A Template-based Information Extraction System for Text Understanding

Dania Sagheer PhD Student Artificial Intelligence and Natural Language Dept Aleppo University, Syria

ABSTRACT

This paper presents a template-based information extraction system for Arabic descriptive text understanding. The system depends on knowledge base. The knowledge base contains facts and rules. The facts are derived from AL Khalil lexicon, Al Ramous lexicon and a Stanford model. The rules represent the designed templates. The templates are helpful for detecting the meaning of the text. the inference engine depends on the hybrid chaining to fill the slots in templates from the text. The semantic criterion is augmented to the templates. the criterion calculates the frequency of the template in the text. the system is tested on Arabic texts taken in oil production domain from Arabic news website as Arabic CNN, and Arabic BBC. The system implements good response in getting the goal of descriptive text. Text understanding is made efficiency, and high accuracy is obtained.

General Terms

Artificial Intelligence, Natural Language Processing.

Keywords

Text Understanding, knowledge Base, Information Extraction. Template.

1. INTRODUCTION

Recently, text understanding has become a very important task of natural language processing. Amount of text is increasing on the websites, and no time to read all text, so the need to automatic text understanding system is increased.

The descriptive text is a text which shows a given problem and the reader needs to read the whole of the text to understand what the goal of this text.

Information Extraction is an important strategy in text understanding task. The researches study how the knowledge is extracted from unstructured text. the researchers conducted many approaches for extracting of information from huge textual data as classification and clustering [1]. In [2] information extraction becomes more challenging because the text in social media is not structured according to the grammatical rules. The research work in [3] shows an information extraction system to link entities in the text with their corresponding entities in a knowledge base. The text can be represented in many approaches as a bag of words and a graph, query representation can link the entity in the text with query for build information retrieval system [4].

Information extraction can be performed by template based approach [5].

A template based approach is used in many researches for text understanding [6]. A template is a set of slots or variables that Fadel Sukkar Professor Artificial Intelligence and Natural Language Dept Aleppo University, Syria

describe an event in a specific domain [7] and presents helpful information for getting the goal of the text.

The research [8] presents a joint entity model to extract templates and slots from raw text.

To get matching between the entities in the raw text and the templates, information rich lexicons are needed. research teams in Arabic language present AL Khalil lexicon [9], [10], Al Ramous lexicon [11] and a Stanford model to give lexical vocabulary [12].

The paper is organized as follows: second section illustrates preparing of texts, third section shows designed templates. The fourth section presents the knowledge based system.in the fifth section the semantic inference rules are presented. Whereas the sixth section presents the work algorithm, in seventh section the results are shown, and finally in the eighth section the conclusion is illustrated.

2. PREPARING OF TEXTS

Descriptive texts are collected from formal news websites Arabic BBC, Arabic CNN, Syrian Arab News Agency SANA.

Collected texts describe the state of the oil production and explain oil organizations opinions about production of the oil.

The texts are segmented into paragraphs according to punctuation marks as point and comma. The paragraphs are split into their words.

3. DESIGNED TEMPLATES

The templates are designed according to the morphological, syntactical and vocabulary components of text sentences.

The designed templates are created for helping to get the goal of oil production text. the goal of oil production text is the knowledge about the organizations agreement to increase the oil production or decrease it.

The texts are studied and analyzed for finding the appropriate entities and designing the templates.

The entities of templates are choosed to detect the appropriate goal of the descriptive text.

The oil minister plays role in making decision about oil production, it is remarked that oil price is the important factor to detect amount of oil production, oil is measured by barrel. The organizations measure its production by amount of barrels or by financial values. Each country or organization sets a price for oil barrel.

The designed templates are:

		0		
		Entities		
Verb "فعل"	Oil minister وزير نفط"	Organization "منظمة"	Product -ion "الإنتاج"	
Produ- ction "الإنتاج"	Prices "السعر "	Oil availability "وفرة النفط"		
Verb "فعل"	Organiz- ation "منظمة"	Production "الإنتاج"	Amount "کمیة"	Unit "واحدة "
Verb "فعل"	Organiz- Ation	Production "الإنتاج"	Value "قيمة"	Unit "و احدة "

Barr-

el

'بر میل

Unit

اه احدة"

Table 1. Designed Templates

4. KNOWLEDGE BASED SYSTEM

To implement the template based information extraction system, the knowledge based system is done.

