2.1 Medical Diagnosis & its prerequisites

Medical diagnosis is the process of determining which disease or condition explains a person's signs and symptoms. The information required for diagnosis is typically collected from the history of patient and also by physical examination of the patient seeking medical care. Often, one or more diagnostic procedures, such as diagnostic tests, are also conducted during the process. Correct diagnosis is often challenging, because many signs and symptoms are nonspecific and overlapping with multiple diseases. A diagnosis, in the sense of diagnostic procedure, can be regarded as an attempt at classification of an individual's condition into separate and distinct categories that allow medical decisions about treatment and prognosis to be made. A diagnostic does not necessarily involve elucidation of the etiology of the diseases or conditions of interest, that is, what caused the disease or condition. Such elucidation can be useful to optimize treatment, further specify the prognosis or prevent recurrence of the disease or condition in the future. There are a number of methods or techniques that can be used in a diagnostic procedure, including performing a differential diagnosis or following medical algorithms etc. (John et al., 2002). In reality, a diagnostic procedure may involve components of multiple methods.

2.2 Differential diagnosis

The method of differential diagnosis is based on finding as many candidate diseases or conditions as possible that can possibly cause the signs or symptoms; followed by a process of elimination or at least of rendering the entries more or less probable by further medical tests and other processing until, aiming to reach the point where only one candidate disease or condition remains as probable. The final result may also remain a list of possible conditions, ranked in order of probability or severity. The resultant diagnostic opinion by this method can be regarded more or less as a diagnosis of exclusion. Even if it doesn't result in a single probable disease or condition, it can at least rule out any imminently life-threatening conditions. Unless the provider is certain of the condition present, further medical tests, such as medical imaging, are performed or scheduled in part to prove or disprove the diagnosis. If unexpected findings are made
during this process, the initial hypothesis may be ruled out and the provider must then
consider other hypotheses.

2.3 Pattern recognition

In a pattern recognition method of diagnosis, the medical expert uses experience
to recognize a pattern of clinical characteristics (John et al., 2002). It is mainly based on
certain symptoms or signs being associated with certain diseases or conditions, not
necessarily involving the more cognitive processing involved in a differential diagnosis.
This may be the primary method used in cases where diseases are "obvious", or the
diagnosis provider's experience may enable him or her to recognize the condition quickly.
Theoretically, a certain pattern of signs or symptoms can be directly associated with a
certain therapy, even without a definite decision regarding what is the actual disease, but
such a compromise carries a substantial risk of missing a diagnosis which actually has a
different therapy so it may be limited to cases where no diagnosis can be made.

2.4 Diagnostic criteria

The term “diagnostic-criteria” designates the specific combination of signs,
symptoms, and test results that the clinician uses to attempt to determine the correct
diagnosis. Some examples of diagnostic criteria, also known as clinical case definitions,
are:

- Amsterdam criteria for hereditary non-polyposis colorectal cancer
- McDonald criteria for multiple sclerosis
- ACR criteria for systemic lupus erythematosus
- Centor criteria for strep throat

2.5 Medical Diagnosis using Clinical decision support system

Clinical decision support systems are interactive computer programs designed to
assist health professionals with decision-making tasks. Their formal definition as
provided by Robert Hayward is: "Clinical Decision Support systems link health
observations with health knowledge to influence health choices by clinicians for improved health care”. A clinical decision support system has been defined as an "active knowledge systems”, which use two or more items of patient data to generate case-specific advice. The clinician or medical expert interacts with the software utilizing both the clinician’s own knowledge and experience and the software to make a better analysis of the patient’s condition. Typically the system makes suggestions for the doctor to look through and the he picks useful information and removes erroneous suggestions (Musen et al., 2014). Typically the clinical decision support systems have been designed to facilitate physicians and other medical health care professionals with the assistance to reach at the expert decision using the patients history, present signs and symptoms along with the computer based intelligent algorithms.

2.6 Prerequisites of a good medical diagnosis / Diagnostic procedure

The physician's ability to sharply limit the number of hypotheses under active consideration at any one time is a key element in expert performance (Kassirer et al., 1978, Elstein et al., 1978, Pauker et al. 1976). Computer programs that use the strategies of experts can accomplish this same goal and devote the bulk of their computational resources to the sophisticated evaluation of a small number of hypotheses. Any program designed to serve as a consultant to the physician must contain certain basic features. It must have a store of medical knowledge expressed as descriptions of possible diseases. Depending on the extent of the clinical domain, the number of hypotheses in the database can range from a few to many thousands. In the simplest conceivable representation of such knowledge, each disease hypothesis identifies all of the features that can occur in the particular disorder. In addition, the program must be able to match what is known about the patient with its store of information. Even the most sophisticated programs typically depend on this basic strategy. The simplest version of such programs operates in the following fashion when presented with the chief complaint and when later given additional facts.

1. For each possible disease (diagnosis) determine whether the given findings are to be expected.

2. Score each disease (diagnosis) by counting the number of given findings that would have been expected.

3. Rank-order the possible diseases (diagnoses) according to their scores.

   The power of such a simple program can be greatly enhanced through the use of a mechanism that poses questions designed to elicit useful information. Take, for example, an expansion of the basic program by the following strategy:

4. Select the highest-ranking hypothesis and ask whether one of the features of that disease, not yet considered, is present or absent.

5. If inquiry has been made about all possible features of the highest-ranked hypothesis, ask about the features of the next best hypothesis.

6. If a new finding is offered, begin again with step 1; otherwise, print out the rank-ordered diagnoses and their respective supportive findings and stop.

   Steps 1 through 3 contain a primitive evaluation of the available information, and steps 4 through 6 contain an equally simple information-gathering strategy that determines what information to seek next. But such a program fails to capture many of the techniques responsible for expert performance. For example, the ranking process does not take into account how frequently particular features occur in a given disease. The program, furthermore, has no knowledge of patho-physiology and is not able to take stock of the severity of an illness. The most serious problem is that each new finding sets into motion a search process tantamount to considering all disease states appearing in a textbook of medicine. Even for a high-speed computer this is not a practical diagnostic strategy and for this reason research has turned to the study of how experts perform.

