Machine Learning with Pattern Recognition in Natural Language Processing without any Grammar and Syntax Knowledge

Taiebeh Mohammadi[1], Mehdi Sadeghzadeh[2] & Javad Akbari Torkestani[1]

[1] Department of Computer Engineering, Arak Branch, IUA, Arak, Iran

[2] Department of Computer Engineering, Dezful Branch, University of Applied Science and Technology Dezful, Iran

Correspondence: Taiebeh Mohammadi, Department of Computer Engineering, Arak Branch, IUA, Arak, Iran. E-mail: alireza_karbalaei_2007@yahoo.com

Abstract

In this system we are able to classify the components of sentences into a parse tree and gradually grow the tree with extra information about the subjects. Then the recognized pattern will be investigated with an interpreter and will be translated to understandable command for machines. After passing elementary education, further training will be taught indirectly just based on the previous orders without any supervisors. Thus after teaching the machine capable to extract concept from the classified data and act with natural language. These training just like primary education to kids without any knowledge about the language syntaxes and grammar only due to the pattern recognition in sentences.

Keywords: machine learning, NLP, pattern recognition, parse tree

1. Introduction

Today, various sciences and technologies are implementable and exploitable in such a way in computers. As regards transition the science to the computer required the computer programmer skill familiar to the science, the cost of this implementation and transition is too high, we decided to teach the science to the computer in writing, in this paper using the natural language processing the knowledge and science will be taught to the machine in writing and speech. The main core of learning will be improved and grows over time then deduction is expected. As noted the natural language processing (NLP) is the main kernel of this paper (Dale et al., 2000; Dale et al., 2005; Grosz et al., 2010) we will try to use this technique to separate the words and figure out the purpose of the sentences.

2. Purpose of the Study

The main target of this paper is implementing an intelligent system to teaching the machines using natural language. As we know the first stage of relation between the humans is learning a common language. The language is a defined protocol to make an interactive communication. And we know to make a communication between humans and computer we need a protocol that called programming language. Despite the programming languages follow a series of special principles and few peoples are predominate on it, as a result to have a communication to the computer we need programmers who act as a translator between us and computers. In this system we try to remove the translation role of programmers and design a protocol based on human natural language. There is more future goals in this context, for instance, consider an online store that has a section for user comments in relation to products, Due to the growth of the information in this section, reading the all description is difficult for other buyers, Using this current system, useful and efficient content will be fetched from the comments, so users only need to read the contents of the summary and conclusion, the obtained conclusion can be shown as graphical chart to the users.

3. Method

Due to the use of natural language processing algorithms that produce a tree structure of sentences components, this research is based on parse tree and search algorithm that leading to a structure classifieds dictionary of words in the form of tree.

3.1 NLP Parser

A natural language parser is a program that works out the grammatical structure of sentences, for instance, which groups of words go together (as "phrases") and which words are the subject or object of a verb. Probabilistic parsers use knowledge of language gained from hand-parsed sentences to try to produce the most likely analysis of new sentences. These statistical parsers still make some mistakes, but commonly work rather well. Their development was one of the biggest breakthroughs in natural language processing in the 1990s. You can try out our parser online (de Marneffe et al , 2006).

3.2 Entering Classified Data

This section including the classified creature, as we know our knowledge is the set of classified and categorized data from the Environment, we used this technique to manage the knowledge of machine with growing gradually. Step by step the classified tree will be filled and grow by data entering from the teacher and learning system that might be a self-learning system or a supervisor.

Human: man, woman

Mother is woman

Mother is parent

Father is man.

Father is parent
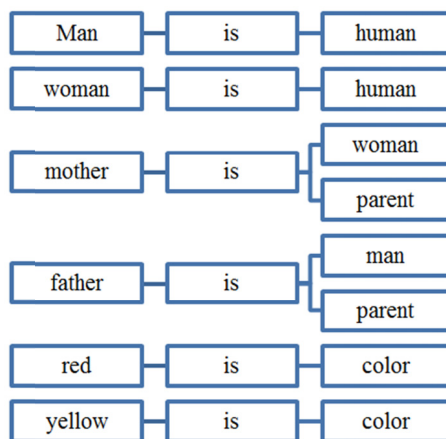
Red is color

Yellow is color

Figure 1. Inserting the words to the parse tree

3.3 Interpreter Commands

Consider the following sentences

"Jim's mother is Sarah = Sarah is Jim's mother"

