Knowledge-Based Systems 234 (2021) 107578

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Intelligent knowledge consolidation: From data to wisdom

Musarrat Hussain [a], Fahad Ahmed Satti [a], Syed Imran Ali [a], Jamil Hussain [b], Taqdir Ali [a], Hun-Sung Kim [c], Kun-Ho Yoon [c], TaeChoong Chung [a,*], Sungyoung Lee [a,*]

[a] Department of Computer Science and Engineering, Kyung Hee University, South Korea
[b] Department of Data Science, Sejong University, South Korea
[c] The Catholic University of Korea, Seoul St. Mary's Hospital, South Korea

ARTICLE INFO

Article history:
Received 18 December 2020
Received in revised form 4 October 2021
Accepted 4 October 2021
Available online 6 October 2021

Keywords:
Knowledge consolidation
Incremental knowledge model
Ripple down rules
Knowledge based system

ABSTRACT

Knowledge based systems have accomplished remarkable achievements in assisting evidence based decision making for complex problems. However, machine learning-driven, intelligent systems of today are dependent on the underlying knowledge model, which is acquired from domain experts, or the available datasets in a structured or unstructured format. Most of the existing literature utilized a single modal, while very few have combined multi-modalities (mainly two) for knowledge acquisition. In order to achieve a strong Artificial Intelligence, multi-domain and multi-modal knowledge acquisition, and consolidation is required. This paper presents the research work, driving the realization of such a comprehensive framework, in the field of healthcare. Using area specific, state-of-the-art machine learning techniques, we first extract knowledge from structured and unstructured data, which is consolidated with expert knowledge and managed through ripple down rules. Our presented technique shows an accuracy of 92.05%, which is much higher than single modal deep learning at 78.20%, naive bayes at 69.70%, logistic regression at 61.20%, expert driven knowledge at 86.02%, and naive knowledge combination at 70.86%. Thus, through the application of our proposed technique, we provide the foundations for an accurate and evolvable knowledge-base, that can greatly enhance decision making in the healthcare domain.

1. Introduction

The availability of healthcare data and advancement in computing technologies have enabled access to knowledge from a plethora of sources. Artificial intelligence plays an important role in knowledge acquisition and intelligent decision making. However, knowledge acquired from single modality produce limited perspective and quality compare to multi-modalities [1,2]. This limitation also affects the state of the art machine learning approaches, which mimic typical information processing flows, by creating a set of interpretable rules, which conform to some initial conditions and restrictions. The created rules are then used to convert the data (such as user input or sensory feedback) into information (such as predicted labels) and extract knowledge (which can be fed back to the system, as in case of active learning techniques) from them. However, in order to break these shackles, it is necessary to look at various perspectives and utilize scientific ideas to deduce the true form of the data. Such an approach would then remove the adhoc restrictions that prevent the usage of real world data beyond information and basic knowledge to deep knowledge extraction and wisdom creation.

Consider, as an example the previous paragraph; Looking at each of the sentences in the paragraph independently can create confusion for the reader. However, in its true form, and within the context of this paper, this paragraph enhances the traditional definition of knowledge, from "justifiable true belief based on the subject acquired from education or experience of a person" [3] to "justifiable true belief based on the subject acquired from education **and** experiences of **many**". Over time, and through the application of many feedback loops, such knowledge can become trust-worthy to create Artificial Wisdom, which will in-turn form the basis for a strong Artificial Intelligence.

Such an AI would require the enhancement of both depth and breadth of knowledge sources, producing comprehensive algorithms for handling multiple domains and scenarios. Many state-of-the-art machine learning software is now focused on

The code (and data) in this article has been certified as Reproducible by Code Ocean: (https://codeocean.com/). More information on the Reproducibility Badge Initiative is available at https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals.

* Corresponding authors.
E-mail addresses: musarrat.hussain@oslab.khu.ac.kr (M. Hussain), fahad.satti@oslab.khu.ac.kr (F.A. Satti), imran.ali@oslab.khu.ac.kr (S.I. Ali), jamil@sejong.ac.kr (J. Hussain), taqdir.ali@oslab.khu.ac.kr (T. Ali), 01cadiz@hanmail.net (H.-S. Kim), yoonk@catholic.ac.kr (K.-H. Yoon), tcchung@khu.ac.kr (T. Chung), sylee@oslab.khu.ac.kr (S. Lee).

https://doi.org/10.1016/j.knosys.2021.107578
0950-7051/

resolving the breadth of knowledge sources by producing generic models, which are able to give good results in specific cases. Recent examples such as BERT [4] and GPT-3 [5] have proved the applicability of these approaches in the Natural Language Processing tasks. Depth of knowledge bases can be enhanced through the application of knowledge consolidation.

Knowledge consolidation is a fusioning process by which knowledge from various sources can be merged to provide a consistent and inclusive decision making workflow, for improving the quality and scope of results. This process takes into account the heterogeneity of data sources, resolves their conflicts, and reconciles the discovered value into true values for ensuring consistency [6] and durability [7] of the knowledge bases. In this way, the depth of the knowledge base can be enhanced, providing a cache for knowledge and automated wisdom.

In the domain of healthcare, while the creation of Hospital Information Management Systems (HMIS), Clinical Decision Support Systems (CDSS), and other tools have seen massive improvements, the creation of an Artificial Medical Intelligence (AMI) is still far away. Literature points to a multitude of reasons behind this delay, including the heterogeneity of data sources [8], difficulty in expanding current solutions [9], the necessity for decision transparency [10], privacy of patient data [11], and many others. Resultantly, existing healthcare solutions suffer from a lack of quality and applicability, unless substantial effort, in terms of expert's time and energy is spent on identifying and creating knowledge [12]. Even then, an expert-only sourced solution would be subjectively biased based on the expert's limited experiences and preferences [13]. Additionally, there is no guarantee of the solution getting long term acceptance [14], unless the knowledge base is self-maintainable and evolve-able.

Alternatively, data driven decision making provides a low quality but evolvable knowledge source [2]. However, due to the very stringent nature of decision making in healthcare, data driven solutions must be merged with expert driven solutions (at least to the extent of verification and validation). The relationship between levels of intelligence and the decision risk associated with it is shown in Fig. 1 [15]. As we move from data to wisdom, while the possible options for reaching a conclusion decrease, the confidence of decision making increases. With enough information processing and reusability of established insights, wisdom allows the computing system to make faster and accurate decisions. Therefore, this process is strictly dependent on how well the raw data has been processed to become information, information processed into knowledge, and wisdom determined through knowledge. Since computing has already witnessed great progress in the first two levels (data $\rightarrow$ information and information $\rightarrow$ knowledge), it is imperative that the focus should now shift towards the third level (knowledge $\rightarrow$ wisdom). As with intelligence, there is no clear and agreed upon definition of the wisdom. Generally its "the right use of knowledge" [16]. Baltes et al. [17] noted that "Wisdom is associated with good judgment and actions that contribute to living well". For this study, we considered wisdom as the ability to determine a preferable recommendation among all valid results using Ripple Down Rules (RDR) methodology as explained in Section 3. In the absence of this assistive, wisdom based decision making tool, the physicians utilize traditional methods stemming from 19th century specifications, which requires a very long diagnosis process, which is dependent on the experience and intuition of the primary physician or their close peers [18]. A data driven decision making solution, (like a recent study for detecting diabetic retinopathy using a deep learning algorithm [19]) has already started to outperform clinicians.

In order to build wisdom from knowledge, the first step is to design, create, and evaluate a framework for knowledge consolidation through the use of multi-modal sources and novel algorithms. The framework should employ state-of-the-art techniques for acquiring knowledge from multi-model data sources including image, structured, unstructured, and expert heuristic for knowledge acquisition, and consolidate the acquired knowledge to a single easy-to-use and maintainable knowledge model. In this paper, we present such a framework, which acquires rules from both expert and data, consolidates them using a novel algorithm, and provides maintenance and evolution opportunities through the use of ripple down rules (RDR). The applications of the framework is evaluated in the domain of diabetes diagnosis. The final result produced by this framework is a high quality, accurate, and transparent knowledge base, which can be used for identifying insights and create human like artificial wisdom. The novel contribution of our proposed research work is twofold.

- Firstly, we designed an integrated architecture of multi-modal knowledge acquisition including structured, unstructured, legacy, and expert sources and their consolidation.
- Secondly, the knowledge acquired through the previous process has been consolidated into a uniform knowledge base using RDR with a novel implementation strategy, which is presented in Section 4.2.

The detailed source code[1] of knowledge acquisition, consolidation, and inferencing are being made available publicly to assist interested readers in reproducing the system and can re-purpose it to other related domains and applications. The rest of the paper is structured as follows. Section 2 provides related work, Section 3 describes the detail about RDR knowledge representation, inferencing and evolution. Section 4 explains the detailed methodology used for knowledge acquisition and consolidation, followed by results achieved and their significant in Section 5. Finally, Section 6 concludes the paper.

2. Related work

The digital healthcare system contains a plethora of loosely coupled data and information management systems. Systems, such as HMIS, Medical Internet of Things (IoT) related applications, Physiological sensors, and Medical Picture Archiving and Communication System (PACS) have been engineered to provide largely foundational support to managing the physician and patient encounters. Patients Data, clinical practice guidelines, online published papers, and clinical notes, are also used to acquire clinical knowledge and provide decision making support for the medical experts. The clinical knowledge can be acquired from patients data (data driven), domain experts (expert driven) or the combination of both (hybridized/consolidated knowledge). Montani et al. [1] studied the prevalence of data driven AI in CDSS focusing on published material between 2017–2018. The authors have evaluated 75 articles, where 65% (49 studies) of all the studies were focused on data-driven AI, 20% were purely expert driven while 6% were hybrid approaches. The study concluded that the hybrid systems which consolidate/fusion expert-driven approach with data-driven, performed better as compared to other approaches. A plenty of knowledge hybridization systems are available in may domains now, in healthcare, transparency of knowledge is still a major requirement [10]. As a result, most current systems forego knowledge consolidation due to its complexity, maintenance problems and difficult verification. However, these systems provide the foundation for initiating and/or progressing a knowledge consolidation model and are necessary to ve reviewed before moving ahead. A summary of some of the existing approaches, their domain, fusion scope, features and limitations are shown in Table 1.

[1] https://github.com/Musarratpcr/KnowledgeConsolidation.