   Artificial intelligence based algorithms are increasingly being used to analyze and interpret large volumes of data for solving problems in medical domains (Sarwar et al., 2013), as is evident from a considerable amount of research done in this field during
recent past. A detailed review of literature regarding the application of artificial intelligence in the field of medical diagnosis has been done through different sources like journals, research papers, books and articles.

2.7 Literature review with respect to Application of artificial intelligence in Heart diseases

(Avci et al., 2009) have developed an expert diagnosis system for interpretation of the Doppler signals of the heart valve diseases based on the pattern recognition. Their study deals with the feature extraction from measured Doppler signal waveforms at the heart valve using the Doppler Ultrasound technique. Wavelet transforms and short time Fourier transform methods are used to feature extract from the Doppler signals on the time–frequency domain and Wavelet entropy method is applied to these extracted features. For classification of extracted features they have used the back-propagation neural network. The performance of the expert system has been evaluated using 215 samples for which the correct classification rate was about 94% for normal subjects and 95.9% for abnormal subjects. The test results have shown that this system was effective to detect Doppler heart sounds. (Kukar et al., 1999) have worked on the improvement of the diagnosis of Ischaemic heart disease using machine learning. Ischaemic heart disease is predominantly the leading cause of mortality worldwide; as such improvements and rationalization of diagnostic procedures would be very useful. The four diagnostic levels of the said disease consist of a) evaluation of signs and symptoms of the disease and ECG (electrocardiogram) at rest, b) sequential ECG testing during the controlled exercise, c) myocardial scintigraphy, and d) coronary angiography. The authors have used machine learning for objective interpretation of all available results of all the diagnostic levels for the increased diagnostic accuracy at each step. They have conducted multiple experiments with various learning algorithms and achieved the performance level comparable to that of clinicians. The results obtained were evaluated using ROC analysis which showed significant improvements of sensitivity and specificity compared to the performance of the clinicians.
Tantimongcolwat et al., 2008) have proposed an automatic method for the interpretation of Ischaemic heart disease pattern of magnetocardiography (MCG) recordings using machine learning approaches. Early detection of IHD may effectively prevent severity and reduce mortality rate. Recently, magnetocardiography (MCG) has been developed for the detection of heart malfunction. Although MCG is capable of monitoring the abnormal patterns of magnetic field as emitted by physiologically defective heart, data interpretation is time-consuming and requires highly trained professional. They used two types of machine learning techniques, namely back-propagation neural network (BNN) and direct kernel self-organizing map (DK-SOM), to explore the Ischaemic heart disease pattern recorded by MCG. For training the system the prepared a database of sequential measurement of magnetic field emitted by cardiac muscle of 125 individuals. This data was divided into training set and testing set of 74 cases and 51 cases, respectively. The performance of the techniques was analyzed using the sensitivity, specificity and classification accuracy. The BNN exhibited sensitivity of 89.7%, specificity of 54.5% and accuracy of 74.5%, while the DK-SOM provided relatively higher prediction performance with a sensitivity, specificicity and accuracy of 86.2%, 72.7% and 80.4%, respectively. This finding suggests a high potential of applying machine learning approaches for high-throughput detection of IHD from MCG data.

Avci, 2009) has presented an intelligent system based on genetic-support vector machines (GSVM) approach for classification of the Doppler signals of the heart valve diseases. The system deals with combination of the feature extraction and classification from measured Doppler signal waveforms at the heart valve using the Doppler ultrasound. Genetic-support vector machines is used for detection of the heart valve diseases, which selects the most appropriate wavelet filter type for problem, wavelet entropy parameter, the optimal kernel function type, kernel function parameter, and soft margin constant C penalty parameter of support vector machines (SVM) classifier. The performance of the GSVM system is evaluated for 215 samples. The averaged rate of correct classification rate was about 95% which shows that this GSVM system is effective to detect Doppler heart sounds. Colak et al., 2008) used eight different learning algorithms for creating artificial neural network models for the prediction of coronary artery disease (CAD). The work was carried out as a retrospective case-control study.
Overall, 124 consecutive patients who had been diagnosed with CAD by coronary angiography (at least 1 coronary stenosis > 50% in major epicardial arteries) were enrolled in the work. Angiographically, the 113 people (group 2) with normal coronary arteries were taken as control subjects. Multi-layered perceptrons artificial neural network architecture were applied. The artificial neural network models trained with different learning algorithms were performed in 237 records, divided into training (n=171) and The results have demonstrated that ANN models trained with eight different learning algorithms are promising because of high (greater than 71%) sensitivity, specificity and accuracy values in the prediction of CAD. Accuracy, sensitivity and specificity values varied between 83.63%-100%, 86.46%-100% and 74.67%-100% for training, respectively. For testing, the values were more than 71% for sensitivity, 76% for specificity and 81% for accuracy, and accuracy values based on standard definitions. The authors have proposed that the use of different learning algorithms other than backpropagation and larger sample sizes can improve the performance of prediction. The proposed ANN models trained with these learning algorithms could be used as promising approach for predicting CAD without the need for invasive diagnostic methods and could help in the prognostic clinical decision. (Khan et al., 2013 used ANN in diagnosis of Acute Nephrites and Heart disease. The ANN trained by feed forward back propagation algorithm used disease symptoms for the Acute Nephrites and Single Proton Emission Computed Tomography (SPECT) images for the heart disease. The trained ANN then classified the patient’s datasets into infected or non-infected cases

