Informatics in Medicine Unlocked 10 (2018) 117–125

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

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

ST-ONCODIAG: A semantic rule-base approach to diagnosing breast cancer base on Wisconsin datasets

O.N. Oyelade [a,*], A.A. Obiniyi [a], S.B. Junaidu [a], S.A. Adewuyi [b]

[a] Department of Computer Science, Faculty of Physical Science, Ahmadu Bello University, Zaria, Nigeria
[b] Department of Radiotheraphy and Oncology, Ahmadu Bello University Teaching Hospital, Zaria, Nigeria

ARTICLE INFO

Keywords:
Reasoning algorithm
Select and test
Diagnoses
Breast cancer
Instances
Attributes

ABSTRACT

Breast cancer is a major terminal disease that occurs largely among females. This disease stems from abnormal mutations in the genes of normal cells, thereby resulting in development of cancerous cells. Though there have being several research breakthroughs in the field of medicine in taming this disease, however, computer aided diagnosis on the other hand has proven very supportive in the quest. Techniques such as Machine Learning (ML) and Medical Expert Systems (MES) algorithms have added impetus to the use of artificial intelligence in detecting and diagnosing breast cancer. While MES may seem promising in machine based diagnostic systems, their accuracy is often impaired by inefficient medical reasoning algorithms employed. This paper therefore seeks to address the limitation of one such reasoning algorithm known as Select and Test (ST). The approach in this paper is to first create an efficient input mechanism that enables the system to read, filter and clean input from datasets. Secondly, semantic web languages (ontologies and rule languages) were used to create a coordinated rule set and a knowledge representation framework was created to aid the reasoning algorithm. As a result, the reasoning structures of ST were modified to accommodate this enhancement. Thereafter, the input generating mechanism was used to transform instances of the databases of Breast Cancer Wisconsin Data set retrieved from UCI Learning Repository. The generated inputs were passed into the improved ST algorithm to diagnose breast cancer in patients captured in the datasets. Experiments were carried out, and result show that 26.60%, 56.17%, and 54.05% were diagnosed of breast cancer in Wisconsin Breast Cancer Database (WBCD), Wisconsin Diagnostic Breast Cancer (WDBC), and Wisconsin Prognostic Breast Cancer (WPBC) respectively.

1. Introduction

The Center for Managing Chronic Diseases (CMCD) in United States of America defined a chronic disease as a long standing condition that can be controlled but not cured [1]. Cancer in particular is the uncontrolled growth and spread of cells. It can affect almost any part of the body, and there are varying organs that this disease affects, thereby leading to the types of cancer known so far. The World Health Organization (WHO) [2,3] enumerated some basic facts about cancer. These facts are: firstly, cancer causes approximately 14 million new cases and 8.2 million cancer related deaths in 2012, with this new cases rising to about 70% between the years 2030–2040. Secondly, among the men folk, cancer is usually sited around the lung, prostate, colorectum, stomach, and liver, while the women folk is usually sited in the breast, colorectum, lung, cervix, and stomach. The observation of WHO major causes of cancer deaths are centered around high body mass index, low fruit and vegetable intake, lack of physical activity, tobacco use, alcohol use. With tobacco use noted to be the most important risk factor for cancer causing around 20% of global cancer deaths and around 70% of global lung cancer deaths. Lastly on these facts, more than 60% of world's total new annual cases occur in Africa, Asia and Central and South America. It is expected that annual cancer cases will rise from 14 million in 2012 to 22 within the next 2 decades. In the meantime, the WHO's definition of health is not merely the absence of disease but the attainment of a state of physical, mental, emotional and social wellbeing [4]. However, our perspective of ill health is the presence of disease – breast cancer in particular.

Pervasive computing and artificial intelligence put together will result into a stronger framework for deployment of accessible and approximate medical expert systems. Really, Artificial Intelligent (AI) systems are supposed to support health workers in tasks that rely on data and knowledge manipulation [5]. Systems resulting from this are usually termed Clinical Decision Support Systems (CDSS), and are useful in alerts

* Corresponding author.
E-mail addresses: olaide_oyelade@yahoo.com (O.N. Oyelade), aaobiniyi@yahoo.com (A.A. Obiniyi), abuyusra@gmail.com (S.B. Junaidu), sadewuyi2003@yahoo.com (S.A. Adewuyi).

https://doi.org/10.1016/j.imu.2017.12.008
Received 6 October 2017; Received in revised form 19 December 2017; Accepted 20 December 2017
Available online 28 February 2018

and reminders, therapy critiquing and planning, diagnostic assistance, information retrieval, image recognition and interpretation and prescribing decision support system. In Ref. [6], the authors revealed that CDSS significantly improved clinical practice by 69% of trials. Diagnoses Decision Support System (DDSS) a type of CDSS, requires patient to enter some required information and then intelligently carry out diagnoses and then respond to patient with appropriate set of diagnoses, and then a physician may select the diagnosis relevant to the patient in question [7]. These decision support systems are driven by reasoning algorithms such as Select and Test (ST).

This paper demonstrates the use of an improved medical reasoning algorithm called ST. Furthermore, the improved ST was enabled to automatically generate input base on clinical protocols from benchmarking datasets from of Breast Cancer Wisconsin Data set retrieved from UCI Learning Repository. Then, experimental results were generated, presented, and compared for performance measurement. The remaining part of this paper consists of related works, the improved reasoning algorithm, the input generating procedure explained, implementation of the proposed system, result and discussion, and finally conclusion.

2. Related works

In reviewing related works, this paper organizes this review into three namely: medical expert system, medical diagnostic algorithm, and an overview of breast cancer.

2.1. Medical expert system

Medical expert systems (MES) continue to play varying roles, such as electronic medical record management, medical diagnoses, and medical image processing. This paper however focuses on the diagnostic capabilities of MES. Some classical health informatics applications includes MYCIN - an early expert system for the diagnosis and treat meningitis and bacteremia [8], INTERNIST-I is an experimental computer-based diagnostic consultant for general internal medicine [9], CADIAG-1 and CADIAG-2 - expert system for diagnosing disease relating to internal medicine [10,11], designed an Expert System for Thyroid Disease Diagnosis (ESTDD), and PUFF [12].

On the other hand, some modern DDSS includes the following: the authors [13] developed an information system that stores knowledge on breast cancer and cases of treatment applied in previous treatment of patients. This system is expected to serve medical personals with relevant knowledge to speed up their tasks. Meanwhile, the authors were able to add breast cancer staging capability to the system through the use of data of the Siriraj Hospital algorithm. Also [14], employed the use of iterative algorithm in gathering features, extracted from their electronic health record (EHR) that exposes the patients to breast cancer. The resulting system from that algorithm became a computer aided screening system. The authors in Ref. [43] studied different methods for detecting cancer. Their data was sourced by using Matlab software to pre-process histological breast images to extract first and second order statistical features. Thereafter, different classifiers (Support Vector Machine, linear discriminant analyzing and neural network) were employed in classifying the input. In Ref. [44], the authors proposed Computer Aided Diagnosis (CAD) System for breast cancer using mammography images. Using statistical features of a mammogram, the extractions were made using simple image processing techniques to determine if a breast tissue is either benign or malignant.

