opción

Revista de Antropología, Ciencias de la Comunicación y de la Información, Filosofía, Lingüística y Semiótica, Problemas del Desarrollo, la Ciencia y la Tecnología

Año 35, 2019, Especial N° 20
Revista de Ciencias Humanas y Sociales
ISSN 1012-1587/ ISSNe: 2477-9385
Depósito Legal pp 198402ZU45

Universidad del Zulia
Facultad Experimental de Ciencias
Departamento de Ciencias Humanas
Maracaibo - Venezuela

Opcion, Año 35, Especial Nº 20 (2019):*2899-2921*
ISSN 1012-1587/ISSNe: 2477-9385

Unsupervised-Based Information Extraction from Unstructured Arabic Legal Documents

Khudhair J.Kadhim, Ahmed T. Sadiq, Hasanen S. Abdulah

Computer Sciences Department, University of Technology, Baghdad, Iraq.

Abstract

In order to make the unstructured or semi-structured traditional legal texts that are meet the requirements of high-level application such as A.I applications in legal, must overcoming on challenge how to extract and analyze structured information from the legal documents automatically. This paper proposes architecture that using a combined approach that utilizes features, lexical and rules based approaches to extract the needed information from traditional legal documents. This research uses a dataset that is collected from Iraq federal court of cassation decisions documents to extract two sets of information, the first is a set of general information, including reference law category, date of decision, court of jurisdiction name, and document no., decision type that are called valuables attributes information, and the document essence is a focused legal information that include principle, arguments, opinions legal, and facts of the case which can used in any analysis phase. This research is a part of big project entitled "The Arabic documents opinion extraction using argumentation mining", and the preliminary results were quite promising.

Extracción de información basada en supervisión de documentos legales árabes no estructurados

Resumen
Para que los textos legales tradicionales no estructurados o semiestructurados que cumplan con los requisitos de la aplicación de alto nivel, tales como las aplicaciones de A.I en legal, deben superar el desafío de cómo extraer y analizar información estructurada de los documentos legales automáticamente. Este documento propone una arquitectura que utiliza un enfoque combinado que utiliza características, enfoques léxicos y basados en reglas para extraer la información necesaria de los documentos legales tradicionales. Esta investigación utiliza un conjunto de datos que se recopila de los documentos de decisiones del tribunal de casación federal de Iraq para extraer dos conjuntos de información, el primero es un conjunto de información general, que incluye la categoría de ley de referencia, la fecha de la decisión, el nombre del tribunal de jurisdicción y el documento no. , el tipo de decisión que se denomina información de atributos de valores, y la esencia del documento es una información legal enfocada que incluye principios, argumentos, opiniones legales y hechos del caso que pueden usarse en cualquier fase de análisis. Esta investigación es parte de un gran proyecto titulado "La extracción de opinión de documentos árabes utilizando la minería de argumentación", y los resultados preliminares fueron bastante prometedores.

1. Introduction

Texts included in the legal documents are express about conditions what is permitted, forbidden or mandatory, which written in natural language in the context they regulate [1]. These traditional decisions and judgments of court or legal texts are unstructured documents, which cannot meet the requirements of high-level application of legal such as A.I applications in legal and other. On the other side, automated legal applications need to process formal information to achieve its desired aims.

The growing area of applications that are concerned with legal domain, such as AI systems which are backing coming to legal inferences or formulating legal argument or legal advisory systems and other, which aims to overcome difficulties that facing decision makers in recognize and analysis problems. The search operations in law area is time consume. Whatever

their purpose, whether to retrieval information from texts of legal documents or analyze the content of those documents to support final decisions or both [2]. In addition, the manually achievement any of those aspirations is a very costly job, therefore became motivate to highlight the importance of Information Extraction (IE) as a key Phase or critical factor in enhance and improve the performance of those applications.