2.8 Literature review with respect to Application of artificial intelligence in Lung diseases

(Rao et al., 2007) have proposed an artificially intelligent system (LungCAD) that helps in the detection of Lung cancer. They have applied a classification algorithm for detecting solid pulmonary nodules from CT thorax studies. The LungCAD system was clinically tested by a number of radiologists and was found to deliver significantly greater accuracy both in detecting the affected nodules and in identifying the potentially actionable nodules. LungCAD was approved by FDA in 2006. (Zhou et al., 2002) have proposed an automatic pathological diagnostic procedure based on the ensemble of
neural networks for identification of lung cancer cells in the images of the specimen of needle biopsies obtained from the bodies of the subjects to be diagnosed. They named this system as neural ensemble based detection (NED). They have build the ensemble on a two level based architecture among which the first level ensemble is used to judge whether the cell is normal or cancerous with high level of confidence. The second level deals with the cells which are marked as cancerous by the level one ensemble. The second level ensemble has five outputs viz adenocarcinoma, squamous cell carcinoma, small cell carcinoma, large cell carcinoma and normal. The authors have reported that the neural ensemble based detection (NED) has not only shown a high rate of overall identification but also a very low rate of false negative identification. (Soda et al., 2011) developed an Artificial intelligence based system that helps in the diagnosis of Systemic Lupis Erythematosus, an autoimmune disease in which our body’s own immune system by mistake attacks the healthy tissue, causing a long term chronic inflammation. The traditional recommended method involves the indirect immunofluorescence based on Crithidia Luciliae substrate. This method suffers from number of issues which challenge the reliability of the results. The authors have proposed a system based on three main decision phases: threshold based classification, cell location & feature extraction and cell classification. The performance of the Artificial intelligence based system when compared with annotated database of immunofluorescence based on Crithidia Luciliae substrate was found to be 99.4%, 98.6% and 99.6% of the three respective phases. (Kodaz et al, 2009) have proposed a medical application of a new artificial immune system which minimizes the overheads of taking into account all attributes for calculating Euclidean distance in shape-space representation used in many artificial immune systems. They called this application as the information gain based artificial immune recognition system (IG-AIRS). Authors have tested the system using the thyroid disease data set. The proposed system expressed an efficiency of 95.90% for ten-fold cross validation. The results suggest that the information gain based artificial immune recognition system (IG-AIRS) would be helpful in diagnosing thyroid function based on laboratory tests.
2.9 Literature review with respect to Application of artificial intelligence in Thyroid disorders

(Sharpe et al., 1993) studied the application of Artificial Neural Networks for devising a helpful aid to doctors in diagnosis of Thyroid related disorders. They used a secondary data set of 392 patients, having the complete set of results of six laboratory tests viz total thyroxine, free thyroxine, tri-iodothyronine, tri-iodothyronine uptake test, thyrotropin, and thyroxine-binding globulin. Using the data set two types of artificial neural networks were made viz multilayer perceptron with back-propagation, and a learning vector quantization network. The training data subsets for the two types of networks were selected according to two strategies: either training data based on cases with extreme values for the laboratory tests with randomly selected non-extreme cases added, or training cases from very pure functional groups. Both network architectures were found to perform efficiently irrespective of the type of training data. The authors found that the correct allocation of cases in test data subsets was 96.4-99.7% when extreme values were used for training and 92.7-98.8% when only pure cases were used. (Keleş et al., 2008) have designed and developed an expert system for the diagnosis of thyroid disorders. They have called this system as Expert System for Thyroid Disease Diagnosis (ESTDD). The thyroid gland is one of the most important organs in the body as thyroid hormones are responsible for controlling metabolism. As a result, thyroid function impacts on every essential organ in the body. When the thyroid produces too much hormone, the body uses energy faster than it should. This condition is called hyperthyroidism. When the thyroid does not produce enough hormones, the body uses energy slower than it should. This condition is called hypothyroidism. Thyroid disease can be difficult to diagnose because symptoms are easily confused with other conditions. If thyroid disease is diagnosed in early stage, treatment can control the disorder even before the onset of symptoms. Authors have found fuzzy rules by using neuro fuzzy method, which will be emplaced in ESTDD system. ESTDD could diagnose thyroid diseases with 95.33% of accuracy. (Azar et al., 2012) have proposed a novel classifier for diagnosis of Thyroid disorders, which they named as Linguistic Hedges Neural-Fuzzy Classifier with Selected Features (LHNFCFSF). The metrics for evaluation of performance
considered by the authors is the comparison of classification results with the actual real
time results and the k-fold cross validation. The results indicated that the classification
accuracy without feature selection was 98.6047% and 97.6744% during training and
testing phases, respectively with RMSE of 0.02335. After applying feature selection
algorithm, LHNFCSF achieved 100% for all cluster sizes during training phase. However,
in the testing phase LHNFCSF achieved 88.3721% using one cluster for each class,
90.6977% using two clusters, 91.8605% using three clusters and 97.6744% using four
clusters for each class and 12 fuzzy rules. (Polat et al., 2007) have proposed a hybrid
machine learning method for diagnosing thyroid disease using classification technique
which includes fuzzy weighted pre-processing. Proper interpretation of the thyroid gland
functional data is an important issue in the diagnosis of thyroid disease. Artificial immune
systems (AISs) is a new but effective branch of artificial intelligence. Among the systems
proposed in this field so far, artificial immune recognition system (AIRS), which was
proposed by A. Watkins, has shown an effective and intriguing performance on the
problems it was applied. The authors have evaluated the robustness of this method with
regard to sampling variations using a cross-validation method, on datasets obtained from
UCI machine learning repository. They have obtained about 85% of classification
accuracy for 10 fold cross validation.