Given that there are different structures for sentences, assisted by interpreter, semantically similar situation will be obtained.

$$\_/NNP \_/POS \_/NN \_/VBZ \_/NNP =$$
$$\_/NNP \_/VBZ \_/NNP \_/POS \_/NN \qquad (1)$$

As you see the first part of equation included the parser pattern of first sentence and the second part is another semantic similar situation, According to this equation every second part pattern will be saved like first one. This section play a compiler role that will be design by the programmer to handle and execute the orders on feed data. In some sentence the ownership is visible e.g.

The leaves of tree are green =>

The/DT leaves/NNS of/IN tree/NN are/VBP green/JJ ./.

As you see in the sample we can find that the leaves belong to the tree, so the interpreter can figure out the

ownership with this template

$$\underline{A1}/NNS \ \underline{of}/IN \ \underline{B1}/NN=>$$

$$(\underline{A1}).parent = \underline{B1} \qquad\qquad (2)$$

As you saw in the examples, the interpreter works based on POS-tags, due to the input commands analyze the sentences and categorized it in a proper position.

A Part-Of-Speech Tagger (POS-Tagger) is a piece of software that reads text in some language and assigns parts of speech to each word (and other token), such as noun, verb, adjective, etc., although generally computational applications use more fine-grained POS tags like 'noun-plural' (Toutanova and Manning,2000; Toutanova et al, 2003) .

Table 1. POS tag list

Tag	Description
CC	Coordinating conjunction
CD	Cardinal number
DT	Determiner
EX	Existential there
FW	Foreign word
IN	Preposition or subordinating conjunction
JJ	Adjective
JJR	Adjective, comparative
JJS	Adjective, superlative
LS	List item marker
MD	Modal
NN	Noun, singular or mass
NNS	Noun, plural
NNP	Proper noun, singular
NNPS	Proper noun, plural
PDT	Predeterminer
POS	Possessive ending
PRP	Personal pronoun
PRP$	Possessive pronoun
RB	Adverb
RBR	Adverb, comparative
RBS	Adverb, superlative
RP	Particle
SYM	Symbol
TO	To
UH	Interjection
VB	Verb, base form
VBD	Verb, past tense
VBG	Verb, gerund or present participle
VBN	Verb, past participle
VBP	Verb, non-3rd person singular present
VBZ	Verb, 3rd person singular present
WDT	Wh-determiner
WP	Wh-pronoun
WP$	Possessive wh-pronoun
WRB	Wh-adverb

3.4 Entering the Raw Data

After entering the raw data, the parse tree will start parsing process and the sentence will be decomposed to the components and their roles (POS-tags).

The POS-tags parts will be inserted to the main tree continuously, e.g.

Jim is man. =>

Jim/NNP is/VBZ man/NN ./.

Jim's parent is Sarah. =>

Jim/NNP 's/POS parent/NN is/VBZ Sarah/NNP ./.

Sarah is woman. =>

Sarah/NNP is/VBZ woman/NN ./.

Apple's color is red. =>

Apple/NNP 's/POS color/NN is/VBZ red/JJ ./.

Apple's taste is sweat. =>

Apple/NNP 's/POS taste/NN is/VBZ sweat/NN ./.

She is writer. =>

She/PRP is/VBZ writer/NN ./.

He is programmer. =>

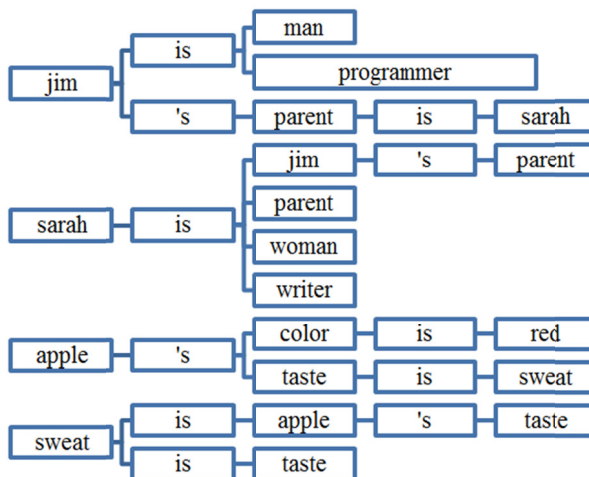*He/PRP is/VBZ programmer/NN ./.*

Figure 2. Interpret iterpereted sentences to the parse tree

As you see after arrival sentence it will be inserted in the pars tree word by word continuously with the POS-tag for every word and it will grows gradually. The important point is that saving same semantic sentence with different structure

