JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A descriptive analysis on Natural Language Processing models for Automating Customer Service

[1]Zahra Jabeen, [2]Khushboo Mishra, [3]Binay Kumar Mishra

[1]Research Scholar, [2]Research Scholar, [3]Professor
[1]University Department of Computer Science,
[1]Veer Kunwar Singh University, Ara, India

Abstract : Natural Language Processing has already given rise to technologies like translators, voice assistants and chatbots. After 2019 pandemic hit there has been a tremendous increase in support tickets across all industries, from travel to finance. Several NLP model break down voice and text data in ways that help the computer understand of what it is ingesting. This paper will present the role of NLP in automating the customer service assistance for businesses and several other related sectors. A brief study of different NLP working models has also been displayed as per their use cases. Methods to check ac-curacy of Natural Language Processing model are also discussed in the article.

***Index Terms* - NLP, Artificial Intelligence, Customer Support, Smart Assistance.**

I. INTRODUCTION

The branch of Artificial Intelligence that gives computers the ability to under-stand text and voices in the manner a human brain does is termed as Natural Language Processing. The first attempt of automated translation was done in 1950s i.e. from Russian to English, laying the groundwork for research in natural language processing. Smart assistants, which were once in the realm of science fiction, are now a common thing. As unstructured data grows, NLP technologies are growing better to understand the nuances, context, and ambiguities of human language to provide fast responses to urgent queries. More than ever, companies are in need to automate simple customer service tasks so that they can handle customer queries in a faster and more effective way.

Integrating NLP tools with help desk software, would automate tedious tasks like routing customer support tickets, freeing agents from time-consuming and tiresome tasks, and providing them time to focus on higher-value tasks. Chatbots are usually able to hold a basic conversation and successfully satisfy end users in the way that they have been trained for. Customer service benefits businesses as a great deal in both time and money, especially during growing periods.

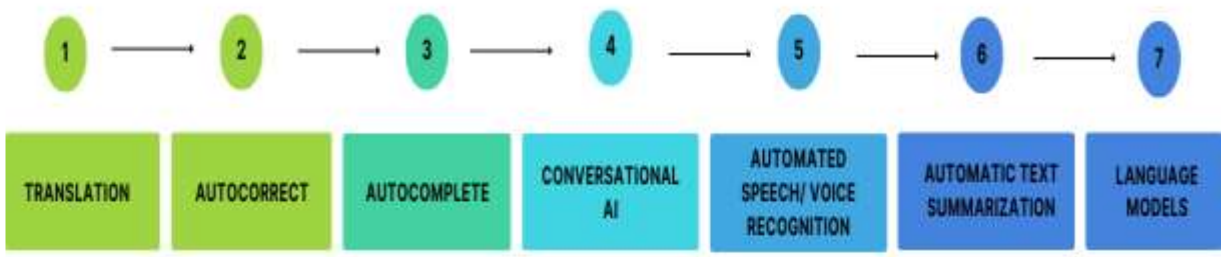

Figure 1. The picture above represents General Application of NLP.

II. ROLE OF NLP IN CUSTOMER ASSISSTANCE

NLP has been around for decades, but the popularity has exploded recently due to the arrival of pre-trained models (PTMs). There are several cloud services and open-source libraries that get pre-trained models accessible for NLP, each tailored to a specific type of NLP task. Some of these tasks include the following:

2.1 Speech Recognition

Speech-to-text is the process of accurately converting voice data into text data. Speech recognition is in demand for any application that answers spoken questions or follows voice commands. Major challenge of speech recognition is especially the way people talk, slurring words together, quickly with varying emphasis and intonation, in various accents and often using incorrect grammar.

2.2 Part of Speech Tagging

Also referred as grammatical tagging, it is the process of determining the part of speech of a particular piece of text based on its context. Part of speech identifies 'make' as a verb in 'I can make a good stew,' and as a noun in 'what make of bike do you own?'

2.3 Word sense disambiguation

It is the selection of the meaning of a word with multiple meanings by a process of semantic analysis that determine the word that is most acceptable in the given context.

2.4 Named entity recognition

NEM, recognizes words as useful entities. NEM will identify 'Bangalore' as a location or 'Einstein' as a name.

2.5 Co-reference resolution

It is the process of understanding if and when two words refer to the same item. One of the many common examples, is determining the person or object to which a certain pronoun refers, i.e. 'she' = 'Margret', but it can also involve identifying a metaphor or an idiom in the text, as an instance in which 'bear' isn't an animal but a large hairy person.

