Deep Learning and Cognitive Computing An emerging AI perspective for NLP

Prof. Aparna Gurjar

Department of Computer Application Shri Ramdeobaba College of Engineering and Management Nagpur, India

Abstract-*Increasing human computer interaction in recent years has necessitated an intelligent semantic based approach for processing of natural language so as to capture the abstract knowledge and meaning present in the spoken or written text. It can be achieved through deep neural networks and cognitive computing techniques. Both these techniques are modelled on the functioning of the human brain and aim to incorporate multi dimensional and abstract semantics for understanding of natural language. They are taking NLP to the next level of intelligent computing by bridging the gap between artificial intelligence and the functioning of the human brain.*

Keywords—NLP, natural language, deep learning, cognitive computing, machine learning, semantics

1. INTRODUCTION

Natural Language Processing (NLP) has seen phenomenal growth in last few years as the scope of Human Computer Interaction (HCI) has increased. NLP basically gives machines the ability to understand spoken and written human language. Traditionally NLP has relied on statistical models to make probabilistic predictions using the Markov property; which holds that there is dependence only between adjacent events, in a chain of linked events. In other words the next state depends on the current state of the system and can be predicted with reasonable accuracy. Among the various Part of Speech (POS) taggers available today, many have elegantly used this model for successful tagging of sentences.

But lately the focus of NLP has shifted more towards machine learning algorithms which work on non-annotated data. This brings into picture Artificial Intelligence (AI) which has been a part of the computing field for a prolonged period of time. AI has always strived to solve computing problems by taking cues from the working of the human brain. It is therefore no surprise that the confluence of AI and NLP have helped develop intelligent language processors. Machine learning (ML) is an important component of AI. Deep learning, which comes under the domain of ML helps achieve the goal of learning with minimal explicit programming of the machines. Deep learning techniques use artificial Neural Network to extract hidden information from the language representations and combine them to create new information through concepts like feature learning using supervised and unsupervised methods.

The ease with which the humans understand natural language, recognise objects etc. through perception, knowledge and inference is very difficult for machines to achieve. It is currently being accomplished through the use of complex Natural language processing and pattern recognition algorithms. The NLP algorithms however work on the linguistic components of the text like words, lexemes etc. through parsing, and fall short of understanding the abstract semantics associated with the text. So in order to build machines that are capable of understanding and responding to inherent abstractness in natural languages, innovative techniques have to be developed. Techniques those which not only handle words and plain semantics but also incorporate multi-layered abstract knowledge into the processing task. Cognitive computing techniques help create such models and design algorithms which actually mimic the various human cognitive processes.

Both Deep Learning and Cognitive Computing are fuelling the growth of AI. The next section explores few concepts related with deep learning and the cognitive computing; and investigates how both are steadily bridging the gap between functioning of artificial intelligence and the human brain.

2. DEEP LEARNING TECHNIQUES

2.1 DeepNL Architecture

The basic way in which deep learning gets implemented is by having multiple intermediate layers between the input and output layers of an artificial neural network. A new feature is learnt at each layer by applying filters which is then combined with previously learned features for generating new learning. This enables the network to handle high dimensionality of input data.

In case of natural language, the text or speech may contain many complex features. They can be either contextual information, regional flavour of languages, inflections or abstract semantics which contribute to high dimensionality and thereby affect the meaning implied by the sentences. In such conditions, in each layer a particular feature can be extracted, recombined with multiple other feature sets to implement regression or classification in a hierarchical way. Convolution neural networks (CNN) are generally used for this purpose. In [1] the authors describe a deep learning architecture using CNNs which is capable of creating tools for word embeddings. Word embedding is a technique where words or phrases in a vocabulary are mapped to vectors which contain real numbers. These vectors help in representing data that has high dimensionality.

For e.g. In case of digital image processing if a pixel is to be represented in a gray scale then the availability of number of bits determines the richness of the grey colour which can be represented in an image. In a single bit only two colours viz. black and white can be represented but a nibble (4 bit) can represent 16 shades of grey. However in case of NLP, the words which act as the basic unit are treated as mere symbols and are given an id. This process makes it impossible to associate a single word with more information. Considering the example of the word "cat", then cat not only means a four legged animal but can also be associated with the concept of pet. This can be handled using word embeddings. Word embeddings can provide a method of vector spaces for representation of words where a particular vector in a particular dimension can represent similar semantic properties of that word.

