International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

A Noval Model to Extract Medical Records Processing & Identification using Multi Agent System

Nishakar Kankalla[1], Shubhaker Bandari[2]

[1]Associate Professor, ECE Dept, St Martin's Engineering College, Secunderabad. T.S. -500100.
[2]Assistant Professor, ECE Dept, BVRIT College of Engineering for Women's, Hyderabad - 500090.

Abstract: *Medical files and perception papers have consistently been a significant wellspring of information. Shockingly, the majority of the occasions they are as yet put away as actual reports either printed or manually written, subsequently making it difficult to move this valuable data starting with one spot then onto the next, or concentrating and removing new information from it. However, with these days headways in software engineering this issue might be taken care of. In this paper we propose a patient focused Multi-specialist framework project that can extricate pertinent data from the patients' wellbeing records and store that information on a concentrated information store, in light of a predefined philosophy conspire. The framework's motivation is to normalize and enhance the information by performing different mining undertakings on the given content. The final reason for existing is to give diverse clinic divisions an apparatus which they can question for valuable data about the patients' drug, therapies as well as booked tests and tests and get ideas in regards to their treatment goals.*

***Keywords*- Multi-agent System, eHealth, Knowledge-based System, Ontology, Text-mining.**

1. INTRODUCTION

For quite a long time specialists in emergency clinics have been saving their perceptions on paper, either printed or manually written. This has been making it difficult to move valuable data starting with one medical clinic then onto the next, and ordinarily even starting with one office then onto the next, on account of patients with a few conditions.

Be that as it may, with the advancement of PCs and rapid Internet, the actual exchange quit being an issue, however the extraction, comprehension and normalization of the informa-tion in the records turned out to be increasingly more an emphasized need.

In this paper we attempt to handle a piece of this complex challenge, by proposing a specialist based framework project equipped for changing crude information, into normalized designs, and afterward separating significant information like medication connection or connections among sicknesses and side effects.

Our examination attempts to help the requirements of the clinical work force from the Romanian emergency clinics, who have the vast majority of their patient records on paper, and might want to have them concentrated and organized. We have been given, by the clinics from Timisoara, with examined clinical records from patients, as jpeg and pdf. Every one of the archives are written in Romanian with or without diacritics.

In the first stage the framework gets the records and changes them to plain text[1]. The subsequent stage is described by the normalization utilizing an area specific metaphysics produced for this specific case. The specialist parses the report searching for specific examples, in view of rules and normal articulations, as these sorts of records are semi-organized, and afterward in the final stage it takes care of them to an appropriate data set and parts of the data to the files utilized in text mining. The issue is taken care of by a specialist that fulfills different assignments like general connection perceptions, measurable relationships and term affiliations.

2. BACKGROUND AND RELATED WORK

Text mining or Knowledge-Discovery in Text (KDT) tech-niques [2] are applied to various composed sources, to naturally remove valuable examples in the unstructured text based information. These examples, for our situation patient's clinical information, can be utilized in programmed clever dynamic cycle.

Utilizing triplestores, for putting away the information removed from the patients' clinical records, was propelled as a result of the benefits that proposals over the standard data sets [3]. Among which we can specify:

- flexibility - triplestores are very flexible, as they do not require complex schema development and the data can be easily loaded and queried.

- Standards - all triplestores share the same standards, thus allowing the easy access from a

different health facilities.

- provenance - stored data contains the semantics factsabout the data sources.

An approach for ontology-driven system for clinical information exchange is described in [4]. The authors propose an information model instantiated as a middle layer between the Electronic Health Record (EHR) and Personal Health Record(PHR).

The information model addresses the semantic and syntactic interoperability issues between these two systems.

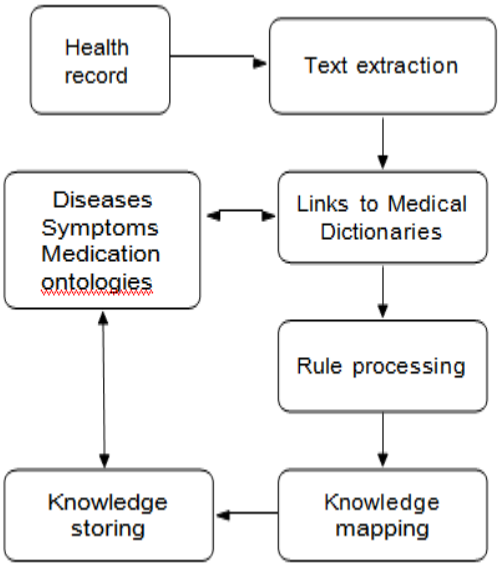

Figure 1. System workflow.

