NAME: Timothy-Jacob marvelous miyenseigha

DEPARTMENT: Computer Engineering

MATRIC NO: 18/ENG02/094

COURSE TITLE: Structured Programming

COURSE CODE: ENG 224

WEB-BASED MEDICAL ASSISTANT SYSTEM FOR COVID-19 DIAGNOSIS AND THERAPY

Introduction

Telemedicine is a way of delivering health information across distances. Telecommunication devices such as internet and telephone help lines are mainly used to break the barrier of distance. Though telemedicine has been regarded as one solution by which necessary health care services can be provided to everybody at a reasonable cost, the practice should be treated with caution so that the quality of treatment is not compromised. Economic savings alone should not be a justification for the use of telemedicine if the quality of health care provided deteriorates**.** The telephoned countries as leading countries in technological advancement have exploited these ICT potentials in bringing hospital to homes. Diagnosis is the identification of abnormal condition that afflicts a specific patient, based on manifested clinical data or lesions. If the final diagnosis agrees with a disease that afflicts a patient, the diagnostic process is correct; otherwise, a misdiagnosis occurred. Medical diagnosis is a categorization task that allows physicians to make prediction about features of clinical situations and to determine appropriate course of action. It involves a complex decision process that involves a lot of vagueness and uncertainty management, especially when the disease has multiple symptoms. Medical diagnosis has undergone different phases of research from statistical methods, which saw the application of Bayesian inference, utility theory, Boolean logic and discriminant analysis. When it was evident that statistical tools could not deal with most complex medical problems, Artificial Intelligence (AI) principles were applied.

The **2019–20 coronavirus pandemic** is an ongoing pandemic of coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The outbreak was identified in Wuhan, China, in December 2019, declared to be a Public Health Emergency of International Concern on 30 January 2020, and recognized as a pandemic on 11 March 2020. As of 10 April 2020, approximately 1.61 million cases of COVID-19 have been reported in 210 countries and territories, resulting in approximately 97,000 deaths. About 364,000 people have recovered. The case fatality rate has been estimated to be 4% in China, while globally ranging from 13.04% in Algeria to .08% in New Zealand.

The virus is mainly spread between people during close contact, often via small droplets produced during, sneezing, or talking. While these droplets are produced when breathing out, they usually fall to the ground or surfaces rather than being infectious over large distances. People may also become infected by touching a contaminated surface and then their face. The virus can survive on surfaces for up to 72 hours. Coronavirus is most contagious during the first three days after onset of symptoms, although spread may be possible before symptoms appear and in later stages of the disease.

Common symptoms include fever, cough and shortness of breath. Complications may include pneumonia and acute respiratory distress syndrome. The time from exposure to onset of symptoms is typically around five days, but may range from two to fourteen days. There is no known vaccine or specific antiviral treatment. Primary treatment is symptomatic and supportive therapy.[25]

Recommended preventive measures include hand washing, covering one's mouth when coughing, maintaining distance from other people, and monitoring and self-isolation for people who suspect they are infected. Authorities worldwide have responded by implementing travel restrictions, quarantines, curfews, workplace hazard controls, and facility closures.

The pandemic has led to severe global socioeconomic disruption, the postponement or cancellation of sporting, religious, political and cultural events, and widespread shortages of supplies exacerbated by panic buying. Schools and universities have closed either on a nationwide or local basis in 193 countries, affecting approximately 99.4 percent of the world's student population. Symptoms of COVID19 can be relatively non-specific and infected people may be asymptomatic. The two most common symptoms are fever (88%) and dry cough (68%). Less common symptoms include fatigue, respiratory sputum production (phlegm), loss of the sense of smell, shortness of breath, muscle and joint pain, sore throat, headache, chills, vomiting, hemoptysis, diarrhea, or cyanosis.

Further development of the disease can lead to severe pneumonia, acute respiratory distress syndrome, sepsis, septic shock and death. Some of those infected may be asymptomatic, with no clinical symptoms but with test results that confirm infection, so researchers have issued advice that those with close contact to confirmed infected people should be closely monitored and examined to rule out infection. Chinese estimates of the asymptomatic ratio range from few to 44%. The usual incubation period (the time between infection and symptom onset) ranges from one to 14 days; it is most commonly five days.

