Developing a Web based Support System to Monitor Symptoms of Brain Cancer Patients

Shubham R.
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Nigdi, Pune, India

Juby Babu
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Nigdi, Pune, India

Gouri Kolape
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Nigdi, Pune, India

Pratiksha Yadav
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Nigdi, Pune, India

Sonal Gore
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Nigdi, Pune, India
Professor

ABSTRACT
A brain tumor is an abnormal growth of normal cells in the brain. There are three possible treatments suggested by the oncologist are Chemotherapy, Radiotherapy and Operation. These treatments have an adverse affect on patients body and symptoms experienced are similar to the symptoms of brain tumor so the patients' panic. The proposed system provides a support for such patients. In the proposed system, Decision Tree (a supervised learning algorithm) is used to create a decision models to predict patient's health condition based on symptoms. Severity of symptoms, provided as input by the patient, is compared with the training model and it is predicted if patients' condition has altered. The current geographical location of the patient is obtained and nearest oncologists is suggested to the patient. Later, the patient can decide to send appointment request to any of the suggested oncologist. The oncologist can see patient's request and view patient's condition as predicted by system. The location and volume of the tumor based on previous Magnetic Resonance Imaging (MRI) record is entered by oncologist and condition of patient is again determined by system. Graphical view of comparison between volume and location of tumor of that patient over time can be viewed by doctor. Patient receives appointment via mail. Dietary tips are also provided to patient to improve his/her Quality Of Life (QOL) during the course of treatment.

Keywords
Brain Tumor, Geolocation API, Decision Tree, Symptom Monitoring

1. INTRODUCTION
According to American Brain Tumor Association, only in USA, there were nearly 7,00,000 people living with a primary brain tumor in year 2017 [10]. A brain tumor is an abnormal growth of normal cells in the brain. There are 4 grades of tumors as per World Health Organization (WHO) classification: grade I and grade II (benign tumors) and grade III and grade IV (malignant or cancerous tumors). Cancerous tumors can start within the brain (primary tumors), or can reach brain from somewhere else (secondary tumors or metastasis tumors). The most common symptoms occurring with brain tumors patients include headaches, changes in speech, vision, or hearing, etc. Also there are side-effects of treatments given to patients like nausea, vomiting, hair loss, loss of energy and these are similar to the Brain Tumor symptoms. The patients panic when they cannot decide if the symptoms they're experiencing are side effects of medication or actually the tumor is spreading. In case of grade IV tumor, sometimes it is not possible to cure the patients. It is not feasible to perform surgery or any other form of treatment and the death of patient is inevitable. So in this case, the aim of treatment is not only to prolong life but also to improve quality of patient's life. Some symptoms like nausea, diarrhea, constipation and fatigue, may make it difficult for patients to eat. Malnutrition may interfere with the patient's' ability to heal by interrupting or delaying treatment. The system recommends dietary changes and supplements approved by doctor to help prevent malnutrition and increase energy and strength of patients.

2. EXISTING SYSTEM
2.1 Literature Survey
2.1.1 Computer-based Quality Of Life monitoring of brain tumor patients
The system specified in this [4] used 110 primary brain tumor patients. The patients were asked questions related to their physical functioning, social functioning, role functioning, emotional functioning, and cognitive functioning using computerized questionnaires. Then software that scores and graphically represents questionnaires and medical data was used. The graph marks patients that need further medical and psycho-oncologic interventions. **The paper suggests that monitoring of quality of patient's life should be an integral part of oncologic care, which can contribute to individualization of treatment.**

2.1.2 Estimation of Symptom Severity during Chemotherapy based on passively sensed data
In the 4-week long study specified in this [2], 14 patients undergoing chemotherapy for gastrointestinal cancer were considered. Patients completed daily symptoms severity ratings of 12 common symptoms using Android phones. Also phone usage data was collected using mobile sensors. **The study concluded that long-term monitoring of patients can help in management of worsening and severe symptoms.**

2.2 Conclusion of literature survey

The symptoms mentioned above highlight the use of monitoring of symptoms to manage their severity and decide further treatment plans. But they do not use machine learning algorithms to classify the stage of tumor using the ratings of severity symptoms experienced by the brain tumor patients.