Value

"قبمة"

4.1 Knowledge Base

"منظمة"

Price

'سعر "

#

1

2

3

4

5

Verb

"فعل"

The knowledge consists of Al Khalil Lexicon, Al Ramous, and Stanford model.

Al Khalil lexicon contains facts about nouns, verbs, special words, prefixes and suffixes. The nouns include the vowled, unvoweled patterns and roots. The verbs also include the vowled, unvoweled patterns and roots, whereas the special words consist of tools and proper nouns. The prefixes are classified to nouns-prefixes, verbs-prefixes and nouns-verbsprefixes. The suffixes are classified as prefixes classification.

Al Ramous lexicon includes vocabulary and set of morphological rules applied to it. The number of vocabulary reaches to 50000 distributed to among 10000 verbs, and 40000 nouns and letters. each vocabulary contains morphological properties as root, type, is female or not, is plural or not, plural type, definition and usability of it.

In this research, some facts are added to Al Ramous as currency type, countries of oil production, and definitions to some vocabulary.

Stanford model presents part of speech of Arabic Language. The data set of the model contains close class included tools. pronouns and prepositions, and contains open class included verbs, nouns, adjectives which link to the position of the sentence. The model is trained by machine learning techniques as classification using forward and backward chaining.

The vocabulary and their morphological and syntactical properties represent the facts of the knowledge base, whereas the rules are represented by the designed templates.

The form of rules is if-then rules.

Example1: if ((word from text explains verb is true) and (word from text explains oil minister is true) and (word from text explains organization is true) and (word from text explains production level is true) then (template1 is true).

Example2: if (word from text explains Production is true) and (word from text explains prices is true) or (word from text explains Oil availability is true) then (template2 is true).

4.2 Control Strategy

Hybrid strategy of forward and backward chaining is used for retrieve the words of texts to the facts in the knowledge base. in order to fill the entities of templates with appropriate words from the text. to implement this purpose, the morphological rules are used as extract stem, root, check plural state and check type of the word.

For each template represented rule, because the texts contain many verbs, it is searched for the entities which refer to noun, then backward to look for the verb.

some entities as prices, it is searched for the morphological property stem then forward to link other entities.

5. SEMANTIC INFERENCE RULES

The semantic interface rules are created for expressing about extracted designed templates from text.

the semantic criterion is the frequency of the same template in the text. Principally the template contains selected entities for detect meaning. so recurrent same template in the text explains that the template is important.

The goal of the text is either increasing oil production or decreasing it. For helping in detecting this goal, some facts and rules are set.

Facts:

Fact1: decrease oil production from minister.

Fact2: decrease oil production outcome price.

Fact3: decrease oil production with amount

Fact4: decrease oil production with value.

Fact5: increase oil production from minister.

Fact6: increase oil production outcome price.

Fact7: increase oil production with value.

Fact8: increase oil production with amount.

The frequency of templates in text is taken with each template as semantic attribute.

Rules:

If (templat1 is true) then (Fact1(frequency of template1)) is true) or (Fact5(frequency of template1) is true).

If (template2 is true) then (Fact2(frequency of template2) is true) or (Fact6(frequency of template2) is true).

If (template3 is true) then (Fact3(frequency of template3) is true) or (Fact7(frequency of template3) is true).

If (template4 is true) then (Fact4(frequency of template4) is true) or (Fact8(frequency of template4) is true).

If (Fact1(frequency of template1) is true) or (Fact2 (frequency of template2) is true) or (Fact3 (frequency of template3) is true) or (Fact4 (frequency of template4) is true) then (decrease oil production is true), (frequency of decrease templates= frequency of template1+ frequency of template2+ frequency of template3+ frequency of template4).

.....

If (Fact5 (frequency of template1) is true) or (Fact6 (frequency of template2) is true) or (Fact7 (frequency of template3) is true) or (Fact8 (frequency of template4) is true) then (increase oil production is true), (frequency of increase templates=frequency of template1+ frequency of template2+ frequency of template3+ frequency of template4).

If (decrease oil production is true) and ((increase oil production is true) and (frequency of decrease templates>= frequency of increase templates) then the goal is decrease oil production.

If (decrease oil production is true) and ((increase oil production is false) then the goal is decrease oil production.

If (decrease oil production is false) and ((increase oil production is true) then the goal is increase oil production.

If (decrease oil production is true) and ((increase oil production is true) and (frequency of decrease templates<frequency of increase templates) then the goal is increase oil production.