The IE problem is referring to the automatic extraction of structured information from noisy unstructured textual sources. IE problem from text documents is an important and a research topic in different computer science communities, such as: databases, information retrieval, and artificial Intelligence [4].The effectiveness of IE methods from documents greatly are affect by layout, content and relations between them of the documents under consideration[3]. Although this is a general domain problem, it has a special relevance in the legal domain.

The nature of writing documents that using with proposed approach in this paper is impose an additional challenge to the known challenges (legal domain and Arabic language) in IE, plus the sentence boundary detection challenge. Target documents in this paper are written as paragraphs each one in a single sentence only without punctuation indicate for boundaries of sentences or phrases except rarely and many different layouts used in each paragraph (will be explained later in 3).

Extracting, distinguish, and identifying particular information from cassation decisions documents which handling in this paper, is possible for the expert, but in same time is range limited and time exhaustion. The task seems impossible or at least difficult from point of view who are outside the specialty.This paper tries to answer the following question: How to exploit combines the document structure with multi-techniques of NLP to extract a set of general information and important parts of legal text such as principle, opinions, legal reasoning, and facts of the case automatically in court cassation decisions?

The answer will be through the approach which followed in this paper.The remaining paper sections is structured as follows. Section 2, relevant previous works discussed. Section 3, highlights about the paper application domain. Section 4, presented the proposed approach design. Section 5, discusses the proposed approach implementation. In last two sections, a brief explaining about evaluation of the proposed system and the conclusions.

2. Related Works

The task of extracted and identify information in legal documents treatment automatically is an urgent need for law applications area, which received

investigated extensively in the literature. But, noteworthy is the scarcity in similar works which related to Arabic legal documents if compared with their counterparts in English or other Languages and is still open problem. There are several approaches that related to the proposed approach which are listed as follows:

In 2010, [8] approach try to develop a mechanism that support effective search to identified and extracted the relevant information to characters' role and events in a case semi-automatically from collection of legal documents, which support lawyers' activities in corporate litigation. In information extraction phase from this approach is using NLP tool and information of entity related and identifying other factors (topic, time and location of event and character role in event). The approach based on Xerox Incremental Parser (XIP) as a for linguistic processing such as chunking, dependencies, NER, morphological analyzer, POS and patterns over chunks sequences where event recognition. In 2016, the proposed approach [7] is utilize the IE techniques to improving legal information retrieval from Arabic legal texts (Tunisian laws). Approach is relying on ontology domain with tools of NLP in the process of IE. Approach is used a general structures schema to identify general class for legal document and the structure elements of class with taken into consideration the relations between these types, because the references to other texts is exist. By applied the syntactic grammar for extracting "table of contents" of document, that allow to recognize the key terms in document's headers, which in the end is information about the document's main topics. In order to overcome challenges of the natural language in legal texts, this approach used an especially NLP tool called "NooJ" combine with morphological a dictionary for Arabic legal terms to extract the matching sequences and builds a concordance. In 2017, approach in [1] proposed unsupervised technique, which combines different NLP techniques to automated extraction of rules from Australian legal "Telecommunications consumer protections code" documents. Exploits WordNet to process variability in natural language expressions of legal texts, and relying on the Stanfordparser to obtain the grammatical representation of the sentences. In other words, uses methodology which combines syntactic-based together with a logic-based, that exploited the logical dependencies between chunks of text. Results of evaluation for this work are rely on compared with manually rules which produced by an analyst. In 2017, approach in [3] proposed an environment called (CLIEL). The main idea in this approach is organize the document text in a way accessible manner, then easy extraction the information. This work