The authors [1] propose multi-layer neural network architecture. The first layer transforms the input into a feature vector representation. Each feature has many parameters which need to be learnt. It has hidden layers and activation function. At each layer of the neural network, an activity can be performed. It takes in annotated sentences as input and creates a probability distribution of outputs. They have trained the DeepNL sequence tagger on different features like "caps feature" and the suffix feature. "The cap feature" identifies whether the word is in lower case, upper case or starts with capital; and the suffix feature identifies the ending of a token as to whether they belong to the given test set or not.

2.2 The SEBLA Approach

In [2] the authors propose a semantic based machine learning algorithm for semantic driven processing of natural language. Their algorithm uses human like approach as it tries to understand semantics and derive new knowledge from the existing by using a semantic and logic driven learning paradigm. This algorithm is based on deep learning.

They have developed a Semantic Engine using Brain-Like approach called (SEBLA) that uses Brain-Like algorithms to solve the natural language understanding problem. The algorithm considers each word as an object having different features at the standalone level of single word. The word-semantics generated at this stage are used to derive new semantics for a complete sentence which is done by taking the context provided by the sentence itself. This process is then applied to multiple sentences in order to generate the meaning of the complete paragraph. For each word in a sentence the function words (WF) and world knowledge (WK) are found out. For e.g. the function word for apple would be {red, green, sweet, fruit, health, juice.... }. Thus each word gets associated with multiple other words based on semantics as well as world knowledge and is able to process it more deeply. New knowledge is found by identifying causal relationship between sentences.

3. COGNITIVE COMPUTING TECHNIQUES

3.1Perception cognition model

In [3], the authors' base their algorithm on the perception-cognition model (P-C model) of human brain proposed by Sherwin. They propose a layered unified cognition framework. Refer figure 1. In this model different layers perform different activities. The lower layers receive the physical signals from the real world and its processing for generating cognition is done in the higher layers. The unified framework enables the unification of different types of perception features like auditory, visual and sense of touch. It is argued that cognition is the end result of creation of concept space of which the concept is a basic cognition unit. In the human mind all the concepts are logically related and can be retrieved as per need. According to this model each word in a sentence creates a concept in the concept space and is associated with numerous features. These features themselves are also mini concepts. Thus there is a hierarchy of concepts. If all the features are perceived as having the same value, the word is not conceptualized but behaves as an instance. Conceptualization happens only when some features show different values. The word then is abstracted to a higher level as a concept and its associated features provide the relevant semantics. For e.g. If an apple has only red colour then it will behave as an instance, but when the colour property shows two values; "Red" and "Green", then apple becomes a concept which has many properties among which colour is one. A green coloured apple could have associated semantics of belonging to a particular geographical region.

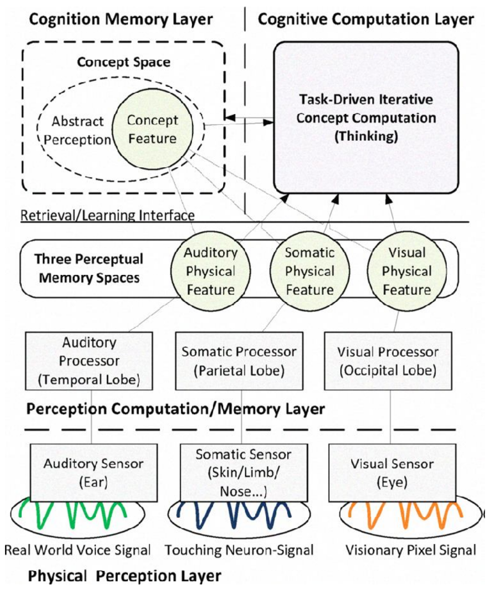

Figure 1 Unified cognition framework [3]

Ref: Unified hierarchical iterate model of human Conceptualization and cognition

This is exactly the way in which the human brain processes information while generating cognition. Knowledge is stored as concepts and through perception the brain tries to fit incoming information to those concepts in order to understand it.

3.2 Inference based model

The act of reasoning and drawing inferences are important constituents of cognition in humans. Inference is defined as a conclusion which is formed as a result of known facts or is arrived at, by the process of reasoning. The conclusion could also be in a form of opinion. In either case it is something which is not explicitly written down or spoken but is inferred implicitly. From the current computational perspective the problem of drawing semantic inference from texts is limited to recognizing textual entailment (TE) between two given sentences.