2.6 Sentiment analysis

It will attempt to extract subjective qualities like different attitudes, confusion, emotions, sarcasm, or suspicion from text.

2.7 Natural language generation

It may be sometimes explained as the opposite of speech-to-text or speech recognition; it is the task of putting structured/ processed information into human language.

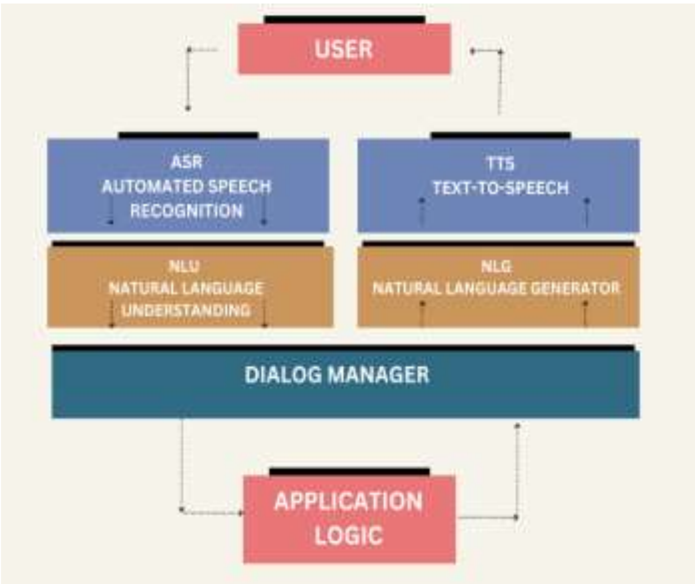

Figure 2. The flowchart above represents automated speech/ voice recognition.

III. PRE-TRAINED MODELS USED IN CUSTOMER SERVICE

Pre-trained models (PTMs) for natural language processing (NLP) are those deep learning models that have been trained on large datasets to perform specific NLP tasks such as transformers. Pre-trained models can be easily loaded into NLP libraries such as Tensorflow , PyTorch, etc. that can be used for performing NLP tasks with minimal extra effort. Several cloud services and open-source libraries that provide pre-trained models accessible for NLP are present among which some of the most popular ones are listed below:

3.1 Google BERT

BERT is an abbreviation for Bidirectional Encoder Representations from Trans-formers. It is a machine learning model which is a state-of-the-art. BERT was developed at Google in 2018 by Jacob Devlin and his colleagues and as a part of the TensorFlow project it was made open source in March 2019, to make it helpful for developers and data scientists to build AI models using existing state-of-the-art algorithms similar to BERT. BERT has been trained on NLP tasks like sentence segmentation, NER, part-of-speech tagging etc.

3.2 CodeBERT

CodeBERT is a bimodal pre-trained model for natural language (NL) and programming languages (PL). Microsoft's NLP engineers have published their NLP pre-trained model CodeBERT on GitHub. CodeBERT learns general-purpose representations that support applications such as code documentation generation, natural language code search, etc. Transformer-based neural architecture is used to develop it and then trained with a hybrid objective function that inculcates the pre-training task of replaced token detection, which then finds out plausible alternatives sampled from generators.

3.3 Huggingface transformers

Huggingface will provide pipeline APIs for grouping together different pre-trained models for their respective NLP tasks.

3.4 OpenNMT

OpenNMT was started in December 2016 by the Harvard NLP group and SYSTRAN, it is an open-source ecosystem for neural machine translation and neural sequence learning. The project has been used in several industrial applications and research projects. SYSTRAN and Ubiqus are presently in maintenance job for this model. NER models are usually used as part of the machine translation pipeline for pre-processing input data before sending it over to the Translate Model which performs actual sentence translations using Neural Machine Translation (NMT) models.

3.5 Facebook RoBERTa

Facebook's NLP engineers have published their NLP pre-trained model RoB-ERTa on GitHub. It has been used in NLP applications like Facebook Messenger etc. RoBERTa uses Bidirectional Encoder Representations from Transformers or BERT, a self-supervised method released in 2018 by Google. RoBERTa builds on BERT's language masking strategy, wherein the system learns to foresee intentionally hidden sections of text within otherwise unannotated language examples.

3.6 ELMo

ELMo refers to "Embeddings from Language Models". NLP scientists at Allen AI research center developed this NLP pre-trained model and as part of the Tensor-Flow project it was made open source in March 2019, to make it simpler for data scientists and developers to build AI models using existing state-of-the-art algorithms like ELMo. It is a deep contextualized word representation that models both complex characteristics of usage of words e.g. syntax and semantics and how these uses vary across linguistic contexts i.e. to model polysemy. They can be easily added to presently working models.