According to the production entity state in each template, the appropriate fact be true.

6. WORK ALGORITHM

1- Knowledge base design.

- Facts: Al Khalil, Al Ramous, and Stanford model part of speech.
- Rules: designed (hand crafted) templates.
- 2- Preparing of texts.
- Collecting texts and preprocessing of it.
- 3- Control strategy
- Extract the appropriate words (information) from texts to fill the entities of templates.
- Retrieve the words of texts to the facts through morphological rules.
- Get the true templates for the text.
- 4- Design semantic inference rules for get the goal of the text.
- Determine true facts to reach to the true rule containing the appropriate goal of the text.

7. RESULTS

7.1 First Text (BBC-22/1/2017)

''أكدَ وزير الطاقة السعودي خالد الفالح أن منتجي ألنفط من داخل منظمة الدول المصدرة أوبك وخارجها ملتزمون بتعهداتهم بخفض الإنتاج

وأعرب الفالح عن اعتقاده بأن المخزون العالمي من النفط قد يعود إلى متوسط خمس سنوات بحلول منتصف العام الجاري مع الامتثال الكامل بالخفض. وقال الوزير السعودي: " لا يوجد ما يدعونا للقول على نحو مفاجئ مع دخول يناير، أننا بحاجة إلى خفض أكبر أو فترة أطول. هل هذا سيتغير في الربع الثاني؟ هذا ممكن، لكنه غير مناسب اليوم".

وجاء حديث الفالح في تصريحات للصحفيين قبيل الاجتماع الأول للجنة المعنية بمراقبة التزام الدول المنتجة للنفط باتفاق خفض الإنتاج.

ويبحث الاجتماع مدى التقدم الذي أحرز في الاتفاق المهم المبرم في ديسمبر كانون الأول الماضي بهدف رفع أسعار النفط.

يلزم الاتفاق 11دولة من خارج أوبك وفي مقدمتهم روسيا ب خفض الإنتاج بواقع 558 ألف برميل يوميا. وأوضحت موسكو أنها ستخفض الإنتاج بواقع 300 ألف برميل بحلول أبريل نيسان أو مايو أيار حسب ما ذكرت شبكة بلوم بيرغ الأمريكية.''

"Saudi Oil Minister Khalid Al-Faleh has condemned oil producers from within and outside the Organization of Petroleum Exporting Countries (OPEC) as being committed to cutting output.

Al-Faleh believed that global oil reserves could return to an average of five years by the middle of this year with full compliance with the reduction. "There is no reason for us to say abruptly as January enters, that we need a bigger cut or a longer period, will this change in the second quarter?

Al-Faleh told reporters before the first meeting of the committee to monitor the commitment of oil producing countries to the agreement to reduce production.

The meeting will discuss the progress made in the important agreement signed in December to raise oil prices.

The agreement requires 11 non-Opec countries, notably Russia, to cut production by 558,000 bpd.

Moscow said it would cut output by 300,000 barrels by April or May, according to Bloomberg."

Text is split into paragraphs as in table 2.

Table 2. paragraphs of first text

Paragraph num	Paragraph			
1	Saudi Oil Minister Khalid al-Faleh confirmed oil producers from within and outside the Organization of Petroleum Exporting Countries (OPEC) are committed to cutting output أكد وزير الطاقة السعودي خالد الفالح أن منتجي النفط من داخل منظمة الدول المصدرة أوبك وخارجها ملتزمون بتعهداتهم ب خفض الإنتاج			
2	Al-Faleh said that global oil reserves could return to an average of five years by the middle of this year with full compliance with the reduction. أعرب الفالح عن اعتقاده بأن المخزون العالمي من النفط قد يعود إلى متوسط خمس سنوات بحلول منتصف العام الجاري مع الامتثال الكامل بالخفض			
19	Moscow has said it will cut production by 300,000 barrels by April or May, according to Bloomberg وأوضحت موسكو أنها ستخفض الإنتاج بواقع 300 ألف برميل بحلول أبريل نيسان أو مايو أيار حسب ما ذكرت شبكة بلوم بيرغ الأمريكية			

The paragraphs are split into words. The system retrieves the words of the text to the entities in the templates.