Also an ontology-based framework for the medical records is proposed in [5] by the authors with the intent to facilitate the exchange of the information and knowledge between different EHR systems and hospital departments. But in our case patients observation files are recorded manually on paper or in a text file, then later inserted into a database. This is a laborious task, so our proposed system is trying to automate this process, thus maximizing the medical personnel's time with the patients, reducing expenses and increasing the quality of the healthcare services.

3. ARCHITECTURE

The proposed Multi-Agent framework (see Figure 1) has four sorts of specialists which can deal with information extraction, putting away and questioning: Text Mining Agent, Knowledge Store-Agent, Hospital Departments specialists (in this work we use Radiology Agent) and Knowledge Discovery Agent.

TextMiningAgent, we utilize this kind of specialist for extricating information from clinical records (see Figure 2). It stacks the area specific cosmology created for this specific case which contains semantic comment for infections, indications, medicines and prescriptions. Data extraction depends on Named Entity Recognition (NER) approach. Removed information is put away locally and communicated to the KnowledgeStore-Agent.

KnowledgeStoreAgent, this specialist gets messages, that contain patients' data, from the TextMiningAgent, using the developed ontology schema it uploads that knowl- edge, as a RDF, to the triplestore. Agent's interaction with the ontology is achieved using SPARQL query language.

***HospitalDepartments agents*,** one of these type of agents exists for each hospital department. Their responsibility is to query centralized triplestore, in order to obtain useful information. For example, *RadiologyAgent* can check if there are newly hospitalized patients scheduled for CT scan.

KnowledgeDiscoveryAgent, this agent is addressed for Remote Patient Monitoring Systems (RPMSs), and it's main responsibility is to connect to the hospital knowledge base in order to update personal system with the new knowledge about the patient's health, which can be important during the treatment and/or recovery process.

Knowledge Base

The proposed framework utilizes area specific ontologies for information portrayal. These ontologies were com-bined into a solitary one, HospitalOnt, which is utilized by the TextMiningAgent to perceive clinical documentation for sicknesses (ex. Parkinson, Depression), side effects (ex. postural shakiness, quake), meds (ex. Pk Merz, Aspirin), from the clinical records. The framework additionally utilizes a cosmology model, HospitalKBSchema, for taking care of and putting away the gathered information as RDF in triplestore.

As you can find in Figure 3, the metaphysics information is minuscule right now (just 6 infections with their cor-reacting manifestations, paraclinical assessments and prescription ication are accessible), however it is continually evolved as data from the specialist is given to us. The cosmology has a job in advancing the substance too. For instance in the event that we get designs that have "and" ("s, I" in Romanian),

it will realize that there are two sicknesses or two side effects and separate them appropriately: "strolling and balance issues" ("tulbura˘ri de mers s, I echilibru" in Romanian) will be put away in the information base as "strolling jumble" ("tulbura˘ri de mers"), "balance issue" ("tulbura˘ri de echilibru"). This underlying information will give the vital contribution to a managed learning calculation (for example Contingent Random Fields [6]).

4. SYSTEM IMPLEMENTATION AND EVALUATION

The framework model has been carried out utilizing the Scala programming language [7], Akka Framework [8] for specialists' displaying, Apache Jena Framework [9] for the collaborations with the ontologies and triplestore, and Fuseki Server [10] for putting away the information extricated from the clinical records.

The crude information comes as a jpeg picture, or pdf design, and is changed through OCR into text. We should comment that here we had two cases, the one wherein the report machine composed, and the OCR changed them effectively, and the manually written (or composed and overlaid penmanship) records that required really preparing, and for the present were set to the side for additional work.