Description of Datasets

The data were collected from a reputable hospital in Ikorodu, Lagos State, Nigeria under the supervision of the Chief Medical Director of the hospital. The data collected were mainly records of patients with COVID-19 cases comprising the symptoms observed by the medical practitioner and complaints made by the patients. The data were collected in two phases, the first set of data was ninety nine in number and it was used as training set. The second phase consists of fifty sets of data and was used as testing set. The two sets of data were collected at different time. The training set was collected in March, 2020 while the testing set was collected in April, 2020.

All the data were assigned classes by a medical practitioner and the medical expert grouped the COVID-19 cases into five classes according to the level of severity- Very High, High, Moderate, Low and Very Low using the symptoms of COVID--19 of each patient. There are nineteen conditional attributes (symptoms) and one decision attribute, shown in the table below.

Attributes of COVID-19

S/N	ABREVIATION	ATTRIBUTE	ATTRIBUTE TYPE

A1	WKN	Weakness	Discrete
A2	APB	Abdominal Pain	Discrete
A3	COH	Cough	Discrete
A4	BOP	Body Pain	Discrete
A5	FVR	Fever	Discrete
A6	RGR	Rigor	Discrete
A7	COD	Cold	Discrete
A8	FAT	Fatigue	Discrete
A9	HEC	Headache	Discrete
A10	CAH	Catarrh	Discrete
A11	INS	Insomnia	Discrete
A12	LSM	Loss of Smell	Discrete
A13	VOM	Vomiting	Discrete
A14	JOP	Joint pain	Discrete
A15	DSN	Dizziness	Discrete
A16	ILL	Ill-looking	Discrete
A17	COV	Convulsion	Discrete
A18	BOT	Body Temperature	Discrete
A19	DIA	Diarrhea	Discrete
A20	COVID-19 DIAG	COVID-19 Diagnosed	Discrete

.

<u>Rough Set</u>

Basic Concept of Rough Set

Rough set theory (RST) is a useful mathematical tool to deal with imprecise and insufficient knowledge, find hidden patterns in data, and reduce dataset size. Also, it is used for evaluation of significance of data and easy interpretation of result. RST contributes immensely to the concept of reducts. Reducts is the minimal subset of attributes with the most predictive outcome [1]. Rough Set is a machine learning method which generates rules based on examples contained within an information table. Rough set theory has become well established as a mechanism for solving the problem of how to understand and manipulate imprecise and insufficient knowledge in a wide variety of applications related to artificial intelligence.

Let = (U,C) be an appropriate space, where U is a non-empty , finite set called the universe; A subset of attributes $R \subseteq C$ defines an equivalent on U. Let $[X]_R$ $(X \in U)$ denote the equivalence class containing x.

Given $R \subseteq C$ and $X \subseteq U$. X can be approximated using only the information contained within R by constructing the R-lower and R-upper approximations of set X defined as:

$$\underline{R}X = \{x \in X | [x]_R \subseteq X\}$$

$$\overline{R}X = \{x \in X | [x]_R \cap X \neq 0\}$$

Where $\underline{R}X$ is the set of objects that belong to X with certainty, and $\overline{R}X$ is the set of objects that possibly belong to X. The R-positive region of X is $POS_R(X) = \underline{R}X$, the R-negative region of X is $NEG_R(X)$

$= U - \overline{R}X$, and the boundary or R-borderline region of X is $BN_R(X) = \overline{R}X - \underline{R}X$. X is called Rdefinable if and only if $\overline{R}X = \underline{R}X$. Otherwise $\overline{R}X \neq \underline{R}X$ and X is rough with respect to R iff $\underline{R}X \neq \overline{R}X$.

The approximation measure $\alpha_R(X)$ is defined as

$$\alpha_R(X) = \frac{|\underline{R}X|}{|\overline{R}X|}$$

where $X \neq \varphi$, and $|X|$ denotes the cardinality of set X.

Algorithm LEM2 below developed by Grzymala-Busse in 1997 was used in building the classification model for COVID-19 diagnosis classes.