2.10 Literature review with respect to Application of artificial
intelligence in Cancer

(Vander Gaag et al., 1992) developed a decision-support system for patient-
specific therapy selection for oesophageal cancer. The system predicts the correct stage of
cancer, which helps the oncologist to start with the correct treatment plan for the patient.
The kernel of the system is a probabilistic network that describes the presentation
characteristics of cancer of the oesophagus and the patho-physiological processes of
invasion and metastasis. Authors state that the construction of the graphical structure of
the network was relatively straightforward, probability elicitation with existing methods
proved to be a major obstacle. To overcome this obstacle, they designed a new method for
eliciting probabilities from experts that combines the ideas of transcribing probabilities as
fragments of text and of using a scale with both numerical and verbal anchors for marking
assessments. The new method developed allowed authors to elicit many probabilities in reasonable time. To gain some insight in the quality of the probabilities obtained, they conducted a preliminary evaluation study of our network, using data from real patients. Upon testing the system it, was found that for 85% of the patients, the network predicted the correct cancer stage. (Liu et al., 1996) prepared a database on 2692 dyspeptic patients over the age of 40 which consisted of 73 epidemiological and clinical attributes. A tree-based machine learning algorithm (PREDICTOR) was applied to this database, in order to attempt to find rules which would classify patients into 2 groups, i.e., those suffering from gastric or oesophageal cancer, and the remainder. The results were encouraging. The cross-validated classification performance figures showed that, by classifying 61.3% of the patients as high risk, a sensitivity of 94.9% and a specificity of 39.8% could be achieved. They have also constructed an expert system based on the rules produced by the machine learning algorithm, in order to provide preliminary screening for cancer in dyspeptic patients. (Hammond et al., 1998) analyzed the application of machine learning techniques for prognosis of oral cancers. Although Oral cancer has relatively low incidence but it is potentially very serious if not identified at preliminary stages. For economic reasons, widespread screening for the disease is not appropriate and feasible, as such opportunistic, computer aided screening in primary and secondary healthcare highly needed for preliminary screening. The low incidence and disproportionately low number of positive diagnoses arising in oral screening programs gives rise to a very sparse dataset. Even so, the authors have found the performance of artificial intelligent systems to be at par with the medical specialist in prognostic screening of high risk oral cancer patients. Once screened by the artificial intelligent system, the selected patients can be recalled for a detailed examination of their oral mucosa and lifestyle counseling. (Sarkar et al., 2010) studied the breast cancer diagnosis and found its correlation with the pattern classification problem of computer science. They applied the Wisconsin- Madison Breast Cancer data set for detailed analysis of the problem. To address this problem they applied the machine learning approach called K-Nearest Neighbour and compared the results with the other techniques used for addressing the same problem. In the analysis it was found that K-Nearest Neighbour algorithm offers a potentially significant aid to the diagnosis of breast cancer and was also found to perform 1.17% better than the best known solutions.
for this problem. (Delen et al., 2005) applied multiple Artificial Intelligent algorithms for predicting the survivability of breast cancer patients. They used two popular data mining algorithms (artificial neural networks and decision trees) along with a most commonly used statistical method (logistic regression) to develop the prediction models. For the purpose of training and testing the secondary data was used from a large dataset having more than 200,000 cases. Authors also used 10-fold cross-validation methods to measure the unbiased estimate of the three prediction models for performance comparison purposes. The results of the study indicated that the decision tree (C5) is the best predictor with 93.6% accuracy on the holdout sample (this prediction accuracy is better than any reported in the literature), artificial neural networks came out to be the second with 91.2% accuracy and the logistic regression models came out to be the worst of the three with 89.2% accuracy. This study provides an insight into the relative prediction ability of different data mining algorithms. (Nandy, 2013) did an analytical study of supervised and un-supervised machine learning technique for the classification of breast cancer. ANN and SVM from the supervised section showed better classification accuracy as compared to the KNN from the unsupervised learning category. The paper also showed the effect of feature selection on the classification accuracy as the performance dropped sharply when used a wrong set of input features. (Arihito Endo et al., 2008) have performed a comparative analysis of various artificial intelligence and statics based models to predict the survival rate of breast cancer patients in five years. They have examined 37,256 follow-up patients by 2002 that were diagnosed as breast cancer and registered in the SEER program from 1992 to 1997. They implemented seven common algorithms like Logistic Regression model, Artificial Neural Network, Naive Bayes, Bayes Net, Decision Trees with naive Bayes, Decision Trees (ID3) and Decision Trees (J48)) besides the most widely used statistical method (Logistic Regression model) to develop the prediction model. The accuracy in the prediction was 85.8±0.2%, 84.5±1.4%, 83.9±0.2%, 83.9±0.2%, 84.2±0.2%, 82.3±0.2%, 85.6±0.2% for the Logistic Regression model, ANN, Naive Bayes, Bayes Net, Decision Trees with naive Bayes, ID3 and J48, respectively. Authors have reported that in their study, Logistic Regression model showed the highest accuracy. The J48 had the highest sensitivity and the ANN had the highest specificity. The Decision Trees models tend to show high sensitivity. And the Bayesian
models were apt to show the accuracy going up. (Abbod et al., 2007) compared the predictive accuracy of neuro-fuzzy modeling, Artificial Neural Networks and Statistical Regression in predicting the behavior of bladder cancer. Authors developed separate models for all the three methods that predicted the presence and the timing of tumor progression. The comparative analysis showed that neuro-fuzzy modeling and artificial neural networks showed good accuracy ranging from 88% to 100% as compared to the comparative low performance of statistical regression. Among neuro-fuzzy modeling and artificial neural networks, the former proved to be good predictor for predicting the timing of progression.

(Chiu et al., 2009) studied the application of artificial neural networks in predicting the skeletal metastasis in patients suffering from prostate cancer. Authors analyzed whole body bone scintigraphies of patients with prostate cancer who underwent the technethium-99m methylene diphosphate (tc-99m MDP) for a period between 2001 to 2005. For the analysis they used an artificial neural network having four layers of perceptrons which was trained with the data set of 111 prostate cancer patients. For evaluating the performance of ANN, Receiver Operating Characteristics (ROC) analysis was used. The area under the ROC curve (0.88±0.07) revealed excellent discriminatory power (p<0.001) with the best simultaneous sensitivity (87.5%) and specificity (83.3%). Authors concluded that that an ANN, which is based on limited clinical parameters, to be a promising method in forecasting of the skeletal metastasis. (Chang et al., 2009) developed an artificial intelligence based system for early screening of pancreatic carcinoma. The earliest symptoms of pancreatic carcinoma are non-specific, causing the doctors’ diagnostic decisions to differ based on individual experiences, knowledge of the disease, and influence of the mental conditions during examination. Certain diagnostic errors are thus inevitable to occur and mislead the following treatment plans. Authors used artificial neural network, genetic algorithm of artificial intelligence, and logistic regression of statistics to construct three types of screening models for pancreatic cancer and acute pancreatitis. In addition to this, they used the ROC curves to compare and analyze discriminations of the above-mentioned three screening models. Authors have used 234 case patient data for training and 117 cases for testing. It was found that the AI based system performs significantly better than traditional approach. (Kononenko et al.,
presented a view on the use of Machine learning techniques 1) in the past for the interpretation of medical data 2) for intelligent analysis of medical data in the current scenario and 3) For assistance of physicians in diagnosis of medical disorders, in the future. The authors suggested integration of machine learning techniques with the existing instrumentations for the acceptance of machine learning in medicine. (Sarhan et al., 2009) Developed an artificial neural network (ANN) model for classification of tumor or non-tumor stomach cancer (gastric cancer) based on the gene expression data. Discrete carrier transform (DCT) is used for the reduction of micro-array images. The micro-array image is first pre-processed and converted into a matrix of DCT coefficients. The Stomach cancer is most common in the old male people as compared to females and most of them develop the cancer after 55 years of age. The ANN model is trained on 50 microarray images (25 of tumour, 25 of non-tumour used). The system is then tested on 30 microarray images (15 of tumour, 15 for non-tumour). The data for experimentation is accessed from the Stanford microarray database (SMD).