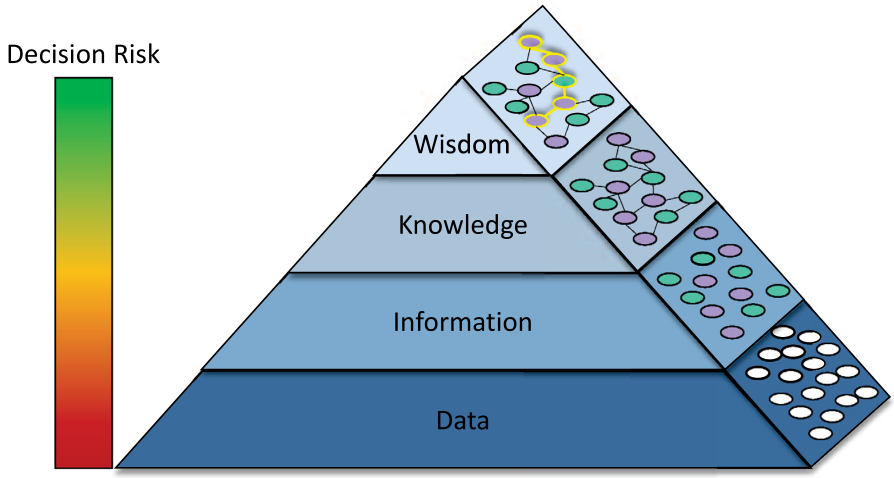


Fig. 1. Decision risk associated with DIKW pyramid [15].

Table 1
Summary of related studies.

Study	Domain	Fusion scope	Key techniques	Limitations
S.L. Ting et al. [20]	Medical prescription	Case base reasoning + Bayesian reasoning	• Similar patients identification using case based reasoning (CBR) • Gather physicians experience in prescribing drugs • Produces the best prescription recommendations	Utilizes two separately generated recommendations to produce the final recommendation
M. Berkan et al. [21]	Lung cancer care	Rules based + Probabilistic decision support	• Similar patients identification using CBR • Physicians' experience identification using Bayesian reasoning • Final prescription based on commonalities between the identified recommendation	The system generates drug recommendation using the results of two independent reasoning techniques
N larburu et al. [22]	Breast cancer	Guidelines knowledge + Experts knowledge	• Extracts breast cancer-related knowledge from CPGs • Convert experts decision for successfully treated patients into production rules • Augment CPG knowledge with expert knowledge taken	The system acquires and combines knowledge from two sources.
Ying Shen et al. [23]	Antibiotics prescription	Combined various disease ontologies into single ontology	• Generated a consolidated ontology from infectious diseases and antibiotic therapy ontologies. • Infers potential infectious disease from patient data • Proposes a relevant antibiotic therapy	Ontology maintenance and evolution is a well known problem
Soufi et al. [24]	Triage management	Rule based + Fuzzy logic	• Categories emergency severity index using a combination of rule based and fuzzy logic.	• Knowledge acquisition from two sources • Rule based approach used for emergency severity index and fuzzy logic for triage assessment

As mentioned earlier, the results achieved by hybrid CDSS are better than single model [1]. However, most of the existing systems either focused on data-driven or expert-driven knowledge for CDSS. There are very few systems that utilize hybrid and consolidated knowledge mainly because of its complexities and difficulties in multi-model knowledge acquisition and consolidation. As shown in Table 1, most of the existing hybrid knowledge base studies generate multi-step recommendations followed by a final recommendation. In the multi-step recommendations, the initial results are produced independently, which may not be as effective as in the case of knowledge consolidation. In knowledge consolidation, multi-model knowledge is being processed to produce a single knowledge model and generate more appropriate recommendations in a single step.

On the other hand, some of the studies shown in Table 1 have consolidated expert knowledge with data-driven knowledge and produce promising results. The performance of the system can be boosted by adding diverse modalities and state-of-the-art knowledge extraction techniques. Therefore, this study focuses on extraction knowledge from various modalities including image data such as Fundus images, x-ray, structured data such as EMR, EHR, unstructured data such as CPGs, clinical notes, legacy knowledge in the form of production rules, and expert knowledge and consolidate it to a single easy to use and maintainable knowledge model (RDR).

3. Ripple Down Rules (RDR)

Knowledge bases provide automated systems with a rich cache of memories and patterns, which are essential for making intelligent decisions. When more efforts are put into building and evolving a knowledge base from rich data sources or experts, then it improves the final decision output. However, the problems of knowledge transparency, maintenance, validation, verification, and consolidation, are just some of the major hindrances to the aim of building a medical AI [10]. Ripple down rules provides a knowledge acquisition and maintenance model based on the incremental modeling approach [25]. As mentioned in [26], the traditional knowledge base systems (KBS) mainly focused on knowledge acquisition, while ignored knowledge maintenance and its contextual representation. As said "rules are made to be broken" [26], therefore, RDR adopts incremental knowledge acquisition methodology [26]. It combines rule-based and case-based reasoning and unifies knowledge acquisition with knowledge maintenance, maintains the context, and converts the phrase into "rules are made to be corrected" [26]. It uses a failure driven approach, where knowledge is contextually patched to the rule that produced incorrect result for the case in hand. The case in hand is attached to new patched rule as cornerstone case, the case that cause the addition of the new rule to the knowledge base. RDR stores the acquired knowledge as a finite binary tree [25,27], where each node represents a rule of the form **IF** condition(s) **THEN** conclusion and can be expended by two distinct branches called except and if not [28]. An example of the RDR tree is shown in Fig. 2.

Starting from the initial situation where nothing is known, RDR defines a default rule (Node 1), which is always satisfied. Normally, the default rule represents the most common decision of the domain. The new rules will be added as child rule (except branch) on each incorrect result produced by a particular rule (node). The RDR follows the inferencing order from top to bottom and left to right during recommendation generation and knowledge evolution. The RDR inference engine starts from the top node (Node 1) and tests whether rule conditions are satisfied or not. If satisfied, it then evaluates the child nodes and will continue until there is no child node or no satisfied child node of the last satisfied node. The final result will be produced from the last satisfied rule. While, the new rule will be added to the Except branch of the final matched rule during the knowledge evolution.

The example of inferencing order during recommendation and knowledge evolution is shown in Fig. 2. let suppose the knowledge model in Fig. 2 contains a sample of five nodes (from Node 1 to Node 5), initially. A new case with conditions C_2, C_3, C_9 and C_{10} arrive to the system. Based on the current knowledge the RDR inference engine will produce "Class 2" as the final recommendation because the last satisfied rule is Node 3 of the model. However, the domain expert disagrees with this system generated recommendation, citing the reason that due to conditions C_9, and C_{10} the final recommendation should be "Class 5" using his intelligence and wisdom. Therefore, the system will evolve the knowledge model and will add a new rule **IF** (C_9 and C_{10}) **THEN** "Class 5" at "EXCEPT" branch of the Node 3. However as there is already an "EXCEPT" branch (Node 5) of the rule Node 3 therefore, the new rule (Node 6) will be added at the "IF NOT" branch of the Node 5. The input case (C_2, C_3, C_9, C_{10}) will be appended as cornerstone case for the newly added rule Node 6. The same process will continue on each incorrect result, evolving the knowledge, and improving the model performance in terms of accuracy, over time. In this way, the model incorporates human intelligence and wisdom and reflects it in making human-like intelligent, and wise decisions. The RDR knowledge acquisition and inferencing process flow can be found here [26].

4. Materials and methods

The presented framework, as shown in Fig. 3, has three main modules, knowledge acquisition, knowledge consolidation, and output interfaces.

4.1. Knowledge acquisition

The knowledge acquisition module caters for variety in healthcare systems by reusing existing tools for acquiring knowledge from structured, unstructured, expert based, and legacy sources. A summary of all the knowledge acquisition techniques used in the study is given in Table 2. Structured data such as Electronic Medical Records (EMR) are an abundant source of patient information that typically contain their demographics, lab test results, vital signs, diagnosed diseases, treatments, and follow up plans. Using advanced information processing technologies such as semantic query processing and machine learning, this data can produce some meaningful knowledge. In order to achieve this, we have to first look at the format of elements in the structured data. Normally, structure data contains disease related images, which can clarify disease status. Therefore, first we utilize deep learning for identifying the labels corresponding to each image and its associated feature map. Deep learning techniques have already proven applications of better interpretation, classification, and pattern identification in clinical images [29–31]. These can also be used to extract valuable features, insights, and labels for an image with human-like intelligence. We utilize the Skip-connection Deep Network technique based on a pretrained Convolution Neural Network (CNN) [32] to extract the classification labels and features map of each image. The labels are then augmented into the EMR data, which is analyzed using our previous work on cost-sensitive ensemble feature ranking and automatic threshold selection methodology [33]. This technique is applied on the serialized form (flat comma separated values, CSV based file with keys and attributes forming the header and values in cells) of the augmented EMR data. Then we select all attribute names as features of the data and allow the expert to associate a cost to each or any feature, optionally. Then we apply feature scoring using tree based modeling approaches and ensemble learning to identify cumulative relevance score based on its given cost, if any, for each feature. We then identify the feature selection threshold, dynamically, to find most relevant features. Using this set of relevant features, we create a cumulative set of interpretable rules that have transparent decision making logic [34,35] obtained from various other white box techniques such as C4.5, Random forest, Ripper, CART, and PART. These techniques can be further enhanced or specialized based on expert's intuition. An abstract overview of knowledge acquisition and consolidation is shown in Fig. 4. The final output of the service is a set of production rules, which are a restructured form (**IF** conditions **THEN** Conclusion) of the interpretable rules.

Unstructured and semi-structured data such as clinical notes and standard practice guidelines are a second source for the knowledge acquisition module. Using an active transfer learning based natural language processing methodology, we identify the conditions and action parts of the clinical text that are used to extract knowledge rules. The technique used to convert descriptive text into rules and the addition of Active Transfer learning is a part of our previous work, presented in [36] and [37], respectively. As shown in Fig. 5, the process starts with well-known

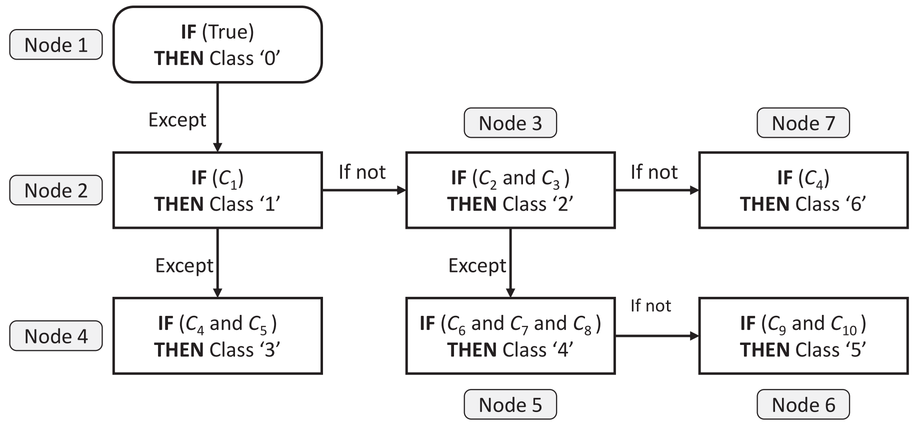


Fig. 2. An example of ripple down rules knowledge tree.