TE can be explained as follows: If the truth of the hypothesis (H) can be understood from the text (T), then there is an entailment relationship between the two sentences. i.e. T would entail H. For e.g. consider the statement “Today the umbrella saved me from rain”. Considering Text (T) and Hypothesis (H) as,
T: Today the umbrella saved me from rain.
H: It rained today.
Here the truth of the hypothesis (H) can be understood from the text (T), by the fact that “If the umbrella saved someone from getting wet due to rain today, then it must have rained today”, therefore T entails H. We can thus draw inference that it rained today even though the sentence does not say so explicitly. In context of the statement, if any question answering system is asked. “Did it rain today?”, then the answer would be “yes”, because the given text entails the hypothes is There are many other inferences which can be inferred from the above statement T, which could be based on the latent knowledge of words contained in the sentence and can entail multiple hypotheses. These kinds of inferences are very easily inferred by the human brain but present a computationally intensive task to the machine. In [4] the authors approach this semantic entailment issue by proposing a concept based knowledge representation model which defines the concept by three values viz. attribute, relation and behaviour. Attribute is the defining feature of a concept. Relations show interrelatedness of concepts. Behaviour is the interaction between them. Analyzing the sentences based on such a methodology, enables the system to cluster together sentences having dissimilar forms but same semantic meaning. It also helps draw inferences. The associative memory present in human brains works in similar way so as to establish relationships between two entities in order to remember them.

4 OBSERVATIONS

From the above concepts related with deep learning and cognitive computing few observations can be made.

- The deep learning methods incorporate processing of abstract semantics through feature learning and its extraction. The combination of various features at each level gives rise to rich information and ability to understand complex semantics.
- The cognitive methods learn abstract semantic through creation of concept hierarchy and its representation. The process of inference generation using first order logic and textual entailment also helps generate new information which is not explicitly contained in the input.
- Both these techniques take cues from the functioning of the brain and aim to incorporate multi dimensional semantics for understanding of natural language. This is completely different from the shallow learning techniques which can process and work only on limited semantic information.

5 CONCLUSION

It is thus clear that the processing of natural language has progressed from lexical parsing and plain semantic processing, to a more abstract knowledge processing; in order to grasp the deeper semantics represented by the natural language. It can be done by using supervised/unsupervised deep neural networks which traditionally achieve deep learning through multiple hidden layers between the input and output layers. It can also be done by representing concepts as hierarchical knowledge, where they act as a basic unit of semantic knowledge. Conceptual semantics, entailment, etc. are building blocks of cognitive computing which aim to solve the knowledge representation problem. Thus deep learning and cognitive computing are taking NLP to the next level of intelligent computing by bridging the gap between artificial intelligence and the functioning of the human brain.

References

[1] Giuseppe Attardi “DeepNL: a Deep Learning NLP pipeline” Proceedings of NAACL-HLT 2015, pages 109–115, Denver, Colorado, May 31 – June 5, 2015. © 2015 Association for Computational Linguistics
[2] Dr. Emdad Khan, “Machine Learning Algorithms for Natural Language Semantics and Cognitive Computing” 2016 International Conference on Computational Science and Computational Intelligence
[3] Jiayu Zhou, Lifeng Jin, Sherwin Han, “Unified hierarchical iterate model of human Conceptualization and cognition” Proceedings of 8th IEEE Int. Conf. On cognitive Informatics ICCI’09
[4] Zhao Meijing, Ni Wancheng, Zang Haidong, Yang Yiping, “A Concept-Based Knowledge Representation Model for Semantic Entailment Inference” Proceedings of the 33rd Chinese Control Conference July 28-30, 2014, Nanjing, China
[5] J. Octavio Gutierrez-Garcia, Emmanuel López-Neri, “Cognitive Computing: A brief survey and

open research challenges", 2015 3rd International Conference on Applied Computing and Information Technology/2nd International Conference on Computational Science and Intelligence

[6] Marc Moreno Lopez, Jugal Kalita "Deep Learning applied to NLP", arXiv:1703.03091v1 [cs.CL] 9 Mar 2017

[7] Denis A. de Araujo, Sandro J. Rigo , Carolina Muller, Rove Chishman "Automatic information extraction from texts with inference and linguistic knowledge acquisition rules" 2013 IEEE/WIC/ACM International Conferences on Web Intelligence (WI) and Intelligent Agent Technology (IAT)

[8] Yann Lecun, Yoshua Bengio, Geoffrey Hinton, "Deep Learning" Review pages 436 -444 NATURE Vol 521 28 MAY 2015

[9] Yoshua Bengioy, Aaron Courville, and Pascal Vincenty, "Representation Learning: A Review and New Perspectives" arXiv:1206.5538v3 [cs.LG] 23 Apr 2014

[10] Eric Cambria, Bebo White, "Jumping the NLP Curves: A review of Natural Language Processing Research" pages 48-57 IEEE COMPUTATIONAL INTELLIGENCE MAGAZINE | MAY 2014