CHAPTER 1

INTRODUCTION

This thesis is aimed to develop a computer based aid for physicians. Physicians are prone to making decision errors, because of high complexity of medical problems and due to cognitive limitations. Computer based aids aim to reduce physician’s errors by providing guidance/support for decision making. In this chapter, we present an overview of the need for clinical decision support system, and provide a brief description of research undertaken in this thesis.

1.1 Need for Clinical Decision Support System

In the medical domain, a vast amount of knowledge is required even to solve seemingly simple problems. A physician is required to remember and apply knowledge of a vast array of documented disease presentations, diagnostic parameters, combination of drug therapies and guidelines. However the physician’s cognitive abilities are restricted due to factors like multi-tasking, limited reasoning and memory capacity. Consequently, it is impossible for an unaided physician to make the right decision every time. Ironically, the increasing rate of information generated by medical advances has aggravated the physician’s task.

The ideal decision making process contains the process of knowledge discovery. Many researchers consider data mining programs as a way to make decision making tools intelligent. The potential of computer based tools to address the medical decision making problems are realized half a century ago and several algorithms have been developed to construct Clinical Decision Support Systems for a variety of medical specialities. In the following sections we explore the concept of Clinical decision support system.
1.2 Clinical Decision Support System

Clinical Decision Support System (CDSS) is an interactive Decision Support System (DSS) computer software, which is designed to assist physicians and other health professionals with decision making tasks, as determining diagnosis of patient data [1].

A working definition has been proposed by Dr. Robert Hayward of the Centre for Health Evidence "Clinical Decision Support Systems link health observations with health knowledge to influence health choices by clinicians for improved health care". This definition has the advantage of simplifying Clinical Decision Support to a functional concept.

1.3 Purpose of CDSS

The main purpose of modern CDSS is to assist clinicians at the point of care. This means that a clinician would interact with a CDSS to help determine diagnosis, analysis, etc. of patient data. Previous theories of CDSS were to use the CDSS to literally make decisions for the clinician. The clinician would input the information and wait for the CDSS to output the “right” choice and the clinician would simply act on that output. The new methodology of using CDSS to assist forces the clinician to interact with the CDSS utilizing both the clinician’s knowledge and the CDSS to make a better analysis of the patients data than either human or CDSS could make on their own. Typically the CDSS would make suggestions of outputs or a set of outputs for the clinician to look through and the clinician officially picks useful information and removes erroneous CDSS suggestions.

Another important classification of a CDSS is based on the timing of its use. The doctor uses these systems at point of care to help them as they are dealing with a patient, with the timing
of use as either pre-diagnosis, during diagnoses, or post diagnosis. Pre-diagnosis CDSS systems are used to help the physician prepare the diagnosis. CDSS used during diagnoses help review and filter the physician’s preliminary diagnostic choices to improve their final results. Post-diagnosis CDSS systems are used to mine data to derive connections between patients and their past medical history and clinical research to predict future events. Clinical decision support systems are broadly classified into two main groups namely Knowledge based CDSS and Non-knowledge based CDSS.

1. 4 Knowledge Based CDSS

The knowledge based clinical decision support system contains rules mostly in the form of IF-Then statements. The data is usually associated with these rules. The knowledge based generally consists of three main parts. Knowledge base, Inference rules and a mechanism to communicate. Knowledge base contains the rules, inference engine combines rules with the patient data and the communication mechanism is used to show the result to the users as well as to provide input to the system. They are the commonest type of Clinical Decision Support System used in clinics and hospitals. They can have clinical knowledge about a specially defined task, or can even be able to work with case based reasoning. The knowledge within expert system is generally represented as set of rules. Sometimes the knowledge based is used with variance management to execute patient care process and provide high quality health care services dynamically. Knowledge based CDSSs are then further divided into three main categories. They are Fuzzy Logic Rule Based CDSS, Bayesian Network, Rule-based systems and Evidence based Systems.
1.4.1 Fuzzy Logic Rule Based

Fuzzy logic is a knowledge based approach that facilitates solutions to resolve vagueness in decision support system. Fuzzy logic rule based approach can be a very useful approach for describing vagueness and imprecision in precise mathematical language, explicitly representing clinical vagueness[2]. The Fuzzy Logic Rule based classifier is very effective in high degree of positive predictive value and diagnostic accuracy.

1.4.2 The Bayesian network

The Bayesian network is knowledge based graphical representation that shows a set of variables and their probabilistic relationships between diseases and symptoms. They are based on conditional probabilities, the probability of an event given the occurrence of another event, such as the interpretation of diagnostic tests. Baye’s rule helps us compute the probability of an event with the help of some more readily available information and it consistently processes options as new evidence is presented. In the context of CDSS, the Bayesian network can be used to compute the probabilities of the presence of the possible diseases given their symptoms [3].