In the research presented by Ref. [45], the authors demonstrated novelty by creating an amalgamation of artificial immune recognition system (AIRS) and real tournament selection mechanism to increase the accuracy of the diagnoses case-studied. Though, AIRS has been used in diagnosing various diseases [45], specifically employed their improvement in using the resulting system (RAIRS2) to diagnose Tuberculosis (TB). Now, their aim was to improve the classification accuracy of the diagnosing algorithm. This aim was objectively targeted at controlling the population size of the model and to curtail the difficulty in selection from the population. Evidently, their result showed that the accuracy of the hybridized algorithm was adjudged perfect with root-mean square error of zero. On the other hand [46], carried out a survey on different computational intelligence techniques for the functionality of protein. Basically, their review borders around the following categories: using sequence and structure, protein interaction network, gene expression data, pathway analysis from gene expression data, and using sequence derived properties. The need for this comparison was necessitated by the fact that protein functions seems to appear changing just when one computational intelligence techniques assumes to have determined its functionality. In fact [46], noted that categorically mentioned the failure of the traditional homology based approach triggering the birthing of new techniques for protein function prediction. However, even these techniques when used in isolation have yet to prove sufficient in helping to certify and identify the functionalities of proteins. Hence [46], concluded their review with a discovery that hybridizing multiple computational techniques and different data sources will most likely proof sufficient in ascertaining protein functions. Moreover, considerations must be made so as to factor in each protein functionality discovery. This arises from the fact that a particular discovery peculiar to one protein's functionality will not apply to another.

2.2. Medical diagnostic algorithm

Meanwhile, these MES are driven by different underlying clinical reasoning algorithms. These algorithms may be classified into three: probabilistic, model based and rule based. Some of the clinical reasoning algorithms of the rule based includes: Scheme inductive reasoning [15] is based on adding characteristics of the syndrome to narrow the list of potential diagnoses. In scheme inductive reasoning, schemes are drawn to resemble that of road maps. It helps clinicians break down information into chunks, storing them in their memory and then retrieving them subsequently for problem solving task. Another algorithm is pattern recognition which is employed in machine learning for assigning some outputs to some inputs based on the coordination of a given algorithm [16]. Furthermore, hypothetico-deductive a reasoning algorithm involves the self-reflection and informed clinical decision making process of generating and testing hypotheses in association with the patient's presenting symptoms and signs [17]. Also, forward chaining system, includes writing rules to manage sub goals. Whereas, backward chaining systems automatically manage sub goals [18]. While forward reasoning is efficient and fast, backward reasoning can be employed to resolve the conflict between two competing hypotheses. A combination of the two reasoning method – backward and forward – with increased experience leads to increased coordination of hypothesis and evidence [19]. Next is the fuzzy logic based clinical reasoning algorithm. Here, linguistic variables are used to represent operating parameters in order to apply a more human-like way of thinking [20]. Some of the main factors affecting fuzzy logic model performance is data clustering for membership function generation. Processing model for diagnostic reasoning is another reasoning algorithm [21]. Lastly in this category is the ST algorithm, adjured to be the most approximate, which was proposed by Refs. [22,23]. However, the ST algorithm is limited by the following features: the abstraction layer omits an open-ended symptom elicitation method or input mechanism; data modeling was done strictly on tabular or entity relational form; the implementation of the induction stage is omitted in their work; there is no way to model the severity of symptoms elicited from patient.

Furthermore, on model based reasoning algorithms, parsimonious covering theory (PCT) works on the basis of associating a disorder to a set of manifestations. It uses two finite sets (disorders & manifestations) to define the scope of diagnostic problems [24]. Two basic limitations of PCT are: domain knowledge is not associated with time (*atemporal*) and one or more disorder or manifestation in the cover (Parsimonious criterion) could easily adversely affects the explanation. Another model based algorithm is certainty factor (CF), which is used for managing uncertainty cases in a rule based system [25] and can be interpreted as measures of

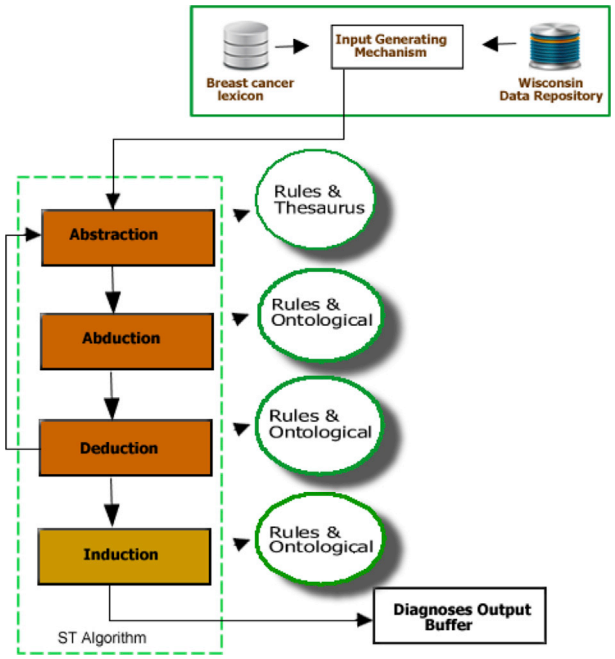

Fig. 1. The modified ST model with its algorithms and the input generating interface.

change in belief within the theory of probability [26].

Finally, the third category which is the probabilistic category of clinical reasoning algorithms includes Bayesian networks: an oriented acyclic graphs consisting of nodes (circles), which represent random variables; arcs (arrows), which represent probabilistic relationships among these variables [27] and this helps in dealing with uncertainties.

2.3. Overview of breast cancer

Breast cancer is a malignant tumor that starts in the cells of the breast. Malignant tumor is a group of cancer cells that can grow into surrounding tissues to distant areas of the body [28]. Though this disease occurs almost entirely in women, but men can be diagnosed of it too. The female breast is made up mainly of lobules (milk-producing glands), ducts (tiny tubes that carry the milk from the lobules to the nipple), and stroma (fatty tissue and connective tissue surrounding the ducts and lobules, blood vessels, and lymphatic vessels). Most breast cancers begin in the cells that line the ducts (ductal cancers). Early detection means using an approach that lets breast cancer get diagnosed earlier than otherwise might have occurred.