Apple's taste is sweat = sweat is apple's taste

In this equation both of sentence have same mean with different structures, that both of them will be saved to facilitate searching for conclusion, this role is responsibility of the interpreter.

3.5 Words Dictionary

Each word before adding to the parse tree will be saved in a queue with unique code. The queue items including a listed structure to keeping the addresses of usage of the word in the main parse tree. In this case in the next returning the usage addresses of the word can be fetched immediately. It will be filled with new addresses too.

3.6 Pattern Recognition

Pattern recognition is a branch of machine learning that focuses on the recognition of patterns and regularities in data, although it is in some cases considered to be nearly synonymous with machine learning. Pattern recognition systems are in many cases trained from labeled "training" data (supervised learning), but when no labeled data are available other algorithms can be used to discover previously unknown patterns (unsupervised learning).

Pattern recognition can great help us in supervising machine learning, due to the POS-tag in NLP parser most of senteces can have same method to understand the subject, for example consider the example in below

Apple is a fruit

With the NLP Parser result the pattern is

A1/NNP is/VBZ a/DT A2/NN

Now it can conclude that the subset of A2 is A1.

With discovering this models classification can be done in a way that is significant for machines.

These sentences patterns are that stamp in our memories so that as soon as we hear the sentences we can understand it rapidly without knowing any syntaxes and react. Pattern recognition not only in the language but in most subjects, especially visual play an important role

The humans by seeing a simple image of an object can detect deformed shape of it easily, for example a chair with different shapes detectable for us because of its main pattern carved in our memory. The pattern recognition causes increasing the speed of understand, learning and inference.

3.7 Pronoun

There is a buffer structure to saving the last used mentioned nouns in the sentences, this structure is separate from the parse tree and individually acts and help to get the related nouns for each pronoun. in the sentences whenever have a name the related pronoun in the buffer will be filled.

3.8 Learning

Learning is the act of acquiring new, or modifying and reinforcing, existing knowledge, behaviors, skills, values, or preferences and may involve synthesizing different types of information. The ability to learn is possessed by humans, animals, plants and some machines. Progress over time tends to follow a learning curve. It does not happen all at once, but builds upon and is shaped by previous knowledge. To that end, learning may be viewed as a process, rather than a collection of factual and procedural knowledge. Learning produces changes in the organism and the changes produced are relatively permanent (Daniel et al, 2011).

Human learning may occur as part of education, personal development, schooling, or training. It may be goal-oriented and may be aided by motivation. The study of how learning occurs is part of educational psychology, neuropsychology, learning theory, and pedagogy. Learning may occur as a result of habituation or classical conditioning, seen in many animal species, or as a result of more complex activities such as play, seen only in relatively intelligent animals.

Learning may occur consciously or without conscious awareness. Learning that an aversive event cannot be avoided nor escaped is called learned helplessness.

Let to see an example, a kid doesn't have any information around the environment, in this stages the parents will teach them, for instance by repeating the "Go" word and helping the kids to walking they are trying to carving the "Go" word in its mind.

Then even we tell to them to go the "Go" word will be fetched then they start to walk. Whit this work we taught them to staring going after hearing the "Go" word.

Now consider this scenario for a robot, we tell to the robot "Go" this word have to be assigned to going action in robot. For this section we can use a console to do it and make it meaningful for the machine.

Go : this.move(forward)

This command plays the supervisor teaching role just like in kids training (Mohri et al ,2012). This kind of training continues until they do not require direct instruction between teacher and student, then the education is trained only with language.