Some of the advantages of Bayesian Network include the knowledge and conclusions of experts in based on the probabilities that are applicable to many models. The disadvantages of Bayesian Network include the difficulty to get the probability knowledge for possible diagnosis and not being practical for large complex systems given multiple symptoms. The Bayesian calculations on multiple simultaneous symptoms could be overwhelming for users.
1.4.3 Rule-Based Systems & Evidence Based Systems

They tend to capture the knowledge of domain experts into expressions that can be evaluated as rules. When a large number of rules have been compiled into a rule base, the working knowledge will be evaluated against rule base by combining rules until a conclusion is obtained. It is helpful for storing a large amount of data and information. However it is difficult for an expert to transfer their knowledge into distinct rules.

For closing the gap between the physicians and CDSSs, evidence based appeared to be a perfect technique. It proves to be a very powerful tool for improving clinical care and also patient outcomes. It has the potential to improve quality and safety as well as reducing the cost.

1.5 Non Knowledge Based CDSS

CDSS without a knowledge base are called as non knowledge based CDSS. These Systems instead used a form of artificial intelligence called as machine learning. Non knowledge based CDSSs are then further divided into four main categories. They are Artificial Neural Networks, Genetic Algorithms, statistical methods and Hybrid systems.

1.5.1 Artificial Neural Network

Artificial Neural Networks (ANN) is non-knowledge based adaptive CDSS that uses a form of artificial intelligence, also known as machine learning, that allows the systems to learn from past experiences and recognizes patterns in clinical information [4]. It consists of nodes called neuron and weighted connections that transmit signals between the neurons in a forward or looped fashion. An ANN consists of 3 main layers: Input (data receiver or findings), Output (communicates results or possible diseases) and Hidden (processes data). The system becomes
more efficient with known results for large amounts of data. The advantages of ANN include the elimination of needing to program the systems and providing input from experts. The ANN CDSS can process incomplete data by making educated guesses about missing data and improves with every use due to its adaptive system learning. Some of the disadvantages are that the training process may be time consuming leading users to not make use of the systems effectively. The ANN systems derive their own formulas for weighting and combining data based on the statistical recognition patterns over time which may be difficult to interpret.

Neural Networks have been widely applied to non-linear statistical modeling problem and for modeling large and complex databases of medical information. Goal of training is to optimize performance of network in estimating output for particular input space. Back propagation training algorithm, a popular approach used with medical databases adjusts weight of an ANN to minimize a cost function. The ANN maintains correct classification rates and allows a large reduction in complexity of the systems.

1.5.2 Genetic Algorithms

A Genetic Algorithm is a non-knowledge based method based on Darwin’s evolutionary theories that dealt with the survival of the fittest. These algorithms rearrange to form different recombinations that are better than the previous solutions. Similar to neural networks, the genetic algorithms derive their information from patient data. An advantage of genetic algorithms is these systems go through an iterative process to produce an optimal solution. The fitness function determines the good solutions and the solutions that can be eliminated. A disadvantage is the lack of transparency in the reasoning involved for the decision support systems making it undesirable for physicians. The main challenge in using genetic algorithms is in defining the
fitness criteria. In order to use a genetic algorithm, there must be many components such as multiple drugs, symptoms, treatment therapy and so on available in order to solve a problem.

1.5.3 Statistical Method

It is one of most simple and useful method used for data collection. It can be in the form of a survey, experiment result or questionnaire. Development of clinical decision support systems using statistical method as an integral part is very common. Data can be collected as a questionnaire mentioning the status of patients how he looks like, their way of talking, what he feels and many more.

1.5.4 Hybrid Systems

A combination of two or more methodologies within a design of single system results into a hybrid system. Hybrid systems extract the best from all methodologies and provide an optimal solution for clinical decision support systems. Meta reasoning method such as hybrid systems consists of different reasoning methodologies. It can consist of a rule based, case based and model based reasoning. That finally results into an overall improvement of the system performance.

1.6 Benefits of CDSS

The potential benefits of using clinical decision support systems fall into three broad categories

1. Improved patient safety through reduced medication errors and improved test ordering.
2. Improved quality of care by increasing clinicians’ available time for direct patient care, increased application of clinical pathways and guidelines, facilitating the use of up-to-date clinical evidence, improved clinical documentation and patient satisfaction

3. Improved efficiency in health care delivery by reducing costs through reductions in test duplication and changed patterns of drug prescribing favouring cheaper but equally effective generic brands.

1.7 Clinical Decision Support System for appendicitis

According to Department of Measurement and Health Information, World Health Organization, data from Death and Age-standardized disability-adjusted life year (DALY) rates for Appendicitis in India is 9 per 100,000 inhabitants.