(Cheng et al., 2010) have develop an efficient approach to improve the accuracy of medical diagnosis of breast cancer. The proposed approach is a combination of support vector machine & artificial immune system. For the experimentation they have used the dataset from UCI Machine Learning Repository on which they have applied standardization and normalization pre-processing procedure. A simple artificial immune algorithm with various affinity criteria is investigated for comparison. Furthermore, the grid-search with 10-fold cross-validation is applied to choose two parameters of $C$ and $\gamma$ for AIS-based machine learning classifier. Through grid-search technique, the proposed classifier could yield the best results.

(Ye HQ et al., 2003) have analyzed the expression profiles of hepatocellular carcinoma (HCC) samples without or with intra-hepatic metastases. Using a supervised machine-learning algorithm, they generated for the first time a molecular signature that can classify metastatic HCC patients and identified genes that were relevant to metastasis and patient survival. Hepatocellular carcinoma (HCC) is one of the most common and aggressive human malignancies. Its high mortality rate is mainly a result of intra-hepatic metastases. They found that the gene expression signature of primary HCCs
with accompanying metastasis was very similar to that of their corresponding metastases, implying that genes favoring metastasis progression were initiated in the primary tumors. Osteopontin, which was identified as a lead gene in the signature, was over-expressed in metastatic HCC; an osteopontin-specific antibody effectively blocked HCC cell invasion in vitro and inhibited pulmonary metastasis of HCC cells in nude mice. Thus, osteopontin acts as both a diagnostic marker and a potential therapeutic target for metastatic HCC.

### 2.11 Literature review with respect to Application of artificial intelligence in Diabetes

(Humar et al., 2008) have proposed a system called Design of hybrid system for the diabetes for the diabetes and heart disease. In this system data is classified according to its properties. Classification is implemented by developing model with existing records by using sample data. The data was obtained from the University of California at Irvine (UCI) machine learning repository. In order to evaluate the performance of proposed method accuracy, sensitivity and specificity performance measures that are used commonly in medical classification studies. The classification of accuracy of these dataset were obtained by K-fold cross validation and the proposed method achieve accuracy values 84.24% and 86.8% for Pima Indian Diabetes Dataset and Cleveland Heart Disease Dataset respectively. It has been observed that these results are one of the best results obtained from related previous studies and reported in UCI websites. (Nahla Barakatet.al 2010) have proposed intelligible Support Vector Machines for Diagnosis of Diabetes Mellitus. In this system it has been shown that 80% of type 2 diabetes complications can be prevented by early identification of people at risk by using several data mining and machine learning methods for the diagnosis, prognosis and management of diabetes. In this support vector machines (SVM) is used which turns into ‘black box’ model of an SVM into an intelligible representation of SVM diagnostic. Results on real-life dataset show that intelligible SVM provide a promising tool for prediction of diabetes, with prediction accuracy of 94%, sensitivity of 93% and specificity of 94%.

(Oguz et al., 2012) have proposed a system for diagnosing diabetes using neural networks on small mobile devices in which Pervasive computing is being used in
improving healthcare. In this system a novel approach is used for diagnosing diabetes with the help of neural networks and pervasive healthcare computing technologies. In this new software techniques and mobile services have been used for pervasive healthcare computing in which initial results for simple client (patient’s PDA) and server (powerful desktop PC) the computations of neural network operations are used on both client and server sides whose applications are based on artificial neural network (ANN) techniques. The result shows that the computations of ANN operations on both client and server sides and wireless network communications between them are optimized for real time use of pervasive healthcare services. The client mobile application tries to make its ANN and other complex calculations locally and shows the results to the patient without contacting to the server for a range of illness including diabetes. (Shu-ChenCheng et al., 2003) have proposed a system based on the fractal characteristic of retinal images in which a novel diagnostic scheme is used to develop quantitative indexes of diabetes. The fractal dimension of the vascular distribution is estimated because the fractal dimension of a severe diabetic patient’s retinal vascular distribution appears greater than that of normal human’s. Clinically, the incidence of diabetes retinopathy (DR) is prevalent in patients diagnosed as having diabetes. DR has features of both micro-vascular occlusion, retinal hypoxia which causes blindness. As a result an effective diagnostic scheme is being used in which four different approaches are compared to achieve higher accuracy to develop quantitative indexes for diabetes. With these quantitative indexes, diabetics can maintain good metabolic control to retard the development of Diabetes retinopathy.