FOAIE DE OBSERVATIE CLINICA GENERALA

Nume: Ionescu Ioan CNP:1234567890123_ Sex: M Varsta: 67 Diagnostic:

1. Boala Parkinson
2. Dementa mixta
3. Hipertensiune

Motivele internarii: tulburari de memorie, tremor la nivelul membrelor inferioare, declin cognitiv, insomnie, tulburari de echilibru si mers

Paraclinic:

EKG

Ex.psihologic

RMN si Rx

Lab. biochimic

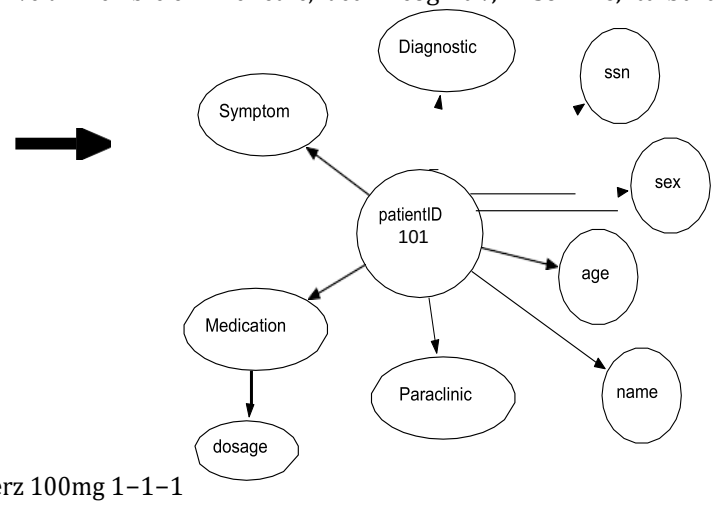

Figure 2. Knowledge extractions from the medical records.

Anamneza:

a. Antecedente personale: Boala Parkinson

Medicatie de fond administrata inaintea internarii: Pk merz 100mg 1–1–1

Madopar 250mg 1–1–1

Frontan 0,25mg 0–1–0

Aspirin protect 100mg 1–0–0

Istoricul bolii: Pacient in varsta de 67 de ani cunoscut cu Boala Parkinson std II de aprox 3 ani, hipertensiv si cu dementa mixta se interneaza in Clinica Maria pentru tulburari de memorie, tremor la nivelul membrelor inferioare, declin cognitiv, insomnie, tulburari de echilibru si mers accentuate in ultimele 2 luni. Se efectueaza investigatii clinico–paraclinice si stabilirea conduitei terapeutica de specialitate.

Along these lines, as information comes out as plain content, it should be curated. To do message investigation, we use are executing a clinical acknowledgment substance

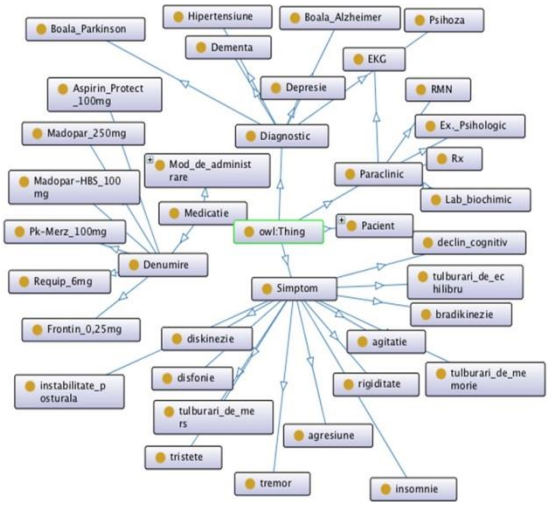

Figure 3. Hospital Ontology Model

framework, in a mixture structure (both AI and space information), as it has been appeared to give the best outcomes [11].

We have additionally experienced some semantics issues (thing phrase recognition is not the same as the English one, and furthermore the manner in which the words are framed. For instance, "hipertensiune" ("hypertension" in English) comes from the root "tensiune" and the prefix "hiper", yet then there are likewise varieties like ”hipertensiv/hipertensivă” (”hypertensive” in English), and medics use all forms, and sometimes even abbreviations like ”htensiv” or ”hta”, that must be taken into account. We intend to solve them by using the Romanian WordNet Ontology [12].

The triplestore is used for two purposes: the first one is for simple querying (see Figures 4 and 5); while the second one is to deliver data for analysis

	patientID	name
1	A:patientID101	"Ionescu Ioan"
2	A:patientID102	"Ionescu Ana"

Figure 4. Patients scheduled for CT scan. Query performed by the Radiology Agent.

and machine learning. By having enough information the system will be able to make suggestions to the doctor. When he types his observations, if two drugs he wrote interact he will be notified.

	medication	dosage
1	A:Pk_Merz_100mg	A:1-1-1
2	A:Aspirin_Protect_100mg	A:1-0-0
3	A:Madopar_250mg	A:1-1-1
4	A:Frontin_0_25mg	A:0-1-0
5	A:Pk_Merz_100mg	A:1-1-1

Figure 5. Patient’s medication treatments. Query performed by the Knowledge Discovery Agent.