used mechanism relying on documents layout detection phase and a set of JAPE grammar rules to extract information from commercial law documents. First, applies a rule-based NLP model to split documents based on section boundaries of text (e.g. titles and indexes), and Gazetteer for names of entities to produces annotate documents file in XML format in a hierarchical data structure. Each text unit is representing by node, where the sections headings in level, the subsections headings in the next level and so, the leaves contain the actual text. The second, are applied to parsed the annotations for specific information which will extracted and stored. In 2017, the work in [5] proposes a model that exploit knowledge-supporting of a domain in IE process as a key factor, call it controlled domains, which refer to relatively a free and open text domain e.g. news articles. This work focus on handling Arabic texts to extraction the atomic concepts and relations in the beginning, while concepts and relation composite extract using description logic (DL) later. The proposed model uses pattern-matching techniques to employs the syntactic elements and semantic elements to extracting concepts and them relations and to overcome language variation which can be appear from syntactic pattern, thus used description logic in validated of information accuracy which extracted.

3.Application Domain of Proposed Approach

The absence of Arabic legal texts corpus and unavailable support sources for law Arabic documents, that give a motivation to build corpus concerned with one of legal documents types, which will besharing it in the future after increase the number of documents.

This paper targets a set of 60 paper documents of Iraq Federal Court of Cassation decisions paper documents. They are collected randomly while ensuring diversity of cases, references laws, and formats. Manually converted to digital copies. The corpus of documents used increases complexitydue the format, writing style, and language factors. Different format of documents context is arising from difference in reference laws (cross-references). In terms of the writing way these decisions are written, each section is represented by a paragraph, written Modern Standard Arabic (MSA) in a free human language and does not use punctuation in most of the document sections, so each section can be visualized as a long sentence.IE suffers from challenges of Arabic texts like other technique from high inflectional, rich morphology, diversity and other features in Arabic language, that could be considered a reason for lack of Arabic-based related work.

Although most cassation decisions documents tend to be organized in

semi-standard sections, but the variety in the ways of present, arrangement, and content of information depending on the type of court of jurisdiction and law. cassation decisions document is often divided into four basic sections with variable lengths, in addition document number. First section includes information about category of governing or reference law and date of commission formation. Second section includes names of plaintiff and defendant, has been overlooked in this work (The privacy is the main reason behind its ignoring. Its belong to the first type of information extracted "general information". That can be directly extracted, it is clear and free of filler text). The third, its contain the name and decision of the court of jurisdiction along with the main reason for litigation and the facts of the case. The last, for the cassation commission opinion and the legal principle based, supported by legal arguments.

4. The Proposed Approach design

The proposed work methodology is adopting combine document's structure and NLP techniques to extract the legal information from legal texts in unstructured documents. In particular, approach implemented exploits many elements to achieve its goals, which are:

i) Document's structure to identify types of information extracted. document's structure is one of the valuable sources to information extracted process from court of cassation decisions. In actuality, distinguishing between the layout or hierarchical of document helps to automatically extraction for general information types and other information derived from these documents texts.

The order of parts in document's structure is the feature exploited by this proposed approach, which used as a guide in determine the type of information will be extracted from each part. e.g. the part or first paragraph text of the document includes information on the category of the reference law of the case and date of the Commission formation in court of cassation. This approach cited the work [3] which used layout feature of documents but this approach represents each part in document as a node in a tree include only one level after the root, due the nature of the documents content organization as indicated in advance, whose parts tend to be semi-independent. It supports parallel processing of the document. Although there is type of correlation between the last two parts of decisions documents, which is represented by description of some events of the case that occasionally referred to in the decision of the cassation Commission. The proposed approach ignores this relationship for several reasons. Their presence in the documents is not guaranteed, to reduce complexity, and

importantly the proposed approach is interested with " The main reason for litigation " of the case as a more general representation for type of legal cases Which are belong to the same legal subdivision rather than focusing on details. e.g. the cases belong to "إبطال حجة الإذن" Revoke argument of permission.

ii) Small lexical dictionary contains general specific terms (not necessarily legal terms) that are extracted manually. Generally used as guidance indicators for candidate text zone to extract information or to sentences and phrases separation in legal texts. The location and use of these terms varies from section to section in the document structure. It can be at the beginning or end of the sentences, it can be used in patterns and information extraction rules or only as a lamp in the path. with considerations the paradigms of inflectional and derivational of terms to bypass the semantic complexity. e.g. the term "المحكمة .... تصدر الحكم", "The court issued ..... a judgement" in third section of the document structure is a tag refers to the end of the facts of the case and beginning text that containing the jurisdiction court decision.

iii) Techniques and libraries of NLP for parsing tree, part of speech (POS) tag for grammatical representation of sentences in order to recognize them. Stanford library is used for doing the syntactic part.

