Gamification: "The System Beats Human Huge Critical Thinking"

D.Asha
Research Scholar,Vels University, Tamilnadu, India.

P.Mayilvahanan
Professor, Vels University, Tamilandu, India.

Abstract – The NLP(Natural language processing) is supporting Deep AQ system .The Deep AQ is a Large scale of NLP. It is respond across the structured and unstructured data. There are hundreds of analytics that will be combined with more than 50 models. That all are trained through the machine learning process. The milestone of 2011 history was the jeopardy game show .The game show was defeating the two best human players. The IBM Watson technology was used in this game. Deep AQ is depended on real word business problem gamification. Watson is a mixture of training transformation, functional and content for nongame play. This is the gamifying a business domain for Watson .The business domain gamification cannot follow the same machine learning process. That will be chance effects the machine learning for system. For example medical.Financial ,government or any other business domain. The orignial Watson game show as opposed the positive to negative to convert the ratio for example 1:100 the question and answer pairs are transformed to true or false format.The positive to nagattiveraio is very low means the ratio of 1:100 00.The big challenge of Watson machine learning pipeline are imbalance the initial extreme class in domain gamification.The combination of NLP algorithms and configuration settings the question and answer paris date sate are very challenging. We can used several data engineering technique to set the answer key, question set modification oversampling, vetting and expansion they are increase the positive lables.The Newton Raphsonis the addition algorithm engineering to used implement the logistic regress with a regular trem. It is used to imbalance of class training transformatiom.Finally we conclude by demonstranting the data and algorithm are compermentary.

Index Terms – Natural language processing, deep AQ system, IBM Watson.

1. INTRODUCTION

Natural language processing (**NLP**) is a field of computer science, artificial intelligence, and computational linguistics concerned with the interactions between computers and human (natural) languages. As such, NLP is related to the area of human–computer interaction. Many challenges in NLP involve natural language understanding, that is, enabling computers to derive meaning from human or natural language input, and others involve natural language generation.

- The process of computer analysis of input provided in a human language (natural language), and conversion of this input into a useful form of representation.
- The field of NLP is primarily concerned with getting computers to perform useful and interesting tasks with human languages.

The field of NLP is secondarily concerned with helping us come to a better understanding of human language.

2. NLP TASKS

2.1 Automatic summarization:

Produce a readable summary of a chunk of text. Often used to provide summaries of text of a known type, such as articles in the financial section of a newspaper.

2.2 Coreference resolution:

Given a sentence or larger chunk of text, determine which words ("mentions") refer to the same objects ("entities"). Anaphora resolution is a specific example of this task, and is specifically concerned with matching up pronouns with the nouns or names that they refer to. The more general task of coreference resolution also includes identifying so-called "bridging relationships" involving referring expressions. For example, in a sentence such as "He entered John's house through the front door", "the front door" is a referring expression and the bridging relationship to be identified is the fact that the door being referred to is the front door of John's house (rather than of some other structure that might also be referred to).

2.3 Discourse analysis:

This rubric includes a number of related tasks. One task is identifying the discourse structure of connected text, i.e. the nature of the discourse relationships between sentences (e.g. elaboration, explanation, contrast). Another possible task is recognizing and classifying thespeech acts in a chunk of text (e.g. yes-no question, content question, statement, assertion, etc.).

2.4 Machine translation:

Automatically translate text from one human language to another. This is one of the most difficult problems, and is a member of a class of problems colloquially termed "AI-complete", i.e. requiring all of the different types of knowledge that humans possess (grammar, semantics, facts about the real world, etc.) in order to solve properly.

2.5 Morphological segmentation:

Individual morphemes and identify the class of the morphemes. The difficulty of this task depends greatly on the complexity of the morphology (i.e. the structure of words) of the language being considered. English has fairly simple morphology, especiallyinflectional morphology, and thus it is often possible to ignore this task entirely and simply model all possible forms of a word (e.g. "open, opens, opened, opening") as separate words. In languages such as Turkish or Manipuri, a highly agglutinated Indian language, however, such an approach is not possible, as each dictionary entry has thousands of possible word forms.

