**Chapter 2 Literature Survey & Proposed Work (Phase wise)**

2.1 Introduction

Analysis of the current market scenario and technological capabilities is essential to the building of a project. Here we analysed various prevalent research papers pertaining to our project as a part of the literature survey. The project aims at creating an intelligent but lightweight healthcare assistant. This application will be able to help healthcare and other hospital staff work efficiently. This also aims at having a basic machine learning capability for disease prediction.

Our literature survey was focused on the capabilities of the choices we made for the core part of the project, namely artificial neural networks and the use of wearable technology for data collection.

The initial phases focused on the planning and design of the project. We analysed extant products, customer interests and so on. For analysis, we looked at the current iterations of popular and cutting-edge algorithms used in machine learning, such as artificial neural networks and convolutional neural networks. The analysis was openly available in numerous research papers. The design involved comparing the project with some similar projects.

2.2 Literature Survey Table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ref.  No. | Year | Author | Key findings (Results) | Research gaps |
| 2 | April 2017 | Chen, hao, hwang, wang, wang | 1) Big data has grown in healthcare, leading to potential for accuracy in medical analysis.  2) Wearable tech provides new ways to obtain data from patients and monitor their health.  3) A convolutional neural network was used to extract text characteristics.  4) 20 million records from 31 thousand patients were used in the analysis. Records were of three types   * Structured data * Text data * Structured and text   5) The text data is represented in vector format i.e. Word embedding.  6) The multi-model disease risk prediction algorithm gives an accuracy of 94.8%. | 1) Application method for real life healthcare tasks.  2) Methods to improve data structuring for better accuracy. |
| 3 | January 2017 | Chen, Ma, Li, Wu, Zhang, Youn | 1) Washable smart clothing consists of sensors, electrodes, and wires.  2) Advanced methods of data retrieval are combined with advanced methods of data analytics and cognitive computing.  3) Body temperature sensor is put in the underarm seam; a set of ECG sensors are mounted on the chest, shoulders, and ribs; the myocardial sensor is embedded in the left part of the chest; and the SpO2 sensor is deployed on the triceps.  4) The data is dependent on the emotional state of the subject.  5) 2 layers of storage are used, the first on local devices and the second on the high-level clouds for advanced processing. | 1) A proper method to collect data without making the user uncomfortable.  2) Usage of algorithms to handle data processing and compression. |
| 4 | March 2015 | S. Vijayarani, S.Dhayanand | 1) The healthcare industry collects huge amounts of healthcare data  2) This data was used to train a system running SVM and ANN algorithms. | 1) Methods to improve classification accuracy |
| 5 | November 2016 | Abadi, Barham, Chen, Dean, Devin, Warden, Ghemawat, Irving, Isard, Kudlur, Chen, Monga, Moore, Murray, Davis, Vasudevan | 1) Tensorflow is a large-scale, heterogeneous machine learning system, which uses dataflow graphs to represent computation.  2) Tensorflow is a descendant to DistBelief, which Google has used since 2011  3) Many neural networks are trained using stochastic gradient descent  4) Tensorflow works across a distributed GPU architecture to maximize efficiency. | 1) Methods to improve classification accuracy |
| 6 | June 2016 | Yong-Young Kim, Mi-Hye Kim | 1) Software developers & medical experts believe that health monitoring and the prevention of diseases should be explored.  2) Absence of certifications for healthcare software is still a concern | 1) Societal acceptance of automized healthcare  2) Specification of anticipated problems |
| 6 | October 2014 | Jürgen  Schmidhuber | 1) A standard neural network (NN) consists of many simple, connected processors called neurons  2) Re-cyclic neural networks are the deepest of all neural networks | 1) Constraints in application |
| 7 | November  2014 | Kouroua,  Exarchos,  Exarchos,  Karamouzis, Fotiadis | 1) Machine learning techniques are being utilized to model the progression and treatment of cancerous conditions.  2) It uses methods like Artificial neural networks, Support vector machine, Decision trees and Bayesian networks. For evaluating the performance, it uses Holdout method, random sampling, Cross validation and Bootstrapping. | 1) Application method for real life cancer detection.  2) Validation accuracy of algorithm. |

2.3 Problem definition (Phase wise)

We have divided the project into 2 phases, which are further divided into 5 and 2 sub-phases each. The problem definitions for each phase and sub-phase are as follows.

