**Research Review and Project Proposal Worksheet**

**Team Name: CDS 15**

**Students Name: M.Sathvika**

**J.Naveen**

**B.Danish**

**Date:25/10/2023**

**Research Topic: Disease Prediction System using Machine Learning**

**Section 1: Research Summary**

**1. Research Summary**

Provide a concise summary of your research topic, including the main objectives and scope.

|  |
| --- |
| In this digital world, most people are prone to diseases, due to a lack of healthy food, proper sleep, and daily exercise. It is crucial to know if we are suffering from a disease, at an early stage rather than discovering it later. The traditional way of diagnosis may not be sufficient in the case of a serious ailment. Our research includes developing a medical diagnosis system based on machine learning (ML) algorithms for the prediction of any disease can help in a more accurate diagnosis than the conventional method. |

**2. Key Findings and Insights**

List the most significant findings and insights from your literature review. Include relevant citations.

|  |
| --- |
| There is a need to groundwork and evolve a system that will enable end users to predict diseases without visiting a physician or doctor for a diagnosis. To identify various diseases by observing the symptoms of patients and applying various Machine Learning Model techniques. There is no proper procedure for handling text and structured data. Both structured and unstructured data would be considered by the proposed framework. |

**3. Research Gaps**

Identify gaps or areas in the existing research that your project aims to address. Explain why these gaps are significant.

|  |
| --- |
| Heart disease prediction has been done using various algorithms which are based on the expedition of lab reports. Algorithms such as naïve Bayesian, random tree classifier, and other classifiers have been used and compared to each other To find out the best classifier for prediction of diseases and have concluded naïve Bayesian to be the most efficient one from the bunch. Currently, we use traditional methods of diagnostics like visiting a doctor and getting a basic diagnosis by providing symptoms and examining visually if it can be. But it raises a problem if a doctor is not available or isn't in a position to respond. |

**Section 2: Project Proposal**

**4. Project Title**

Propose a descriptive and catchy title for your project.

|  |
| --- |
| DISEASE PREDICTION SYSTEM USING MACHINE LEARNING |

**5. Project Objectives**

List specific and measurable objectives that your project aims to achieve.

|  |
| --- |
| The objective of the project is to design a machine learning-based web application capable of taking in required attributes of lab reports obtained through diagnosis and the symptoms observed by the effecting party and drawing out a conclusion as to whether the patient has a disease or not with the highest accuracy possible along with a user-friendly UI that can be used to consult, contact and talk to the concerning doctor and carry on with the medication and further instructions to take care of the disease. This helps keep track of previous consultations and if the current doctor who is being consulted is unavailable at the moment this record of consultation can be used to treat the disease more efficiently based on previous medication and diet routine |

**6. Target Audience**

Describe the intended audience or users of your project. Include demographics and user needs.

|  |
| --- |
| The target audience is the people themselves. As they are busy in their day-to-day lives they may not have time to consult a doctor.  Some people may waste a lot of money with simple problems. They can diagnose their issues on the website itself. |

**7. Problem Statement**

Clearly define the problem your project seeks to solve. Explain its significance.

|  |
| --- |
| At present, when one suffers from a particular disease, the person has to visit a doctor which is time-consuming and expensive. Also if the user is out of reach of doctors and hospitals it may be difficult for the user as the disease can not be identified. So, if the above process can be completed using an automated program which can save time as well as money, it could be easier for the patient which can make the process easier. |

**8. Solution Overview**

Provide an overview of the proposed solution, including its novelty and how it addresses the problem.

|  |
| --- |
| This proposed system is a web-based application that provides you with a platform to identify your disease based on the symptoms that you have. This uses logical regression and decision tree classifier to parse through the structured and unstructured data obtained through several sources to learn and identify the disease. |

**9. Key Features and Functionality**

List the main features and functionalities your project will include. Explain how each feature contributes to solving the problem.

|  |
| --- |
| User: The patient will log in with a username and password to get diagnosed.  ♣ Doctor will log in with a username and password to register his expertise.  ♣ Admin will have unrestricted access to data but credentials will be encrypted through SHA-256 algorithm.  PC: It hosts the database that constitutes all of the data related to doctors, patients, etc.  ♣ A web browser is used to emulate a user interface for the said web application.  Database: Stores all the information about the diseases, patients, and doctors. |