3. The proposed select and test (ST) model and the input generation interface

The medical expert system proposed in this paper is based on the model illustrated by Fig. 1. This diagnostic system employs the use of an input generating mechanism to produce input. The mechanism became necessary giving the need to translate attribute values from the datasets into tokens recognized by the proposed diagnostic model in Figure can understand. This token are then passed on to the reasoning structures (discussed in Section 3.2) of ST algorithm. This algorithm intelligently carries out the diagnostic task based on some modularized ontological knowledgebase in Ref. [30]. Thereafter, the result of the diagnoses are formatted and output for necessary decision making.

3.1. Acceptable tokens input generating mechanism

The ST model in Fig. 1 collects input in the form of symptoms, clinical examinations, laboratory observations and other reserved words in the domain of breast cancer. The basis for this facility is to provide an interface between the 'unformatted' dataset/database and the proposed reasoning algorithm. The use of the word 'unformatted' is meant to state that the Wisconsin datasets are not formatted for the need of this proposed reasoning algorithm. Meanwhile, the pattern of generating tokens by this mechanism is discussed in Section 4. The process of generating this input covers the abstraction level modeled in Fig. 1.

3.2. ST reasoing structures

The ST algorithm as modified by these authors in Refs. [29,30], consists of four reasoning levels namely: abstraction, abduction, deduction, and induction.

The aim of the abduction level is to retrieve all diagnosis that relates to some given symptoms. Actually, following the order of execution, the output of abstraction level becomes an input into the abduction level. Therefore, to elicit diagnosis base on the list of symptoms sent in from abstraction algorithm, the reasoning pattern of abduction is used. Abduction reasons by hypotheses: studying the available facts and creating a means to explain it. Normally, in medical facilities, while patients are expected to present symptoms, medical personnel on the other hand seeks out for signs to support their medical reasoning process, which usually help in deciding on a particular which disease being suspected. This is why the abduction and the deductive algorithms can be seen as the process of reasoning clinically to differentiate and drop out unlikely disease leaving the likely ones. So, let it be noted that the abduction algorithm aims to generate a list of reasonable diseases with respect to presented symptoms and observed signs.

In our improved ST algorithm, deductive reasoning acts as a process for using general premises to obtain specific inference. It is a form of logic that identifies particular disease by its semblance to a set of accepted facts or pattern. Deduction process supports its conclusions with TRUE result. The aim of deduction algorithm is to elicit symptoms from diagnosis: it gets all the symptoms modeled by an expert, shown in Ref. [30], to actually relate with each of the disease in the knowledge base.

Reasoning by induction is the last point of reasoning in the model we presented in Fig. 1. While the abstraction module is considered to be the point of input collection from the input generation mechanism, abduction and deduction modules as the process of carrying out clinical reasoning, induction reasoning can be seen as point where clinical decision is taken. In our proposed algorithm, reasoning by induction will be used to select the disease (out of the likely diseases) that meets the diagnostic criteria.

Now this entire processes (abstraction, abduction, deduction, and induction) are further modeled by ST algorithm in Section 3.4. This paper implements this cyclic pattern of abstraction-abduction-deduction and induction module algorithms.

3.3. Knowledge representation

Again, these authors in Ref. [30] have developed an ontological knowledge framework for the improved ST algorithm. As a result, the knowledge bases developed under the framework were used in executing this proposed medical expert system.

It will be observed from Fig. 1 that each of the sub-algorithm or level (abstraction, abduction, deduction, and induction) has its modularized knowledge base. Each modularized knowledge base consists of facts and rules. The facts were modeled as ontologies and the rules represented using semantic web rule language (SWRL) and Java expert system shell (JESS) language.

3.4. ST algorithm

The four levels of reasoning described in Section 3.2 above are hereby modeled using our improved ST algorithm outlined by Algorithm 1.

Algorithm 1. ST Algorithm

Algorithm 1: ST Algorithm

```
Let
  symptomsFound; symptomsToBeElicited;
  symptomsAlreadyElicited;
  diagnosesToBeElicited;
  diagnosesAlreadyElicited;
  likelihoodThreshold,  criticalityThreshold,
  symptomsWeight;
 be variables initialized to default values

 WHILE diagnoses not done
 //ABDUCTION: all the diagnoses related to symptoms
  get all related_rules from rule system.
  get facts from domain causal ontology
  FOR EACH symptom in symptomsFound DO
    declare diagnoses as temp linkedlist
    get all the diagnoses related to symptom
    diagnosis ← reasoner(engine, rule, symp)
    IF   explanation is true  THEN
       store diagnosis  in diagnoses
    END IF
    FOR EACH diagnosis in diagnoses DO
     IF   diagnosis   is   NOT   already   in
(diagnosesToBeElicited                    OR
diagnosesAlreadyElicited) THEN
      store diagnosis in diagnosesToBeElicited
      END-IF
    END FOR EACH
  END FOR EACH
//DEDUCTION: get all the symptoms related to diagnoses
  WHILE diagnosesToBeElicited is NOT empty DO
   declare symptoms as a temporary linked list
  IF deductionRuleEngine (symptom)
    THEN  get all the symptoms related to the
  current diagnosis in diagnosesToBeElicited
  END IF
  FOR EACH symptom in symptoms DO
   IF symptom is NOT already in (symptomsFound
        OR symptomsAlreadyElicited) THEN
       store symptom in symptomsToBeElicited;
  END-IF
END-FOR
    remove the current diagnosis from
diagnosesToBeElicited and store it in
diagnosesAlreadyElicited;
    next diagnosis in diagnosesToBeElicited
becomes the current diagnosis
END-WHILE
//ABSTRACTION; Check if the expected symptoms in likely
diagnoses are found in patient
    semantic_input_token =read user input
from input model
    patientProfile =
createSummaryStatement(semantic_input_token)
    symptomsToBeElicited=
semantic_input_token
   IF exist(patient profile file)
      update patient profile
   ELSE
   create patient profile file
   add new info to patient profile
   WHILE symptomsToBeElicited is NOT empty DO
   IF the current symptom in
symptomsToBeElicited is found in
patientProfile THEN
     store the current symptoms in
symptomsFound;
   END-IF
   store the current symptom in
symptomsAlreadyElicited
   remove the current symptom from
symptomsToBeElicited
   next symptom in symptomsToBeElicited
becomes the current symptom;
END-WHILE
//INDUCTION: Check if the likely diagnoses meet their
diagnostic criteria
  FOR EACH diagnosis in
diagnosesAlreadyElicited DO
    criteria  = diagnosticCriteria(
symptomsFound [observation], rule, KB)
    IF criteria is satisfied
       THEN  diagnosis is concluded
          computeCriticalityThreshold(
symptomsFound, symptomsFoundWeight,
monitoring_log)
       Advice treatment
       show criticalityThreshold
    ELSE
       diagnosis is not concluded;
    END-IF
  END-FOR
 END
```

```
symptomsFound=AppContainer.userInputModelsymptomsFound;//CALL SUB-PROGRA
AbstractionAlgorithm abstractalgol=new AbstractionAlgorithm(this);
DeductiveAlgorithm deductalgol=new DeductiveAlgorithm(this);
AbductiveAlgorithm abductalgol=new AbductiveAlgorithm(this);
QueueAsLinkedList tmpSymptomsAlreadyElicited, tmpDiagnosesToBeElicited,
while(!symptomsFound.isEmpty()){ //WHILE we are still treating initial s
  tmpDiagnosesToBeElicited=abductalgol.algorithm(); //CALL SUB-PROGRAM->
  showstore.updateAbductiontionOutput(tmpDiagnosesToBeElicited);
  tmDpdiagnosesAlreadyElicited=deductalgol.algorithm();//CALL SUB-PROGRA
  showstore.updateDeductionOutput(tmDpdiagnosesAlreadyElicited);
  tmpSymptomsAlreadyElicited=abstractalgol.algorithm();//CALL SUB-PROGRA
  showstore.updateAbstractionOutput(tmpSymptomsAlreadyElicited);
  symptomsFound.dequeue();
} //END-WHILE
InductiveAlgorithm inductivealgol=new InductiveAlgorithm(this);//CALL SU
inductivealgol.algorithm();
showstore.setDiagnosesResult(alignedResult);//show output
```