Phase I – Phase I will focus on forming a base for the project. With thorough planning, analysis and design, we will ensure that the actual implementation is smoother.

1. Planning: Applying agile methodology for planning our healthcare software project, and to achieve better results in the time frame given to us, along with better flexibility.
2. Analysis: Prepare an analysis of our own project. Prepare a detailed analysis on present healthcare assistants and overcoming its limitation and performance benchmarks.
3. Design: Integration of data and designing of event system app.
4. Coding: Writing the whole system app code and taking help of open source.
5. Implementation: Giving user to test alpha testing and gathering centric analysis of performance, feedback, and try to improve the quality of the result/output

Phase II – Phase II will focus on fine tuning the project and ensuring that the intended features will work as we wanted them to.

1. Testing: Doing the various test on system app like as Unit testing, Integration testing, Regression testing, system testing, etc, of test case data to check if the Integrated system functions are as desired by the client.
2. Deployment: Conduct beta testing for identifying any further errors, bugs and improvements that can be performed. After testing and approval, deploy the proposed system.

2.4 Feasibility study

Economic feasibility

The project relies on open source software, such as python, R, NumPy and SciPy. The functionality provided by these software’s is enough to construct a product capable of usage for healthcare applications.

For the machine’s learning and training, we will use freely available datasets. These sets run into thousands of lines, and thus can provide enough learning capability to the software. Due to these 2 being freely available, the project is very low cost and thus feasible from an economic standpoint.

The expected total cost is thus 0 rupees. As such, we do not need any outside funding. The project will be expected to make minimal sales, due to the presence of a variety of other similar tool and the slow adaption of software’s by the medical community.

Technical

The capabilities required for the project are fairly in the feasible range. Most of the planned feature hinge around processing datasets. A GUI is required, which will be done with the pyGUI framework. The GUI will be kept simple and minimalist. For the hospital searching and insurance selection, we will use readily available data. We also have a feature for users to set personal data, preferences.

For the machine learning, we will use Neural networks to train our datasets, both of which are available on open source repositories.

Operational

On an operational level, the software will need to run intensive processing tasks while it’s in its learning phase. Since the learning phase is a very short process, it is doable without the need for extra hardware. The software will be deployed as a Desktop application and will need users to download it. It will also need internet to operate.

Socio-cultural

This application will have many positive applications on society. It will enable users to manage their healthcare needs in a cheap and easy way. Users with chronic illnesses will find it especially helpful. Those with financial problems will also find it helpful.

Some users may not be willing to accept automated healthcare, but most would be willing. Since it reduces the need for health monitoring, users living in areas with costly healthcare (e.g. most of the USA) will find it helpful. Also, the sharing of instantly available, accurate information can make the difference between life and death.

Legal

Since the project deals with a sensitive topic with potential for major harm, it will carry a disclaimer for improper usage. The software is not meant to be a substitute for doctors or hospital provided healthcare. Users should ideally only use it to supplement professional healthcare. However, the potential for misuse, whether intentional or not, still exists. Much has been said about the ethical advantages of free software in general, and it is particularly true in a profession in which the sharing of instantly available, accurate information can make the difference between life and death. As medical software begins to offer decision support, risk management, performance rating, and analytic features, physicians should learn to be more careful with its usage.

Ethical

While the software is meant to help users with their health, there still is potential for improper usage leading to damage. The software should not be used as a substitute for doctors or professionally provided healthcare. Users must only use it to supplement the same. Much has been said about the ethical advantages of “free” software in general, and it is particularly true in a profession in which the sharing of instantly available, accurate information can make the difference between life and death. As medical software begins to offer decision support, risk management, performance rating, and analytic features, physicians should not accept black boxes and secret formulas that constrain sharing and intimately affect patient care and reimbursement.

2.5 Methodology used

The methodology that we chose for our project is Agile. While waterfall model is a tried and tested framework, it is not suitable for our project, due to the changeable nature and small team size. Agile allows us to be flexible, manage changing requirements, manage the ever-increasing scope as well as get consumer perspective.