LEM2 Algorithm

Start
Input: k set of objects Output:
R set of rules
begin
G=K;
R=φ;
While G $\neq \varphi$ do
begin C $\neq \varphi$

C(G)= $\{c : [c] \cap G \neq \varphi\}$;

While $(C \neq \varphi)$ or $(!([C] \subseteq K))$ do begin

select a pair c $\in C(G)$ such that $|[c] \cap G|$ is maximum;

if ties, select a pair c $\in C(G)$ with the smallest cardinality $[c]$; if further ties occur, select the first pair from the list;

$C = C \cup \{c\}; G = [c] \cap G;$

$C(G) = \{C : [c] \cap G \neq \varphi\};$

C(G)=C(G)-C;
end;
for each elementary condition c$\in C$ do

if $|C - c| \subseteq K$ then $C = C - \{c\}$;

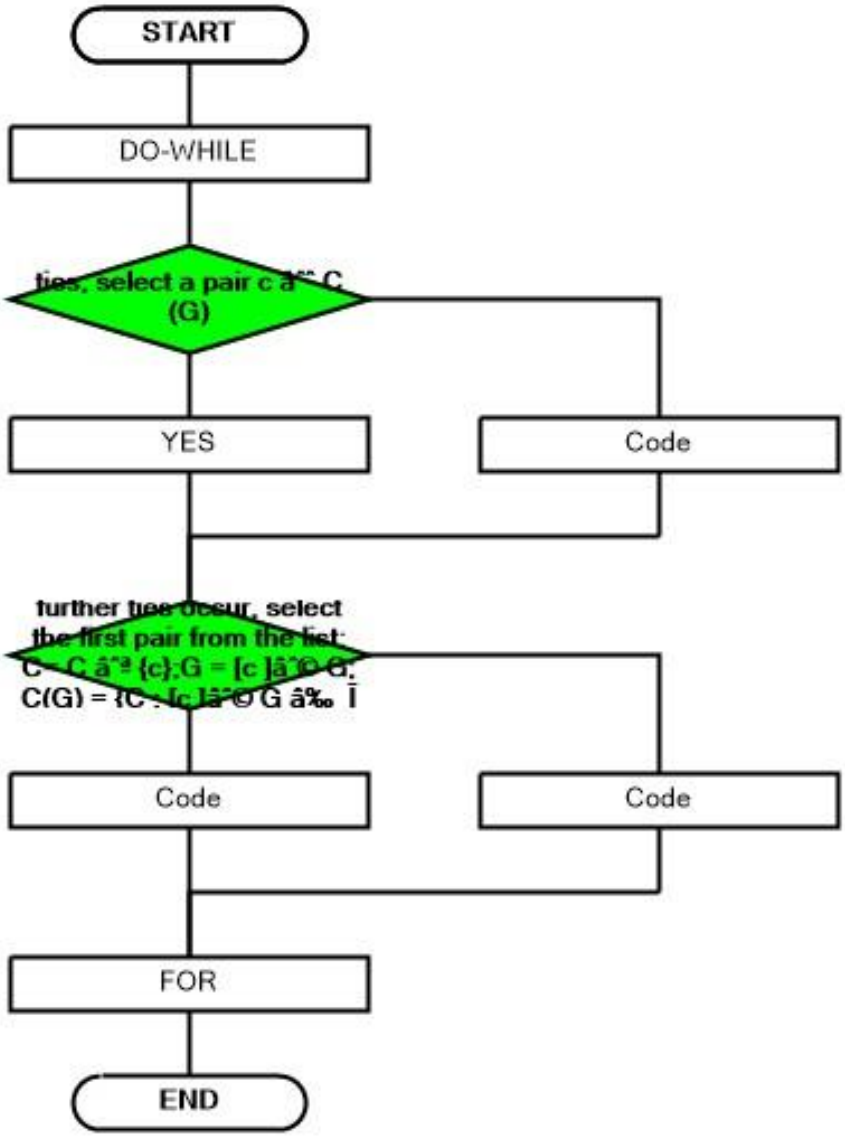
create rule r basing the conjuction C and add it to R;

$G = K - \bigcup_{r \in R} |R|$

end; for each $r \in$
R do

if $\bigcup_{s \in R - r} |S| = K$ then R = R-r
end

Flowchart Process



Application Set Up and Results

Rough Set used the training data set to build a classification model. For easier programming, it is easier to work around with numbers since the number of bytes that will be reserved for integer numbers will be smaller. For the decision attribute, the data Very Low, Low, Moderate, High and Very High are thus converted to integer numbers 1,2,3,4 and 5 respectively. For the conditional attributes (symptoms) A1 to A19, each symptom is classified as either high or low. Value high is converted to integer 1; Value Low is converted to integer 2 while default exists for symptom not applicable and takes a value of 0.