Appendicitis is diagnosed when there is an inflammation of the appendix. The appendix is a small finger-shaped pouch attached to the large intestine and is generally found on the lower right side of the abdomen. This organ of body does not have any function. Appendicitis creates a medical emergency and doctors have to perform immediate surgery to remove the appendix. If treatment is not given at crisis time, inflamed appendix will ultimately burst, or perforate, spilling infectious materials into the abdominal cavity. This can lead to peritonitis, a severe inflammation of the abdominal cavity's lining that can be deadly unless it is treated rapidly with strong antibiotics. Due to these facts, cases of appendicitis require surgery. An appendectomy is an emergency surgical procedure to remove an inflamed or infected appendix, a condition known as appendicitis [5-8].

Patient suffering from Appendicitis exhibit many symptoms that include: Dull pain near the navel or the upper abdomen that becomes intense as it moves to the lower right abdomen.
This is considered as a first sign of appendicitis. Patients feel nausea and/or vomiting soon after abdominal pain begins. There is loss of appetite, abdominal swelling, fever range from 99° F to 102° F, failure to pass gas almost half the time. Other symptoms are sharp pain anywhere in the upper or lower abdomen, back or rectum. Patients feel pain in passing urine. Patient vomits that go before the abdominal pain. Patients have severe cramps, and complain for constipation or diarrhea with gas. It is advised to patients that they must contact doctor if they have pain with above symptoms.

Diagnosis of appendicitis can be difficult because patients show similar symptoms that match with other illness. For example, patients who suffer from gallbladder problems, bladder or urinary tract infection, Crohn's disease, gastritis, intestinal infection and ovary problems show parallel symptoms. Patient must visit expert medical professional for correct diagnosis. Some tests are frequently done for the diagnosis of appendicitis. There are abdominal exam to detect inflammation, urine test to rule out a urinary tract infection, rectal exam, blood test to observe if body is fighting infection, CT scans and/or ultrasound

For the treatment of appendicitis, surgery is done to remove the appendix, which is termed as an appendectomy in medical language. It is the normal treatment for appendicitis. Antibiotics are prescribed before an appendectomy to fight possible peritonitis. When surgery is done, patient is given general anesthesia and the appendix is removed through a 4-inch incision or by laparoscopy [9-11].

1.8 Problem Statement

Design and development of efficient feature selection and classification techniques for Clinical Decision Support System is identified as a critical research problem. Hence it is
proposed to develop a Clinical Decision Support System that shall meet out the following requirements:

- To propose a Genetic Algorithm based feature selection approach and compare the performance of existing and the proposed feature selection algorithms on clinical datasets.
- To determine which feature selection technique and classification algorithm is best suited for clinical datasets under consideration.
- To determine which feature extraction technique and classification algorithm is best suited for clinical datasets under consideration.
- To design a Fuzzy Logic Rule based Classifier for appendicitis diagnosis and to predict the severity of appendicitis in patients presenting with right iliac fossa (RIF) pain.
- To propose a scoring system to find the severity of appendicitis for patients who are suffering with right iliac fossa (RIF) pain.

1.9 Thesis Organization

The thesis report is organized as follows. The chapter 1 discusses the background of the research work. It starts with explaining the clinical decision support system and the two types of clinical decision support system namely knowledge based clinical decision support system and non-knowledge based clinical decision support system. The motivation and the main aim of research are also discussed in this chapter.

The chapter 2 presents a detailed literature survey done for the development of clinical decision support system. First it discusses the research issues in developing a clinical decision
support system for clinical data and CDSS for appendicitis. Then it explains the survey of the feature selection algorithms and classification techniques used in CDSS.

The chapter 3 explains the data mining model for CDSS. Here different types of Feature selection techniques used in our project namely Information gain, Chi-Square algorithm, Fast Correlation-Based Filter Approach and Sparse Logistic regression with Bayesian Regularization. It also discusses the different types of and classification techniques used in our project namely Decision Tree induction, Bayesian classifiers and k-Nearest Neighbor (k-NN) Classification.

The chapter 4 explains the proposed scheme titled “Feature Selection using Genetic algorithms”. First Relief Algorithm based feature selection is implemented in Genetic algorithm. Then different classification techniques are used to test the performance. After this, experimental results and analysis and critical findings of Genetic algorithms based feature selection are discussed.

The chapter 5 presents the proposed scheme titled “Implementation of Support Vector Machines for classification of clinical datasets”. Here the performance of different feature extraction techniques namely Principal Component Analysis (PCA) and Linear Discriminant Analysis are compared by a Support Vector Machine based classifier.

The chapter 6 discusses the proposed scheme titled “A Novel Neuro Fuzzy logic rule based classifier for diagnosing appendicitis” to predict the presence and severity of appendicitis among the Patients in a hospital.

The chapter 7 presents the experimental results and analysis of all the existing and proposed feature selection and classification techniques.

The chapter 8 discusses the conclusion of this research and future enhancements.