2.6Named entity recognition (NER)

Given a stream of text, determine which items in the text map to proper names, such as people or places, and what the type of each such name is (e.g. person, location, organization). Note that, although capitalization can aid in recognizing named entities in languages such as English, this information cannot aid in determining the type of named entity, and in any case is often inaccurate or insufficient. For example, the first word of a sentence is also capitalized, and named entities often span several words, only some of which are capitalized. Furthermore, many other languages in non-Western scripts.

2.7 Natural language generation:

Convert information from computer databases into readable human language.

2.8 Natural language understanding

Convert chunks of text into more formal representations such as first-order logic structures that are easier for computer programs to manipulate. Natural language understanding involves the identification of the intended semantic from the multiple possible semantics which can be derived from a natural language expression which usually takes the form of organized notations of natural languages concepts. Introduction and creation of language metamodel and ontology are efficient however empirical solutions. An explicit formalization of natural languages semantics without confusions with implicit assumptions such as closed-world assumption (CWA) vs. open-world assumption, or subjective Yes/No vs. objective True/False is expected for the construction of a basis of semantics formalization.

2.9Optical character recognition (OCR)

Given an image representing printed text, determine the corresponding text.

2.10 Part-of-speech tagging

Given a sentence, determine the part of speech for each word. Many words, especially common ones, can serve as multiple parts of speech. For example, "book" can be a noun ("the book on the table") or verb ("to book a flight"); "set" can be a noun, verb or adjective; and "out" can be any of at least five different parts of speech. Some languages have more such ambiguity than others. Languages with little inflectional morphology, such as English are particularly prone to such ambiguity. Chinese is prone to such ambiguity because it is a tonal languageduring verbalization. Such inflection is not readily conveyed via the entities employed within the orthography to convey intended meaning.

2.11Parsing:

Determine the parse tree (grammatical analysis) of a given sentence. The grammar for natural languages is ambiguous and typical sentences have multiple possible analyses. In fact, perhaps surprisingly, for a typical sentence there may be thousands of potential parses (most of which will seem completely nonsensical to a human)

2.12 Question answering:

Given a human-language question, determine its answer. Typical questions have a specific right answer (such as "What is the capital of Canada?"), but sometimes open-ended questions are also considered (such as "What is the meaning of life?"). Recent works have looked at even more complex questions.

2.13 Relationship extraction

Given a chunk of text, identify the relationships among named entities (e.g. who is married to whom).

Sentence breaking (also known as sentence boundary disambiguation):

Given a chunk of text, find the sentence boundaries. Sentence boundaries are often marked by periods or other punctuation marks, but these same characters can serve other purposes (e.g. marking abbreviations).

2.14 Sentiment analysis:

Extract subjective information usually from a set of documents, often using online reviews to determine "polarity" about specific objects. It is especially useful for identifying trends of public opinion in the social media, for the purpose of marketing.

2.15 Speech recognition:

Given a sound clip of a person or people speaking, determine the textual representation of the speech. This is the opposite of text to speechand is one of the extremely difficult problems colloquially termed "AI-complete" (see above). In natural speech there are hardly any pauses between successive words, and thus speech segmentation is a necessary subtask of speech recognition (see below). Note also that in most spoken languages, the sounds representing successive letters blend into each other in a process termed coarticulation, so the conversion of the analog signal to discrete characters can be a very difficult process.

2.16 Speech segmentation:

Given a sound clip of a person or people speaking, separate it into words. A subtask of speech recognition and typically grouped with it.

2.17 Word segmentation

Separate a chunk of continuous text into separate words. For a language like English, this is fairly trivial, since words are usually separated by spaces. However, some written languages like Chinese, Japanese and Thai do not mark word boundaries in such a fashion, and in those languages text segmentation is a significant task requiring knowledge of the vocabulary and morphology of words in the language.

2.18 Word sense disambiguation

Many words have more than one meaning; we have to select the meaning which makes the most sense in context. For this problem, we are typically given a list of words and associated word senses, e.g. from a dictionary or from an online resource such as WordNet.