For template1, it is searched for entity minister" تفزير "if the matching is then forward for the stem of oil "نفظ" or any of synonyms of oil from Al Ramous lexicon. The matching is got with the synonyms power "خلقة" then backward for searching for verb, AL Khalil morphological analyzer gets probability for be the word verb or noun. And the part of speech in Stanford model helps for determine the word is verb or noun. The verb" "أكذ" is detected. It is looked for oil organization or country in Al Ramous lexicon. "Opec" "أوبك" is determined so the template 1 is true.

For template2, it is searched in text for finding the root "ستعر" "price" through Al Khalil lexicon, then the" price" associate word "تراجع" "drop" is gotten, so the entity "prices" is filled by "تراجع" "drop", then the root of "production" "تراجع" is searched to link the entity "production" "الإنتاج" with "production" associate word "زيادة" "increase" it is remarked that oil availability isn't explained in the text and the template 2 is true.

For template3, the oil organization "Opec" "أوبك" is found through Al Ramous lexicon, then backward to search for verb, the entity "verb" is filled by "يلزم" "impose", the "production" entity is found in the same method in template1, so "production" entity is filled by "خفض" "decrease", then the amount of production is searched through ASCII code in order to look for numbers express about amount of oil production. Al Ramous is needed to detect the state of number hundreds or thousands through the definition, stem and form of plural and singular. In this text, "558 thousand" is linked to "amount" entity in the template. After that the measurement unit of oil production is looked for, "barrel" "بر ميل" is linked to through Al Khalil lexicon. So template 3 is true.

For template 4, the words in the text retrieve to the facts as in the previous templates and through Al Ramous lexicon for getting the currency vocabulary. And the template 4 is true.

The words of texts are linked to entities as follows in table3

Table 3. the words of the text linked to the entities

Entity	Word from text	Entity	Word from text
Verb "فعل"/	Impose "يلزم"/	/Minister "وزير "	Power minister "وزير الطاقة"/
/Value "قيمة"	558 thousand/ "ألف358"	organization ''منظمة''	''أوبك''/ Opec
Unit/ "واحدة"	Barrel ''بر میل''/	/Production "إنتاج"	"خفض"/Decrease
/Verb "فعل"	"تأمل"(/Hope	Currency/"عملة"	''دولار ''/Dollar
/Value ''قيمة''	1.8 milion/"1.8 "مليون	/Production "إنتاج"	''تز اید''/Increase
prices "سعر "/	drop "تراجع"/	''فعل''/Verb	"أعرب"/Said

The templates of this text is are illustrated in table4

Table 4. templates in the first text

#	Entities				
1	Verb	Oil	Organization	Product-	

	"فعل"	minister	"منظمة"	ion	
				"-15:NI"	
		ورير نفط"		, مِ تَتَاج	
		Power		Decrease	
	Said	minister	Opec	''خفض''	
	"أعرب"	"وزير الطاقة"	"أوبك"		
	Confir-	Power		Decrease	
	m	minister	Opec	''خفض''	
	"أكدّ"	"وزير الطاقة"	"أوبك"		
	Produ-	Prices	Oil		
2	ction	I Hees	availability		
	"الإنتاج"	السعر ا	"وفرة النفط"		
$ \top$	increase	drop			
	"زيادة"	"تراجع"			
	Maula	Organiz-	Due du etiere	Amount	Unit
3	verb	ation	Production	"الكمية"	"واد
5	"فعل"	"منظمة"	"الإنتاج"		دة"
				550	Dom
	Impose	Opec	Decrease	338	-el
	·· :ل:	المراي"	"· · · · · · · · · · · · · · · · · · ·	thousand	· · · "
	پترم	اوبت		558 ألف	بر د یل"
				221	Barr
	Drop	Opec	Drop	thousand	-el
	''تراجع''	"أوبك"	''تراجع''	221 ألف	"برم
					یل"
	Explain	Mosco	Decrease	300 thousand	Barr -el
	"أه ضب ه "	" 5	"خفض [»]	ممد ألف	"
	,ويسن	موسير	حصن	<u> </u>	بر۔ یل"
	Vorb	Organiz-	Production	Value	Unit
4	verb	Ation	r toudetion	"قيمة"	"واد
	''فعل''	"منظمة"	"الإنتاج"		دة"
				1.8	Doll
1			1		
	Hope	Opec	Decrease	million	-ar
	Hope "تأمل"	Opec "أوبك	Decrease ''خفض''	million 1.8"	ar- "دو لا

From table 3, It is noted that template1 is repeated twice, and template3 is repeated 3 times.