3. PROPOSED SYSTEM

3.1 Brief

The System is designed for medically diagnosed brain tumor patients. The patient is asked to input around 17 symptoms along with its severity like headache, nausea and vomiting, seizures, dizziness, balance and walking problems, vision problems like double vision & blurred vision, cognitive problems like learning, language, attention problems, hearing problem, behavioral changes like irritability, mood swings and withdrawal problems, sleepiness, weight changes like weight loss or weight gain, delayed development (in case of children), weakness, numbness and stiffness in the arms or neck. The patient is also asked to input his/her age, gender and history/hereditary information like presence of cancer in the family. The system should determine if the stage of tumor is normal (tumor is of WHO grade I or II) or critical (tumor is of WHO grade III or IV) using the above mentioned symptoms and personal information. Also the system helps the patients to book an appointment with his oncologist online and also helps in finding any nearby oncologist to book an appointment in case of emergency. The Geographical location of patient is obtained using Global Positioning System (GPS). Malnutrition and dehydration is often caused due to cancer treatments like radiotherapy and chemotherapy. So the system also provides the patient with dietary tips so that the patient's health does not deteriorate. Graphical representation of patient's history of tumor stage, its location and volume can be viewed by oncologist.

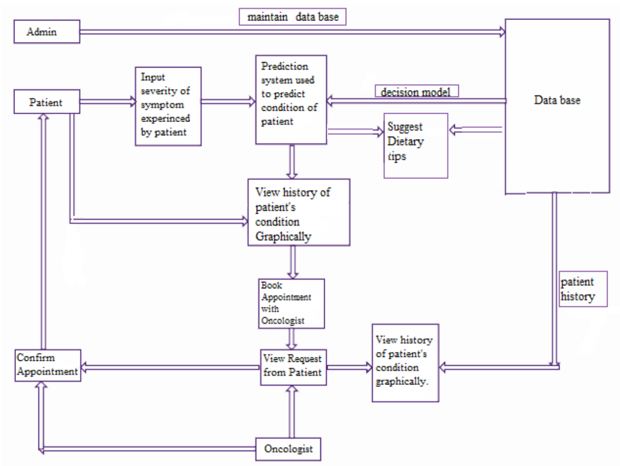

Fig 1: Architecture of the Proposed System

The system has 3 users: Admin, Oncologist and Brain Tumor Patient. The user interacts with the system using web browser over the web server. This web server is linked with the relational database that maintains training dataset, list of dietary tips, records on patients, administrator and oncologists, appointment requests.

Functions of Admin:

a. Approve registration of Oncologists and Patients.
b. Update database.
c. Update appropriate dietary tips after consulting the oncologist.

Functions of Patient:

a. Input severity of symptoms experienced and view condition of tumor.
b. Request for appointments with oncologist if interested.

Functions of Oncologist:

a. View his patient history and condition graphically.
b. Confirm an appointment as per request by patient.
c. Inputs his patient's latest MRI report details like tumor location and tumor volume.

3.2 Algorithms

3.2.1 Decision Tree Algorithm

Decision Tree is a Supervised Machine Learning Algorithm where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes which are the symptoms (headache, nausea and vomiting, etc.) or personal information (age, gender, family history) and leaves which are the decisions or the final outcomes i.e. normal state of tumor or critical state of tumor. And the decision nodes are where the data is split. The branches in the decision tree are the class labels for that decision node. There are many algorithms out there which construct Decision Trees, but the proposed system uses ID3 (Iterative Dichotomiser 3).

ID3 uses Entropy, Information Gain and Gain concepts for decision of splitting of dataset.