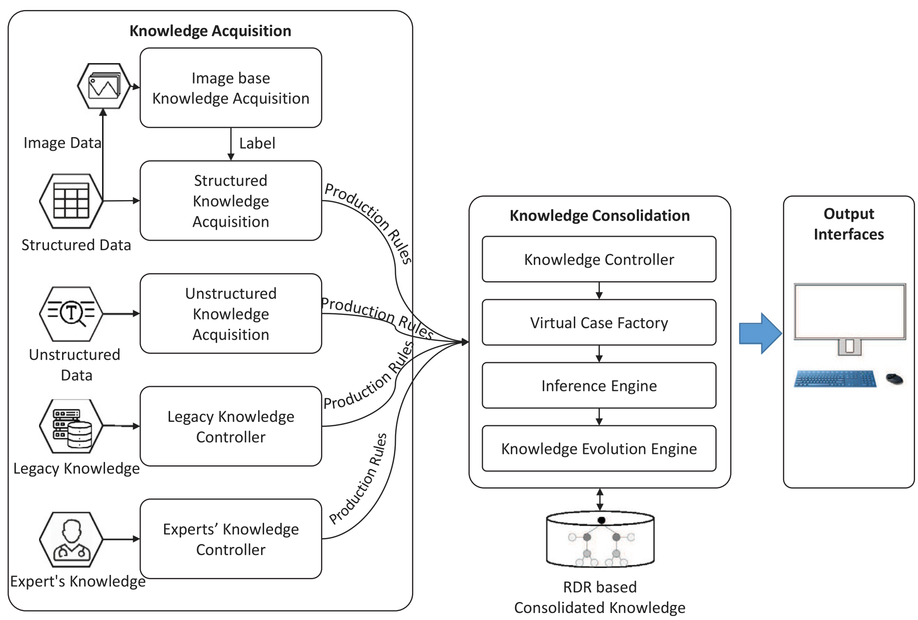


Fig. 3. Knowledge extraction and consolidation workflow.

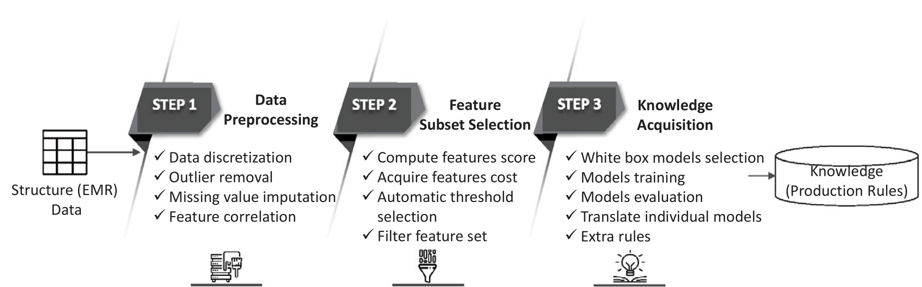


Fig. 4. Structured knowledge acquisition process flow.

text preprocessing of clinical corpus, which identifies the sentences, words, and parts of speech in the text. We then identify the relevant qualifier in each sentence by calculating its weight. This is calculated using the average weight by correlation, Gini Index, Information Gain, and Information Gain Ratio. The top 20 qualifiers are then extracted in the sentence, which are expanded using semantic and syntactic techniques [36]. Then the direction of the qualifier is determined to check the relationship, if any, between the conditions and action parts of the sentence. At the same time, we use the pre-processed text to also identify the clinical and wellness terms of each participating sentence. Each term once identified through the Unified Medical Language System (UMLS) API [38], is augmented with its associated semantic type, also queried through UMLS. The semantic type helps to identify the category of the term, such as "Sign or Symptom", "Acquired Abnormality", "Clinical Drug", and others. These categories are used to classify each term as a condition or an action (consequence). The conditions are then used to identify a probable triple representation of each condition, by using the form <Term, Qualifier, Value>. These are converted into an embedding

Table 2
A summary of knowledge acquisitions methods used in proposed study.

Knowledge acquisition method	Input	Key techniques	Output	Limitations
Image based knowledge acquisition	Fudus image EMR data	Trilogy of skip-connection deep networks	Diabetic retinopathy risk progression	• Requires full set of pre-specified fundus images and EMR-based attributes • Retinal modalities such as Fluorescein Angiography (FA) and Optical Coherence Tomography Angiography (OCTA) are not considered.
Structured knowledge acquisition	EMR data	• Cost based features selection • Ensemble modeling on white box classification models	Productions rules	• Cost is an external factor which may not be available. Therefore, a sudo cost needs to be assigned to features. • The automatic threshold selection, selects a high threshold value which filtered out most of the features.
Unstructured knowledge acquisition	Clinical text document	• Transfer learning for concept expansion and similarity identification • Active learning for knowledge base enhancement	Productions rules	• The rules extraction is heavily dependent on concept identification. • The methodology process document sentence wise while rule conditions and corresponding action may be described in more than one sentence.
Legacy knowledge controller	Production rules	• SNOMED CT and vMR for concept alignment to domain clinical model (DCM) and MLM representation • Semantic reconciliation model (SRM) for concept mapping	Productions rules	The DCM used for mapping local concepts with standard terminologies is a manual task and required tremendous efforts and time.
Experts' knowledge controller	–	Mind map and Iterative decision tree (IDT) base knowledge acquisition	Productions rules	• The expert base knowledge acquisition is dependent on experts experience and heuristics, which may various between various experts. • The expert knowledge may not be compliant with standard medical procedures.

vector using transfer learning technique applied on a pre-trained BERT [4] generic model. The newly generated embedding vector is then matching with our pre-trained triple model, which contains a large set of valid clinical causal relationships. The results of this triple matching then determine if we should use the conditions and actions to create a rule of the form <IF condition THEN Action>. From the set of all possible rules, we finally amalgamate the results by removing duplicates and combining rules, where the actions are exactly the same. The final filtered set of rules are then reformatted into interpretable rules and stored in a repository, temporarily for further processing.

The third source of knowledge is acquired by interfacing with legacy knowledge sources and providing an interoperability framework for it. The legacy HMIS has diverse schemas for presenting their data and knowledge models that cause hindrance in knowledge interoperability. Here, we re-purpose our previous work [39] to map concepts of the legacy knowledge to our target concepts for consolidation. The knowledge interoperability is achieved via a semantic reconciliation model (SRM) that generates maps between legacy concepts via domain clinical model (DCM), and decision support standards such as virtual medical record (vMR), and SNOMED CT. The SRM enables us to acquire, understand and supplement the legacy knowledge bases into our consolidated knowledge model. The SRM model enables us for inter (legacy) and intra (structure, unstructured, and expert) knowledge base communication to achieve the interoperability.

Finally, we provide the domain experts with a smart and user-friendly interface for creating knowledge. The Intelligent Knowledge Authoring Tool (IKAT) [39], provides a web based interface, where the medical expert can add the various conditions and a conclusion of a rule, using intelli-sense based on clinical concepts. Leveraging the authors' expertise in the field of endocrinology, 14 rules, based on 12 attributes (further explained in Section 5.1.1 as Seoul Saint Mary-Data Set; SSM-DS), and 4 conclusions were produced. These rules combine both theoretical knowledge and practical experience of the physicians, working actively to treat diabetic patients. Using an iterative process these rules have undergone numerous refinements, eventually producing the workflow diagram as shown in Fig. 6.

The same interface is also used for verification and validation of the rules before they are added to RDR knowledge base. This is further explained in the following section.

4.2. Knowledge consolidation

State-of-the-art solutions based on multi-modal knowledge acquisition have found acceptance in the healthcare domain but the accuracy of these independent (using only one modality) or in some cases partially dependent (using a combination of two to three modalities) solutions is not at par with the high standards of healthcare service delivery in the real world [40]. Knowledge consolidation can resolve this problem with the help of RDR, which provides a well-managed and evolvable knowledge base. However, the acquired knowledge in the form of interpretable rules has a different format from the one required by RDR as input. These rules can contain inequalities (less than, greater than, not equal to), intervals (closed or open), and other indicators, which provides a general requirement. However, as explained in Section 3, RDR requires cornerstone case for each rule. Therefore, we used the Virtual Case Factory to generate virtual case related to each rule.

A concrete example of the complete knowledge consolidation workflow is shown in Fig. 7. The various services involved in this process are explained next, with a focus on describing our novel contributions.

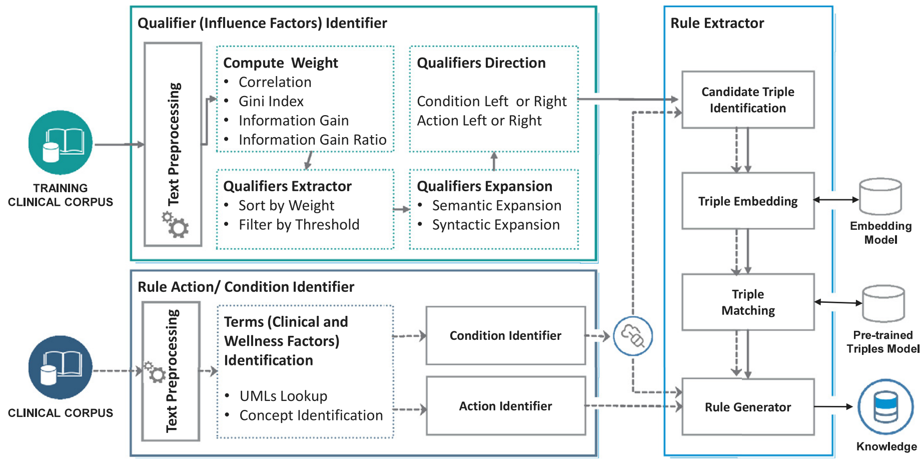


Fig. 5. Unstructured knowledge acquisition process flow.

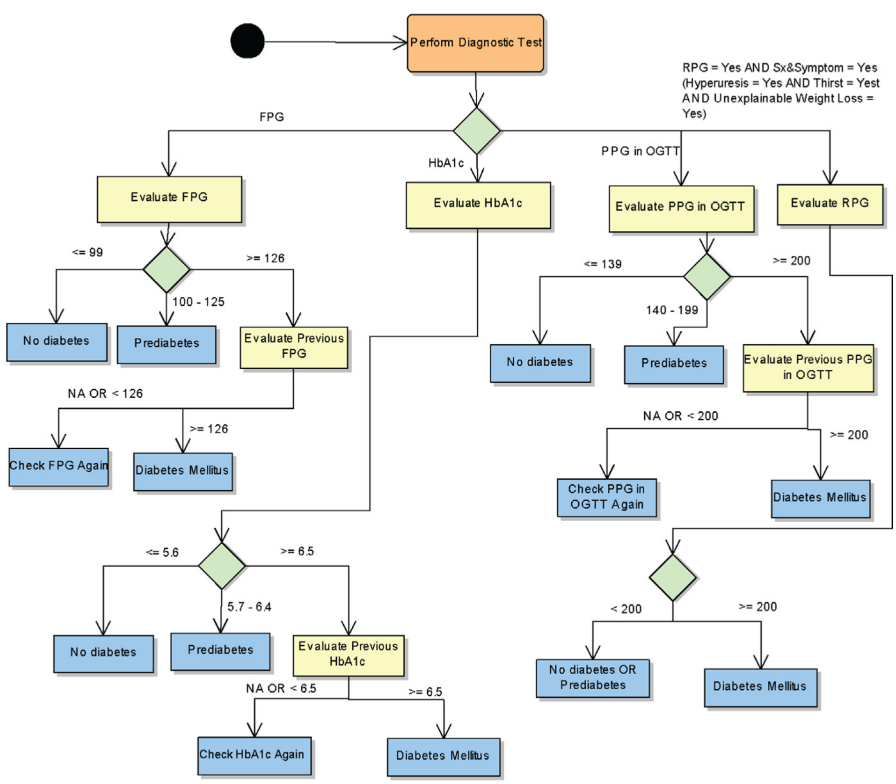


Fig. 6. Physician's knowledge workflow diagram for diabetes.