Rough Set was then used on the training set to give a diagnosis classification model for the five labels (five cases of COVID-19) which is in form of explainable rules displayed in the table below.

Generated rules of the five cases of COVID-19 by rough set

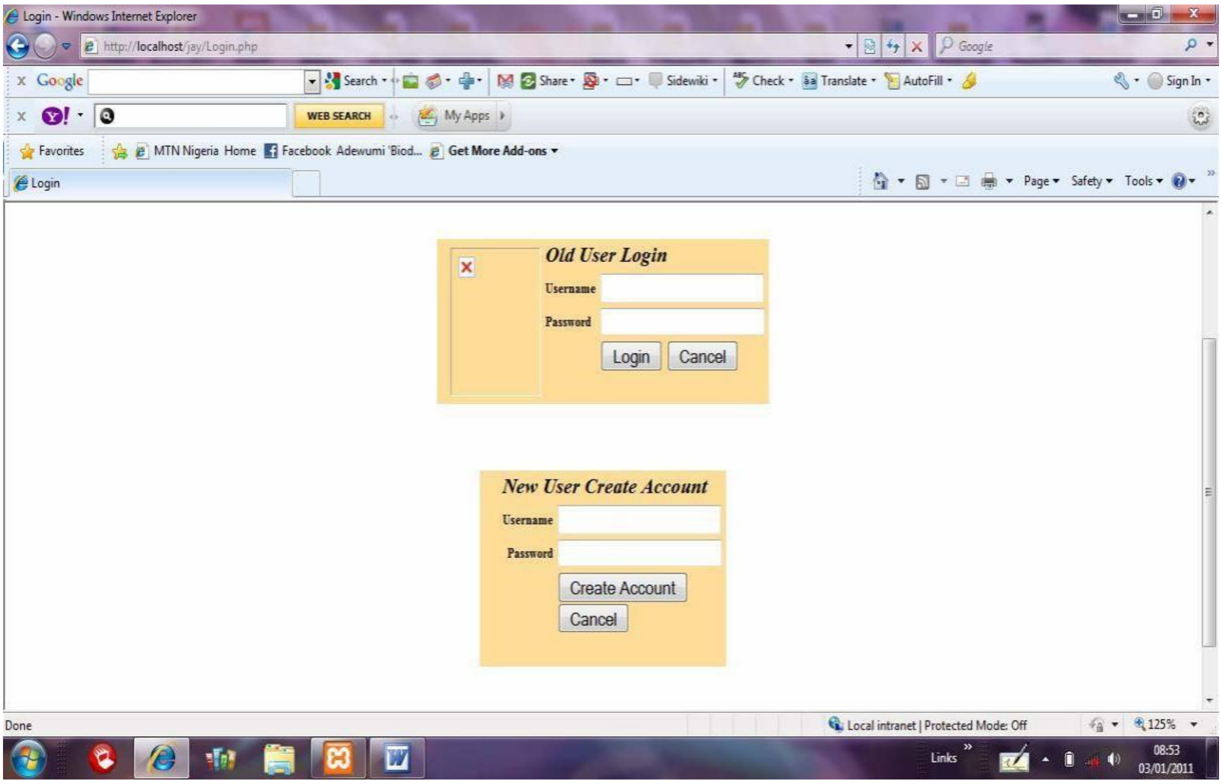
Rule No	Rule in details
Rule1	(A2=2)&(A5=2)&(A7=2)&(A9=2)&(A11=2)&(A13=2)&(A17=2) => (Apr=1)
Rule2	(A1=2)&(A2=2)&(A3=2)&(A4=2)&(A5=1)&(A9=2) => (Apr=1)
Rule3	(A3=1)&(A8=2)&(A9=2)&(A11=2)&(A12=2)=> (Apr=2)
Rule4	(A6=1) => (Apr=2)
Rule5	(A1=1)&(A2=2)&(A5=1)&(A9=2) => (Apr=2)
Rule6	(A1=2)&(A2=2)&(A8=2)&(A13=1)&(A15=2) => (Apr=2)
Rule7	(A3=2)&(A5=2)&(A8=2)&(A9=1)&(A11=2) => (Apd=2)
Rule8	(A1=2)&(A5=1)&(A8=2)&(A9=1)&(A10=1)&(A11=2)&(A15=2) => (Apr=3)
Rule9	(A4=2)&(A5=1)&(A9=1)&(A11=2)&(A15=2)&(A16=2)=> (Apr3)
Rule10	(A5=1)&(A9=1)&(A13=1)&(A15=2)&(A19=2) => (Apr=3)
Rule11	(A3=1)&(A4=2)&(A9=1)&(A10=2)&(A12=2) => (Apr=3)
Rule12	(A2=2)&(A4=1)&(A5=1)&(A7=2)&(A8=2)&(A10=2)&(A13=2) => (Apd=3)
Rule13	(A1=2)&(A11=1)&(A15=2) => (Apr=4)
Rule14	(A1=2)&(A8=1)&(A15=2)&(A17=2) => (Apr=4)
Rule15	(A5=2)&(A8=1)&(A10=2) => (Apr=4)
Rule16	(A1=1)&(A2=2)&(A5=2)&(A9=1)&(A18=2) => (Apr=4)
Rule17	(A8=1)&(A9=1)&(A13=2)&(A19=2) => (Apr=4)
Rule18	(A1=2)&(A2=1)&(A4=1)&(A13=2)&(A15=2) => (Apr=4)
Rule19	(A5=2)&(A10=1)&(A11=2) => (Apr=4)
Rule20	(A1=1)&(A2=1)&(A10=1) => (Apr=5)
Rule21	(A17=1) => (Apr=5)
Rule22	(A15=1) => (Apr=5)
Rule23	(A1=1)&(A9=2)&(A11=1) => (Apr=5)