3.7 GPT-3

GPT-3 is an autoregressive language model that produces human-like text using deep learning. Created by OpenAI, it is the third-generation language prediction model in the GPT-n series (and the successor to GPT-2).

3.8 XLNet

Based on a novel generalized permutation language modeling objective, it is a new unsupervised language representation studying method. Moreover, XLNet employs Transformer-XL as its backbone model, displaying extraordinary per-formance for language tasks involving lengthy context.

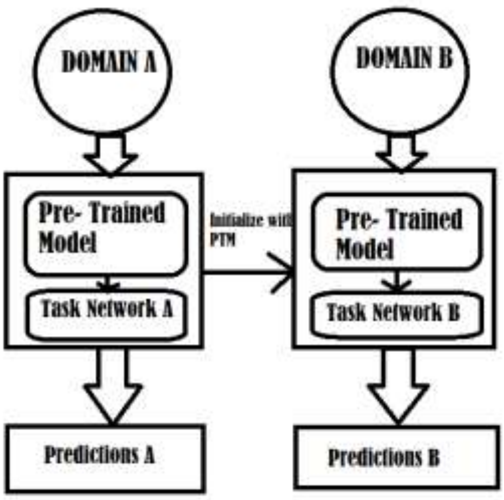

Figure 3. The picture above represents how a common pre-trained model can be reused (based on transfer learning) for different tasks resulting in task specific network.

IV. METHODOLOGY

To ensure the accuracy of any NLP model one should start by using a sufficient and representative amount of data for training and testing. The set of data that is used to train NLP model to learn the features and patterns of the task is known as Training data. The data that is used to evaluate how well NLP model generalizes to new and unseen data is called Testing data. Some common metrics to find out the accuracy and quality of NLP models are mentioned below:-

4.1 Accuracy and Error Rate

Accuracy is the percentage of correct predictions made by the used model out of the total number of predictions made. Error rate is the percentage of incorrect predictions made by the used model out of the total number of predictions made. Both accuracy and error rate are useful for multiclass or binary classification tasks, such as sentiment analysis, spam detection or identification of topic. Accuracy and Error rate can be calculated by following method:

Accuracy = (number of correct predictions) / (total number of predictions)

Error rate = (number of wrong predictions) / (total number of predictions)

4.2 Precision, recall and F1-score

Precision, recall and F1-score are metrics suitable for evaluating NLP models which deal with multiple labels, partial matches or ranking. Precision, recall, and F1-score are useful for performing tasks such as named entity recognition. For example, if your model extracts 10 entities from a text, but only 8 of them are correct, and there are 12 entities in the text, then your precision is 0.8, your recall is 0.67, and your F1-score is 0.73. It can be calculated as mentioned below:

Precision = (number of true positives) / (number of true positives + number of false positives)

Recall = (number of true positives) / (number of true positives + number of false negatives)

F1-score = 2 * (precision * recall) / (precision + recall)

4.3 BLEU and ROUGE

BLEU and ROUGE are metrics that are suitable for evaluating NLP models that deal with generating text such as summarization and machine translation. BLEU stands for Bilingual Evaluation Understudy and it will measure how close the generated text is to the reference text in terms of n-gram overlap. ROUGE stands for Recall-Oriented Understudy for Gisting Evaluation and it will measure amount of the important information in the reference text that is captured by the generated text in terms of n-gram recall. It can be calculated as follows:

BLEU = geometric mean of n-gram precisions * brevity penalty

ROUGE = harmonic mean of n-gram recalls

These are used for tasks such as machine translation or summarization, where the production of fluent and informative texts is measured, that matches the style and meaning of the reference texts. For example, if your model translates a sentence from English to Spanish, and there are four reference translations available, then you can compare the n-gram overlap and recall among your translation and the reference translations, and the BLEU and ROUGE scores can be computed accordingly.

4.4 Perplexity and log-likelihood

An alternate method to evaluate NLP models and systems is to measure the pre-diction capability of the next word or token given in a sequence of tokens or words. This is especially relevant for tasks such as text generation, language modelling and speech recognition. Perplexity and log-likelihood are the metrics that quantify the uncertainty of a model when they face fresh data. Lower perplexity and higher log-likelihood indicates to better performance and fit of the model.