The knowledge controller acts as a mediator between knowledge acquisition and knowledge consolidation modules. while the Virtual Case Factory (VCF) transforms an interpretable rule to a case or instance for RDR. Based on the factory software design pattern, the VCF identifies the minimum and maximum values related to each conditional part of the rule. An overview of these conversions is shown in Table 3. In particular, the inequalities and the intervals are converted into their nearest, minimum, valid equalities. As an example, the rule shown in Eq. (1) is converted into the case as shown in Eq. (2). The inference engine evaluates the generated case of each rule with the underlying knowledge modal and produces a recommendation. The human expert evaluates the system generated recommendation, and evolve the knowledge via the knowledge evolution engine for unsatisfied results.

$$\forall\,(FPG >= 126\ \&\ Symptom = Yes\ \&\ previousFPG > 125) \rightarrow DiabetesMellitus \tag{1}$$

$$[FPG = 126,\ Symptom = Yes,\ previousFPG = 126] \tag{2}$$

The created rules are then stored in the Knowledge Base (KB) to support inference and evolution. In traditional RDR model, rules are stored based on their acquisition, with nodes at lower depth providing general recommendations and those at higher depth providing specific recommendations. While this implementation is correct, but it does not achieve knowledge consolidation. Instead, we updated the traditional RDR evolution model with the algorithm shown in Algorithm 1, which also changes the

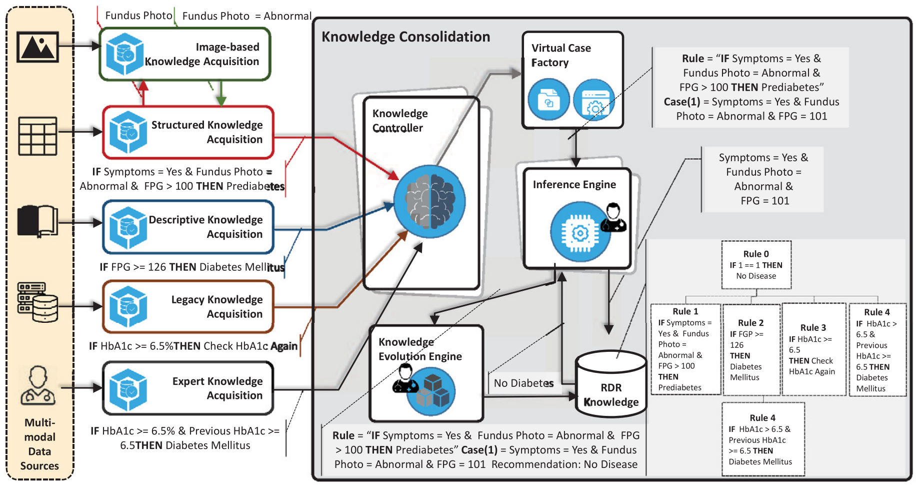


Fig. 7. Knowledge consolidation example.

Table 3
Overview of rule conditions conversion by virtual case factory.

Operator	Conversion	Example
=	–	Rule: IF Age = 30 THEN No Diabetes Case: [Age = 30]
>	Float: Increase by 0.1 Numeric: Increase by 1	Rule: IF Age > 60 AND HbA1c > 5.6 THEN Prediabetes Case: [Age = 61, HbA1c = 5.7]
≥	Drop >	Rule: IF Age ≥ 30 AND HbA1c ≥ 5.6 THEN Prediabetes Case: [Age = 30, HbA1c = 5.6]
<	Float: Decrease by 0.1 Numeric: Decrease by 1	Rule: IF Age < 60 AND HbA1c < 5.6 THEN No Diabetes Case: [Age = 59, HbA1c = 5.5]
≤	Drop <	Rule Age ≤ 60 AND HbA1c ≤ 5.6 THEN No Diabetes Case: [Age = 60, HbA1c = 5.6]

shape of the resulting RDR tree. The algorithm starts by first creating a ChainedHash of generated rules (genRule) coming from the knowledge acquisition module. This process, adds the hash of conclusions from the genRule into the keys of the "recommendations" (ChainedHash) data structure; the values of which hold a one dimensional linked list of all genRules, having the same conclusion. A triple loop then uses the conclusion value of the "recommendations" data structure, and for each rule in its associated linked list, looks up the conclusion value in the KB. Through this process, the algorithm identifies all sub trees, which hold the selected conclusion and the currently iterated genRule is attached as a leaf node to these sub trees. Additionally, the genRule is also added as a leaf to the default node of the KB, which is the root node. In this manner, this extended implementation of RDR for our consolidated KB holds m rules with similar conclusion in m subtrees, where the tree in the first subtree hold m children, the tree in the second subtree hold $m-1$ children, and so on, until the mth subtree holds only 1 child (leaf) node. For j conditions then the tree will have $m * j$ nodes at depth 0. The resulting tree is spatially bigger than the traditional RDR but holds consolidated rules with the highest encapsulation of probable conditions producing a conclusion on the left and lower on the right.

Using the traditional RDR inference input case from Eq. (3) will select the rule as highlighted in Fig. 8.

$$[FPG = 90, Symptom = Yes, PPG = 126] \tag{3}$$

The consolidated knowledge-based system assists healthcare experts in making appropriate clinical decisions. The quality of the generated recommendation is greatly affected by the quantity and quality of the underlay knowledge model. Our proposed system hides technical details and complexity from healthcare experts by providing easy-to-use interfaces. The involvement of experts is limited to knowledge validation and verification while the system is in use at real practice. During real time usage of the system, if experts found any inappropriate recommendation, they need to provide correct recommendation along with key features of the patient case that caused the update in the recommendation. By providing these two pieces of information, the system implicitly evolves the underlying knowledge model to reflect the change for further usage.

Algorithm 1 Knowledge Consolidation Algorithm

```
    Input: RuleSet, KB
    Output: KB
ChainedHash recommendations;
for genRule in ruleSet do:
    if !recommendations.hasKey(genRule.conclusion) then:
        new recommendations[genRule.conclusion]
    end if
    recommendations[genRule.conclusion].add(genRule)
end for
for genRuleConclusion in recommendations do:
    for genRule in recommendations[genRuleConclusion] do:
        ArrayList conclusionGroupsInKB
        for rdrRuleRef in KB do:
            if rdrRuleRef.conclusion==genRuleConclusion then:
                conclusionGroupsInKB.add(rdrRuleRef)
            end if
        end for
    end for
    for rdrRuleRef in conclusionGroupsInKB do:
        rdrRuleRef.addLeaf(genRule)
    end for
end for
KB.getRoot().addLeaf(genRule)
return KB
```

4.3. Output interfaces

The output interfaces for our knowledge consolidated system includes a custom Electronic Health Record (EHR) management system, CDSS, and a Clinical Practice tool, provided

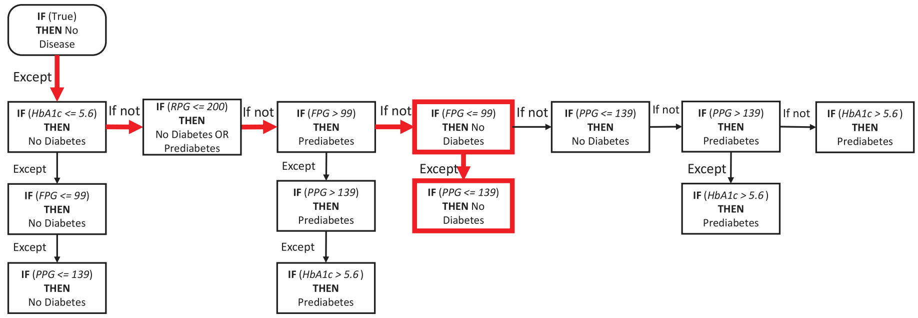


Fig. 8. Example of RDR inferencing for input case shown in Eq. (3) (selected path and rule is highlighted in red).

through various services such as Patient Encounter Management (PEM), Inference, knowledge visualization, and verification and validation.

PEM provides the interfaces for collecting demographic and clinical data relating to a patient, during their encounter with a physician. Through this system, we collects the necessary data for managing EHRs. In our prototype implementation the EHR is stored in a relational database. Based on the recorded patient data in PEM, we apply RDR inferencing on our KB and provide the most relevant recommendation. In case the physician agrees with the recommendation, it is stored in the EHR DB. On the other hand, if the generated recommendation is not correct, the verification and validation services start up, which following the above mentioned Knowledge Evolution process can create new rules or update existing ones in the KB. The same KB is used to create read only scenarios, whereby the user is presented with a set of attributes, corresponding to a randomly selected KB rule. The user can then use this interface to either provide the values, leading a particular recommendation, or provide a recommendation, against randomly allocated values for the attributes, which are evaluated against correct results. Through this process, the user can view the correct attribute values and recommendations in the form of automatically generated visual trees (as shown in Fig. 8) but cannot update the existing rules. Further details of the output interfaces are beyond the scope of this manuscript and will be presented in the future.