Development of Website for COVID-19 Diagnosis and Therapy

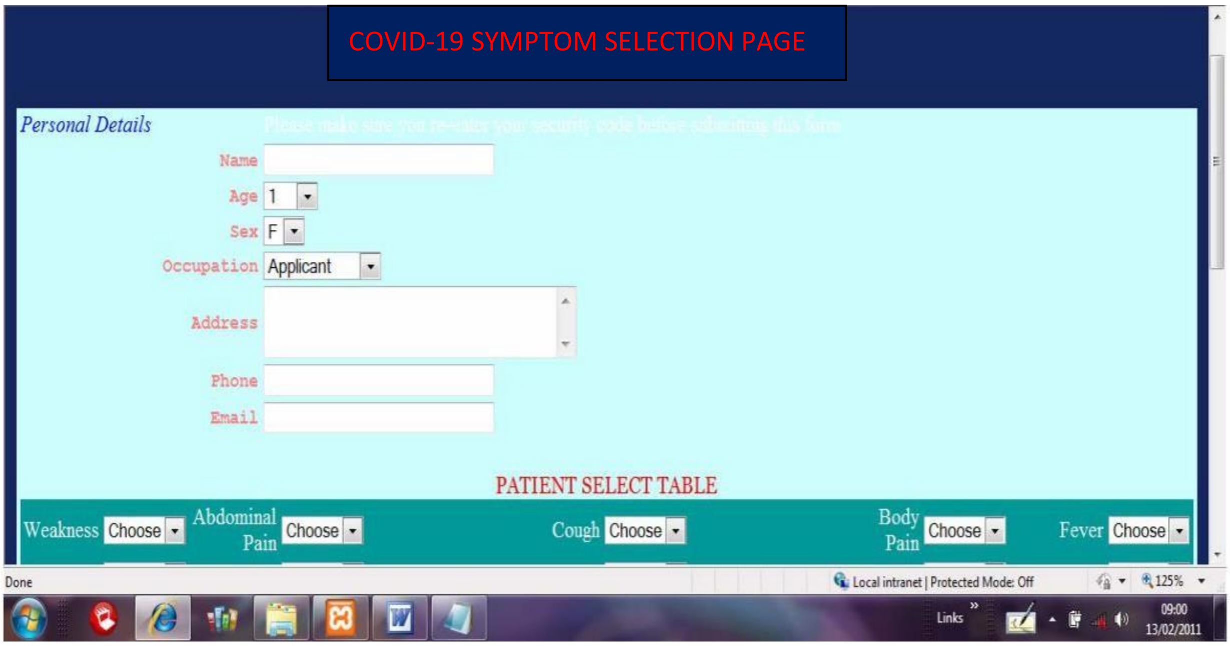
The website for COVID-19 diagnosis and therapy was developed using HTML and PHP as front end and MySQL as the backend.

