CHAPTER 1

INTRODUCTION

1.1 AUTISM

The main aim of research in medical diagnosis is to develop near perfect exact, cost effective, easy-to-use systems, procedures, methods or model for supporting medical practitioners. The advancement in computer technology has encouraged the researchers to develop a model for assisting pediatricians, psychologist, special educators and occupational therapists in better assessment of neurological disorders like Autism and Attention Deficit Hyperactivity Disorder (ADHD) etc., Autism is a developmental disorder characterized by a central triad of deficits (American Psychiatric Association, 2004). These deficits are in social interaction, communication, and imagination (Wing et al 1979). In addition, other features have been associated with the syndrome. These include a restricted repertoire of interests, an obsessive desire for sameness, savant abilities, excellent rote memory, a preoccupation with parts of objects, improved perceptual discrimination, and an impaired ability to form abstractions or generalize knowledge to new situations (Happe, et al 1996). Autism has multiple biological causes (Gillberg, 1996), and a disposition to the disorder is probably inherited (Simonoff et al 1996).

Autism is characterized by impaired social interaction, problems with verbal and nonverbal communication, and unusual, repetitive, or severely limited activities and interests. Experts estimate that three to six
children out of every 1,000 will have autism. Males are four times more likely to have autism than females. There are three distinctive behaviours that characterize autism. Autistic children have difficulties with social interaction, problems with verbal and nonverbal communication, and repetitive behaviours or narrow, obsessive interests. These behaviours can range in impact from mild to disable level (American Academy of Pediatrics, 2001). A child with autism may appear to be normal and then withdraw and become indifferent to social engagement (Dosreis et al 2006).

Children with autism may fail to respond to their name and often avoid eye contact with other people. They have difficulty in interpreting what others are thinking or feeling because they can’t understand social cues, such as tone of voice or facial expressions, and don’t watch other people face for the clues of appropriate behaviour. Even they lack empathy.

Many children with autism engage in repetitive movements such as rocking and twirling, or in self-abusive behaviour such as biting or head-banging. Children with autism don’t know how to play interactively with other children. Some speak in a sing-song voice about a narrow range of favorite topics, with little regard for the interests of the person to whom they are speaking. Even many of them have a reduced sensitivity to pain, but are abnormally sensitive to sound, touch, or other sensory stimulation. These unusual reactions may contribute to behavioural symptoms such as a resistance to being cuddled or hugged.

Children with autism appear to have a higher than normal risk for certain co-existing conditions, including fragile X syndrome (which causes mental retardation), tuberous sclerosis (in which tumors grow on the brain), epileptic seizures, Tourette syndrome, learning disabilities, and attention deficit hyperactivity disorder. For reasons that are still unclear, about 20 to
30 percent of children with autism develop epilepsy by the time they reach adulthood (Heidgerken et al 2005).

Autism varies widely in its severity and symptoms and may go unrecognized, especially slightly affected children or when it is masked by more debilitating handicaps. Doctors rely on a core group of behaviours to alert them to the possibility of a diagnosis of autism. These behaviours are:

- impaired ability to make friends with peers
- impaired ability to initiate or sustain a conversation with others
- absence or impairment of imaginative and social play
- stereotyped, repetitive, or unusual use of language
- restricted patterns of interest that are abnormal in intensity or focus
- preoccupation with certain objects or subjects
- inflexible adherence to specific routines or rituals

Doctors will often use a questionnaire or other screening instrument to gather information about a child’s development and behaviour. Some screening instruments rely solely on parent observations; others rely on a combination of parent and doctor’s observations. If screening instruments indicate the possibility of autism, doctors will ask for a more comprehensive evaluation.

Autism is a complex disorder where a comprehensive evaluation requires a multidisciplinary team including a psychologist, neurologist, psychiatrist, speech therapist, and other professionals who diagnose children with autism. The team members will conduct a thorough neurological
assessment and in-depth cognitive and language testing. Since hearing problems can cause behaviours that could be mistaken for autism, children with delayed speech development should also have their hearing tested. After a thorough evaluation, the team usually meets with parents to explain the results of the evaluation and present the diagnosis.

In general, children who have difficulty in understanding the literal meaning of words are considered to have “traditional” speech and language disabilities. Children with autism also include non-spoken communication problems in particular with socialization/empathy (Gustafsson et al 2004). The following categories are considered to be related to autism:

1.1.1 Categories of the pervasive developmental disorders (PDD)

- Autistic Disorder – Severely disordered verbal and non-verbal language, unusual behaviours.
- Asperger’s Syndrome - Relatively good verbal language, with little non-verbal language problems; restricted range of interests and relatedness.
- PDD-NOS - Non-verbal language problems not meeting strict criteria for other PDD disorders.