Fig. 2. Implementation of the improved ST algorithm.

```
DataSetFileFormater dsff=new DataSetFileFormater(699, 11, 100, 1);
dsff.initOutputFile(ConstantParams.BC_FORMATED_DATA);
dsff.readFile(ConstantParams.BC_WISCONSIN_DATA_ATTR_NAMES, 0);
dsff.retSetAttribute();
dsff.readFile(ConstantParams.BC_WISCONSIN_DATA, 1);
dsff.lauchOncoDiag();
dsff.write2MSWordFileTable();
dsff.closeOutputFile();
isFirstDataset=false;
DataSetFileFormater dsff1=new DataSetFileFormater(198, 8, 50, 2);
dsff1.initOutputFile(ConstantParams.WPBC_FORMATED_DATA);
dsff1.readFile(ConstantParams.WPBC_DATA);
dsff1.lauchOncoDiag();
dsff1.write2MSWordFileTable();
dsff1.closeOutputFile();
DataSetFileFormater dsff2=new DataSetFileFormater(569, 5, 100, 2);
dsff2.initOutputFile(ConstantParams.WDBC_FORMATED_DATA);
dsff2.readFile(ConstantParams.WDBC_DATA);
dsff2.lauchOncoDiag();
dsff2.write2MSWordFileTable();
dsff2.closeOutputFile();
```

Fig. 3. Implementation of the input generation mechanism.

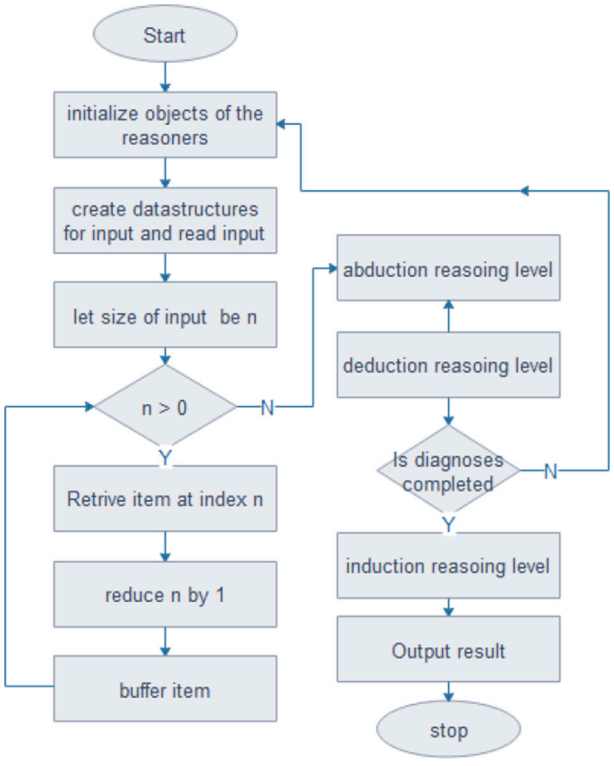

Fig. 4. Flow chart of the improved ST algorithm.

The novelty of this proposed algorithm lies in its ability to reasoning using semantic web tools (pellet and Jess), represents its knowledge base as ontologies and rule languages, and detecting the criticality of diagnosis, staging it and advice on treatment plan.

In conclusion, the complexity of the improved ST algorithm is proven to be $O(n^2)$. This was deduced from the complexity of abstraction, abduction, deduction and induction reasoning levels which are O(n), $O(n^2)$, $O(n^2)$, and O(n) respectively.

4. Parsing the datasets for input generation

There are three datasets/databases sourced as input into the model described the previous section. These are Wisconsin Breast Cancer Database (WBCD), Wisconsin Diagnostic Breast Cancer (WDBC), and Wisconsin Prognostic Breast Cancer (WPBC) [37–42].

4.1. Sourcing WBCD as input into improved ST algorithm

The following are the attributes in this dataset. Sample code number, clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, mitoses, class (for benign or malignant). The approach of this paper in interpreting the values of each these attributes for all instances. For each attributes (except the first and the last), the clinical significance of the attributes were identified alongside the all the symptoms, signs, risk factors of breast cancer associated with the particularly. However, the severity of such interpreted attributes is assigned the corresponding numerical value of the attribute for each instance in the dataset.

For example, given an instance with the following values for the list of attributes above: 1000025,5,1,1,1,2,1,3,1,1,2, the first and the last attributes have already being differentiated. But the second attribute (clump thickness) has the numerical value 5. Hence, a list of symptoms, signs, risk factors, or any clinical terms on breast cancer which are associated with this attribute are generated. Then, these terms/words will be assigned the likelihood/weight value 5 before sending them as input into the system in Fig. 1.

4.2. Sourcing WDBC as input into improved ST algorithm

This dataset/database also contains the following attributes: ID number, diagnosis (malignant or benign), ten real-valued features are computed for each cell nucleus: a) radius (mean of distances from center to points on the perimeter), b) texture (standard deviation of gray-scale values), c) perimeter, d) area, e) smoothness (local variation in radius lengths), f) compactness (perimeter2/area - 1.0), g) concavity (severity of concave portions of the contour), h) concave points (number of concave portions of the contour), i) symmetry, j) fractal dimension ("coastline approximation" - 1). The mean, standard error, and largest (mean of the three largest values) of these features were computed for each image, resulting in 30 features. For instance, field 3 is Mean Radius, field 13 is Radius SE, and field 23 is Worst Radius.