4.5 Human evaluation

Human evaluation includes the interference of human judges or experts to com-pare the outputs of NLP models according to other criteria, such as relevance, readability, accuracy, coherence or usefulness. Human evaluation would provide more nuanced and qualitative feed-back on NLP models, as well as recognize faults or limitations that may not be identified by the numerical metrics. Human evaluation is done via surveys, experiments, interviews or crowdsourcing plat-forms depending on the budget of the project.

V. RESEARCH FINDINGS

5.1 According to recent analysis, North America is the biggest market for NLP, whereas the East Asia region invests heavily on NLP solutions.

5.2 A research by Accenture says more than 75% of CEOs want to completely modify their approach in managing customer relationships in order to keep up with changing consumer needs.

5.3 Deloitte says around 75% of CEOs think that the Great Resignation poses the greatest threat to their businesses. Sentiment analysis can be used by HR departments to pinpoint the prime reason of employee attrition.

5.4 On an average, JioMart sold 1500 orders every day via WhatsApp bot. Using this technology JioMart was able to reduce client response times by 60%.

5.5 Global Adoption Index by IBM cited that almost half of businesses that were surveyed globally are using some kind of NLP powered application.

5.6 In spite of several challenges faced by machines in understanding human language, the global NLP market was estimated at ~$5B in 2018 and is expected to reach ~$43B by 2025. The vast use cases of NLP could be attributed to such exponential growth.

5.7 The North American NLP in Finance Market is projected to be valued at USD 5,870.3 million in 2028; growing at CAGR of 24.7% during the forecast period.

VI. FUTURE SCOPE

Top five expectations regarding the future of NLP is as listed:
6.1. Investments in NLP will continue to rise- Market & Market report suggests an average annual growth rate of over 25% as the NLP market size was around $16 billion in 2022 and might reach to an approximate of $50 billion in 2027.
6.2 Conversational AI tools will be smarter- AI tools used in applications such as Conversational Commerce (Whatsapp Business API, Facebook Market Place, Mobile Application of Companies etc.), Conversational Banking (Banking Chatbots, Wealth management chatbots, Mortgage chatbots), Intelligent Automation (Writing e-Mails, Extracting data from CRM and ERP etc.) are enhancing their capability day by day.
6.3. Companies will use NLG to generate text Using Natural Language Generator, content market has gained 60% new customers.
6.4. More companies from various sectors implement sentiment analysis- Sentiment Analysis uses big data as source working majorly for Finance (Stock prices, Commodities, Coin values), e-Commerce, Human Resource department etc.
6.5. Common use of Voice Biometrics- People's voices, pronunciations, tones, pitches are unique characteristics that will help in providing more tight security than passwords which will increase the use of Voice Biometrics.

VII. CONCLUSION

Even though NLP technology is still in constant development mode, it has already taken huge leaps in establishing overall customer satisfaction. NLP not only addresses the simple task of rerouting agent calls, but also provides high level data analysis, all while being able to interact with users in their preferable language. In order to smoothen specific areas of the business and minimize labor-intensive human work, it's beneficial to harness the power of artificial intelligence. Eventually more and more companies have begun to see the benefits of NLP to draw out insights from huge pile of data and automate repetitive tasks like question answering and ticket routing. NLP is a core piece of machine learning that could be used in customer service departments. They're able to communicate with client in a way that suits the later and these companies can also reduce operational costs without compromising service quality.

REFERENCES

[1] www.ibm.com/topics/natural-language-processing
[2] Nikita Sachdeva, : daffodilsw, : Insights & Blogs Around Software Engineering - Learn, Develop, Grow
[3] research.aimultiple.com/future-of-nlp/
[4] MonkeyLearn Homepage, https://monkeylearn.com/blog/nlp-trends/
[5] Taha Junaid, D. Sumathi, , A.N. Sasikumar , S. Suthir , J. Manikan-dan , Rashmita Khilar , P.G. Kuppusamy , M. Janardhana Raju , A comparative analysis of transformer based mod els for figurative lan-guage classification. Book Title : Computers and Electrical Engineer- ingVolume 101, July 2022, 108051 ,pp. 1–13. ,ScienceDirect, Else-vier (2022). Link : https://doi.org/10.1016/j.compeleceng.2022.108051
[6] www.linkedin.com/advice/1/how-can-you-measure-natural-language-proess-ing#:~:text=Another%20set%20of%20metrics%20that,%2F%20(TP%20%2B%20FP)
[7] CHRIS REABURN, : Nextiva Blog : How to Use NLP in Customer Service & Why It's Important