5. Results and discussion

5.1. Structured knowledge acquisition

5.1.1. Structured dataset description

A thorough evaluation of the presented framework necessitates the application of a multipronged approach for data acquisition, rule formulation, and comprehensive evaluation. Following this aim, we have utilized three datasets to cover the breadth of knowledge acquisition. The first dataset, henceforth referred as Seoul St. Mary's Dataset (SSM-DS), corresponds to the real patient data, collected from past records of Diabetes Mellitus (DM) evaluations performed at The Catholic University of Korea, Seoul St. Mary's Hospital,[2] South Korea. This dataset has been anonymized at source and includes clinical test results for current and past values of Fasting Plasma Glucose (FPG), Hemoglobin A1C (HbA1c), Oral Glucose Tolerance Test (OGTT), Postprandial Plasma Glucose (PPG), Sign and Symptoms (Sx), and Random Plasma Glucose (RPG). Finally, it contains a status of DM recommendation such as qualified as Diabetes, Pre-diabetes, and No Diabetes. The second dataset, Pima Indian Diabetes dataset (PID-DS) [41] is a publicly available corpus of DM diagnostics performed at the National Institute of Diabetes and Digestive and Kidney Diseases. This dataset contains 768 evaluations performed on female patients of Pima Indian heritage. It uses general wellness terms to identify the correlation between DM and other factors, including pregnancies, Glucose, Blood Pressure (BP), Skin Thickness, Insulin, BMI, Diabetes Pedigree Function (corresponding to the family history of DM), Age, and Outcome (state of diabetes corresponding to yes or no). For the third dataset, the Coronary heart disease dataset (CHD-DS) we have switched the domain to include coronary heart disease (CHD) patient data publicly available on Kaggle [42]. The CHD-DS consists of 462 instances evaluated for CHD diagnosis. The factor considered for taking CHD diagnosis decisions includes patients systolic blood pressure (SBP), tobacco use per year in kilogram, low density lipoprotein (ldl), adiposity, family history, type A personality score (typea), obesity measure via body mass index, alcohol usage, and age of the patient. The detail of the datasets is shown in Table 4

5.1.2. Dataset correlation analysis

Datasets correlation analysis is used to understand the relationship among various attributes and assist in predicting one attribute from another. A comparative view of the various attributes and their correlation analysis, within these three datasets are shown in Fig. 9. In SSM-DS, an increased correlation was observed between recommendations and previous values of PPG, HbA1c, and FPG. On the other hand, current evaluation of RPG has a higher impact of 0.49. OGTT shows very small correlation to the factor of −0.07 for previous results and −0.01 for current ones. The highest correlation was observed between previous values of PPG and the recommendation. This indicates that at this stage (before the application of SKA process) and with a threshold of 0.7, DM recommendation can be based on previous PPG values followed by previous HbA1c results.

For PID-DS, the correlation heatmap shows factors such as BP and Skin Thickness having the lowest impact at 0.07, while the highest impact is observed between outcomes and Glucose at 0.47. Therefore, intuitive analysis with a threshold value at 0.7, would not select any attribute from this dataset. Yet, it is clear from the dataset, that these variables have been used by medical experts to evaluate DM status. While in CHD-DS, age has shown the highest correction of 0.37 with CHD, and the lowest correction of 0.10 was observed for typea, and obesity factors. Attributes that have high correlation are more important than others and the knowledge extraction should consider these attributes compare to attributes having less correlation.

5.1.3. Knowledge extraction

In order to acquire knowledge from structured data, our methodology, as presented in Section 4.1, utilizes tree and list based decision algorithms such as Decision Trees, Random Forest, PART, C4.5, CART, and Ripper. An overview of the accuracy achieved by independent application of these algorithms is shown in Fig. 10.

[2] https://www.cmcseoul.or.kr/en.common.main.main.sp.

Table 4
Details of the datasets.

Datasets	Total features	Class labels	Instance ratio	Total instances
Seoul St. Mary's (SSM-DS)	12	4 (No DM, Recheck, Pre DM, DM)	63:49:60:130	302
Pima Indian Diabetes (PID-DS)	8	2 (Yes, No)	268:500	768
Coronary Heart Disease (CHD-DS)	9	2 (0 = No 1 = Yes)	302:160	462

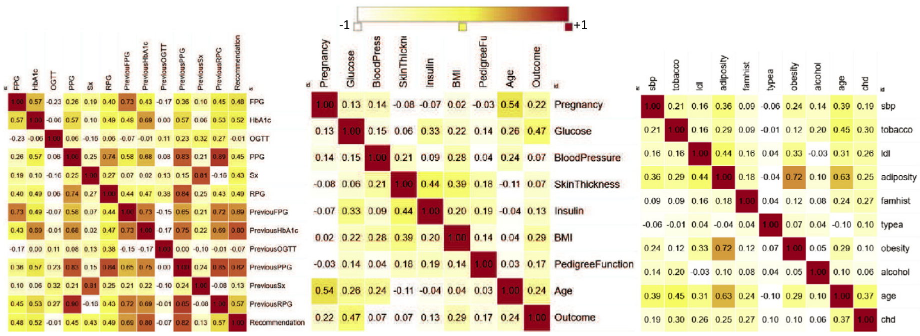
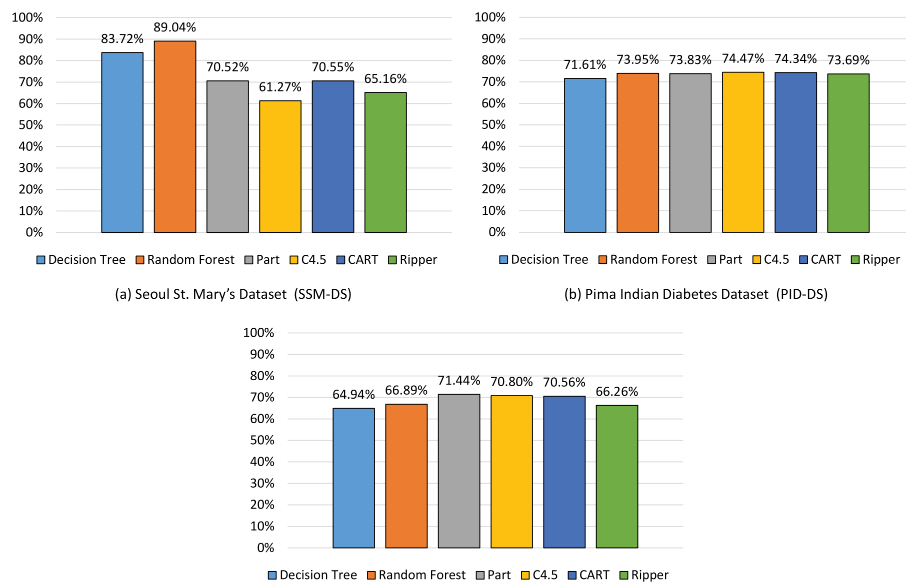

(a) Seoul St. Mary's Dataset (SSM-DS) (b) Pima Indian Diabetes Dataset (PID-DS) (c) Coronary Heart Disease Dataset (CHD-DS)

Fig. 9. Datasets correlation analysis.

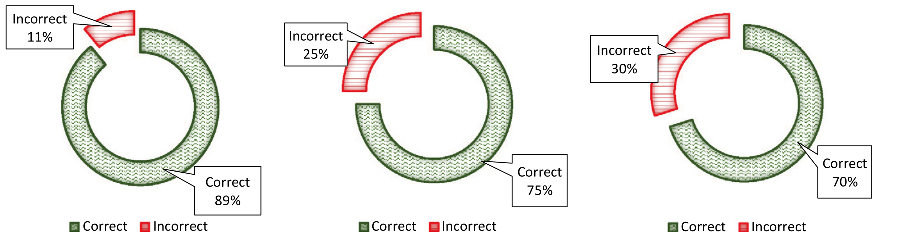


(a) Seoul St. Mary's Dataset (SSM-DS) (b) Pima Indian Diabetes Dataset (PID-DS) (c) Coronary Heart Disease Dataset (CHD-DS)

Fig. 10. Accuracy of executing algorithms for structured knowledge acquisition.

On the SSM-DS dataset, shown in Fig. 10(a), Random Forest, produces the highest accuracy at 89.04%, followed by Decision Tree with an accuracy of 83.72%. For the PID-DS, shown in Fig. 10(b), the resultant accuracy of all algorithms are very close to each other, with a 2.6 percentage point difference between Decision Tree (at 71.61%), and C4.5 (at 74.27%). Similarly, the CHD-DS results, shown in Fig. 10(c), the maximum percentage point difference of 6.47% is evident between Decision Tree and PART. While these results are far from the state-of-the-art, independently they can only moderately identify the set of implicit data rules, which can classify each of these datasets with 10-fold cross validation. It is also pertinent to note that algorithm families such as deep learning and probabilistic can be utilized to improve this learning accuracy but they lack the transparency, which is required for this research work. In fact, the knowledge acquisition process is logically different from the traditional classification process in terms of feature selection and model generation. In order to support the classification process, most interpretable machine learning algorithms rely on the compactness of the model, with the shortest paths required to go from source (data) to target (labels/classes/information). As a result, the decision making process in identifying the individual total functions for producing such path ways ignores the alternatives and reduces the information available for the decision maker to extract important information. Ensemble learning can provide a solution to this problem by expanding the set of possible pathways produced from various algorithms and then limiting them based on some sort of voting strategy, which in our case is majority based. The expanded set of pathways provides a better cover for the rules and their alternatives, which are essential for producing knowledge.

The results for majority voting based ensemble learning, employing Decision Trees, Random Forest, PART, C4.5, CART, and Ripper, is shown in Fig. 11. The SSM-DS, showed the correct



(a) Seoul St. Mary's Dataset (SSM-DS) (b) Pima Indian Diabetes Dataset (PID-DS) (c) Coronary Heart Disease Dataset (CHD-DS)

Fig. 11. Ensemble model prediction results.

classification at 89%, which is same as the best case achieved for independent application of Random Forest. The PID-DS, achieved an accuracy of 75%, while CHD-DS achieved 71%. The results of PID-DS, shown in Fig. 11(b), are little improved from the highest accuracy achieved by the independent application of C4.5, while CHD-DS, shown in Fig. 11(c), is dropped to 70% against the best accuracy achieved by PART at 71.44%. However, as discussed above, the impact of these classification results is somewhat limited and of particular importance to this research work, is the number and precision of the generated rules.

5.2. Unstructured knowledge acquisition

5.2.1. Dataset description

In order to evaluate the unstructured knowledge acquisition process, unstructured textual corpus in the form CPGs were used. We selected three recent CPGs which provide a state-of-the-art scaffolding for augmenting the theoretical and practical proficiency of Endocrinologists. In guideline [43] (G1), the American Diabetes Association (ADA), provides the characteristics and applicable methodologies for diagnosing and managing diabetic patients. In guideline [44] (G2) the National Institute for Health and Care Excellence (NICE), provides a detailed quality standard for diabetes management, focusing primarily on the prevention, education, and treatment of type 1 and 2 diabetes. Finally, guideline [45] (G3), the Scottish Intercollegiate Guidelines Network (SIGN), provided a detailed guideline for not only diagnosing and managing type 1 and 2 diabetic patients, but also comorbidities, such as kidney diseases, cardiovascular disease, and pregnancy.

5.2.2. Knowledge extraction

The clinical corpus described above was fed to our unstructured knowledge acquisition process flow to extract knowledge in the form of interpretable rules. Using the trained model from [37] with 72,359 total triples, we evaluated each guideline in our corpus to extract <Noun, Verb, Noun> triples, after POS tagging each sentence. As shown in Table 5, from G1, the total 1731 triples were extracted, while G2 produced 1142 triples, and G3 10,226. Unique triples from G1 are 1602, from G2 are 948, and from G3 8872, were produced. This was followed by a lookup from UMLS to further remove those triples, where any of the Noun phrase is not recognized. The reduced set of triples produced then are marked as medical triples, where by, the associated Noun phrases are present in UMLS as a concept. In this step, G1 triples were reduced to 1267, G2 to 831, and G3 to 7765. Finally, the total causal triples were evaluated after comparing these embeddings with our trained model, causing the total causal triples by G1 to further reduced to 541, by G2 to 320, and by G3 to 3215. These causal triples are then converted into interpretable rules, by converting the cause noun phrase, along with any qualitative semantic type into a conditional term, corresponding to the "IF" part of the rule. The effect term is converted into the action term, corresponding to the "THEN" part of the rule. The final count

Table 5
Details of unstructured knowledge extraction.