Similarly, the approach of this paper is as described Section 4.1 above. However, each ten real-valued features are summed up to obtain a single value for the feature they all represents. And as described in Section 4.1, this same procedure is done for all instances in the dataset.

4.3. Sourcing WPBC as input into improved ST algorithm

Lastly, is the WPDC dataset/database which has the following attributes: ID number, outcome (recur or non-recur), time (recurrence time or disease-free time), ten real-valued features (as in WDPC above) are computed for each cell nucleus. Other attributes are tumor size - diameter

of the excised tumor in centimeters, and lymph node status - number of positive axillary lymph nodes observed at time of surgery.

In addition to the way WDPC attributes were interpreted as described above, the last two attributes (tumor size and number of positive axillary lymph nodes) in WPDC were very clear in terms of interpretation – the symptoms and signs to be associated with these attributes are easily known.

5. Implementation of the modified ST algorithm using the Wincosin dataset

Both the improved ST algorithm and the input generating mechanism were implemented with Java programming language, Figs. 2 and 3 shows the implementations of the ST algorithm and the input generating mechanism respectively. In the same vein, Figs. 3 and 4 shows flow charts to buttress the illustration of the code snippets.

6. Result and discussion

In this research, we embarked on validating the proof of the result obtained by using standard or benchmark dataset for comparing the performance of the implemented medical diagnosis system which is based on an improved Select and Test (ST) algorithm. Three datasets where retrieved from their online repositories, and these datasets are namely: Wisconsin Breast Cancer Database (WBCD) authored by Wolberg (1992), Wisconsin Diagnostic Breast Cancer (WDBC) authored by Ref. [31], and Wisconsin Prognostic Breast Cancer (WPBC) authored by Ref. [32]. These databases or datasets have been used in some research works which includes [33–35].

The WBCS database has 699 instances (as of 15 July 1992), 10 numbers of attributes plus the class attribute. The 10 attribute Information are: instance code number, Clump Thickness, Uniformity of Cell, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell, Bare Nuclei, Bland Chromatin, Normal Nucleoli, Mitoses, and Class (benign or malignant). The WDBC database has 569 instances, 32 attributes: instance code number, diagnosis (malignant or benign), 30 real-valued input features). Lastly, the WPDC database has 198 number of instances, 34 number of attributes which consists of instance code number, outcome (recur or nonrecur), Time (recurrence time, or disease-free time), 32 real-valued input features, Tumor size, and Lymph node status (number of positive axillary lymph nodes) observed at time of surgery. The WBCS, WDBC, and WPBC datasets have their class distribution as Benign: 458 (65.5%) and Malignant: 241 (34.5%), 357 benign, 212 malignant, and 151 nonrecur, 47 recur respectively.

The approach of this research in using this datasets is to generate input from each dataset and then feed them inputs into the medical expert system proposed by the research. This source of input became relevant to this research considering the similarity in the ailment (breast cancer) being diagnosed by the providers of the databases and this research. Hence, based on the given attributes in each dataset, this research retrieved corresponding inputs from the database of acceptable tokens/inputs developed in the course of this research. Since the datasets only lists attributes names and their corresponding weights on each instance, this research then embark on the style of sourcing al symptoms, signs, medical terms modeled in our database which matches a given attribute, and then assign the weight of the attribute for each instance to be the likelihood of the symptom or sign to be fed as input into the proposed medical expert system.

In a quest to maintain brevity of the generated input into the proposed medical expert system, acronyms were defined and their meaning captured in Table 1. These medically acceptable terms have earlier being modeled in an ontology file as part of this research contribution. Meanwhile, these acronyms are used as the automatically generated input for the improved ST based medical expert system to use.

As earlier stated, Algorithm 1 shown above was implemented and some code snippets already shown in Figs. 2 and 3. The implementation of Algorithm 1 was named ST-ONCODIAG. Therefore, while carrying out our evaluation, the comparison is done with the result of ST-ONCODIAG when using WBCS, WDBC and WPBC datasets against the class values (results of) that datasets themselves. Therefore, Sections 6.1–6.3 detail these comparisons.

Table 1
List of some acronyms for generated acceptable medical terms.

S/N	Medical terms as modeled in knowledge base	Acronym
1	Breast_Lumps	BL
2	Breast_Swelling	BS
3	Increase_in_Size	IS
4	Lymph_Nodes	LN
5	Oedema	O
6	NippleRetractionOrDeviation	NRD
7	Nipple_Discharge	ND
8	StageIV_Metastasis	SM
9	Tenderness	T
10	Nipple_Retraction	NR
11	Hard	H
12	Cough	C
13	Thickening_of_the_Nipple	TN
14	Invasic_Ductal	ID
15	Pain	P
16	Female	F
17	Lump	L
18	Alchol_Consumption	AC
19	Smoking	S
20	History_of_Hormonal_Replacement	HHR
21	Fatigue	FT

6.1. Evaluating and comparing WBCS with ST-ONCODIAG

Our result in Table 2 shows that we have Non Breast Cancer (benign) diagnosed instances to be 73.40% and Breast Cancer (malignant) diagnosed instances to be 26.60% as against the 458 (65.5%) cases of benign and 241 (34.5%) cases of malignant observed by the authors of the dataset. In addition, the staging of the positive (malignant) diagnosis were also staged using the breast cancer staging standard, and a treatment plan was suggested as shown in Table 3. Furthermore, Table 2 draws a similarity of the result for the benchmark dataset and the result of ST-ONCODIAG. Similarly, Fig. 5 shows a graphical illustration of Table 2. We observed that our proposed ST-ONCODIAG was not as accurate as the projected 458 (65.5%) cases of benign and 241 (34.5%) in the dataset. However, the proposed system was able to realize a considerable good percentage of accuracy. These slight discrepancies can be noticed from the little deviations between the lines of benign and malignant for WBCS and ST-ONCODIAG respectively in Fig. 6.

6.2. Evaluating and comparing WDBC with ST-ONCODIAG

Also, the evaluation of the use of WDBC dataset in testing ST-ONCODIAG was carried out, and a comparison done shows that the Non Breast Cancer (benign) diagnosed instances to be 245 (43.83%) and Breast Cancer (malignant) diagnosed instances to be 314 (56.17%). This is manageably different from the 353 (63.15%) benign and 206 (36.85%)

Table 2
The comparison table of the results of WBCS and the proposed medical expert system result.

No. Instances	WBCS		ST-ONCODIAG	
	Benign	Malignant	Benign	Malignant
100	56	44	67	33
200	116	84	139	61
300	163	137	191	109
400	229	171	266	134
500	303	197	343	157
600	380	220	424	176
699	458	241	513	186

Table 3
Comparison table of the results of WBCS and ST-ONCODIAG for the first 10 instances.