The first interface is a user interface which allows users to create a new account by supplying username and password of interest. The created password and username are then used on the same interface to login into the next interface which is diagnosis interface. This phase contains the personal details of the patient (user), the patient's select symptoms table and command button to check COVID-19 status. The personal details section contains name, age, sex, occupation, address, phone number and e- mail. The patient's symptom select table contains nineteen symptoms (conditional attributes), each of which has the value high, low and default in which a patient must select one depending on the feelings of the patient (user). There is a space for a patient to type in any other perceived symptom not identified in the patient's symptoms select table for future consideration of medical experts, however, this is optional.

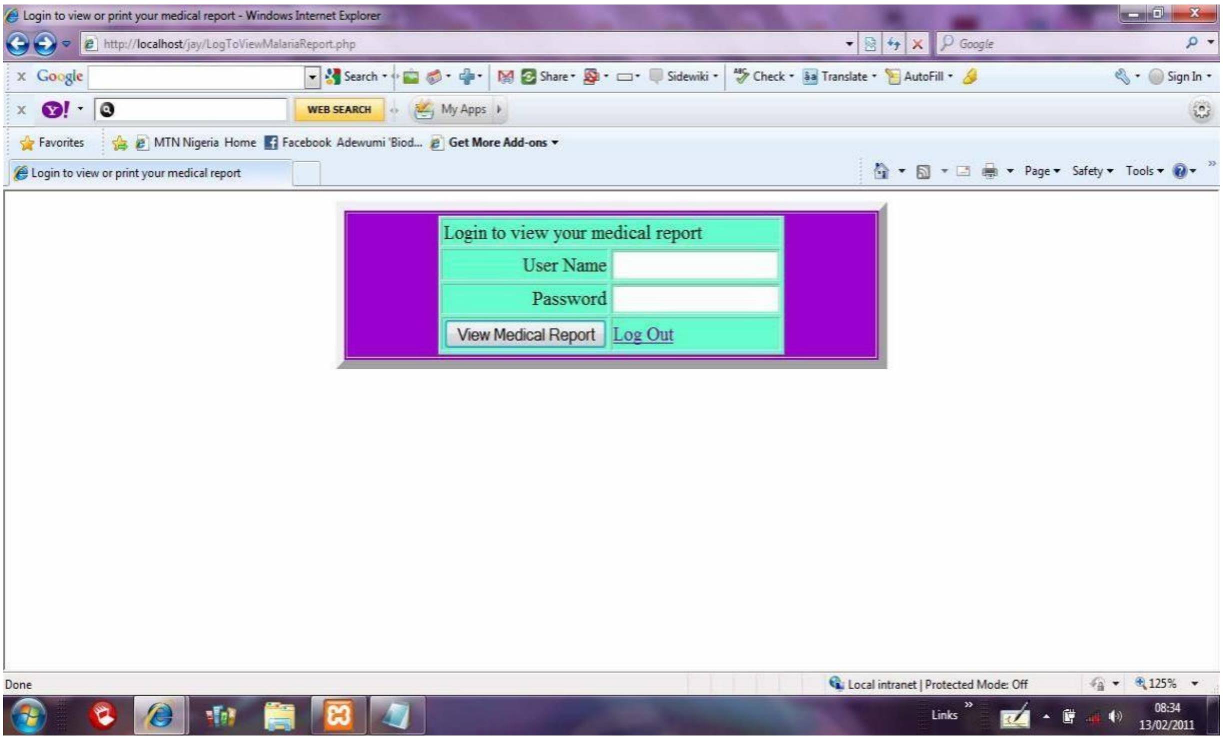
After all the perceived symptoms are selected, a check me button is clicked to take the user to the next phase where the patient views the result of the diagnosis and the appropriate therapy. The rules generated by rough set forms the engine room of the diagnosis. The earlier created username and password are both used at this phase to view medical report and therapy. If these are both correct, the system displays patient's diagnosis and treatment report sheet where a patient knows if he or she has COVID-19 or not. If COVID-19 is detected, the degree of severity is shown and the appropriate treatment is made available. This page also allows a patient who wishes to print the whole report sheet to do so and when a patient is through with the system, he or she clicks on logout to logout.