Process	Guideline 1 (G1)	Guideline 2 (G2)	Guideline 3 (G3)	Total
Extracted triples	1731	1142	10 226	13,099
Unique triples	1602	948	8872	11 422
Medical triples	1267	831	7765	9863
Causal triples	541	320	7765	3215
Extracted rules	29	7	13	49

Table 6
Example rules generated from triples.

S.No	Triple	Rule
1	⟨HbA1c, is, Diabetes⟩	IF HbA1c $\geq$ 6.5 THEN Diabetes
2	⟨FPG, be, Diabetes⟩	IF FPG $>$ 126 THEN Diabetes
3	⟨Greater RPG, leads to, Diabetes⟩	IF RPG $\geq$ 200 THEN Diabetes
4	⟨FPG, be, Prediabetes⟩	IF FPG 100–125 THEN Prediabetes
5	⟨HbA1c, is Prediabetes⟩	IF A1C 5.7–6.4 THEN Prediabetes

of rules by the unstructured knowledge acquisition service for G1 is 29, G2 is 7, and G3 is 13. Table 6 shows five causal triples in their corresponding interpretable rule form, produced by the application of our methodology. As an example the triple <HbA1c, is, Diabetes> is converted into "IF HbA1c >= 6.5 THEN Diabetes", where HbA1c and Diabetes are the Noun Phrases in the triple, relating to cause and effect respectively. The verb "is" represents the causal phrase between these terms, while 6.5 is the qualitative term picked from the text using proximity matching, within the sentence. Multiple qualitative terms in the sentence, if they appear, are converted into ranges for the associated cause phrase, as shown by examples 4 and 5 in Table 6.

The evaluation of the extracted rules on SSM-DS is shown in Fig. 12. The rules 71.79% correctly predicted the true class of instances. The primary reason for the low accuracy could be the "Recheck" class. The extracted rules did not assigned the "Recheck" class to any instance, while the expert has annotated 16.83% of the patients with this label. This is because the guidelines provide a concrete decision, while in real practice some of the patients may not be assigned a particular class and need to re-revalue after some specific time, as evident by the experts provided label in the SSM-DS. Also, there is 6.93%, 3.3%, and 3.96% difference between expert provided labels and predicted labels for "DM", "Prediabetes" and "No Diabetes", respectively.

5.3. Knowledge combination and consolidation

A statistical overview of the rules extracted from each of the aforementioned methods and their datasets is shown in Table 7. Simple summation of the rules, yield 274 instances, which are reduced to 271, after removing 3 duplicates. We then removed the conflicts among the extracted rules, which are based on 13 default conditions, which produce a default conclusion (preset by the knowledge acquisition process, such as if no rule matches

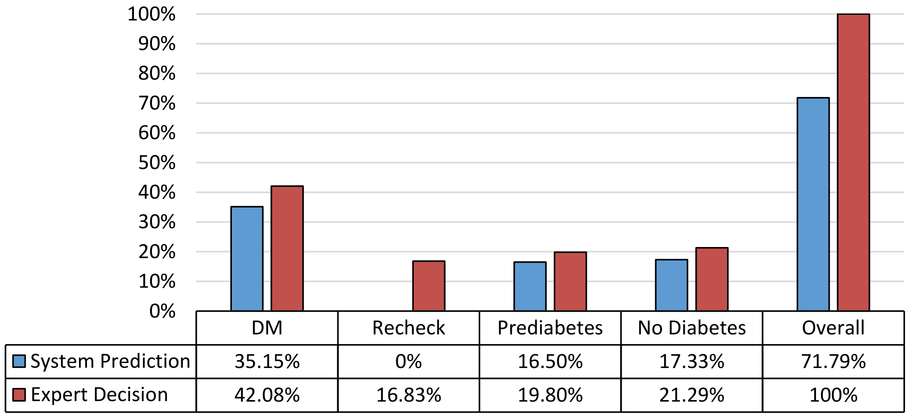


Fig. 12. Unstructured rules evaluation on SSM-DS.

Table 7
Details of extracted rules using our method and datasets.

S. No.	Dataset name	Dataset type	Extracted rules
1.	SSM-DS	Structured data	93
2.	PID-DS	Structured data	86
3.	CHD	Structured data	46
4.	G1	Unstructured data	29
5.	G2	Unstructured data	7
6.	G3	Unstructured data	13
7.	Expert	Expert driven	28
Total extracted rules			274
Total rules without duplicates			271
Total rules without conflicts			258
Consolidated rules			266

Table 8
Confusion matrix of expert source rules.

	Class 1	Class 2	Class 3	Class 4	Precision
Class 1	54	6	0	0	90%
Class 2	0	41	1	16	70.69%
Class 3	9	2	59	8	75.641%
Class 4	0	0	0	106	100%
Recall	85.714%	83.673%	98.333%	81.538%	–

Table 9
Confusion matrix of combined rules.

	Class 1	Class 2	Class 3	Class 4	Precision
Class 1	0	0	0	0	–
Class 2	0	41	1	16	70.69%
Class 3	57	8	59	0	47.581%
Class 4	0	0	0	114	95%
Recall	0%	83.673%	98.333%	87.692%	–

Table 10
Confusion matrix of consolidated rules.

	Class 1	Class 2	Class 3	Class 4	Precision
Class 1	55	0	0	0	100%
Class 2	0	41	1	2	93.182%
Class 3	8	2	59	5	79.73%
Class 4	0	6	0	123	95.349%
Recall	87.302%	83.673%	98.333%	94.615%	–

then it may result prediabetes, or no diabetes). Finally, we got 258 rules, without applying consolidation process. The consolidation process, which relies on the RDR methodology and expert intervention, yields 266 rules, including the default true rule. A partial view of the RDR tree with some of the initial nodes are shown in Fig. 13. Each node contains the condition of the rule, its conclusion, and a source providing backward linkage of the rule with its knowledge source. Since a rule can be sourced from multiple sources, for simplicity we have used random selection.

The evaluation and comparison of expert source rules, combined rules, and consolidated rules provides the basis for establishing the novelty with high accuracy of our proposed methodology. The evaluation was performed on the SSM-DS due to its thoroughness and measure-ability, in terms of standardized medical terminologies usage and the width of conclusions (covering scenarios with No diabetes, prediabetes, recheck, and diabetes). Through expert intervention, (expert rules) was able to correctly identify 54 instances of No diabetes (class 1), 41 of recheck (class 2), 59 of pre-diabetes (class 3), and 106 of diabetes (class 4). Here the precision for class 1 is 90%, class 2 is 70.96%, class 3 is 75.641%, and class 4 is 100%. The recall here for class 1–4 is 85.714%, 83.673%, 98.333%, and 81.538%, respectively as shown in Table 8. The overall accuracy for M0 is 86.093%.

On the other hand, combined rules were unable to classify any instance as class 1. However, for class 2, 41 instances were classified, with a precision rate of 70.69% and recall rate of 83.673%. 59 instances were classified as class 3, with precision rate of 47.581% and a recall of 98.33%. Finally for class 4, 114 instances were correctly classified with a precision rate of 95% and a recall of 87.692%, as shown in Table 9. The overall accuracy for the combined rules is 70.861%. In general, these results indicate that a simple combination of the expert and machine generated rules can diminish the performance of the knowledge application process. It is also pertinent to note here, that in this particular case, the greatest impact to the loss of overall accuracy is from class 1 having 0 classified instances thereby leading to a complete absence of "no diabetes" as a probable medical condition. This is a direct consequence of the unbridled production and accumulation of widely-applicable, machine generated rules, which have not been verified by the expert.

Consolidated rules, identified 55 instance as class 1, with 100% precision and 87.302% recall, 41 as class 2 with 93.182% precision and 83.673% recall, 59 as class 3 with 79.73% precision and 100% recall, and 123 instances as class 4 with 95.349% precision and 94.615% recall, as shown in Table 10. The overall accuracy of the consolidated knowledge is 92.053%. In short, the consolidated rules have achieved an increase of 5.96%, as compared with expert evaluation, and 21.192% from evaluations done by a simple union of multi-sourced rules.

Fig. 14 shows the evaluation and comparison of accuracy achieved by experts' rules, combined rules, and consolidated rules. As a consequence of our earlier discussion on the breadth of

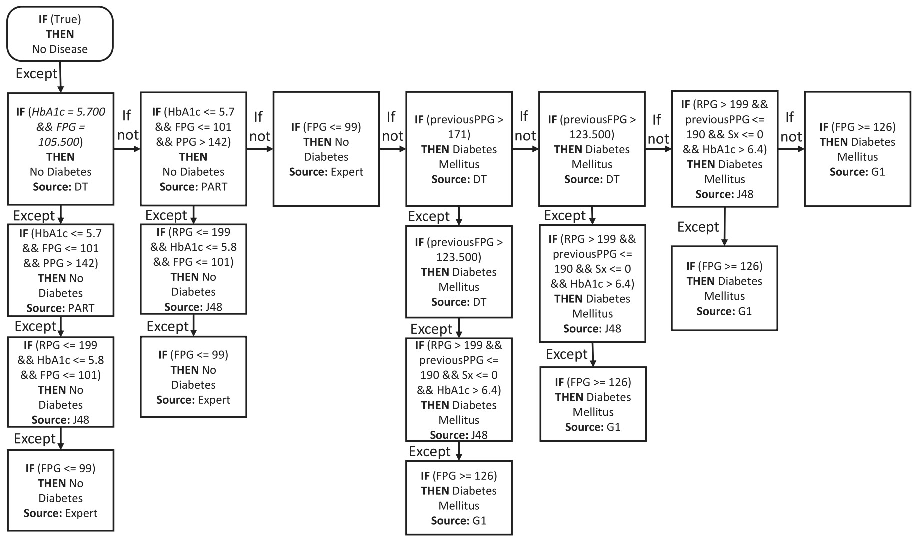


Fig. 13. Consolidated knowledge partial model.