To return to semantic interface rules it is remarked the following true rules:

Fact1(frequency of template1) is true

Fact3 (frequency of template3) is true

Fact4 (frequency of template4) is true

(decrease oil production is true), (frequency of decrease templates= frequency of template1+ frequency of template3+ frequency of template4). Fact6 (frequency of template2) is true

increase oil production is true), (frequency of increase templates=frequency of template2).

decrease oil production is true, frequency of decrease templates=6.

increase oil production is true, frequency of increase templates=1

If (decrease oil production is true) and ((increase oil production is true) and (frequency of decrease templates>= frequency of increase templates) then the goal is decrease oil production.

From the last rule, it is resulted to the goal of the text that the text describes oil production decreasing state.

7.2 Second text (BBC-30/4/2017)

"قال وزير النفط الإيُراني بيجان زنكة أن أعضاء منظمة الدول المصدرة للنفط أوبك ودولاً من خارجها أرسلوا إشارات إيجابية لتمديد خطط خفض الإنتاج التي تدعمها إيران......

وتراجعت أسعار النفط في الأسبوع الماضي، على الرغم من إغلاقها على سعر أعلى يوم الجمعة مدفوعة بتزايد الأمال بموافقة أوبك على تمديد خفض الإنتاج لفترة أطول بما يكفي لتقليص الوفرة العالمية من النفط الخام."

"Iran's Oil Minister Bijan Zanka said members of the Organization of Petroleum Exporting Countries (OPEC) and other countries have sent positive signals to the extension of Iran-backed production reduction plans.

Oil prices fell last week, despite closing at a higher Friday price driven by growing hopes that OPEC would agree to extend output cuts long enough to reduce global crude oil reserves."

The text is separated to paragraphs, and the paragraphs are split into words.

The words in the text are retrieved to the facts in the knowledge base. and the templates for this text are extracted.

The templates of this text are illustrated in the table 5.

Table 5. templates in the second text

#			Entities		
1	Verb "فعل"	Oil minister وزير نفط"	Organization "منظمة"	Product- ion "الإنتاج"	
	Said "قال"	Power minister النفط"	Opec "أوبك"	Decrease ''خفض''	
2	Produ- ction "الإنتاج"	Prices "السعر "	Oil availability "وفرة النفط"		
	decrease "خفض"	higher "أعلى"	reducing ''تقلیص''		

To detect the goal of this text, the semantic inference rules are implemented.

Fact1(frequency of template1) is true

Fact2 (frequency of template2) is true

(decrease oil production is true), (frequency of decrease templates= frequency of template1+ frequency of template2).

decrease oil production is true, frequency of decrease templates=2.

If (decrease oil production is true) and ((increase oil production is false) then the goal is decrease oil production.

From the last rule, it is resulted to the goal of the text that the text describes oil production decreasing state.

7.3 Third text(CNN-13/4/2017)

" قال أحد المسؤولين: لا يوجد افتراح رسمي من المملكة العربية السعودية لكن هناك افتراح يجري تشكيله في لجنة المراقبة.............

خفّضت المملكة العربية السعودية وفق الاتفاق ليتراجع إنتاجها إلى أقل من 10 ملايين برميل يومياً.

و صممت التخفيضات في الإنتاج لدعم الأسعار وتخفيف الضغط على الميزانية الذي تعاني منه الدول المنتجة للنفط. ويبدو أن هذه الاستر اتيجية تنجح إذ استقرت أسعار النفط عند مستوى 50 دولاراً من أجل البرميل، بعد أن انخفضت أسعار النفط بشكل كبير العام الماضي إلى ما يقرب من 26 دولاراً للبرميل."

"One official said: "There is no formal proposal from Saudi Arabia, but there is a proposal being formed in the Oversight Committee.....

Saudi Arabia has cut its production to less than 10 million bpd.

The cuts in production were designed to support prices and ease the pressure on the budget experienced by the oilproducing countries. This strategy seems to be successful as oil prices stabilized at \$ 50 per barrel, after oil prices dropped sharply last year to nearly \$ 26 a barrel."

The templates of this text are illustrated in the table 6.