The proposed approach comprises Three main steps, shown in Figure 1

1- Identification the document structure parts:

This step is responsible any section identification of the document' structure that is being processed currently. To content-based section identification, this step used predefined key terms. the output of this step used by small lexical dictionary to determines key terms that relates to the treated section currently. and determines the types of information extractedfrom the selected section.

2- Text processing: This step is preprocessing includes many sub-steps

A. Sentences and phrases separation:

This process is due to the targeted documents texts properties in the current work, no punctuation in text to allows to splitting. Combine the sentence parser tree which generate using Stanford library and a lightweight grammar rule are exploit in phrases separation.

B. Arabic text normalization, Tokenization, POS Tagger, stemming:

in This step executed set of processes to initialize the text for processing in next steps, include the normalization is important process and focuses on text Arabic, is process to noisy remove from text Arabic and transforming to standard format. As well as for text stemming processing is grouping

different terms with same meaning and forms into standard form. POS Tagger is used by the lightweight sentences patterns

3- Information Extraction: This step to recognize and extracted the phrases that contain Information requiredfrom the previous step output. syntactic grammar is apply by rely on lightweight sentences patterns that alteration depending on section identification and in the presence of the lexical dictionary and the sections information's types identifier.

4- Store extracted information: store extracted information: This step to index each document legal decision and the information which extracted.

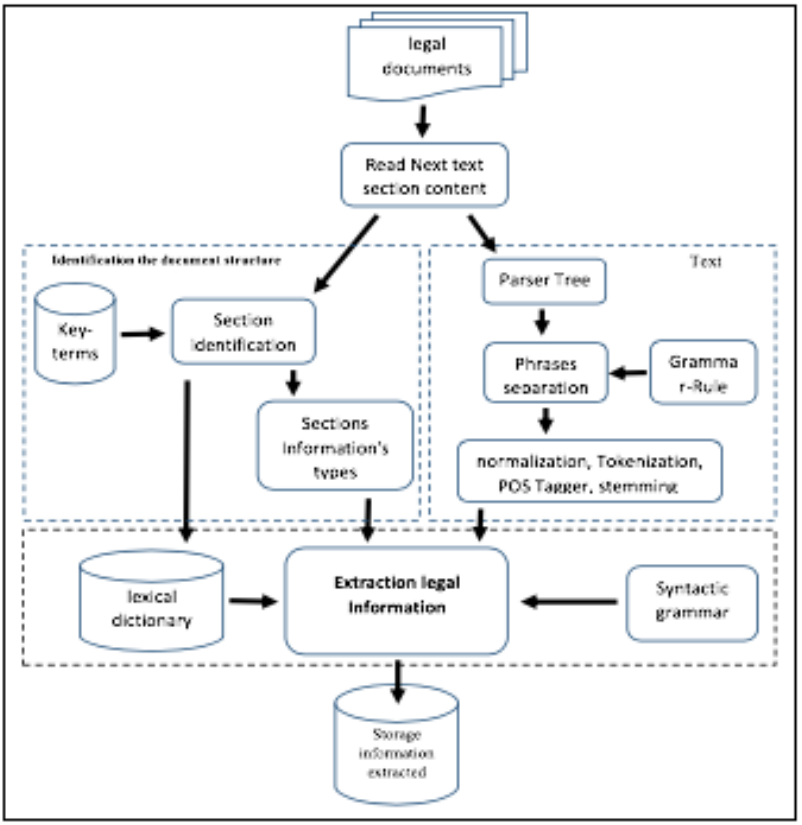

Fig1. The Proposed Approach Architecture

5. The Proposed Approach Implementation

The following algorithm 1 is illustrating a general steps of proposed algorithm, where the input of algorithm consists from lexical dictionary,

Grammar-Rule, Syntactic patterns and Legal document text and the output involve the two types from information extracted Valuable attributes and Analytic information.