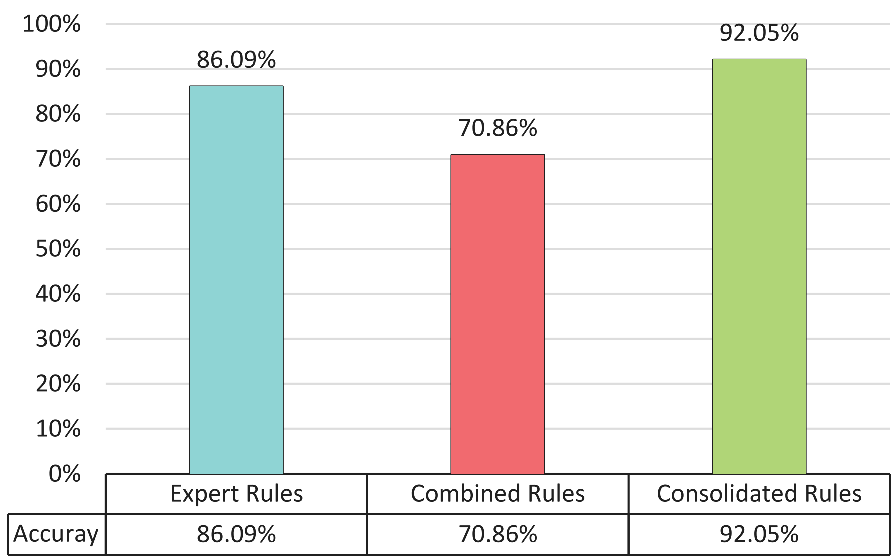


Fig. 14. Evaluation of experts' rules, combined rules and consolidated rules on SSM-DS.

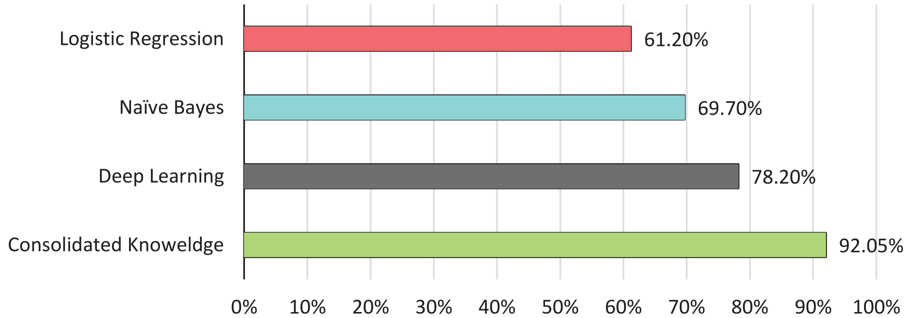
conclusions maintained by a naive combination of rules, leading to 0 classifications of class 1, shows the lowest accuracy of all compared models, while the highest accuracy of 92.05% is achieved by the proposed knowledge consolidation. The consolidated knowledge is initialized by the expert and then extended by consolidating knowledge from various sources. Also, it has the ability to adopt new knowledge which keeps the knowledge model alive and improves its performance over time.

5.4. Comparative analysis

To compare the result of our proposed knowledge consolidation system with the state-of-the-art black-box algorithms, we have used Rapid Miner's auto model feature on SSM-DS. The results achieved by Logistic Regression, Naive Bayes, and Deep Learning compared to our proposed system are shown in Fig. 15. The logistic regression performed verse with an accuracy of 61.20% while Naive Bayes achieved an accuracy of 69.70%, Deep Learning 78.20% compare to our proposed system at 92.05%. The primary reason could be the nature and the size of the dataset. Algorithms like deep learning and others are data-hungry models while in real scenarios the patient data may not be enough for these models. Also, the decision logic is not known to the expert which may reduce the confidence level of human experts on the system-generated recommendation. While application of the improved knowledge base with the ability to improve itself over time like the proposed one best suit the expectation of healthcare service providers and can move towards Artificial Wisdom.

5.5. A step towards wisdom

In order to close the loop and apply the foundational results presented above in practice, we have designed a web based application to allow the medical experts to add patient records and to evaluate these records using our consolidated knowledge in the form of RDR tree. While the details of this system are beyond the scope of this paper, the system utilizes a large stack of technologies to curate data, information, and knowledge at



	Consolidated Knowledge	Deep Learning	Naïve Bayes	Logistic Regression
Accuray	92.05%	78.20%	69.70%	61.20%

Fig. 15. Comparative Analysis of knowledge consolidation with other methods on SSM-DS.

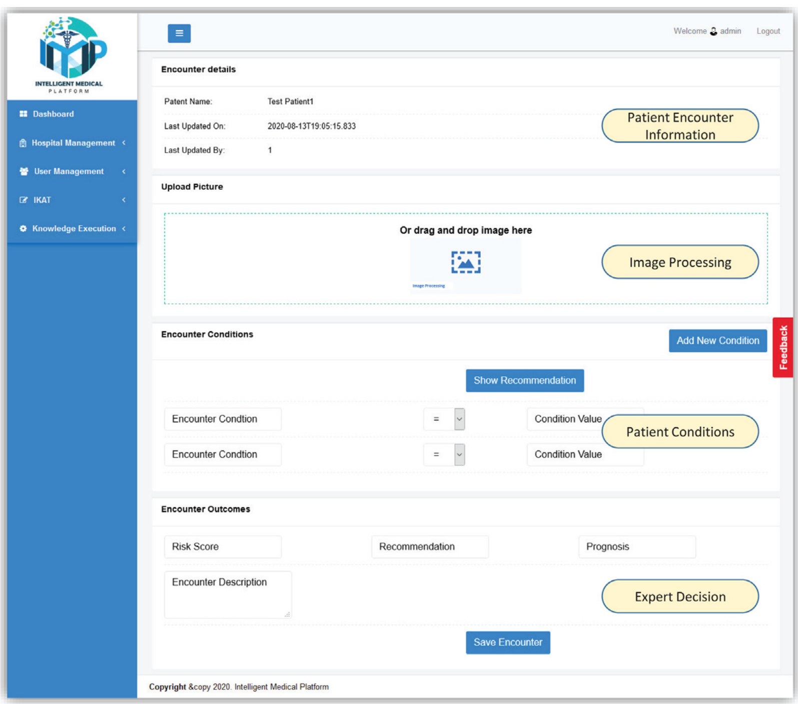


Fig. 16. Patient encounter interface.

various abstractions. While the system is generic enough to deal with any domain, provided the appropriate knowledge repository in the form of RDR is first built through expert intervention, we have restricted its use to the domain of diabetes and chronic heart disease, in this research work. An additional feature of this application is to maintain and extend its knowledge base by utilizing our customized RDR implementation, which provides a feedback loop for actively learning from new rule sources. The in-memory, RDR tree adds new rules at appropriate places and in real-time can extend its inference to include these rules.

Fig. 16 shows a screenshot of the patient encounter creation interface, where the medical expert can add a patient's record. This is followed by Fig. 17, which shows the result of classification from inference on currently available nodes in RDR. In case the classification is incorrect, the expert can then move to the knowledge evolution interface shown in Fig. 18 to add a new rule corresponding to the conditions available in the patient's encounter. Here the expert can modify conditions, conclusions, and other metadata for the new rule, which is immediately made available for future inferencing. These screens are designed with the consensus and feedback of our collaborative hospital's physicians (The Catholic University of Korea, Seoul St. Mary's Hospital, South Korea) to hide the complexity from experts and simplify the knowledge acquisition, consolidation, and inferencing. For inferencing, the expert needs to provide patient conditions, while the system generates recommendations using RDR inferencing techniques and presents the expert with a result along with RDR Knowledge as evidence as shown in Fig. 17. Similarly, the complex process of knowledge evolution is simplified enough so that experts can evolve knowledge with minimum effort and time.

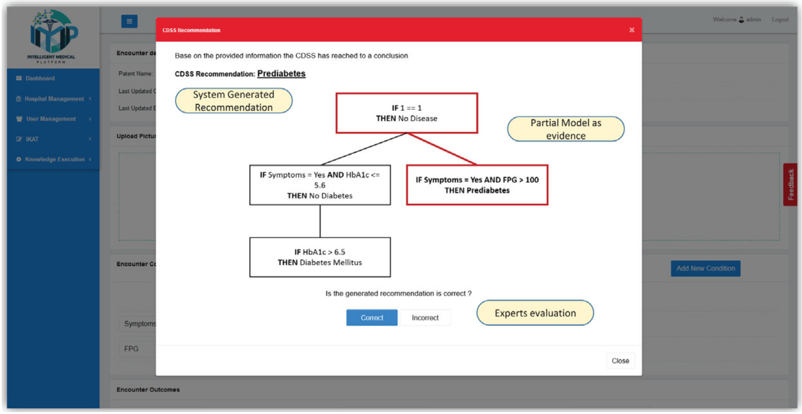


Fig. 17. System generated recommendation interface.

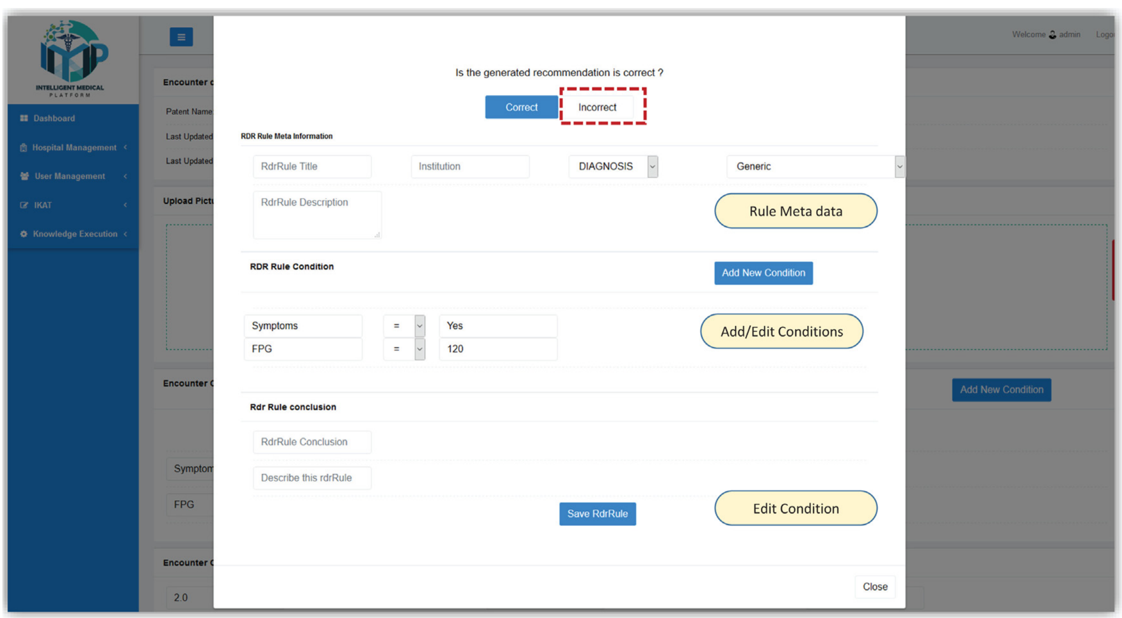


Fig. 18. Knowledge evolution interface.

6. Conclusions

Advancements in Information and Communication Technologies (ICT) in general and AI in particular, is transitioning towards creating an intelligent system, that can surpass human wisdom by making accurate and transparent decisions. Towards this end, in this paper, we present an important aspect of this process, which aims to consolidate knowledge from various sources and provides long term maintenance and evolution supporting services, through the use of RDR. While the generality of this framework can enable its universal applicability, for practical purposes, we have evaluated the same in the healthcare domain, further confined within the field of diabetes diagnostics.