Another example would be the system listing some diseases related to the symptoms he introduced. Visualization is also an important part in understanding results, especially when it comes to people who don’t have a mathematical expertise.

That is the reason our framework is intended to create diagrams and designs at whatever point conceivable (see Figures 6 and 7).

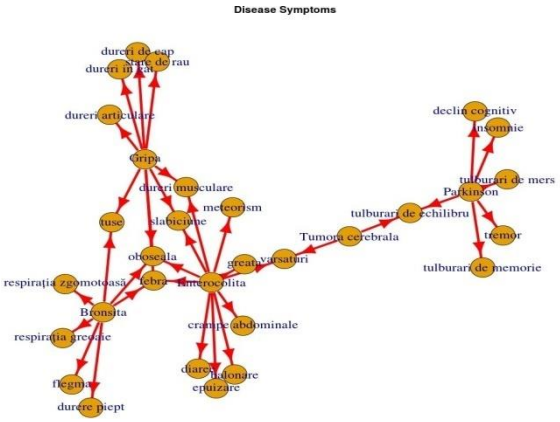

Figure 6. Disease - Symptoms relations.

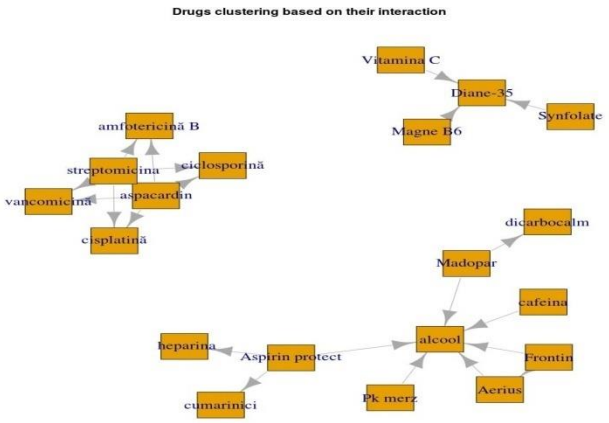

Figure 7. Drugs clustering based on their interaction.

Be that as it may, to give this application to the clinical staff, the framework must first examine the information. That is the reason it will be required for each offered issue to separate its relating exchanges, that will come as a rundown of sets (e.g: [Parkinson, tulburari de memorie], [Enterocolitaˇ, varsaturi], [Parkinson, tremor], and so on) On them we can apply different information mining techniques, still what gives us the most important outcomes in this consideration are affiliation rules in their old style structure for little to medium datasets [13] or improved for enormous (large informational indexes) [14].

The explanation we are utilizing affiliation rules, is that these itemsets (like the one in the model above) structure a characteristic cross section structure (a subset/superset structure): "Parkinson" and "Enterocolitaˇ" ("Enterocolitis") are essential for the illness bunch while "quake" and "vaˇrsaˇturi" ("heaving") are important for side effects bunch, yet they meet framing this net. we are keen on finding the frequencies of happening designs (called support, that must be more prominent than a given edge) and the strength of suggestion (otherwise called confidence). The final object is to recognize successive itemsets (for this we will utilize an Apriori calculation [15]) and recover just the ones that are critical to us (by figuring the "interest", which is the reliance of the given sets [16] and applying pruning calculations [17]).

5. CONCLUSIONS

In this paper we depicted a patient focused Multi-specialist framework design, that can separate important data from the patients' wellbeing records and store that information on concentrated datastore, in light of the predefined philosophy plot, accordingly giving every emergency clinic division a total and modern perspective on their patients status. We persuaded our use of the triplestore approach showing the benefits that it offers over the standard information bases, similar to information interoperability, flexibility and simplicity with regards to questioning. A framework won't ever have the option to supplant the human aptitude, however, as appeared, it tends to be an incredible assistance with regards to recovering data from enormous datasets and summing up or finding significant perceptions.

With respect to future work we expect to improve measurable outcomes and investigate the chance of interfacing other existing ontologies or information bases to our framework. Additionally, we expect to make the framework more autonomous via preparing a profound convolutional neural organization for penmanship recog- nition and change to the normalized mode. This will likewise permit us to deal with data from files and improve understanding from an earlier time, assemble more information and cause the framework to learn new things (counting drug results on patients over the long run, measurements, or side effect illness advancement).