2.19 Information retrieval (IR)

This is concerned with storing, searching and retrieving information. It is a separate field within computer science (closer to databases), but IR relies on some NLP methods (for example, stemming). Some current research and applications seek to bridge the gap between IR and NLP.

2.20 Information extraction (IE)

This is concerned in general with the extraction of semantic information from text. This covers tasks such as named entity recognition,Coreference resolution, relationship extraction, etc.

2.21Speech processing

This covers speech recognition, text-to-speech and related tasks.

2.22.NaturalLanguage Understanding

- Mapping the given input in the natural language into a useful representation.
- Different level of analysis required:

  Morphological analysis,

  Syntactic analysis,

  Semantic analysis,

  Discourse analysis,

2.23Natural LanguageGeneration

Producing output in the natural language from some internal representation.Different level of synthesis required:

Deep planning(what to say),

Syntactic generation

2.24 Morphology:

Concerns how words are constructed from more basic meaning units called morphemes. A morpheme is the primitive unit of meaning in a language.

2.25 Syntax:

Concerns how can be put together to form correct sentences and determines what structural role each word plays in the sentence and what phrases are subparts of other phrases.

2.26 Semantics:

Concerns what words mean and how these meaning combine in sentences to form sentence meaning. The study of context-independent meaning.

3. NLP APPLICATIONS

- Machine TranslationTranslation between two natural languages.
- Information Retrieval Web search system, or a dialogue system.
- Report Generation Generation of reports such as weather reports.
- Some Small Applications Grammar Checking, Spell Checking, Spell Corrector(uni-lingual or multi-lingual).
- Query Answering/Dialogue Natural language interface with a database.

4. CONCLUSION

Finally my dissertaetion lies in frame work of how to design Jeapardy game show used with NLP.

REFERENCES

[1] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J.W.Vaughan, "A theory of learning from different domains," *Mach. Learn.*, vol. 79, pp. 151–175, 2010.

[2] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2006.

[3] M. A. Campbell, J. Hoane, Jr., and F. Hsu, "Deep blue," *Artif.Intell.*, vol. 134, pp. 57–83, 2002.

[4] S. L. Cessie and J. C. Van HouWelingen, "Ridge estimators in logistic regression," *Appl. Stat.*, vol. 41, no. 1, pp. 191–201, 1992.

[5] J. Chu-Carroll, J. Fan, B. K. Boguraev, D. Carmel, D. Sheinwald, and C.Welty, "Finding needles in the Haystack: Search and candidate generation," *IBM J. Res. Develop.*, vol. 56, no. 3/4, pp. 6:16:12,May/Jul.2012,DOI:10.1147/JRD.2012.2186682

[6] J. Chu-Carroll, J. Fan, N. Schlaefer, and W. Zadrozny, "Textual resource acquisition and engineering," *IBM J. Res. Develop.*, vol. 56, no.3/4, pp. 4:1–4:11, May/Jul.2012,DOI:10.1147/JRD.2012.2185901.

[7] S. Deterding, R. Khaled, L. Nacke, and D. Dixon, "Gamification: Toward a definition," in *Proc. GamificationWorkshop*,Vancouver, BC, Canada, 2011 [Online].Available:http://gamification-research.org/wpcontent/uploads/2011/04/02- Deterding

[8] G. Tesauro, D. C. Gondek, J. Lenchner, J. Fan, and J.M. Prager, "Simulation,learning, and optimization techniques inWatson's game strategies," *IBM J. Res. Develop.*, vol. 56, no. 3/4, pp. 16:116:11,May/Jul.2012,DOI:10.1147/JRD.2012.2188931.

[9] S. J. Wright, R. D. Nowak, and M. A.T.Figueiredo,"Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*,vol. 57, no. 7, pp. 2479–2493, Jul. 2009.

[10] T. Zhang and F. J. Oles, "Text categorization based on regularizedlinear classification methods," *Inf. Retrieval*, vol. 4, pp. 5–31, 2000.