A brief Introduction to Natural Language Processing

What is this NLP

According to Wikipedia, It is a subfield of Computer Science, Information Engineering and Artificial Intelligence, Concerned with the interaction between Computers and Human Languages

From a beginners perspective (i.e. my perspective), it is the INTERPRETER that is needed, to convey our expressions and thoughts (expressed in form of words) with appropriate meaning, for a computer program to understand and perform some actions accordingly

Look at how Jarvis(for the DC fans, it is Tony Stark’s “Alfred”, BUT with an Artificial Intelligence, so he can do more than bringing breakfast in the morning) understands everything and even understands the vague instructions and can also read between the lines, AS during his first flight with the mark-II armour, the Iron Man asks Jarvis to check the weather and Air traffic control and listen to the ground control, instruction with which Jarvis understands that Tony is going to try to fly, so he cautions him that there are still terabytes of calculation needed before an actual flight.(which Tony disregards completely, and LUCKILY everything goes okay, or the movie would’ve sucked majorly!)

Of course Jarvis is fictional but our aim as data scientists is to make a computer system so accurate and efficient that it should understand not only the meaning of the words but also humour, sarcasm, sentiments, and intent. So in short to pass the Turing test, we NEED to have Natural Language Processing.

But what IS a Natural Language

Paraphrasing the Wikipedia definition, it is the language that has been naturally established over the period of human evolution.

If only there was a conference in primitive ages where the leader said that “Okay you guys go to China and only use mandarin, here is the structure, grammar and all the words you can use.”

Even English, which is the one of the most spoken language across the world has borrowed many words from different languages, plethora of words originated from Latin and Greek words, and cultural influences change the usage of words altogether, isn’t that cool?(the word cool could mean fashionable or exciting or of course refer to the a thing’s temperature)

Most of today’s textual data is unstructured, there are no “on-the-nose” dialogue in real life, and we use gestures, pauses, the tonality of sentences, and with the understanding of the context of situation, the person next to us understands what we’re trying to convey. “YOU’RE TEARING ME APART LISA!!”

So let’s get back to processing such unstructured text data

What could we achieve if we could accurately find out, the meaning of the text?

1. We can classify and organise chucks of text, (Movie Reviews: Good or Bad based on user comments, no rating system needed)
2. We summarise a huge article (like this one, sorry!) into a paragraph using computer (make that computer read it and make it paraphrase what this bloke has been on about!)
3. Word/ phrase recommendation (Check LinkedIn it gives you a list of plausible and “Professional” replies in the chat window! I never run out of responses now! Wish they could make someone like this for Online Dating apps too! That’s a million dollar idea, right there!)
4. Machine Translation (translate the instruction or just translate to other language using a computer program(google translate))
5. Sentimental Analysis (recommend a sad-face emoji right after you type “Mom, I failed the exam, again!” and an meh-face emoji when your mom says “You don’t say, Mr. GamerBoy1776”)
6. Speech recognition, with the help of a few additional libraries we can convert speech-to-text (Virtual assistants like Siri, Alexa and YouTube’s auto generated closed captions and many more)
7. Topic Segmentation, Separating articles to categories by understanding their content
8. Extracting the most useful Information from the given text (extracting OTP from a text message)
9. Chat bot, the hottest thing of all the NLP application (Currently majority of service industry use domain specific Chat bots for an initial interaction)
10. Advertisement Matching(If I could spy on all the customers, and find out what they like, I can show them targeted advertisements, not that I personally would do that, but If I want to, I could, with NLP! \*\*Evil Laughter\*\* \*\*MuHahahaha\*\*)