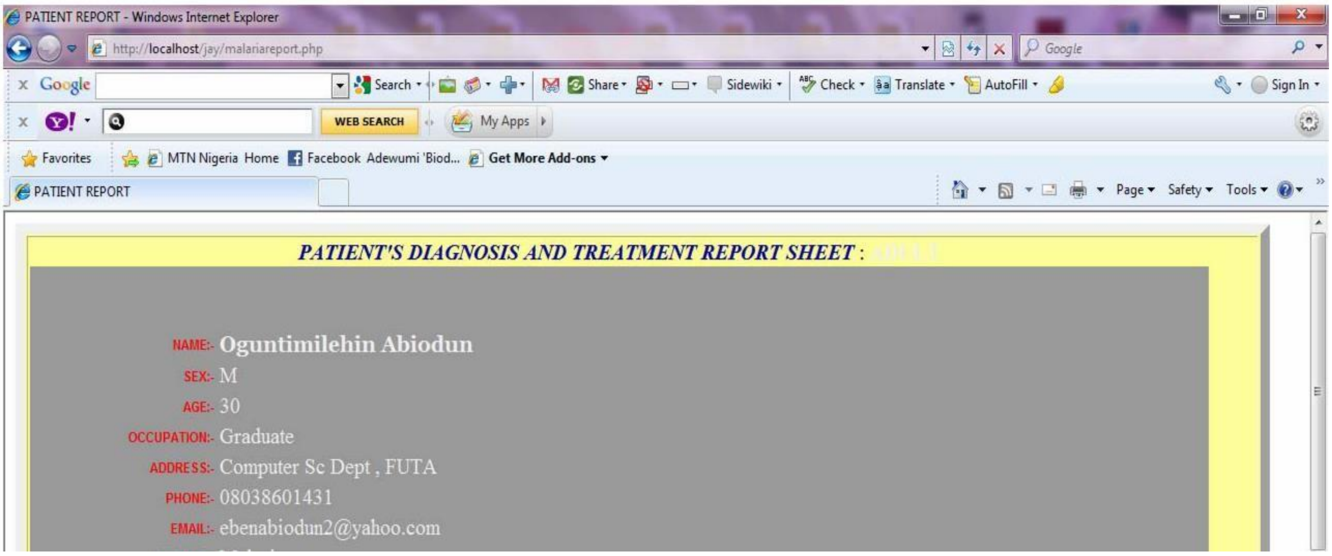
First Interface



Patient Symptom Select Interface



Log in to view medical report interface



Patient Medical Report Interface

Discussion of Results

The performance of the system were measured base on two results generated from the system, one on training set and the other on testing set. The first result was obtained from the training set. All the ninety nine (99) data sets were tested, using the designed web based system according to the twenty three rules generated from rough set and the confusion matrix of the result is given in the table 3 below

Table 3. Confusion matrix for the Training Set

Predicted as Actual	Very Low	Low	Moderate	High	Very High
Very Low (6)	6 (100%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Low (11)	0 (0.00%)	11 (100%)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Moderate (20)	0 (0.00%)	0 (0.00%)	20 (100%)	0 (0.00%)	0 (0.00%)
High (52)	0 (0.00%)	0 (0.00%)	0 (0.00%)	52 (100%)	0 (0.00%)
Very High (10)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	10 (100%)

TP = Class group correctly classified

TN = Class group wrongly classified

$$\text{Detection Rate} = \frac{TP}{TP+TN} = \frac{6+11+20+52+10 = 99}{99+0} = \frac{99}{99} = 100\%$$

All the six labels classified as Very Low were correctly predicted by the system which gives 100% prediction. All the 11 labels classified as low were also actually predicted, given 100%. There are twenty labels classified as moderate and all were correctly predicted, given 100%.. labels classified as High were all correctly predicted and likewise the ten labels classified as Very High. The confusion matrix thus gives 100% Detection Rate.