Table 6. templates in the third text

#			Entities		
2	Produ- ction "الإنتاج"	Prices "السعر"			
	decrease	support			
	"خفض"	"دعم"			
3	Verb "فعل"	Organiz - ation "منظمة"	Production "الإنتاج"	Amount "الكمية"	Unit "واح دة"
	approved "رافق"	Opec "أوبك"	Decrease "خفض"	1 million "1 مليون"	Barr -el بلرم
	Decrease "خفضت"	Opec "أوبك"	Drop "تراجع"	10 million 10 مليون"	Barr -el برم يل"

To return to semantic inference rules, it is noted the following true facts and rules.

Fact2 (frequency of template2) is true.

Fact3(frequency of template3) is true.

(decrease oil production is true), (frequency of decrease templates= frequency of template2+ frequency of template3).

decrease oil production is true, frequency of decrease templates=3.

If (decrease oil production is true) and ((increase oil production is false) then the goal is decrease oil production.

From the last rule, it is resulted to the goal of the text that the text describes oil production decreasing state.

8. CONCLUSION

Template based information extraction system is implemented for text understanding. The texts are collected from formal news websites. The texts describe oil production state and the oil organizations opinions about oil production.

The knowledge base is created from Al Khalil lexicon, Al Ramous lexicon, and Stanford model. The templates are designed for getting the meaning of the text. the words of text are retrieved to the entities of the templates.

The semantic inference rules are designed according to the templates with semantic criterion for getting the goal of the text.

The results of the system are similar to the human experience. Text understanding is made efficiency, and the goal of the text is got. In future work, the templates will be got automatically.

9. REFERENCES

- [1] Sun A., Naing MM., Lim EP., Lam W. 2003. Using Support Vector Machines for Terrorism Information Extraction. In: Chen H., Miranda R., Zeng D.D., Demchak C., Schroeder J., Madhusudan T. (eds) Intelligence and Security Informatics. ISI 2003. Lecture Notes in Computer Science, vol 2665. Springer, Berlin, Heidelberg.
- [2] Susanw. mcroy, Songsak channarukul and Syeds. ali. 2003. An augmented template-based approach to text realization, Natural Language Engineering 9 (4): 381– 420. 2003 Cambridge University Press.
- [3] Nathanael Chambers and Dan Jurafsky. 2011. Template-Based Information Extraction without the Templates Forman, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies

- [4] Lei Sha, Sujian Li, Baobao Chang, Zhifang Sui. 2016. Joint Learning Templates and Slots for Event Schema Induction, arXiv:1603.01333v1 [cs.CL].
- [5] SAID A. Salloum, AHMAD Qasim AlHamad, MOSTAFA Al-Emran, KHALED Shaalan, 2018. A Survey of Arabic Text Mining, Trends and Applications, Studies in Computational Intelligence, Springer International Publishing.
- [6] RIZWANA Irfan, CHRISTINE K. King, DANIEL Grages, SAM Ewen1, SAMEE U. Khan1, SAJJAD A. Madani, JOANNA Kolodziej, LIZHE Wang, DAN Chen, AMMAR Rayes6, NIKOLAOS Tziritas, CHENG-Zhong Xu, ALBERT Y. ZOMAYA, AHMED Saeed Alzahrani, HONGXIANG Li, 2015. A survey on text mining in social networks, The Knowledge Engineering Review, Vol (30:2) 157–170 P. Cambridge University Press.
- [7] Dr. Shubhangi D.C, RAVIKIRAN Mitte-2016. Knowledge based systems text analysis, International Research Journal of Engineering and Technology (IRJET) 0056 Volume.
- [8] CHENYAN Xiong., 2016. Knowledge Based Text Representations for Information Retrieval, Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Language and Information Technologies.
- [9] ABDERRAHIM BOUDLAL.; LAKHOUAJA A, MAZROUI A, MEZIANE A, OULD BEBAH M OULD abdallahi, SHOUL M, 2010. Alkhalil Morpho Sys1: A Morphosyntactic analysis system for Arabic texts, International Arab Conference on information technology, Benghazi, Libya, 1-6p.
- [10] BOUDCHICHE Mohamed, MAZROUI Azzeddine, OULD Bebah Mohamed Ould Abdallahi, LAKHOUAJA Abdelhak, BOUDLAL Abderrahim, 2017. AlKhalil Morpho Sys 2: A robust Arabic morpho-syntactic analyzer, Journal of King Saud University – Computer and Information Sciences, 141–146p.
- [11] http://arramooz.sourceforge.net. Access Date 1/1/2018.
- [12] https://nlp.stanford.edu/software/tagger.shtml, Access Date 1/1/2018.
- [13]