(Elif et al., 2010) proposed a system for automatic diagnosis of diabetes using adaptive neuro-fuzzy inference systems in which a new approach based on adaptive neuro-fuzzy inference system (ANFIS) is presented for diagnosis of diabetes diseases. In this the ANFIS classifiers learn how to differentiate a new case in the domain by being given training set of such records. The ANFIS classifier is used to detect diabetes by defining diabetes symptoms as inputs. The proposed ANFIS model then combines neural network adaptive capabilities and the fuzzy logic approach. The results on diagnosis of diabetes are obtained through analysis of the ANFIS, in which ANFIS model is evaluated in terms of training performance and classification accuracies and the results confirm that the proposed ANFIS model is potential in detecting diabetes disease. (Yasodha et al.,
2011) did analysis of a database of diabetic patients using Weka tool. The authors considered different algorithms such as REP Tree, Bayes Network, J48 and Random Tree classifiers for the study and compared the outputs. The main objective of the study was to develop a Diabetic expert system; inputs being patient’s daily glucose rate and insulin dosages the system would predict the patient’s insulin dosage for the next day. (HasanTemurtas et al., 2009) have proposed a comparative study on diabetes disease diagnosis using neural networks in which diabetes disease diagnosis is done by using a multilayer neural network structure which was trained by Levenberg – Marquardt (LM) algorithm and a probabilistic neural network structure. The results of the study were compared with the results of the previous studies reported focusing on diabetes disease diagnosis and using the same UCI machine learning database.

(Kemal et al., 2007) have proposed a system for diagnosis of diabetes using principal component analysis and adaptive neuro fuzzy inference system. The author’s have used the dataset obtained from UCI machine learning data repository. The system works in two phases, in first phase feature selection algorithm is used to reduce the size of the data set for analysis from eight attributes to four attributes; in the second phase the data set with four attributes is fed to neuro fuzzy inference system for artificial intelligence based diagnosis of diabetes. They have reported that their system has presented an efficiency of 89.47% which is better than the other systems reported in the literature. (Chad et al., 2006) have proposed a system for prediction of diabetes type-II using anthropometrical body. They have used four data mining approaches including dignitary, logistic regression, neural networks and rough set for feature selection and reducing the size of data base for analysis. The results of their study have indicated that the volume of trunk, left and right thigh circumference, waist circumference, volume of right leg and subjects age are the factors /parameters that play key role in the manipulation of disease. They have also reported that the classification by dignitary and rough set is much better as compared to the classification of back propagation and logical regression.

2.12 Literature review with respect to Application of artificial intelligence in Cervical Cancer
Sokouti et al., 2014) have proposed a model (Levenberg–Marquardt feedforward MLP neural network-LMFFNN) for classification of cervical cell images. This model is a novel application based upon the feedforward neural network, and has been trained and tested from the data obtained from 100 patients. The proposed semi-automated cervical cancer diagnosis system is composed of two stages, in the first stage, images are preprocessed to reduce the noise (if any) without compromising on the resolution of images and in the second stage image processing algorithms are applied to cell images to achieve a linear plot, which were then used as LMFFNN inputs for classification of cervical cell images. Their system has shown 100% correct classification rate and was thus found to be in good correlation with the decision made by the medical experts.

Mat-Isa et al., 2008) developed an automated system for diagnosis and screening pre-cancerous cervical cells. Their system was composed of two components one of the components for automatic feature extraction and the other for an intelligent diagnosis. The first component automatically draws out four critical features (i.e. nucleus size, cytoplasm size, nucleus grey level and cytoplasm grey level) from the cervical cells. The authors have developed a novel algorithm called region-growing-based features extraction (RGBFE). This algorithm is used for extraction of features important for diagnosis. The data about all these features when extracted from the cervical cell images are fed to the intelligent diagnostic part. The pre-cancerous stages are predicted using artificial neural network, developed using a novel architecture called hybrid multilayered perceptron (H2MLP) network. The cells are classified into three classes as normal, low grade intra-epithelial squamous lesion (LSIL) and high grade intra-epithelial squamous lesion (HSIL). The capability of the system so developed is assed using 550 clinical cases which were classified as normal cases to be 211, LSIL cases to be 143 and HSIL cases to be 196. For the purpose of evaluation of performance of the system in comparison to the manual extraction by expert cytologist, correlation test was used. The results imply a strong linear relationship between mean of grey level and the estimated size with that extracted the expert cytotechnologist.

Pai et al., 2012) developed an intelligent system for automatic detection of shape of nucleus and cytoplasm of cell of cervix obtained from Pap smear. They called this
system as nucleus and cytoplasm contour detector (NCC detector). They used adaptable threshold decision method to distinguish the cell from the cervical smear image, followed by using the maximal gray-level-gradient- difference method, proposed by them, for extraction of the nucleus from the cell. The comparative analysis of NCC detector with the earlier available methods reveled that NCC detector performs better than the edge enhancement nucleus and cytoplast contour detector model and the then gradient vector flow-active contour model. (Genc-tav et al., 2012) proposed an unsupervised approach for segmentation and classification of cervical cells obtained from Pap smear slides. The segmentation process involves providing an automatic threshold for separating the cell regions from the background, a multi-scale hierarchical segmentation algorithm to partition these regions based on homogeneity and circularity, and a binary classifier to finalize the separation of nuclei from cytoplasm within the cell regions. The proposed procedure constructs a tree using hierarchical clustering, and then arranges the cells in a linear order by using an optimal leaf ordering algorithm that maximizes the similarity of adjacent leaves without any requirement for training examples or parameter adjustment. Performance evaluation using two data sets show the effectiveness of the proposed approach in images having inconsistent staining, poor contrast, and overlapping cells. (Jan Jantzen et al., 2005) have performed a case study and have prepared data and baseline for comparing classification methods. The data collected by them consists of 917 images of Pap-smear cells, classified carefully by cytotechnicians and doctors. Each cell is described by 20 numerical features, and the cells fall into 7 classes. Authors have also done a basic data analysis that includes scatter plots and linear classification results, in order to provide domain knowledge and lower bounds on the acceptable performance of future classifiers.