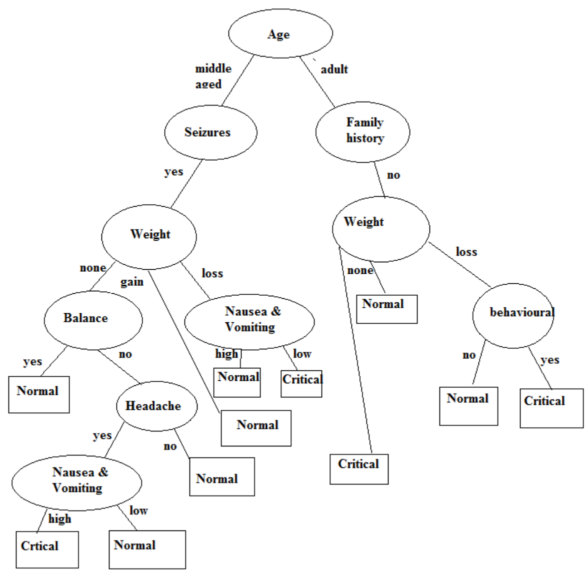

Fig. 2: Part of the Decision model created for Proposed System

Table 1. Schema of Training Dataset

Class	**Labels**
Output	critical, normal
Headache	yes, no
Nausea & Vomiting	high, low

Seizures	yes, no
Dizziness	yes, no
Balance	yes, no
Vision	high, low, none
Cognitive	yes, no
Hearing	yes, no
Behavioral	yes, no
Sleepiness	yes, no
Weight	gain, loss, none
Delayed Development	yes, no
Weakness	yes, no
Numbness & Stiffness	yes, no
Age	adult, middle-aged, children
Gender	male, female
Family History	yes, no

Steps of ID3:

1. Calculate Information Gain using Eq.1 for output class. For output class, entropy E(S) = IG(S). S is whole system.

   $$IG(S) = -\sum_{x \in X} p(x) \log_2 \frac{1}{p(x)} \qquad \text{... (1)}$$

   Where, IG(S) is Information Gain of output class, X = {critical, normal} and p(x) is probability of occurrence of output class labels normal and critical in the dataset.

2. Calculate Information Gain using the Eq. 2 for each feature/class A in the dataset other than the output class. A={headache, nausea and vomiting, seizures, etc}, do the following:

   IG_x will be calculated,

   $$IG_x = -\sum_{x \in X} p(x) \log_2 \frac{1}{p(x)} \qquad \text{... (2)}$$

   Where, X = set of labels in that class. Then p(x) will be probability of occurrence of each label of that feature with respect to output class label.

   E.g. In sample dataset of 10 records, class cognitive problems has labels yes 7 times and no 3 times. When class label is yes, the output has label critical occurs 3 times and normal 4 times. When class label is no, the output has label critical 1 time and normal 2 times. So,

   $$IG_{x=yes} = -\frac{3}{7}\log_2\frac{3}{7} - \frac{4}{7}\log_2\frac{4}{7} = 0.98$$

   $$IG_{x=no} = -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} = 0.92$$

   Then calculate Entropy E for that feature using Eq. 3.

   $$E(A) = \sum_{x \in X} p(x) IG_x \qquad \text{... (3)}$$

   Where, X = labels of that feature.

   So for cognitive,

   $$E(cognitive) = \frac{7}{10} \times 0.98 + \frac{3}{10} \times 0.92 = 0.962$$

   Now, calculate Gain G for that feature using Eq. 4.

   $$Gain(A) = IG(S) - E(A) \qquad \text{... (4)}$$

   So, for cognitive,

   $Gain(cognitive) = 0.97 - 0.962 = 0.01$

3. The feature with maximum Gain is chosen as root/decision node for the decision model. Divide the data set in parts equal to the number of class labels in that feature. In Fig. 2, age is root and it has 3 labels adult, middle-aged, children. So, the 1st part will have records with age = adult, 2nd part will have records with age = middle-aged and 3rd part will have records with age = children. Repeat step 2 for each part of dataset with records in that part leaving the feature according to which it was divided.

4. The process of dividing will stop when the entropy of the feature is 0.

3.2.2 Geolocation API

The Geolocation API defines a high-level interface to location information such as latitude and longitude. Common sources of location information include Global Positioning System (GPS) and location deduced from network signals such as IP address, RFID, Wi-Fi and Bluetooth MAC addresses, and GSM/CDMA cell IDs, as well as the user input. When oncologists register themselves, they mention the location of hospital in which they practice like Pune, Mumbai, etc. Brain tumor patient's current location is obtained using Global Positioning System (GPS). Euclidean distance is calculated between each registered oncologist and the patient. The oncologists are then displayed in the increasing order of the distance between them and the patient.