A more general solution also necessitates the resolution of various other challenges, such as knowledge interoperability, identification and maintenance of its context, peer validation, and self-learning solutions. In particular, identifying the various formats in which knowledge is acquired and stored is only the first step. Heterogeneously formatted knowledge then has to be transformed into RDR compliant form, to take benefit of our proposed framework. Similarly, conversion of the RDR into a knowledge format, which can be consumed by existing applications would ensure their extensibility. Additionally, it is important to identify the context in which knowledge was acquired. In a multi-domain setting, the acquisition context and its usage in inference is all the more important, considering how various devices, measurements, and tools are used in different settings. As an example, if the knowledge base pertains to the manufacturing domain and healthcare, a context-less inference of the rules using "temperature" measurement could cause severe harm to a patient, human workers, or machinery. The same effect may also result from malicious rules or unintentional mistakes from the human expert. In order to reduce the chances and/or effects of such knowledge poisoning, block chain solutions, such as smart contracts and multi-peer verification and validation of knowledge is necessary. Finally, the incorporation of self-learning techniques, once they have matured enough, would further reduce the human expert involvement in knowledge acquisition. The challenges listed above are not exhaustive, however these provide the direction in which we shall focus our future research and development efforts.

CRediT authorship contribution statement

Musarrat Hussain: Conceptualization, Methodology, Software, Investigation, Writing – original draft. **Fahad Ahmed Satti:** Methodology, Validation, Investigation. **Syed Imran Ali:** Validation, Writing – review & editing. **Jamil Hussain:** Visualization, Validation, Writing – review & editing. **Taqdir Ali:** Validation, Writing – review & editing. **Hun-Sung Kim:** Writing – review & editing. **Kun-Ho Yoon:** Supervision. **TaeChoong Chung:** Supervision. **Sungyoung Lee:** Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2017-0-01629) supervised by the IITP(Institute for Information & communications Technology Promotion), by the Korea government MSIT (Ministry of Science and ICT) grant (No.2017-0-00655), by the MSIT Korea, under the Grant Information Technology Research Center support program (IITP-2020-0-01489), (IITP-2021-0-00979) supervised by the IITP (Institute for Information & communications Technology Promotion) and NRF2019R1A2C2090504.

References

[1] S. Montani, M. Striani, Artificial intelligence in clinical decision support: a focused literature survey, Yearb. Med. Inform. 28 (1) (2019) 120.
[2] J. Van Hulse, Data quality in data mining and machine learning, 2007.
[3] F. Hammonds, Is justified true behavior knowledge? Behav. Philos. (2010) 49–59.
[4] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, 2018, arXiv preprint arXiv:1810.04805.
[5] T.B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D.M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language models are few-shot learners, 2020, arXiv:2005.14165.
[6] N. Potyka, M. Thimm, Consolidation of probabilistic knowledge bases by inconsistency minimization, in: ECAI, 2014, pp. 729–734.
[7] X.L. Dong, E. Gabrilovich, G. Heitz, W. Horn, K. Murphy, S. Sun, W. Zhang, From data fusion to knowledge fusion, 2015, arXiv preprint arXiv:1503.00302.
[8] F.A. Satti, T. Ali, J. Hussain, W.A. Khan, A.M. Khattak, S. Lee, Ubiquitous health profile (uhpr): a big data curation platform for supporting health data interoperability, Computing (2020) 1–36.
[9] B. Cánovas-Segura, A. Morales, J.M. Juarez, M. Campos, F. Palacios, A lightweight acquisition of expert rules for interoperable clinical decision support systems, Knowl.-Based Syst. 167 (2019) 98–113.
[10] E.H. Shortliffe, M.J. Sepúlveda, Clinical decision support in the era of artificial intelligence, JAMA 320 (21) (2018) 2199–2200.
[11] W.N. Price, I.G. Cohen, Privacy in the age of medical big data, Nat. Med. 25 (1) (2019) 37–43.
[12] S. Zolhavarieh, D. Parry, Q. Bai, Issues associated with the use of semantic web technology in knowledge acquisition for clinical decision support systems: systematic review of the literature, JMIR Med. Inform. 5 (3) (2017) e18.
[13] L. Muhammad, E. Garba, N. Oye, G. Wajiga, On the problems of knowledge acquisition and representation of expert system for diagnosis of coronary artery disease (cad), Int. J. U- E-Serv. Sci. Technol. 11 (30) (2018) 49–58.
[14] E. Strickland, Ibm watson, heal thyself: How ibm overpromised and underdelivered on ai health care, IEEE Spectr. 56 (4) (2019) 24–31.
[15] A. Intezari, D.J. Pauleen, N. Taskin, The dikw hierarchy and management decision-making, in: 2016 49th Hawaii International Conference on System Sciences (HICSS), IEEE, 2016, pp. 4193–4201.
[16] D.H. Williams, G.P. Shipley, Enhancing artificial intelligence with indigenous wisdom, Open J. Philos. 11 (1) (2020) 43–58.
[17] P.B. Baltes, J. Smith, The fascination of wisdom: Its nature, ontogeny, and function, Perspect. Psychol. Sci. 3 (1) (2008) 56–64.
[18] H. Varonen, T. Kortteisto, M. Kaila, E.S. Group, What may help or hinder the implementation of computerized decision support systems (cdsss): a focus group study with physicians, Family Pract. 25 (3) (2008) 162–167.
[19] K.D. Mandl, F.T. Bourgeois, The evolution of patient diagnosis: from art to digital data-driven science, JAMA 318 (19) (2017) 1859–1860.
[20] S. Ting, S.K. Kwok, A.H. Tsang, W.B. Lee, A hybrid knowledge-based approach to supporting the medical prescription for general practitioners: Real case in a hong kong medical center, Knowl.-Based Syst. 24 (3) (2011) 444–456.
[21] M.B. Sesen, M.D. Peake, R. Banares-Alcantara, D. Tse, T. Kadir, R. Stanley, F. Gleeson, M. Brady, Lung cancer assistant: a hybrid clinical decision support application for lung cancer care, J. R. Soc. Interface 11 (98) (2014) 20140534.
[22] N. Larburu, N. Muro, I. Macía, E. Sánchez, H. Wang, R.J. Winder, J. Bouaud, B. Séroussi, Augmenting guideline-based cdss with experts knowledge, in: HEALTHINF, 2017, pp. 370–376.
[23] Y. Shen, K. Yuan, D. Chen, J. Colloc, M. Yang, Y. Li, K. Lei, An ontology-driven clinical decision support system (iddap) for infectious disease diagnosis and antibiotic prescription, Artif. Intell. Med. 86 (2018) 20–32.
[24] M.D. Soufi, T. Samad-Soltani, S.S. Vahdati, P. Rezaei-Hachesu, Decision support system for triage management: A hybrid approach using rule-based reasoning and fuzzy logic, Int. J. Med. Inform. 114 (2018) 35–44.
[25] P. Compton, G. Edwards, B. Kang, L. Lazarus, R. Malor, P. Preston, A. Srinivasan, Ripple down rules: Turning knowledge acquisition into knowledge maintenance, Artif. Intell. Med. 4 (6) (1992) 463–475.
[26] D. Richards, Two decades of ripple down rules research, Knowl. Eng. Rev. 24 (2) (2009) 159–184.
[27] T. Wada, H. Motoda, T. Washio, Knowledge acquisition from both human expert and data, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2001, pp. 550–561.
[28] P. Compton, L. Peters, G. Edwards, T. Lavers, Experience with ripple-down rules, Knowl.-Based Syst. 5 (19) (2006) 356–362.
[29] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[30] D. Shen, G. Wu, H.-I. Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (2017) 221–248.
[31] G. Lee, H. Fujita, Deep Learning in Medical Image Analysis: Challenges and Applications, Vol. 1213, Springer, 2020.
[32] C.-H. Hua, T. Huynh-The, K. Kim, S.-Y. Yu, T. Le-Tien, G.H. Park, J. Bang, W.A. Khan, S.-H. Bae, S. Lee, Bimodal learning via trilogy of skip-connection deep networks for diabetic retinopathy risk progression identification, Int. J. Med. Inform. 132 (2019) 103926.
[33] S.I. Ali, B. Ali, J. Hussain, M. Hussain, F.A. Satti, G.H. Park, S. Lee, Cost-sensitive ensemble feature ranking and automatic threshold selection for chronic kidney disease diagnosis, Appl. Sci. (Switz.) 10 (16) (2020) http://dx.doi.org/10.3390/app10165663.
[34] S.I. Ali, S. Lee, Ensemble based cost-sensitive feature selection for consolidated knowledge base creation, in: Proceedings of the 2020 14th International Conference on Ubiquitous Information Management and Communication, IMCOM 2020 (2017), 2020, http://dx.doi.org/10.1109/IMCOM48794.2020.9001751.
[35] D.V. Carvalho, E.M. Pereira, J.S. Cardoso, Machine learning interpretability: A survey on methods and metrics, Electronics 8 (8) (2019) 832.
[36] M. Hussain, J. Hussain, T. Ali, F.A. Satti, S. Lee, Ai driven knowledge extraction from clinical practice guidelines: Turning research into practice, 2020, arXiv:2012.05489.
[37] M. Hussain, F.A. Satti, J. Hussain, T. Ali, S.I. Ali, H.S.M. Bilal, G.H. Park, S. Lee, A practical approach towards causality mining in clinical text using active transfer learning, 2020, arXiv:2012.07563.
[38] O. Bodenreider, The unified medical language system (umls): integrating biomedical terminology, Nucleic Acids Res. 32 (suppl_1) (2004) D267–D270.
[39] T. Ali, M. Hussain, W.A. Khan, M. Afzal, J. Hussain, R. Ali, W. Hassan, A. Jamshed, B.H. Kang, S. Lee, Multi-model-based interactive authoring environment for creating shareable medical knowledge, Comput. Methods Programs Biomed. 150 (2017) 41–72.
[40] X. Zhao, Y. Jia, A. Li, R. Jiang, Y. Song, Multi-source knowledge fusion: a survey, World Wide Web 23 (4) (2020) 2567–2592.

[41] J.W. Smith, J. Everhart, W. Dickson, W. Knowler, R. Johannes, Using the adap learning algorithm to forecast the onset of diabetes mellitus, in: Proceedings of the Annual Symposium on Computer Application in Medical Care, American Medical Informatics Association, 1988, p. 261.
[42] Coronary heart disease dataset, 2020, https://www.kaggle.com/billbasener/coronary-heart-disease, accessed: 2020-09-25.
[43] A.D. Association, et al., 2. classification and diagnosis of diabetes: Standards of medical care in diabetes-2020, Diabetes Care 43 (Suppl 1) (2020) S14.
[44] NICE, Diabetes in adults, 2016, no. March, URL http://guidance.nice.org.uk/QS6.
[45] S. I. G. Network, Management of Diabetes-a National Clinical Guideline (No 116), NHS Quality Improvement Scotland, Edinburgh, 2010.