REFERENCES

[1] A. Batbarai and D. Naidu, “Approach for rule pruning in association rule mining for removing redundancy,” Int. J. Innov. Res. Comput. Commun. Eng, vol. 2, no. 5, pp. 4207– 4213, 2014.

[2] F. Berzal, I. Blanco, M. Vila et al., “Measuring the accuracy and interest of association rules: A new framework,” Intelli- gent Data Analysis, vol. 6, no. 3, pp. 221–235, 2002.

[3] M. Al-Maolegi and B. Arkok, “An improved apriori algo- rithm for association rules,” International Journal on Natural Language Computing (IJNLC), vol. 3, no. 1, pp. 21–29, 2014.

[4] G. Qian, C. R. Rao, X. Sun, and Y. Wu, “Boosting association rule mining in large datasets via gibbs sampling,” Proceedings of the National Academy of Sciences, vol. 113, no. 18, pp. 4958–4963, 2016.

[5] C. Zhang and S. Zhang, Association rule mining: models and algorithms. Springer-Verlag, 2002.

[6] D. Tufis¸ , E. Barbu, V. B. Mititelu, R. Ion, and L. Bozianu, “The romanian wordnet,” Romanian Journal on Information Science and Technology, vol. 7, no. 2-3, pp. 105–122, 2004.

[7] A. B. Abacha and P. Zweigenbaum, “Medical entity recog- nition: A comparison of semantic and statistical methods,”in Proceedings of BioNLP 2011 Workshop. Association for Computational Linguistics, 2011, pp. 56–64.

[8] “Apache Jena Fuseki2 - SPARQL Server,” URL: https://jena.apache.org/documentation/ [accessed: 2016- 06-28].

[9] “Apache Jena Framework,” URL: https://jena.apache.org/index.html [accessed: 2016-06-28].

[10] "Akka - Toolkit and runtime for building highly concurrent, distributed, and resilient message-driven applications on the JVM." URL: akka.io [accessed: 2016-06-28].

[11] M. Odersky, L. Spoon, and B. Venners, Programming inscala. Artima Inc, 2008.

[12] C. Sutton and A. McCallum, An introduction to conditional random fields for relational learning. Introduction to statis- tical relational learning. MIT Press, 2006, vol. 2.

[13] C. González, B. Blobel, and D. M. López, "Ontology-based framework for electronic health records interoperability." Studies in health technology and informatics, vol. 169, pp. 694–698, 2010.

[14] P. Plastiras, D. O'Sullivan, and P. Weller, "An ontology- driven information model for interoperability of personal and electronic health records," 2014.

[15] M. Dimitrov, "Semantic technologies and triplestores for business intelligence," in Business Intelligence. Springer, 2012, pp. 139–155.

[16] V. Gupta and G. S. Lehal, "A survey of text mining techniques and applications," Journal of emerging technologies in web intelligence, vol. 1, no. 1, pp. 60–76, 2009.

[17] D. Meyer, K. Hornik, and I. Feinerer, "Text mining infras- tructure in r," Journal of statistical software, vol. 25, no. 5,pp. 1–54, 2008.

AUTHORS:

1. Mr. K. Nishakar received M. Tech in Embedded Systems & VLSI Design from JNTUH Hyderabad, Telangana. in 2010. Currently pursuing Ph.D. degree from Utkal University, Bhubaneshwar in the specialization of Image Processing. He has 12 years of teaching experience.. **Mr. K. Nishakar** is working as an Associate Professor in ECE Department in St. Martin's Engineering College, since 2012. He has published more than 2 0 international journals and conferences, and Life member of IETE, ISTE and IEANG. His research areas Medical Image Processing and Wireless Communications.

2. Mr. B. Shubhaker received M. Tech in Systems and Signal Processing from JNTUH Hyderabad, Telangana. in 2012. Currently pursuing Ph.D. degree from Osmania University, Hyderabad in Radar Signal Processing. He has 10 years of teaching experience.. **B. Shubhaker** is working as an Assistant Professor in ECE Department in BVRIT College of Engineering for Women's, Hyderabad. He has published more than 1 0 international journals and conferences, and Life member of IETE and ISTE. His research areas Radar Signal Processing.