Instance ID	WBCS	ST-ONCODIAG	Staging	Treatment
1000025	Benign	N-BC	N/A	No Treatment Plan
1002945	Benign	N-BC	N/A	No Treatment Plan
1015425	Benign	N-BC	N/A	No Treatment Plan
1016277	Benign	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
1017023	Benign	N-BC	N/A	No Treatment Plan
1017122	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
1018099	Benign	N-BC	N/A	No Treatment Plan
1018561	Benign	N-BC	N/A	No Treatment Plan
1033078	Benign	N-BC	N/A	No Treatment Plan
1033078	Benign	N-BC	N/A	No Treatment Plan

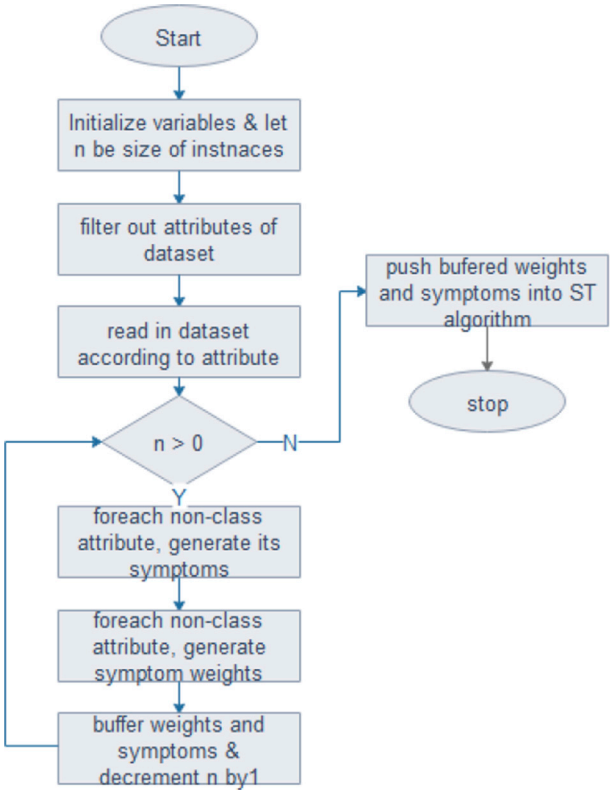

Fig. 5. Flow chart of the input generation mechanism.

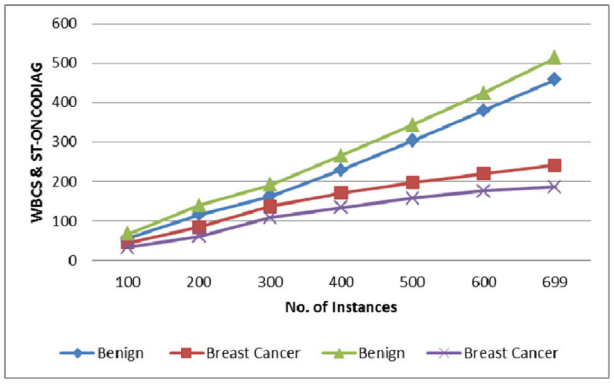

Fig. 6. Graph of comparison of the results of WBCS and the proposed medical expert system result.

malignant recorded in the dataset. Though, we opine that the slight sharpness in the differences could have resulted from noise during the input generation process. Meanwhile, Table 4 shows scaled similarities of the result of the benchmark dataset and the result of this proposed medical expert system (ST-ONCODIAG), and Fig. 7 shows a graphical illustration of Table 4, which confirms the observation on the slight harp differences we noted above. Moreover, the staging and the treatment plan of the breast cancer (malignant) diagnoses were computed and shown in Table 5.

6.3. Evaluating and comparing WPBC with ST-ONCODIAG

Finally, the WPDC dataset was also compared to the output of ST-ONCODIAG and results show that there are Non Breast Cancer (benign) diagnosed instances to be 45.95% and Breast Cancer (malignant) diagnosed instances to be 54.05% against the 151 (76.3%) non-

Table 4
The comparison of the results of WDBC and the proposed medical expert system result.

No. Instances	WDBC		ST-ONCODIAG	
	Benign	Malignant	Not-Breast Cancer	Breast Cancer
100	35	65	51	49
200	96	104	102	98
300	154	146	142	158
400	227	173	179	221
500	305	195	218	282
559	353	206	245	314

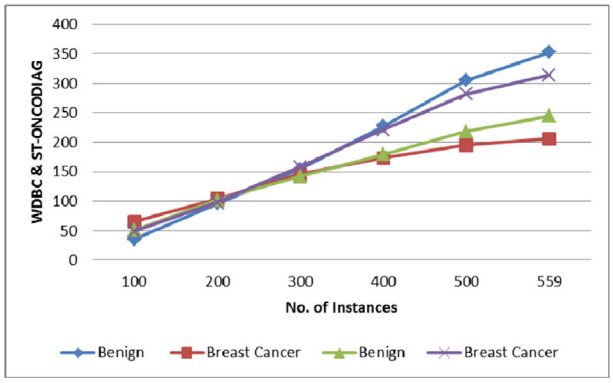

Fig. 7. Graph of similarity table of WDBC result and the proposed medical expert system result.

Table 5
Comparison table of the results of WDBC and ST-ONCODIAG for the first 10 instances.

Instance ID	WBCS	ST-ONCODIAG	Staging	Treatment
842302	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
842517	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
84300903	BC	N-BC	N/A	No Treatment Plan
84348301	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
84358402	BC	N-BC	N/A	No Treatment Plan
843786	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
844359	BC	N-BC	N/A	No Treatment Plan
84458202	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
844981	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
84501001	BC	BC	Stage 0	Endocrine Therapy, Radiation & Chemoteaphy
845636	BC	N-BC	N/A	No Treatment Plan

recur, 47 (23.7%) recur in the WPDC dataset. Similarly, the result of the experiments of ST-ONCODIAG is scaled side by side with that of WPDC dataset in Table 6. Furthermore, Table 4 draws a similarity table of the scaled similarities of the result of the benchmark dataset and the result of this proposed medical expert system. Furthermore, Fig. 8 shows a graphical illustration of the 151 non-recur, 47 recur, and the 91 benign with the 107 malignant of ST-ONCODIAG experiments. In addition, the staging and treatment plan of the positive cases diagnosed are listed in Table 7, though another metric – weight of symptoms on the diagnosis process – was computed.

In summary, Table 8 shows the Sensitivity and specificity indexes for the ST-ONCODIAG algorithm implementation discussed in the previous sections. These computations are on the formula in Eqs. (1) and (2). The specificity of WBCS, WDBC and WPCS databases in ST-ONCODIAG are 0.89, 0.706, and 0.60 respectively, while that of sensitivity in the same databases are 0.81, 1.0, and 1.0 respectively.