Algorithm 1: Legal Information Extraction

Input : Legal document text, Keyword dictionary, Syntactic patterns

Output : Valuable attributes and Analytic information

Begin

1. Read legal document text
2. For each section in legal document text
   3. Section Identification (Document structure, Keyword dictionary)
   4. Parser Tree creation (raw text of section)
   5. Phrases separation (Parser Tree, Grammar-Rule)
   6. For each Phrase in Phrases separation
      NLP-Process Phrases (Phrases)
      Extract-information (Phrases, Syntactic patterns, lexical dictionary)
      Storage of extracted information ( )
   Next
   Next
7. End

The proposed algorithm after read legal document text is executing many procedures for each section from document as follows:

1- Section Identification: this procedure based on document structure to identify the section in document used the Keyword to distinguish them section. other sub procedures are Implicit executed once identify section from document structure. Any information which will be extracted from the selected section by Sections Information's types ( ), and Any Keywords which used to refers to locations of those extracted information by Sections Keyword identification ( ).

2- Phrases separation: this procedure based on document contents. Started by call the NLP libraries to create sentences parser tree using Stan-

ford library (supports the Arabic language) to represent sentences and phrase in text. This approach exploits parser tree to splitting the section's sentences depend on each initial symbol S, sub-S and Coordinating conjunction. In addition, using the Stanford library to produce Tokenization, POS Tagger. This procedure used other techniques of NLP to remove the noisy from text Arabic, using the normalization and stemmer to transformation terms in different groups to unification standard format.

3- Extract-information: These steps are applied on each phrase in each document section. using the keywords that section specify to refer for the zone which can be contains the target information, one or more of keywords can be appear in text zone and can be located at start, end, middle, or both. Therefore, the using syntactic patterns helps in specific what is benefit from text. It's like a shallow parser which take POS of each tokens of text zone (bigrams, trigrams) and looking for a specific pattern such as (NN+JJ) or (NN+NN+JJ) with trigrams.

6. The Proposed Approach result and Evaluation

This paper targets a sample of legal documents texts to experiments consists of 60 documents from Iraq federal court of cassation decisions, collected and distributed randomly over some classes of the laws references. The Figure 2. presents a sample and structure of the court of cassation decisions document and the table. are shows the two type of information which extracted from that document text.

There is no similar previous works in the same specialization or content can be used to results evaluation. Therefore, used to evaluate a website of Iraq Supreme Judicial Council, that offers an online search service to retrieval some cassation court decisions , which are indexed manually.

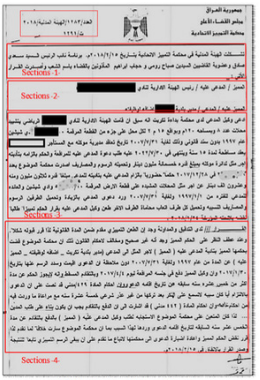

Fig.2. Decisions Sample Court of Cassation

This information extracted manually include the general information (reference law category, Date of decision, judgment issuer), judgment legal principle and decision text (the entire fourth section from any document). The results evaluation of the proposed approach is distributed into two levels. The first level the extracted the valuable attributes information that used the precision and recall measures and the average of results were 93% and 89% respectively.

• Recall: How many relevant valuable attributes are selected? divided by the total number of existing relevant valuable attributes.

