



PREDICTING DISEASE USING DATA MINING BASED ON NAÏVE BAYESIAN CLASSIFIER SYSTEM

Dr.V.Sulochana, Uthra.M

Assistant Professor, Student
Hindusthan college of Arts&Science

ABSTRACT Globally, dyslexia or related learning disabilities affect about 20 percent of the population (National Centre for Learning Disabilities, 2017a). Both private and public schools, the Diagnostic and Statistical Manual of Mental Disorders (DSM), and federal and state laws frequently utilise the term "specific intellectual disability." The phrase "specific learning impairment" refers to a number of learning disabilities, such as dyslexia, dyscalculia, and dysgraphia.

The National Institutes of Health estimates that between 10% and 15% of students in school have a learning problem. According to a 2017 research by the United Nations Educational, Scientific, and Cultural Organisation (UNESCO), every nation needs to make sure that fewer children with intellectual disabilities attend primary or secondary school. A few learning disorders have a particular impact on academic achievement.

A few learning disorders include dyslexia, which affects reading, dysgraphia, which affects writing and drawing, dyscalculia, which affects answering mathematical problems, and so on. The healthcare industry is constantly interested in finding ways to lower the frequency of medical errors. Reducing errors of commission, or incorrect actions, such as amputating the wrong limb, has been a major focus of recent effort.

Errors of omission, or failing to act when patients are not meeting evidence-based goals, have received less attention while being a particularly important field of research in the treatment of chronic diseases like diabetes. Finding behavioural patterns that indicate future errors and choosing the right course of action to prevent errors when error-prone behaviour patterns are discovered are two of the biggest obstacles to preventing errors of omission.

These could involve matching a doctor's skills to a patient whose qualities they are most successful in treating, as well as altering the way each doctor manages their patients.

Most recent data mining research aimed at enhancing care quality has concentrated on identifying outcomes linked to patient/illness characteristics. Few studies have looked closely at patient outcomes and physician behaviour. 1.

INTRODUCTION In contemporary healthcare, early diagnosis is essential since it forms the basis for efficient management and treatment of a wide range of medical disorders. Healthcare practitioners might potentially slow the spread of diseases and improve patient outcomes by quickly intervening when illnesses are still in their early stages.

Furthermore, early identification frequently results in more efficient treatment plans, which lessen the strain on healthcare systems and the financial cost of protracted illnesses. In this situation, it is crucial to spread the word about the value of early diagnosis since it enables people to take preventative measures to protect their health and wellbeing. 2.

EXISTING SYSTEM A large percentage of diagnostic errors are the consequence of cognitive errors brought on by incomplete information, improper data collection, and/or improper verification. According to experts, diagnostic thinking errors by clinicians account for 75% of diagnostic failures. Physicians' diagnostic thinking is based on intricate and fascinating cognitive processes, and in order to improve decision-making, physicians must explicitly understand and apply a variety of cognitive techniques.

A significant field of interest and research is the development and application of solutions for reducing or preventing such diagnostic mistakes. To reduce the occurrence of diagnostic errors, it is recommended that clinicians become more clinically knowledgeable and steer clear of cognitive errors in order to make better recommendations. The utilisation of data sources to raise the standard of patient care is becoming more and more important in the healthcare industry.^{3, 4} Administrative and clinical records are among the available data sources; this is especially true with the introduction of the electronic medical record (EMR).

However, it might be difficult to retrieve useful information from big databases. 3. **PROPOSED SYSTEM** A learning impairment is a type of learning disability known as a neurodevelopmental disorder that can impact a child's reading, writing, cognitive skills, and problem-solving abilities.

These impairments are referred to as specific learning disabilities (SLD) since they primarily disrupt kids' academic performance, notably causing difficulties with reading (dyslexia), writing (dysgraphia), and math (dyscalculia). These pupils must be recognised at an earlier age. With the right support, they can get sufficient experience in a particular area and hone their disability-related abilities.

Testing scale tool has been proposed to diagnose and identify SLD. Students who are suspected of having SLD can take the quiz with the help of the suggested instrument. Certain test questions are repeated three times depending on the kind of learning disability. The decision tree algorithm receives the output data after the test is over.