Agile has focus on customer satisfaction, which is a crucial aspect of any healthcare application. Patients, insurance providers and doctors need to be clued in to the process for the indispensable inputs they can provide.

2.5.1. Agile

Agile software development is a methodology for software projects under whose framework project requirements and solutions evolve through the collaborative effort of self-organizing and cross-functional teams and their customer and intended users. It advocates for flexible modelling and planning, flexible implementation, rapid delivery of code, and constant improvement, and it advocates speedy and flexible reactions to change.

12 Agile Principles

1. Our highest priority is to satisfy the customer through early and continuous delivery of valuable software.
2. Welcome changing requirements, even late in development. Agile processes harness change for the customer's competitive advantage.
3. Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale.
4. Business people and developers must work together daily throughout the project.
5. Build projects around motivated individuals. Give them the environment and support they need and trust them to get the job done.
6. The most efficient and effective method of conveying information to and within a development team is face-to-face conversation.
7. Working software is the primary measure of progress.
8. Agile processes promote sustainable development. The sponsors, developers, and users should be able to maintain a constant pace indefinitely.
9. Continuous attention to technical excellence and good design enhances agility.
10. Simplicity - The art of maximizing the amount of work not done--is essential.
11. The best architectures, requirements, and designs emerge from self-organizing teams.
12. At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behavior accordingly.

2.5.2 Customer interaction details

We used forums as the main source of our customer input. AI forums as well as healthcare forums are buzzing with activity. Users showed interest in Machine Learning based prediction of diseases, as well as a software that can do it for a reasonable cost.

Besides this, we researched interests of customers in wearable technologies. Technologies like Fitbit have already made a splash in the healthcare market. Users of Fitbit displayed interest in software products that could help them better utilise the data that it was gathering.

2.6 Summary

Here we analysed various prevalent research papers pertaining to our project as a part of the literature survey. The project aims at creating an intelligent but lightweight healthcare assistant. This application will be able to help healthcare and other hospital staff work efficiently.

Our literature survey was focused on the capabilities of the choices we made for the core part of the project, namely artificial neural networks and the use of wearable technology for data collection.

We have divided the project into 2 phases, which are further divided into 5 and 2 sub-phases each. The problem definitions for each phase and sub-phase are as follows.

Phase I – Phase I will focus on forming a base for the project. With thorough planning, analysis and design, we will ensure that the actual implementation is smoother.

Phase II – Phase II will focus on fine tuning the project and ensuring that the intended features will work as we wanted them to.

Due to these 2 being freely available, the project is very low cost and thus feasible from an economic standpoint.

The expected total cost is thus 0 rupees. As such, we do not need any outside funding.

The capabilities required for the project are fairly in the feasible range. Most of the planned feature hinge around processing datasets. A GUI is required, which will be done with the pyGUI framework. The GUI will be kept simple and minimalist. For the hospital searching and insurance selection, we will use readily available data. We also have a feature for users to set personal data, preferences. On an operational level, the software will need to run intensive processing tasks while it’s in its learning phase.

This application will have many positive applications on society. It will enable users to manage their healthcare needs in a cheap and easy way. Since the project deals with a sensitive topic with potential for major harm, it will carry a disclaimer for improper usage. The software is not meant to be a substitute for doctors or hospital provided healthcare. Users should ideally only use it to supplement professional healthcare. However, the potential for misuse, whether intentional or not, still exists.

The methodology that we chose for our project is Agile. Agile allows us to be flexible, manage changing requirements, manage the ever-increasing scope as well as get consumer perspective. Agile has focus on customer satisfaction, which is a crucial aspect of any healthcare application.

Agile software development is a methodology for software projects under whose framework project requirements and solutions evolve through the collaborative effort of self-organizing and cross-functional teams and their customer and intended users.

Users showed interest in Machine Learning based prediction of diseases, as well as a software that can do it for a reasonable cost.

Besides this, we researched interests of customers in wearable technologies. Technologies like Fitbit have already made a splash in the healthcare market.

**Citations**

**Wikipedia agile**

**Agile manifesto**