• Precision: How many selected valuable attributes are relevant? divided by the total number of valuable attributes.

The second level the extracted deductive or analytic legal texts, where evaluate results relies on the compared with text extracted manually in website (as shown in Table 2) and present to legal specialists, the results were 85%. The reason is the approach is extract not just the legal principle but the reasons for it from decision text in the form of isolated sentences and pronouns compensate problem in text.

7. Conclusions

This paper is an attempt to founded an efficient approach for information extract from unstructured Arabic legal texts automatic, using unsupervised approach. This paper takes into account the document's structure is a valuable source to determine the relations between target information and using techniques. The combine NLP techniques, lexical terms, and syntactic patterns is supports promising results.

The main idea is exploit document's layout to identify texts sections of legal document, consequently can be determination what the target information in text, and which any keywords which used as a information locations indicators and which any syntactic pattern can be applied to extract the final information from each section in legal document. Can be present the research contributions in brief:

- It's produces two types of extracted information from Arabic legal documents the valuabled attributes and analytic information.
- Improve the indexing, retrieving, and clustering the legal documents automatically.
- To provides a useful foundation and making legal information more available and transparent for legal domain applications.

Type of information will be extracted	Type of information required	values of information extracted
Valuable attributes (specific information called data point)	No. document	1283
	Date of document	2018
	Law References	هيئة الاحوال الشخصية و المواد الشخصية
	Date of commission formation	2018-09-07
	Jurisdiction court Name	الاحوال الشخصية في المشخاب
	Main reason for litigation	بالزامه بالنفقة المستمرة لهاو لابنتها
	Cassation State	قرر قبوله شكلا
	Court Cassation Opinion	وجد انه غير صحيح و مخالف لاحكام الشرع والقانون
Analytic information (part of legal text)	Facts or Brief of the Case	ان المدعى عليه زوجها قد تركها بدون نفقه او منفق شرعي هي و طفلتها لذا
	Jurisdiction court decision	ت محكمة الموضوع حكما حضوريا ب عدد 82/ ش / 2018 و تاريخ 2018/2/14 حكما يقضي اولا / برد دعوى المدعية بخصوص مطالبتها ب النفقة المستمر ه لها و
	Legal principles	i.ذلك ان المدعى عليه لم يهيا بيتا شرعيا للمدعية تتوفر فيها الشروط الشرعية والقانونى. ii.ان الثابت في جلسات المرافعة ان المدعية تسكن مع زوجها المدعى عليه في بيت اهله. iii.لذا فانها تستحق النفقه بنوعيها الماضية والمستمرة