For example telling find the chair, is unclear for a robot, the chair is defined object for machine but the finding command not defined yet. Now for teaching "Find" action it can point to the "go until seeing chair", a sentence that has already been taught to machine before. This training is a speech teaching and do not need any supervisor education. Thus, gradually the unclear sentence will be understandable for machine and robots unless needed to direct command console (supervisor)

Raw entry data divided INto three section

   1) News text

       The sentences after entering will be parsed and insert into the main tree

   2) Imperative text

The end of the command sentences have to be pointed to a real understandable command for the machine directly or indirectly, these command will be defined by the programmer and executable to machine. Purpose of indirect is that the sentence capable to parse to another pointed to real commands sentences. So, the compiler can have an access to the command and execute it.

For example "go" and "see" is pointed to "move" and "capture" command that is understandable and executable for machine, now "finding" action will be trained with "go" and "see" commands like this "go until see the object"

3) Question text

The questions which include discovering inference are some complex. Because in some cases have to search in some branches of tree to find the proper answer.

For instance finding out who is Jim's mother in our tree, required to inserted mother definition in Jim's node in the tree; although this is not defined yet. So based on the interpreter formula the question can be separated to the related questions, for example

Who is Jim's mother = who is Jim's (what is mother)

Who is Jim's mother = who is Jim's (mother is female parent)

Who is Jim's mother = who is Jim's (mother is female and mother is parent)

Now based on this separation the answer will be discovered. Another way is parsing main question pattern to sub question by supervisor and teacher.

Who are Jim's parent?

Which one is woman

Therefore, by answering these questions, the main answer will be fetched easily

3.9 Outcome

The outcome depends on accurate classification and appropriate approach to retrieving the data from parse tree.

There is a simple conclusion instance in bellow.

Banana is a yellow fruit

A1/NNP is/VBZ a/DT A2/JJ A3/NN =>

A1 is A3

A1/NNP is/VBZ a/DT A2/JJ A3/NN =>

A1's Parentof(A2) is A2

Banana is fruit

A1/NNP is/VBZ a/DT A2/JJ A3/NN =>

A1'Parent(A2) is A2

What are the yellow fruits?

What/WP are/VBP the/DT A1/JJ A2/NNS ./. =>

Childs(A2).property(parentof(A1)).value =A1

What is the banana's color?

What/WP is/VBZ the/DT A1/NN 's/POS A2/NN ?/. => A1.property(A2).value

According to the interpreters formula finding answer trained to machine. It should be considered the system responds easily and with high speed inference that the pattern has already been saved. In this case it far more quickly than the machine wants to trace the parse tree branches to finding the relations.

4. Conclusion

In this paper we have presented a new efficient evolution method for machine learning using pattern recognition in natural language processing. cause of the native natural language is reliable and more easy than the second language for all of us is that it doesn't use any syntaxes and grammar like the second language, we know that a kid never use syntaxes and grammars in its native language. In the native language the words and sentences carved as a pattern in our mind and it will be regulated as a syntax structure over time. Investigating this

difference conclude that teaching machine based on the native language approach is efficient and reliable.

References

Dale, R., Moisl, H., & Somers, H. (2000). Handbook of Natural Language Processing. CRC Press, AP - Business & Economics.

Dale, R., Wong, K. F., Su, J., & Kwong, O. Y. (2005). Natural Language Processing – IJCNLP 2005, Proceedings of Second International Joint Conference, Jeju Island, Korea, October, 11-13.

Daniel, L., Schacter, Daniel, T., Gilbert, D., & Wegner, M. (2011). Psychology, 2nd edition. Worth Publishers. p. 264. ISBN 978-1-4292-3719-2.

de Marneffe, M. C., MacCartney, B., & Manning, Ch. D. (2006). Generating Typed Dependency Parses from Phrase Structure Parses. In LREC, Stanford University.

Grosz, B. J., Jones, K. S., & Webber, B. L. (2010). Readings in natural language processing, Kaufman Publishers Inc., Los Altos, CA.

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). Foundations of Machine Learning, The MIT Press ISBN 9780262018258.

Toutanova, K., & Manning, Ch. D. (2000). Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger. In Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000), pp. 63-70.

Toutanova, K., Klein, D., Manning, Ch., & Singer, Y. (2003). Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In Proceedings of HLT-NAACL, 252-259.

Young, M. (1989). The Technical Writer's Handbook. Mill Valley, CA: University Science.

Copyrights