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \quad (1)$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \quad (2)$$

Where: TP (TN) = Number of True Positive (True Negative) and FP (FN) = Number of False Positive (False Negative).

7. Conclusion

In conclusion, this paper presents an improved ST algorithm for carrying out medical reasoning. The reasoning structures of the algorithm were discussed, and the knowledge base was also briefly introduced. Furthermore, an input generating mechanism for sourcing input into the medical expert system was proposed and implemented. This mechanism enables us format benchmarking dataset and then feed them into the expert system for the purpose of diagnostic task. Results shows that the proposed medical expert system was able to compute the specificity of the three databases as 0.89, 0.706, and 0.60 respectively, while that of sensitivity are 0.81, 1.0, and 1.0. Then 26.60%, 56.17%, and 54.05% were the positive diagnoses of breast cancer in WBCD, WDBC and WPCS databases. However, we note that the accuracy of the result in the experiment was limited by input generation mechanism. This arises from the fact that mapping the numeric values of the attributes for each instance in the datasets to the acceptable inputs for the ST-ONCODIAG was complex.

Table 6
The comparison table of the results of WPCS and the proposed medical expert system result.

No. Instances	WPCS		ST-ONCODIAG	
	Non-Recur	Recur	Not-Breast Cancer	Breast Cancer
50	36	14	15	35
100	77	23	29	71
150	112	38	54	96
198	151	47	91	107

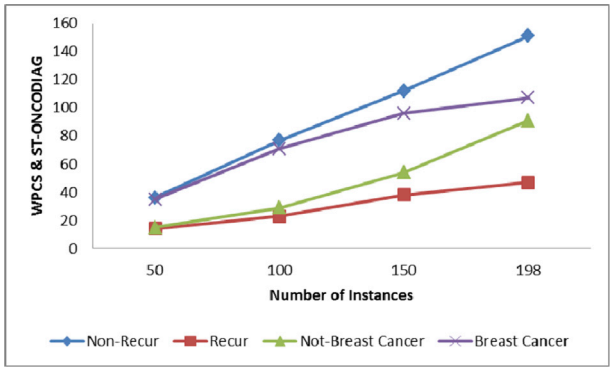

Fig. 8. Graph of comparison table of the results of WPCS and the proposed medical expert system result.

Table 7
Comparison table of the results of WPCS and ST-ONCODIAG for the first 10 instances.

Instance ID	WBCS	Enhanced ST	Staging	Weight of inputs	Treatment
119513	Nonrecur BC	BC	Stage 1B	0.31	Endocrine Therapy, Radiation & Chemoteaphy
8423	Nonrecur BC	BC	Stage 1B	0.55	Endocrine Therapy, Radiation & Chemoteaphy
842517	Nonrecur BC	BC	Stage 1B	0.2927	Endocrine Therapy, Radiation & Chemoteaphy
843483	Nonrecur BC	BC	Stage 1B	0.25	Endocrine Therapy, Radiation & Chemoteaphy
843584	Recur BC	N-BC	N/A	0.0	No Treatment Plan
843786	Recur BC	BC	Stage 1B	0.19	Endocrine Therapy, Radiation & Chemoteaphy
844359	Nonrecur BC	N-BC	N/A	0.0	No Treatment Plan
844582	Recur BC	BC	Stage 1B	0.29	Endocrine Therapy, Radiation & Chemoteaphy
844981	Nonrecur BC	BC	Stage 1B	0.15	Endocrine Therapy, Radiation & Chemoteaphy
845010	Nonrecur BC	BC	Stage 1B	0.15	Endocrine Therapy, Radiation & Chemoteaphy

Table 8
Sensitivity and specificity indexes for the ST-ONCODIAG algorithm implementation.

Metrics	ST-ONCODIAG		
	WBCS	WDBC	WPCS
Positive	186	314	47 (recur)
Negative	513	245	151 (non-recur)
True Positive	241	212	47
True Negative	458	245	91
False Positive	55	102	60
False Negative	55	0	0
Sensitivity	0.81	1.0	1.0
Specificity	0.89	0.706	0.60

Abbreviations

ST	Select and Test
WBCD	Wisconsin Breast Cancer Database
WDBC	Wisconsin Diagnostic Breast Cancer
WPBC	Wisconsin Prognostic Breast Cancer
ST-ONCODIAG	Select and Test oncology diagnostic system
MES	Medical expert system
SWRL	Semantic web rule language
JESS	Java expert system shell
CDSS	Clinical Decision Support Systems
DDSS	Diagnoses Decision Support Systems
AI	Artificial Intelligent

References

[1] CMCD. Retrieved on 29th August 2014, from http://cmcd.sph.umich.edu/what-is-chronic-disease.html.

[2] World Cancer Report. IARC Nonserial Publication; 2014. p. 630. ISBN-13 9789283204299.

[3] de Martel C, Ferlay J, Franceschi S. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol 2008;2012(13):607–15.

[4] Omotosho A, Emuoyibofarhe OJ, Adegboola O. ICT in Health care delivery system: a Framework for developing nations. 2005. Retrieved from, http://www.wikehealth.org/sites/default/files/whitepapers/139/ICT%20in%20Health%20care%20delivery%20system_%20A%20Framework%20for%20developing%20nations.pdf. on January 29, 2016.

[5] Coiera E. The Guide to Health Informatics. second ed. London: Arnold; October 2003.

[6] Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ April 2, 2005;330(7494):765. Epub 2005 Mar 14. Review.

[7] Berner, Eta S, editors. Clinical Decision Support Systems. New York: Springer-Verlag; 2007, ISBN 978-1-4419-2223-6. Edition Number 2.

[8] Shortliffe EH, Fagan LM. Artificial intelligence: the expert systems approach to medical consultation. In: Computer in Critical Care and Pulmonary Medcine. Springer Berlin Heidelberg; 1985. p. 190–7.

[9] Miller RA, McNeil MA, Challinor SM. The INTERNIST-1/QUICK MEDICAL REFERENCE project-Status report, in Medical informatics [Special Issue]. West J Med December 1986;145(816–822):816–22.

[10] Kolarz G, Adlassnig K. Problems in establishing the medical expert systems CADIAG-1 and CADIAG-2 in rheumatology. J Med Syst 1986;10(4):395–405.

[11] Azar AT, Hassanien AE. Expert system based on neural-fuzzy rules for thyroid diseases diagnosis. In: International Conference on Bio-Science and Bio-Technology (BSBT 2012), December 16-19, 2012, Korea. 353(1) of the Communications in Computer and Information Science series. Springer; 2012. p. 94–105.

[12] Aikins JS, Kunz JC, Shortliffe EH, Fallat RJ. PUFF: an expert system for interpretation of pulmonary function data. Comput Biomed Res 1983;16:199–208.