For the testing set, fifty set of data were tested against the twenty three rules using the designed system. There is only one very low label and it was correctly predicted by the system, attaining 100% in this case. Out of seven labels classified as Low, only six were correctly predicted while one was predicted as Very Low, which gives 85.71%. Of the seven labels classified as moderate, only five labels were correctly predicted, one was predicted as high while another one was predicted as Very high, given 71.43%. All the thirty one labels classified as high were correctly predicted, attaining 100%. In the case of the four labels classified as Very High, all were correctly predicted, this gives 100%. The confusion matrix for the testing set is given in table 4 below with Detection Rate of 94%.

Medications and Therapy

Some medical professionals recommend paracetamol (acetaminophen) over ibuprofen for first-line use. The WHO does not oppose the use of non-steroidal anti-inflammatory drugs (NSAIDs) such as ibuprofen for symptoms, and the FDA says currently there is no evidence that NSAIDs worsen COVID-19 symptoms. While theoretical concerns have been raised about ACE inhibitors and angiotensin receptor blockers, as of 19 March 2020, these are not sufficient to justify stopping these medications. Steroids, such as methylprednisolone, are not recommended unless the disease is complicated by acute respiratory distress syndrome**.**

Mechanical ventilation

Most cases of COVID-19 are not severe enough to require mechanical ventilation (artificial assistance to support breathing), but a percentage of cases do. Some Canadian doctors recommend the use of invasive mechanical ventilation because this technique limits the spread of aerosolised transmission vectors. Severe cases are most common in older adults (those older than 60 years and especially those older than 80 years). Many developed countries do not have enough hospital beds per capita, which limits a health system's capacity to handle a sudden spike in the number of COVID-19 cases severe enough to require hospitalisation. This limited capacity is a significant driver of the need to flatten the curve (to keep the speed at which new cases occur and thus the number of people sick at one point in time lower). One study in China found 5% were admitted to intensive care units, 2.3% needed mechanical support of ventilation, and 1.4% died. Around 20–30% of the people in hospital with pneumonia from COVID-19 needed ICU care for respiratory support.

Acute respiratory distress syndrome

Mechanical ventilation becomes more complex as acute respiratory distress syndrome (ARDS) develops in COVID-19 and oxygenation becomes increasingly difficult. Ventilators capable of pressure control modes and high PEEP are needed to maximise oxygen delivery while minimising the risk of ventilator-associated lung injury and pneumothorax.[155] High PEEP may not be available on older ventilators.

	Options for ARDS
Therapy	**Recommendations**
High-flow nasal oxygen	For SpO_2 <93%. May prevent the need for intubation and ventilation
Tidal volume	6mL per kg and can be reduced to 4mL/kg
Plateau airway pressure	Keep below 30 cmH_2O if possible (high respiratory rate (35 per minute) may be required)
Positive end-expiratory pressure	Moderate to high levels
Prone positioning	For worsening oxygenation
Fluid management	Goal is a negative balance of 0.5–1L per day
Antibiotics	For secondary bacterial infections
Glucocorticoids	Not recommended

Experimental treatment

No medications are approved to treat the disease by the WHO although some are recommended by individual national medical authorities. Research into potential treatments started in January 2020, and several antiviral drugs are in clinical trials. Although new medications may take until 2021 to develop, several of the medications being tested are already approved for other uses or are already in advanced testing. Antiviral medication may be tried in people with severe disease. The WHO recommended volunteers take part in trials of the effectiveness and safety of potential treatments. The FDA has granted a temporary authorization to convalescent plasma as an experimental treatment in cases where the person's life is seriously or immediately threatened. It has not undergone the clinical studies needed to show that it is safe and effective for the disease **Conclusion**

With the detection rate of 100% for the training set and 94% for the dataset and medical experts' provision of therapy for COVID-19 according to World Health Organization Guidelines for the treatment of COVID-19, it is hopeful that the web based system will go a long way to reduce the large deaths being caused by COVID-19. Apart from reducing the number of patients waiting for doctors for consultation on COVID-19 cases we also hope that many people will have access to it mostly in rural areas where hospitals are not available or not sufficient; the web-based system is expected to save a lot of lives since people can have access to it through GPRS enabled mobile phones, Cybercafés and telecommunication companies' data link modems.