The naïve bayes algorithm predicts children with learning problems based on student marks and time. With the use of the suggested tool, educators and parents can create a comprehensive, user-friendly tool that accurately diagnoses reading, writing, and math issues as well as recommend the best course of action and educational activities. 4.

IMPLEMENTATION OF NAÏVE BAYESIAN CLASSIFIER We employ the machine learning technique Naive Bayes to address classification issues. The Bayes Theorem forms its foundation. Across numerous industries, this is one of the most effective and straightforward machine learning methods now in use. **MEDICAL DIAGNOSIS** Physicians can use the data provided by the classifier to diagnose patients.

Naive Bayes is a useful tool for medical experts to utilise when determining whether a patient is more susceptible to heart disease, cancer, or other illnesses. The Naive Bayes classifier, which is much faster than other classification algorithms, would be the best course of action in this case. 5. **MODULES DESCRIPTION** ADMIN 1. Login: By entering proper credentials, administrators can gain secure access to the system. 2.

Register Questions: Permits administrators to include new inquiries in the database of the system. These are probably questions from quizzes or tests that cover medical or healthcare-related subjects. 3. View Questions, Edit, Update: Gives administrators access to the system's existing questions for viewing, editing, and updating.

This feature guarantees the accuracy and correctness of the question database. 4. Register Hospital: This feature enables administrators to add new hospitals to the database of the system. This entails including information about the hospital, including its name, address, phone number, and any other pertinent details. 5. View Hospital, Edit, Update: Gives administrators the opportunity to see, modify, and update data on hospitals that have registered.

Accurate maintenance of the hospital database is ensured by this functionality. 6. Register Doctor: Allows system administrators to add additional physicians or medical specialists. This entails including information on the physician, including their name, specialty, contact details, and other pertinent qualifications. 7. View Doctor, Edit, Update: Gives administrators the ability to see, modify, and update data pertaining to licenced physicians or other healthcare providers.

This guarantees that accurate and up-to-date information is maintained in the doctor database. 8. Register Treatment Details: Gives administrators the option to record specifics about the different medical procedures or treatments that hospitals or other healthcare facilities provide. This might cover things like treatment efficacy, prices, and descriptions. **USER** 1.

Login: By entering legitimate login credentials, users can safely access the system. 2. View Questions: This feature enables users to see the questions that are stored in the database of the system. These might be used in tests or quizzes and are probably connected to medical or healthcare-related subjects. 3.

View Hospital: This feature lets users see details about hospitals that have registered with the system. This could contain information about the hospital, including its name, address, phone number, etc. 4. View Doctor: Enables users to see details about physicians or other healthcare providers who are enrolled with the system.

This could contain information on the doctor, like his or her name, specialty, and contact data. 5. Submit Answer: Allows users to enter their responses to the questions that are stored in the database of

the system. Most likely, this feature is a component of an evaluation or quiz feature. 6.

View Prediction Result: Gives people the ability to see the predictions that the system has produced. These forecasts can concern ailments, the efficacy of treatments, etc. DATA FLOW DIAGRAM 6. CONCLUSION A major factor in reducing the death rate from capillariasis is early identification and treatment.

Health care professionals need to be educated about the typical symptoms and indicators of capillariasis, as well as the disease's treatment plans, particularly in endemic areas. The clinical significance of these diseases and the requirement for early detection and comprehensive intervention programmes have been recognised by medical and mental health systems. It has been contested by clinical observations and outcome research that patients with somatoform diseases generally have a bad prognosis.

If the patient can clearly see the benefits of psychotherapy and psychiatric interventions, their acceptability can be increased. There may be a "positive vicious circle" and a sense of mastery associated with compelling models for a better explanation and comprehension of symptoms, the formulation of specific, attainable goals, and the provision of good experiences in symptom reduction and coping ability enhancement.

/ Figure 1 LOGIN PAGE / Figure 2 ADD DOCTOR PAGE / Figure 3 FRONT PAGE / Figure 4VIEW BOOKINGS PAGE / Figure 5USER PAGE / Figure 6DOCTOR PAGE / Figure 7 UPDATE QUESTIONS PAGE 7. FUTURE ENHANCEMENT Poor spelling, trouble naming words, and a lack of reading and/or writing skills are the hallmarks of dyslexia, a learning condition. There are two types of dyslexia: phonological dyslexia and surface dyslexia.