Table 1. The Information that Extracted by Approach

Approach Extraction	Manually Extraction	URL	Raw Text of Decision
لان الزوجة تستحق كل المهر المسمى بالدخول استنادا للمادة (21) من قانون الاحوال الشخصية رقم 188 لسنة 1959 المعدل	تستحق الزوجة كل المهر المسمى بالدخول استنادا للمادة (21) من قانون الاحوال الشخصية رقم 188 لسنة 1959 المعدل.	https://www.hjc.iq/qview.1780/	لدى التدقيق والمداولة تبين ان الطعن التمييزي واقع ضمن المدة القانونية لذا قرر قبوله شكلا ولدى عطف النظر على الحكم المميز فقد وجد انه صحيح ومخالف للقانون ذلك لان الزوجة تستحق كل المهر المسمى بالدخول استنادا للمادة (21) من قانون الاحوال الشخصية رقم 188 لسنة 1959 المعدل وافق للشرع والقانون لذا قرر تصديقه ورد الطعون التمييزية وتحميل المميز رسم التمييز وصدر القرار بالاتفاق في 18/شوال/1433هـ 2012/9/5م.
ذلك ان الضرورة الملجئة حالة خاصة بالموجر تفرض عليه ولا يسعى اليها وان قيام الموجر (المميز عليها) بايجار الدار بتاريخ 2002/10/24 وسكنه في دار شقيقة زوجته منذ عام 2002 وعودتها الى العراق وترغيب السكن في دارها المشغولة من قبله لا يعتبر من قبيل الضرورة الملجئة كما ان تحسن الظروف الامنية هي حالة عامة وليست خاصة بالموجر. وعليه تكون الضرورة الملجئة غير متحققة في دعوى المدعي (المميز)	الضرورة الملجئة حالة خاصة بالموجر تفرض عليه ولا يسعى اليها فسكنى الموجر في دار شقيقته وقيامه بايجار داره للغير فعودة شقيقته وطلبها سكن دارها لا تعد ضرورة ملجئة وكذلك تحسن الوضع الأمني حالة عامة وليست خاصة بالموجر.	https://www.hjc.iq/qview.1222/	لدى التدقيق والمداولة وجد أن الطعن التمييزي مقدم ضمن المدة القانونية قرر قبوله شكلاً ولدى عطف النظر على الحكم المميز وجد بأنه غير صحيح ومخالف للقانون ذلك ان المدعي(المميز عليه) قد استند في دعواه على الضرورة الملجئة المنصوص عليها في الفقرة (12) من المادة السابعة عشرة من قانون ايجار العقار رقم 87 لسنة 1979 من انه يسكن الدار العائدة لشقيقة زوجته الواقعة في حي المغرب منذ عام 2002 لسفرها خارج العراق وكونها عادت وترغب السكن بدارها لتحسن الظروف الامنية وان محكمة البداءة قضت في حكمها المميز بتخلية الدار موضوع الدعوى اذ رأت ان الضرورة الملجئة متوفرة في الدعوى مستندة الى ان عودة شقيقة زوجة المدعي كان بعد واقعة ايجار المدعى الدار بالاضافة الى تحسن الوضع الامني . ان ما ذهبت اليه المحكمة غير صحيح ولا يستند على اساس قانوني سليم ذلك ان الضرورة الملجئة حالة خاصة بالموجر تفرض عليه ولا يسعى اليها وان قيام الموجر (المميز عليها) بايجار الدار بتاريخ 2002/10/24 وسكنه في دار شقيقة زوجته منذ عام 2002 وعودتها الى العراق وترغيب السكن في دارها المشغولة من قبله لا يعتبر من قبيل الضرورة الملجئة كما ان تحسن الظروف الامنية هي حالة عامة وليست خاصة بالموجر. وعليه تكون الضرورة الملجئة غير متحققة في دعوى المدعي (المميز) ولما كانت المحكمة قد خالفت ذلك مما اخل بصحة حكمها المميز . لذا قرر نقضه واعادة الاضبارة الى محكمتها لاتباع ما تقدم على ان يبقى رسم التمييز تابعا للنتيجة وصدر القرار بالاتفاق في 8/رمضان/1429هـ الموافق 2008/9/8م.
ذلك ان الماجور خاضع لاحكام قانون ايجار العقار وان المادة العاشرة فقرة (1) من القانون تنص على تدفع سلفا اجرة العقارات المشمولة باحكام هذا القانون باقساط شهرية بصرف النظر عن مدة سريان عقد الايجار.	ان المادة العاشرة الفقرة (1) من قانون ايجار العقار تنص على تدفع سلفا اجرة العقارات المشمولة باحكام هذا القانون باقساط شهرية بصرف النظر عن مدة سريان عقد الايجار ويقع باطلا كل اتفاق يقضي بخلاف ذلك وعند امتناع المدعى عليه عن دفع اجرة شهر كانون الاول سنة 2013 خلال مدة ثمانية ايام من تاريخ تبلغه بالانذار المعمير اليه فللموجر حق طلب تخلية الماجور.	https://www.hjc.iq/qview.2100/	بعد التدقيق والمداولة تبين ان الطعن التمييزي واقع ضمن المدة القانونية قرر قبوله شكلا ولدى عطف النظر على القرار المميز وجد انه غير صحيح ومخالف للقانون. ذلك ان الماجور خاضع لاحكام قانون ايجار العقار وان المادة العاشرة فقرة (1) من القانون تنص على تدفع سلفا اجرة العقارات المشمولة باحكام هذا القانون باقساط شهرية بصرف النظر عن مدة سريان عقد الايجار ويقع باطلا كل اتفاق يقضي بخلاف ذلك وعند امتناع المدعى عليه عن دفع اجرة شهر كانون الاول سنة 2013 خلال مدة ثمانية ايام من تاريخ تبلغه بالانذار المعمير عن طريق كاتب عدل في الزبير بدفع اجرة الشهر المذكور لذا من حق الموجر طلب تخلية الماجور لسبب المذكور . عليه قرر نقضه واعادة الى محمتها لاتباع ما تقدم على ان يبقى الرسم تابعا الى النتيجة وصدر القرار بالاتفاق في 14/ جمادي الاخر/ 1435 هـ الموافق 2014/4/14م.