**10. Technology Stack**

Specify the technologies, frameworks, and tools you plan to use. Explain why they are suitable for your project.

|  |
| --- |
| 1. HTML- is used to display content in the browser.  2. CSS -To properly align the HTML content.  3. Django - To create a sorted and clean administration.  4. Postgres SQL – can be used to create a local database to store all the data. |

**Section 3: Brainstorming**

**11. Brainstorm Ideas**

Brainstorm additional ideas or concepts related to your project, even if they aren't part of the core proposal.

|  |
| --- |
| Disease Prediction has been already implemented using different techniques like Neural Networks, decision trees, and the Naïve Byes algorithm. Particularly heart-related disease is mostly analyzed. From the analysis, it was found that Naïve Bayes is more accurate than other techniques. So, Disease Predictor also uses Naïve Bayes for the prediction of different diseases |

**12. Feasibility Assessment**

Evaluate the feasibility of your project in terms of:

|  |
| --- |
| The project is technically feasible as it can be built using the existing available technologies which are machine learning algorithms such as Naïve Bayesian Classifier, programming languages capable of creating a web page, a database, and an administration platform. It is a web-based application. The technology required by Disease Predictor is available and hence it is technically feasible.  The project is economically feasible as the cost of the project is involved only in the hosting of the project. As the data samples increase, they consume more time and processing power. In that case better processor might be needed.  The project is operationally feasible as the user has basic knowledge about computers and the Internet. Disease Predictor is based on a client-server architecture where the client is user and server is the machine where datasets are stored. |

**13. Risks and Mitigations**

Identify potential risks or challenges your project may face and propose strategies to mitigate them.

|  |
| --- |
| The average probability accuracy of prediction is 85% so there is a chance of misdiagnosis. So we will at most try to make the prediction 100% by collecting more relevant data. |

**Section 4: Next Steps**

**14. Project Timeline**

Create a detailed timeline outlining the major project milestones and deadlines. Include key activities and their estimated durations.

|  |
| --- |
|  |

**15. Resource Requirements**

List all the resources required for your project, such as hardware, software, datasets, or personnel. Include estimated costs if applicable.

|  |
| --- |
| Hardware Requirements:  ♣ Processor - Pentium –IV  ♣ Speed - 1.1 GHz  ♣ Hard Disk - 20 GB  ♣ Monitor - SVGA and higher Software Requirements:  ♣ Operating System - Windows 7 and later  ♣ Programming Language - Python  ♣ Browser - Chrome, Firefox, Edge, etc.  Cost: The cost of the project is involved only in the hosting of the project. As the data samples increase, they consume more time and processing power. In that case better processor might be needed. |

**16. References**

Provide a comprehensive list of references and sources used in your literature review. Follow a citation style guide.

|  |
| --- |
| K.M. Al-Aidaroos, A.A. Bakar, and Z. Othman have conducted research for the best medical diagnosis mining technique - (Al-Aidaroos, Bakar, & Othman, 2012)  Disease Predictor histories suggest previously unconsidered concerns and facilitate discussion about early testing and prevention. (A.Davis, V.Chawla, Blumm, Christakis, & Barbasi, 2008)  M. Maniruzzaman et al developed a machine learning (ML)-based method to detect disease patients using logistic regression (LR) to identify risk variables based on the odds ratio (OR) and p-value.(Maniruzzaman, M., Rahman, M., Ahammed, B. and Abedin, M., 2020.  Shadab Adam Pattekari and Asma Parveen have conducted research using the Naïve Bayes Algorithm to predict heart diseases where the user provides the data which is compared with a trained set of values.(Adam & Parveen, 2012) |

**Section 5: Reflection**

**17. Reflect on the Worksheet**

Write a reflective paragraph on how completing this worksheet has contributed to the refinement of your project proposal. Identify any areas where you need further clarification or research.

|  |
| --- |
| This worksheet has made our research more on the topics which we are not aware of. This also made our work easy as the same information can be used for the documentation. |

**Section 6: Feedback**

**18. Peer Review**

Share your worksheet with a peer or mentor for feedback and comments. Ask them to provide constructive suggestions and insights.

|  |
| --- |
|  |

**Section 7: Finalizing Your Proposal**

**19. Final Project Proposal**

This project aims to predict the disease based on the symptoms. The project will be designed in such a way that the system takes symptoms from the user as input and produces output i.e. predict disease.