Surface dyslexia is the test of looking at the word, but phonological dyslexia is the problem of looking at a specific section of a word. Due to its greater severity, phonological dyslexia is the area of study that most interests researchers. Most of the time, a child can read and exhibit signs of reading difficulties, and dyslexia is acknowledged.

It would have significant benefits for early reading if phonological indications could be utilised to identify the illness before a child could read it. The present project attempts to create a software tool that parents can use to assess whether their child's dyslexia is at risk. SVM and Grid search CV are the approaches utilised in this, with a 97.42% accuracy rate.

By employing traditional methods of dyslexia prediction, we have increased the accuracy of our predictions. 8. REFERENCES [1] Data Mining Classification Methods for Assigning Job Profile and Personality Type to Individuals

Mridul Sharma, Soumya Sharma, Abhilasha Singh, Mridul Sharma, and Ritu Gupta

Nineth International Conference on Infocom Technologies, Optimisation, and Reliability (Trends and Future Directions) (ICRITO) in 2021 [2]A Survey of Data Minig & Artificial Intelligence Techniques for Disorder Prediction

Power Amogh; Shilvant Seema Shivendra Aswale, Pratiksha Shetgaonkar, Varsha Pawar, and Vishesh Parab

The Vision Towards Emerging Trends in Communication and Networking (VITECoN) International Conference 2019 [3] Classification of Depression Level Using Real-Time Machine Learning and Data Mining Techniques

Arpita Das, Awal Ahmed, Dola Das, and Tamzid Hasan Nahid Fire

ICCIT, the 26th International Conference on Computer and Information Technology, is scheduled for 2023. [4] D. E. Adkins (2017). A paradigm shift: machine learning and electronic health records.

174(2), 93–94 in American Journal of Psychiatry [5] Nair, M. K., and Joshi, S. (2018). Examining Healthcare Classification-Based Prediction Methods. 11(15), 1–19; Indian Journal of Science and Technology.

INTERNET SOURCES:

<1% - <https://www.ncbi.nlm.nih.gov/books/NBK519710/>

<1% - <https://repository.gheli.harvard.edu/repository/12769/>

<1%

<https://my.clevelandclinic.org/health/diseases/4865-learning-disabilities-what-you-need-to-know>

<1% - <https://www.ncbi.nlm.nih.gov/books/NBK43675/>

<1% - <https://pubmed.ncbi.nlm.nih.gov/33332905/>

<1% - <https://pubmed.ncbi.nlm.nih.gov/24070580/>

<1% - <https://link.springer.com/article/10.1007/s00146-021-01154-8>

<1%

https://journals.lww.com/md-journal/Fulltext/2020/02210/The_adoption_of_electronic_medical_record_by.75.aspx

1% - <https://ieeexplore.ieee.org/document/9544617/>

<1% - <https://www.towardsanalytic.com/decision-tree-algorithm/>

<1% - <https://www.linkedin.com/advice/0/what-steps-determine-best-course-action-meeting>

<1% - <https://www.forbes.com/sites/jiawertz/2024/03/16/how-ai-is-uprooting-major-industries/>

<1% - <https://www.analyticsvidhya.com/blog/2021/01/a-guide-to-the-naive-bayes-algorithm/>

<1% - <https://quizlet.com/177466132/chapter-4-medical-records-flash-cards/>

1%

https://www.researchgate.net/publication/5379624_Two_types_of_phonological_dyslexia_-_A_contemporary_review

<1%

https://books.google.com/books/about/2021_9th_International_Conference_on_Rel.html?id=Ony6zgEACAAJ

1% - <https://library.iitd.ac.in/node/24881>

<1%

<https://www.semanticscholar.org/paper/Different-Approaches-in-Depression-Analysis-%3A-A-Krishna-Anju./d187b414a3335ed6020128092ebac7a83c62f6a7>

<1% - <https://pubmed.ncbi.nlm.nih.gov/28142275/>

<1% - <https://ajp.psychiatryonline.org/?cookieSet=1>

1% - <https://i-scholar.in/index.php/indjst>