**Natural Language Processing**

Brandy Davis, Rebecca Leu, and Ronel Sylvester

ALY 6110: Data Management and Big Data

February 8th, 2021

**Natural Language Processing and How it Works**

Natural language processing (NLP) is a branch of artificial intelligence that helps computers understand, interpret and manipulate human language. (*What is natural language processing?*). Computers operate on a much different language that the average human, as its base computer language is just a bunch of 1’s and 0’s that produce actions. NLP is one of the most difficult things in computer science because of its complexities for computers to understand (Garbade, 2018). For anyone that has learned or attempted to learn another language, it's easy to see that a lot of language rules are abstract, and the same word or sentence could mean several different things. Programming computers to understand this language and make meaning of it is a lengthy process because of its ambiguity and imprecise characteristics (Garbade, 2018).

NLP uses a machine learning algorithm to identify and extract language rules that are unstructured and converts it into a form that the computer can understand. NLP is a lot like diagramming sentences like many did in school. Determining the words part of speech, what language it derives from, tokenization, lemmatization and stemming are all part of this process. From there NLP can be used to categorize the information. It can capture meaning and themes in the text and even identify sentiment to understand mood or feelings about a topic. Additionally, it can transform voice to text, one language to another and summarize large bodies of text. (*What is natural language processing?*) These are just some of the modern uses of NLP.

**The Purpose of Natural Language Processing**

NLP is important because there are endless amounts of unstructured data available for many different businesses and disciplines that can be used to help make predictions or solve problems.  Examples of NLP would be extracting information from social media to determine what customers are saying about a product or service. Amazon's Alexa or Apple’s Siri are highly complex examples of NLP, able to understand a person's voice and respond back with information needed or perform a task in a seemingly human way. NLP is used for automated call answering systems like with banks or insurance companies, able to route callers to the right representative, or able to give information to the caller without the need for human interaction at all. It is used for talent recruitment in identifying the skills or potential hires. NLP has also come a long way in the healthcare field who now uses it in the prediction of a disease based off of a patient’s health records. (Yse, 2019). The uses of NLL are seemingly endless, and the better computers get at recognizing and interpreting human language, the more ways this type of technology can work for us in new and unexpected ways.

**How, When, and Why to Use Natural Language Processing**

NLP can be used for many use-cases within the frameworks of linguistics and technology. NLP can simply be used to gather information from an article or book by pointing out key words/phrases that bring about it’s main point. We can even create a word cloud from the findings of the NLP algorithm. This use-case is just hitting the tip of the iceberg of what NLP can do though. NLP can also be used for summarizing long reports or research papers. Now, doing this for one article at a time is quite simple, however, tackling various forms of unstructured text data at high volumes takes a lot of computational power which is why we must use big data tools. “In every domain – say medical, legal, pharmaceutical, sports, education, and so on, large chunks of data are archived daily in the form of documents, customer inputs, sales information, etc.” (EuroSTAR Huddle, 2017). We need big data tools to analyze unstructured text data like these because the data is also constantly growing each day at higher veracities.

Sentiment Analysis is also a huge use-case for companies today to understand the truest overall reaction of how customers perceive their service/product. Comment sections of videos or even checking through customer reviews of a product on Amazon requires big data analysis to sift through all the data. Not only does the algorithm have to crawl through all the data, but it must also classify the sentiment of the comment/review. For example, does this customer have a more negative or positive outlook on the product? What did they like about it? What did they dislike about it? Finally, the algorithm must be able to generate a summary of its findings using Natural Language Generation (NLG, a subsection of NLP).

“NLG, a subfield of artificial intelligence (AI), is a software process that automatically transforms data into plain-English content. The technology can actually tell a story – exactly like that of a human analyst – by writing the sentences and paragraphs for you.” (Narrative Science, 2020). NLG is also integrated into chatbots and speech recognition products such as Siri. NLG is an important part of the process of NLP because a good NLG engine can answer back or summarize the findings of the NLP algorithm. Big data analysis is required for this as well because as enterprises increase in the veracity/volume of their data, processing all of the unstructured data will take more and more processing power and simple analytic tools will not be able to carry the workload.

**Natural Language Processing in the Real World**

One application of NLP that would be beneficial for most people to understand is how it is utilized by the HR departments of companies all around the world.  They are able to harness the power of NLP in the recruitment/hiring process as well as to drive employee satisfaction.

Employers have been using NLP in the application screening process for some time.  Most applicants are now savvy to the fact that they need to update their resume to have as many matching words and phrases to the job description as possible.  Companies use NLP to scan resumes that are submitted and use that info to determine whether the applicant should be put through to the next stage of the screening process.

Beyond the resume, employers are also using NLP as a “listening tool” (Biswas, 2019) to analyze the social media content of employees, or even applicants, to “uncover areas of interest, identify employee potential and talent, identify competence, and track behavior trends.” (Biswas, 2019)  Using social media analytics to extract information (BhandariI, 2020) can be useful in determining whether the attitudes of prospective employees would jive well with the company culture, something you just can’t determine through a resume.  It can also be used to gauge employee loyalty and help to reduce turnover.

HR departments are also able to use survey analytics to understand how employees are feeling about their work environment(Biswas, 2019).  When companies are genuinely willing to make changes according to the feedback their employees provide, then employees feel more empowered to be honest about their opinions in order to make the workplace a safer and happier place to be.  Through feedback from employees, both good and bad, companies are able to create action plans that help to build on the concepts and keywords highlighted through the NLP process.

HR also processes sensitive complaints from employees.  Anonymously tracking the key points in the complaints could also uncover important trends with employees that can create a toxic environment to work in.  If used appropriately, an HR team could drive their company to be on of the most sought after employers on the market.

References

BhandariI, A. (2020, December 23). *Information Extraction From Text Python*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/06/nlp-project-information-extraction/.

Biswas, S. (2019, July 17). *3 Ways Natural Language Processing (NLP) Will Transform HR in 2019*. HR Technologist. https://www.hrtechnologist.com/articles/digital-transformation/3-ways-natural-language-processing-nlp-will-transform-hr-in-2019/#:~:text=The%20Role%20of%20Natural%20Language%20Processing%20in%20HR,amounts%20of%20text%20into%20quantifiable%20insights%20almost%20instantaneously.

Chirag, W. (2017, June 21). NLP for big data Analytics - a guide. Retrieved February 04, 2021, from <https://huddle.eurostarsoftwaretesting.com/nlp-for-big-data-how-nlp-will-revolutionise-big-data-analytics/>

Garbade, D. M. J. (2018, October 15). A Simple Introduction to Natural Language Processing. Medium. https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32.

Kendall, S. (2020, April 25). What is natural language GENERATION (NLG)? Retrieved February 04, 2021, from https://narrativescience.com/resource/blog/what-is-natural-language-generation/

Rea, N. (2018, May 18). *Why is NLP useful for HR Professionals?* NLP World. https://www.nlpworld.co.uk/why-is-nlp-useful-for-hr-professionals/.

What is natural language processing? https://www.sas.com/en\_us/insights/analytics/what-is-natural-language-processing-nlp.html.

Yse, D. L. (2019, April 30). Your Guide to Natural Language Processing (NLP). Medium. https://towardsdatascience.com/your-guide-to-natural-language-processing-nlp-48ea2511f6e1.