasks.

Chatbots will also continue to play a significant role on the frontline of customer service. Though they have some limitations, they're generally able to hold a basic conversation and successfully complete tasks they've been trained for. Now, with advances in NLP and rising demand in customer service, we can expect to see huge strides towards the next generation of chatbots which will be able to self-improve, hold more complex conversations, and potentially learn how to complete new tasks without previous training.

NLP for Social Media Monitoring

Sentiment analysis, or opinion mining, will keep playing an important role in 2021, allowing businesses to monitor social media and gain real-time insights into how customers feel towards their brand or products.

Using NLP tools to gauge brand sentiment can help companies identify opportunities

for improvement, detect negative comments on the fly (and respond proactively), and gain a competitive advantage. Other interesting use cases for sentiment analysis in social media monitoring include analyzing the impact of marketing campaigns, and evaluating how customers react to events like a new product release.

Detecting Fake News and Cyber-Bullying

NLP has become an essential tool to reduce the time and human effort to detect and prevent the spread of fake news and misinformation. This year, with so much false information on Covid-19 making the rounds, we've already seen some interesting approaches towards automatic fake news detection (using transformers, no less), so we'll definitely see more of it during 2021.

Another way NLP is being used for positive impact is cyberbullying detection. Classifiers are being built to detect the use of offensive, and insulting language, or hate speech across social media.

Considering the on-going debate on whether social media content should be regulated, these NLP applications may become even more relevant.

Conclusion

Natural Language Processing (NLP) is rapidly evolving, and applications of natural language processing are growing by the minute. With so much data at our disposal, it's crucial to understand it, monitor it, and, in some cases, censor it.

In the years to come, NLP will become even more widespread thanks to ready-to-use pre-trained models and low-code, no-code tools that are accessible to everyone. Businesses, in particular, will continue to benefit from NLP, from improving their operations and customer satisfaction to reducing costs and making better decisions.

Now, more than ever, businesses need a helping hand, and NLP is just the ticket.

Visit MonkeyLearn to find out how you can implement NLP tools.