4. RESULT

4.1 Confusion Matrix

A confusion matrix is a technique for summarizing the performance of a classification algorithm.

There are 250 records in database. Of which 50 were used for testing purposes. The result of testing is shown in the Table 2.

TN = Normal patients classified as Normal

TP = Critical patients classified as Critical

FN = Critical patients classified as Normal

FP = Normal patients classified as Critical

Table 2. Predictive Model Evaluation using Confusion Matrix

Predicted Classification (no. of records)			
Normal	**Critical**		
TN = 32	FP = 1	**Normal**	**Actual Classification (no. of records)**
FN = 6	TP = 11	**Critical**	

Accuracy=TP+TNTP+TN+FP+FN⋯ (4)

Precision=TPTP+FP⋯ (5)

F−measure=2.Precision.RecallPrecision+Recall⋯ (6)

Recall=TPTP+FN⋯ (7)

Table 3. Performance measures of Proposed System

Precision	Recall	F-measure	Accuracy
0.765	0.929	0.84	0.86

5. CONCLUSION

A Web based Support System is developed for Brain Tumor Patients. Brain Tumor Patient's health condition is predicted based on current health symptoms and its severity using decision tree algorithm ID3. The 86% of accuracy is achieved with proposed system. This system will definitely prove as an effective support system for brain cancer patients to get immediate health status to decide further line of action in terms of booking appointments with oncologist, getting dietary guidelines etc. Such immediate response with proposed system will surely help to improve the quality of patient life.

6. FUTURE SCOPE

Physiological parameters of patient like Blood Pressure, Heart Rate, etc., obtained through body sensors and mobile usage pattern of patient, obtained through call logs, location of mobile, etc., can be used with clinical signs and symptoms of patients, obtained through questionnaires, to have a real-time monitoring of patient's condition and more accurately predict if patient's health condition has changed.

7. REFERENCES

[1] K. Balachandran, Dr. R. Anitha. "Ensemble based optimal Classification model for pre-diagnosis of Lung cancer", Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 2013.

[2] Carissa A Low, et al. "Estimation of Symptom Severity During Chemotherapy from Passively Sensed Data: Exploratory Study", Journal of Medical Internet Research, Vol. 19, No 12, 2017.

[3] Minjoung Monica Kooa, Christian von Wagnera, Gary A. Abelb. "Typical and atypical presenting symptoms of breast cancer and their associations with diagnostic intervals: Evidence from a national audit of cancer diagnosis", The International Journal of Cancer Epidemiology, Detection, and Prevention, Cancer Epidemiology Vol.48, Pg.140-146, 2017.

[4] Astrid Erharter, et al. "Implementation of Computer-Based Quality-of-Life Monitoring in Brain Tumor Outpatients in Routine Clinical Practice", Journal of Pain and Symptom Management, Vol. 39 No.2, 2010.

[5] A Robson, J Sturman, P Williamson, P Conboy. "Pre-treatment clinical assessment in head and neck cancer: United Kingdom National Multidisciplinary Guidelines", The Journal of Laryngology & Otology, 130(Suppl. S2), S13-S22, 2016.

[6] Maurizio Bergamino, et al. "Brain Tumor Database, a free relational database for collection and analysis of brain tumor patient information", Health Informatics Journal, Vol. 21(I), 36-45, 2015.

[7] Fiona M Walter, et al. "Symptoms and patient factors associated with diagnostic intervals for pancreatic cancer (SYMPTOM pancreatic study): a prospective cohort study", The Lancet Gastroenterology & Hepatology, Vol. 1, Pages 298-306, 2016.

[8] B Talwar, R Donnelly, R Skelly, M Donaldson. "Nutritional management in head and neck cancer: United Kingdom National Multidisciplinary Guidelines", The Journal of Laryngology & Otology, 130(Suppl. S2), S32-S40, 2016.

[9] H Cocks, K Ah-See, M Capel, P Taylor. "Palliative and supportive care in head and neck cancer: United Kingdom National Multidisciplinary Guidelines", The Journal of Laryngology & Otology, 130(Suppl. S2), S198-S207, 2016.

[10] Brain Tumor Statistics http://www.abta.org/about-us/news/brain-tumor-statistics.