[13] Intarajak T, Kang SH. Breast cancer decision support system for rural people. In: International Conference on IT, March 2009, Thailand. Special Issue of the International Journal of the Computer, the Internet and Management, vol. 17; 2009. No. SP1.

[14] Alaa AM, Moon KH, Hsu W, van der Schaar M. ConfidentCare: a clinical decision support system for personalized breast cancer screening. IEEE Trans Multimed 2016;2016.

[15] Anderson KJ. Factors affecting the Development of Undergraduate Medical Students' Clinical Reasoning ability. Medicine Learning and Teaching Unit, Faculty of Health Sciences, University of Adelaide; 2006. p. 18.

[16] Umoh UA, Umoh AA, James GG, Oton UU, Udoudo JJ. Design of pattern recognition system for the diagnosis of Gonorrhea disease. Int J Sci Technol Res 2012;1(5):74–9.

[17] Kumar SP, Kumar A, Sisodia V. Clinical reasoning and sports medicine-application of hypothetico- deductive model. J Sports Med Doping Stud 2013;3(1). ISSN: 2161–0673 JSMDS.

[18] Sharma T, Tiwari N, Kelkar D. Study of difference between forward and backward reasoning. Int J Emerg Technol Adv Eng October 2012;2(10).

[19] Hardin LE. Research in medical problem solving: a review. JVME 2002;30(3):227–32. © 2002 AAVMC.

[20] Torshabi AE, Riboldi M, Pella A, Negarestani A, Rahnema M, Baroni G. A clinical application of fuzzy logic. Fuzzy Log Emerg Technol Appl 2013:1–18.

[21] Stausberg J, Person M. A process model of diagnostic reasoning in medicine. Int J Med Inf 1999;54:9–23.

[22] Fernando DAIP, Henskens FA. Select and test (ST) algorithm for medical diagnostic reasoning. Software engineering, artificial intelligence, networking and parallel/distributed computing. Stud Comput Intell 2016;653. https://doi.org/10.1007/978-3-319-33810-1_6.

[23] Fernando I, Henskens FA. ST algorithm for medical diagnostic reasoning. Polibits 2013;(48):23–9. ISSN 1870–9044.

[24] Wainer, J. and Rezender, A. M. (1996). A temporal extension to the parsimonious covering theory.

[25] Heckerman D. Probabilistic similarity networks. Networks August 1990;20(5):607–36.

[26] Heckerman DE, Shortliffe EH. From certainty factors to belief networks. Artif Intell Med 1992;4(1):35.

[27] Gadewadikar J, Kuljaca O, Agyepong K, Sarigul E, Zheng Y, Zhang P. Exploring Bayesian networks for medical decision support in breast cancer detection. Afr J Math Comput Sci Res October 2010;3(10):225–31.

[28] American Cancer Society. Breast Cancer Prevention and Early Detection. 2014. Accessed at, http://www.cancer.org/Cancer/BreastCancer/DetailedGuide/index. on November 13, 2015.

[29] Oyelade ON, Obiniyi AA, Junaidu SB, Kana AFD. A modified select and test (ST) algorithm for medical diagnosis in an ad-hoc network environment. In: IEEE 3rd International Conference on Electro-Technology for National Development (NIGERCON); 2017. p. 19–23.

[30] Oyelade ON, Obiniyi AA, Junaidu SB. An ontological-based knowledge framework for diagnosing breast cancer. In: 13th International Conference Proceedings of Nigerian Computer Society (NCS), vol. 28; 2017.

[31] Wolberg WH, Street WN, Mangasarian OL. Wisconsin Prognostic Breast Cancer (WPBC). University of Wisconsin, 1210 West Dayton St., Madison, WI 53706; 1995. wolberg@eagle.surgery.wisc.edu.

[32] Wolberg WH, Street WN, Mangasarian OL. An inductive learning approach to prognostic prediction. In: Prieditis A, Russell S, editors. Proceedings of the Twelfth International Conference on Machine Learning, pages 522–530. San Francisco: Morgan Kaufmann; 1995.

[33] Wolberg WH. Wisconsin Breast Cancer Database. Madison: Obtained from the University of Wisconsin, Hospitals; 1992.

[34] Mangasarian OL, Wolberg WH. Cancer diagnosis via linear programming. SIAM News September 1990;23(5):1–18.

[35] Wolberg WH, Mangasarian OL. Multisurface method of pattern separation for medical diagnosis applied to breast cytology. Proc Natl Acad Sci U S A December 1990;87:9193–6.

[37] Street WN, Wolberg WH, Mangasarian OL. Nuclear feature extraction for breast tumor diagnosis. In: IS&T/SPIE 1993 International Symposium on Electronic Imaging: Science and Technology, vol. 1905; 1993. p. 861–70. San Jose, CA.

[38] Mangasarian OL, Street WN, Wolberg WH. Breast cancer diagnosis and prognosis via linear programming. Oper Res July-August 1995;43(4):570–7.

[39] Wolberg WH, Street WN, Mangasarian OL. Machine learning techniques to diagnose breast cancer from fine-needle aspirates. Canc Lett 1994;77:163–71.

[40] Wolberg WH, Street WN, Mangasarian OL. Image analysis and machine learning applied to breast cancer diagnosis and prognosis. Anal Quant Cytol Histol April 1995;17(2):77–87.

[41] Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computerized breast cancer diagnosis and prognosis from fine needle aspirates. Arch Surg 1995;130:511–6.

[42] Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computer-derived nuclear features distinguish malignant from benign breast cytology. Hum Pathol 1995;26:792–6.

[43] Mohamed T. CAD Based automated Carcinoma Detection and Classification in Breast Cancer Diagnosis. An M.Sc. Thesis. University Of Eastern Finland, Faculty of Science and Forestry, School of Computing, Kuopio Campus; 2015. Retrived from, http://epublications.uef.fi/pub/urn_nbn_fi_uef-20150882/urn_nbn_fi_uef-20150882.pdf. on December 14, 2017.

[44] Alshamlan H, El-Zaart A. Breast Cancer Computer aided Diagnosis (CAD) System. 2012. Retrieved from, https://www.researchgate.net/publication/264885871_BREAST_CANCER_COMPUTER_AIDED_DIAGNOSIS_CAD_SYSTEM.pdf. on December 14, 2017.

[45] Saybani MR, Shamshirband S, Golzari S, Wah TY, Saeed A, Kiah LM, Balas VE. RAIRS2 a new expert system for diagnosing tuberculosis with real-world tournament selection mechanism inside artificial immune recognition system. Int Fed Med Biol Eng 2015;54(2–3):385–99.

[46] Tiwari KA, Srivastava R. Review Article: a survey of computational intelligence techniques in protein function prediction. Hindawi Publ Corp Int J Proteom 2014;2014:1–22. Article ID 845479, https://doi.org/10.1155/2014/845479.