(Mat-Isa et al., 2008) develop an automated diagnostic system for screening precancerous cells of cervix. The automated system developed by them consists of two parts; an automatic feature extraction part and an intelligent diagnostic part. In the automatic feature extraction, the system automatically extracts four cervical cells features (i.e. nucleus size, nucleus grey level, cytoplasm size and cytoplasm grey level). A new features extraction algorithm called region-growing-based features extraction (RGBFE) is proposed to extract the cervical cells features. The extracted features will then be fed as
input data to the intelligent diagnostic part. A new artificial neural network (ANN) architecture called hierarchical hybrid multilayered perceptron (H2MLP) network is proposed to predict the cervical pre-cancerous stage into three classes, namely normal, low grade intra-epithelial squamous lesion (LSIL) and high grade intra-epithelial squamous lesion (HSIL). Authors have empirically assessed the capability of the proposed diagnostic system using 550 reported cases (211 normal cases, 143 LSIL cases and 196 HSIL cases). For evaluation of the automatic feature extraction performance, correlation test approach was used to determine the capability of the RGBFE algorithm as compared to manual extraction by cytotechnologist. The results show that the estimated size and mean of grey level have strong linear relationship (correlation test more than 0.8) with those extracted manually by cytotechnologist. This project has successfully developed an automatic diagnostic system for cervical pre-cancerous.

(Zhang et al., 2008) developed a computerized system for differentiating cervical lymph nodes on ultrasonography as malignant or benign. For this classification they used a rough margin based support vector machine trained by a dataset of 210 clinical cases, which contained the data for ten quantitative features pertinent to the cervical lymph nodes. For evaluating the performance of the computerized system, Receiver Operating Characteristics (ROC) analysis was used, which showed the normalized area of 0.892, whereas that for that of a radiologist, it showed an area of 0.784. Authors conclude the system to be a potentially good aid to radiologists in the task of distinguishing between malignant and benign cervical nodes on ultrasonography.

(Tsai et al., 2008) have developed a system for automatic segmentation of cytoplasm and nucleus from the images of cervical cells. They called this system as Cytoplasm and Nucleus Contour (CNC) Detector. They have also proposed the bi-group enhancer to make a clear-cut separation for the pixels laid between two objects, and the maximal colour difference (MCD) method to draw the most appropriate nucleus contour. The CNC detector adopts a median filter to sweep off noises, the bi-group enhancer to suppress the noises and brighten the object contours, the K-mean algorithm to discern the cytoplasm from the background, and the MCD method to extract the nucleus contour. The experimental results show that the CNC detector can give an impressive performance.
Besides cervical smear images, these proposed techniques can be utilized in segmenting objects from other images also.

(Marinakis et al., 2009) have proposed a meta-heuristic algorithm in order to classify the cells of cervix. Two databases are used, constructed in different times by medical expert, which consist of 917 and 500 images of pap smear cells, respectively. Each cell is described by 20 numerical features, and the cells fall into 7 classes but a minimal requirement is to separate normal from abnormal cells, which is a 2 class problem. For finding the best possible performing feature subset selection problem, an effective genetic algorithm scheme is proposed. This algorithmic scheme is combined with a number of nearest neighbor based classifiers. Results show that classification accuracy generally outperforms other previously applied intelligent approaches. (Plissiti et al., 2011) have presented an automated method for the detection and boundary determination of cells nuclei in Pap stained cervical smear images. The detection of the candidate nuclei areas is based on a morphological image reconstruction process and the segmentation of the nuclei boundaries is accomplished with the application of the watershed transform in the morphological color gradient image, using the nuclei markers extracted in the detection step. For the elimination of false positive findings, salient features characterizing the shape, the texture and the image intensity are extracted from the candidate nuclei regions and a classification step is performed to determine the true nuclei. They have examined the performance of two unsupervised (K-means, spectral clustering) and a supervised (Support Vector Machines, SVM) classification technique, employing discriminative features which were selected with a feature selection scheme based on the minimal-Redundancy–Maximal-Relevance criterion. The proposed method was evaluated on a data set of 90 Pap smear images containing 10,248 recognized cell nuclei. Comparisons with the segmentation results of a gradient vector flow deformable (GVF) model and a region based active contour model (ACM) are performed, which indicate that the proposed method produces more accurate nuclei boundaries that are closer to the ground truth.

(Sarwar et al., 2015) presented a novel hybrid ensemble technique i.e. ensemble of ensemble methods for improving the predictive performance of Artificial intelligence
based system for screening of cervical cancer by characterization and classification of Pap smear images. Using such a technique, the classification potentials of individual algorithms are fused together to gain greater classification accuracy. In addition to this they have also presented a comparative analysis of various artificial intelligence based algorithms for screening of cervical cancer. For evaluation of the predictive performance of the said technique, authors have used root mean square error and percentage of correct classification as the measures. The results indicate that hybrid ensemble technique is an efficient method for classification of Pap smear images and hence can be effectively used for diagnosis of cervical cancer. Among all the algorithms implemented, the hybrid ensemble approach outperformed & expressed an efficiency of about 96% for 2-class problem and about 78% for 7-class problem. The results when compared with the all the standalone classifiers were significantly better for both 2-class and 7-class problems.

(Sarwar et al., 2016) have developed a novel benchmark database of digitized and calibrated, cervical cells obtained from slides of Papanicolaou smear test. This database can serve as a potential tool for designing, developing, training, testing and validating various artificial intelligence based systems for prognosis of cervical cancer by characterization and classification of Papanicolaou smear images. The database can also be used by other researchers for comparative analysis of working efficiencies of various machine learning and image processing algorithms. Besides developing a rich machine learning database we have also presented a novel artificial intelligence based hybrid ensemble technique for efficient screening of cervical cancer by automated analysis of Papanicolaou smear images. The correct and timely diagnosis of cervical cancer is one of the major problems in the medical world. The contribution of their work is that they have created a rich machine learning database of quantitatively profiled and calibrated cervical cells obtained from Pap-smear test slides. The database so created consists of data of about 200 clinical cases (8091 cervical cells), which have been obtained from multiple health care centres, so as to ensure diversity in data. The Pap-smear slides were processed using a multi-headed digital microscope and images of cervical cells were obtained, which were then passed through various data pre-processing subroutines. After pre-processing the cells were morphologically profiled and scaled to obtain separate quantitative measurements of various features of cytoplasm and nucleus respectively. The
cells in the database were carefully classified in different corresponding classes according to the latest 2001-Bethesda system of classification, by multiple cyto-technicians and histopathologists. In addition to this, we have also pioneered to apply a novel hybrid ensemble system to this database in order to evaluate the effectiveness of both novel database and novel hybrid ensemble technique to screen cervical cancer by categorization of Pap smear data. The paper also presents a comparative analysis of multiple artificial intelligence based classification algorithms for prognosis of cervical cancer. For evaluating the effectiveness and correctness of the digital database prepared in this work, authors implemented this database for training, testing and validating fifteen different artificial intelligence based machine learning algorithms. All the algorithms trained with this database presented commendable efficiency in screening of cervical cancer. For two-class problem all the algorithms trained with the digital database showed the efficiencies in the range of about 93 to 95 % while as in case of multi class problem algorithms expressed the efficiencies in the range of about 69 to 78 %. The results indicate that the novel digital database prepared in this work can be efficiently used for developing new machine learning based techniques for automated screening of cervical cancer. The results also indicate that hybrid ensemble technique is an efficient method for classification of pap-smear images and hence can be effectively used for diagnosis of cervical cancer. Among all the algorithms implemented, the hybrid ensemble approach outperformed and expressed an efficiency of about 98 % for 2-class problem and about 86 % for 7-class problem. The results when compared with the all the standalone classifiers were significantly better for both two-class and multi-class problems.