Table 2. The Information that Extracted by Approach

References

[1] M.Dragoni, S.Villata, W.Rizzi, and G.Governatori, 2016, Combining NLP Approaches for Rule Extraction from Legal Documents,1st Workshop on MIning and Reasoning with Legal texts (MIREL), France.
[2] M.Francine,E.Boiy, R.Mochales, and C.Reed,2007,Automatic Detection of Arguments in Legal Texts,11th International Conference on Artificial Intelligence and Law (ICAIL), CA USA.
[3] M.García,K.Atkinson, D.Bollegala, K.Chapman, F.Coenen, C.Roberts, and K.Robson,2017,Context-Based Information Extraction from Commercial Law Documents ,17th International Conference on Artificial Intelligence and Law (ICAIL), London, United Kingdom.
[4] Cortez.E, Da Silva.A, 2013,Unsupervised Information Extraction by Text Segmentation, Springer, New York Dordrecht London.
[5] M.Fasha, N.Obeid, and B.Hammo, 2017, A Proposed Model for Extracting Information from Arabic-Based Controlled Text Domains, New Trends in Information Technology (NTIT), Jordan.
[6] G. Buey, C.Roman, A.Garrido, C.Bobed, and E.Mena, 2017, Automatic Legal Document Analysis Results of Information Extraction Processes using an Ontology, 23rd International Symposium on Methodologies for Intelligent Systems (ISMIS),Warsaw, Poland.
[7] I.Bouaziz, F.Gargouri, 2016, A Framework for the Annotation of Arabic Legal Documents, 27th International Business Information Management Association Conference (IBIMA), Milan, Italy.
[8] N.Lagos, F.Segond, S.Castellani, and J.O'Neill,2010, Event Extraction for Legal Case Building and Reasoning, Springer, FIP Advances in Information and Communication Technology.
[9] M.AL ACHKAR, R.Mahmoudb, 2012, Arabic Legal Information System for Case-Law Indexing and Retrieval, Egyptian Computer Science Journal(ECS), Egyptian.
[10] T.Tin, J.Leonard, M.Davies, K.Gerard, and R.Edita, 2009, Information Extraction from Legal Documents, IEEE, Eighth International Symposium on Natural Language Processing, Philippines.
[11] I.Bouaziz, F.Gargouri, 2015, Towards an Arabic legal ontology based on documents properties extraction, IEEE.

UNIVERSIDAD
DEL ZULIA

opción
Revista de Ciencias Humanas y Sociales

Año 35, N° 20, (2019)

Esta revista fue editada en formato digital por el personal de la Oficina de Publicaciones Científicas de la Facultad Experimental de Ciencias, Universidad del Zulia.
Maracaibo - Venezuela

www.luz.edu.ve

www.serbi.luz.edu.ve

produccioncientifica.luz.edu.ve