(Sarwar et al., 2014) developed an AI based classification system that after analysis of certain parameters can predict that whether a person is diabetic or not. Diabetes is inability of body to manage the levels of sugar in the blood. It being one of the most chronic diseases around the world causes around 3.8 million deaths every year. Authors have identified 10 parameters that play an important role in diabetes and prepared a rich database of training data which served as the backbone of the prediction algorithm. Keeping in view this training data authors developed a system that uses the naïve-Bayes classification algorithm to serve the purpose. When the parameters of the test data are fed to the system, it anticipates & classifies the test data into one of the two
categories viz diabetic & not diabetic. The performance accuracy of this artificial intelligent system when compared with the actual medical diagnosis of the subjects was found to be 95%. This system can be used to assist medical programs especially in geographically remote areas where expert human diagnosis not possible with an advantage of minimal expenses and faster results.

2.13 Conclusion

Most approaches to computer-assisted diagnosis have, until the past few years, been based on one of three strategies i.e flow-charts (Schwartz, 1970, Lusted, 1968, Bleich, 1972), statistical pattern-matching (Rosati et al. 1975), or probability theory ((Gorry et al., 1973, Kassirer et al., 1973, Dombal et al., 1972, Weinstein et al., 1980)). All three techniques have been successfully applied to narrow medical domains, but each has serious drawbacks when applied to broad areas of clinical medicine. Flow charts quickly become unmanageably large. Further, they are unable to deal with uncertainty, a key element in most serious diagnostic problems. Probabilistic methods and statistical pattern-matching typically incorporate unwarranted assumptions, such as that the set of diseases under consideration is exhaustive, that the diseases under suspicion are mutually exclusive, or that each clinical finding occurs independently of all others (Szolovits et al., 1978). In theory, these problems could be avoided by establishing a database of probabilities that copes with all possible interactions (Pearl, 1986). But gathering and maintaining such a massive database would be a nearly impossible task. Moreover, all programs that rely solely on statistical techniques ignore causality of disease and thus cannot explain to the physician their reasoning processes nor how they reach their diagnostic conclusions.

Programs using artificial intelligence techniques have several major advantages over programs using more traditional methods. These programs have a greater capacity to quickly narrow the number of diagnostic possibilities, they can effectively use pathophysiologic reasoning, and they can create models of a specific patient's illness. Such models can even capture the complexities created by several disease states that interact and overlap. These programs can also explain in a straightforward manner how particular conclusions have been reached (Clancey, 1983, Swartout, 1983). This latter
ability promises to be of critical importance when expert systems become available for
day-to-day use; unless physicians can assess the validity of a program's conclusions, they
cannot rely on the computer as a consultant. Indeed, a recent survey has shown that a
program's ability to explain its reasoning is considered by clinicians to be more important
than its ability to arrive consistently at the correct diagnosis. An explanatory capability
will also be required by those responsible for correcting errors or modifying programs; as
programs become larger and more complicated, no one will be able to penetrate their
complexity without help from the programs themselves. Causal, quantitative reasoning
also leads to programs that can plan and manage therapy. Past events can be used not only
to predict current findings but to anticipate the possible future evolution of an illness and
the consequences of particular therapeutic actions (Long, 1983, Long et al., 1984). Such
capabilities provide the framework for expanding computer programs beyond their
conventional bounds as diagnostic aids. Progress toward developing practical consulting
programs has been slow despite the rapid increase in our understanding of how experts
solve problems. Experience shows that 5 years is required to incorporate a new cognitive
model into an artificial intelligence program and to test it adequately. Two major factors
have prevented more rapid implementation. First, a large amount of detailed medical
knowledge must be gathered even when one is dealing with a relatively narrow clinical
domain. Second, newer cognitive models are so complex that their implementation
typically poses a major technical challenge. Even if the various problems in
implementation can be solved, further obstacles will impede the development of programs
that are ready for routine clinical use. Decisions must be made concerning acceptable
performance levels (Schwartz, 1970) and extensive debugging and in-hospital testing
must be done to assure that the standards are being met. Fortunately, even before the
advent of fully functional computer programs that can act as sophisticated consultants on
the most difficult medical problems, the fruits of artificial intelligence research can be
applied in less taxing medical settings. Two recent programs, for example, combine the
scoring methods of Internist-I (Miller et al., 1982) and databases that link diseases with
their manifestations to generate lists of hypotheses that may be worthy of detailed
consideration (Miller et al., 1986, Barnett et al., 1987 ). Other artificial intelligence
programs applied in narrow medical domains have also proved to have practical value, in
applications ranging from laboratory data interpretation to protocol-based patient management (Weiss et al. 1981, Ailins et al. 1983, Hickam et al. 1985). Although only a few such programs are currently available, the evidence suggests that the continued development of artificial intelligence techniques will eventually give the computer a major role as an expert consultant to the physician.