1.1.2 Other Autistic Spectrum Disorders

- Semantic Pragmatic Communication Disorder - Delay and trouble with the use of language (both semantic and pragmatic), but socialization relatively spared.
- Non-Verbal Learning Disabilities - Trouble integrating information in three areas: non-verbal difficulties causing the patient to miss the major gestalt in language, spatial perception problems, and motoric coordination problems.
- High Functioning Autism - For some authors, synonymous with Asperger’s; for others, implies slight autism without retardation.
- Hyperlexia - Most notable for incredible rote reading skills starting at an early age.
- Some aspects of Attention Deficit Hyperactivity Disorder (ADHD) - Impulse and control difficulties in ADHD may lead to trouble showing their empathy.

1.2 ATTENTION DEFICIT HYPERACTIVITY DISORDER (ADHD)

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common neurobehavioural disorders of childhood and can persist through adolescence and into adulthood. A person with ADHD has a chronic level of inattention, impulsive hyperactivity, or both, such that daily functioning is compromised. The symptoms of the disorder will be present at levels that are higher than expected ones for a person's developmental stage, which interferes with the person's ability to function in different settings (e.g., in school and at home). A person with ADHD may faces lots of difficulties in certain part of life such as peer and family relationships, and school or work performance.

criteria of ADHD

The two main criteria used to make a diagnosis of ADHD are attention symptoms and hyperactivity symptoms.
The key feature associated with symptoms of inattention includes:

- failure to give close attention to details and difficulty in sustaining attention in tasks or play
- not listening when spoken to them
- not following the instructions and failure to finish tasks
- difficulty in organizing tasks and activities
- avoiding, disliking or being reluctant to engage in tasks that are required for sustained mental effort
- losing things that are necessary for tasks or activities
- easy distraction

The key features associated with symptoms of hyperactivity (sometimes known as hyperactivity-impulsivity) include:

- fidgeting with hands or feet, squirming in seat
- leaving seat when remaining sitting is expected
- running about or climbing excessively
- difficulty in playing or engaging in leisure activities and often ‘on the go’
- talking excessively and blurting out answers before a question is completed
- interrupting others

1.3 STATE OF THE ART IN SUPERVISED AND UNSUPERVISED TRAINING ALGORITHMS

Connectionism is a set of approaches in the fields of artificial intelligence, cognitive psychology, cognitive science, neuroscience and
philosophy of mind, that models mental or behavioural phenomena as the emergent processes of interconnected networks of simple units. There are many forms of connectionism, but the most common forms use neural network models which are computer models loosely based on neural information processing. A survey of connectionist models of abnormal cognitive development illustrates how deviations in these constraints can lead to the development of abnormal behaviours. Many researchers have focused on models of development in autistic children. Connectionist models have been used to explore a range of developmental disorders, including autism (Cohen, 1994, Gustafsson, 1997), developmental dyslexia (Harm et al 1999; Manis et al 1996; Plaut et al 1996; Seidenberg et al 1989, 1996; Zorzi et al 1998), Specific Language Impairment (Hoeffner et al 1993), the development of morphology in a damaged language system (Marchman, 1993), and language processing in Williams syndrome (Thomas et al 2003). Models of developmental disorders have assumed that one or more of these elements are typical in the disordered system. Cohen (1994) focused on the pattern of improved perceptual discrimination and impaired generalization found in the syndrome. He concluded that evidence from neuropathological investigations of the brains of affected individuals was suggestive of abnormal wiring patterns in various brain regions, perhaps caused by deficits in neuronal migration during fetal development (Piven et al 1990), by curtailment of normal neuronal growth (Bauman, 1991), and/or by aberrant development (Courchesne et al 1997). Gustafsson (1997) followed a similar line of thought in proposing that the capacity of autistic individuals to form representations of sensory experience may be impaired. He suggested that their cognitive impairment arose in the development of cortical feature maps. These maps are thought to develop in systems using unsupervised learning, where the aim is for a set of processing units to self-organize such that they form a concise representation of the information they receive. This concise representation may then be used to drive higher-order processes. In Gustafsson’s view, the
formation of such feature maps is damaged in autism due to too much of lateral inhibition in this layer of units.

Even though he did not run any computer simulations, Oliver et al (2000) have examined the ways in which just such a process of feature map formation can be disrupted by changes in the initial properties of a self-organizing connectionist network. Both Cohen and Gustafsson’s models tackle associated perceptual impairments of the disorder. But the above models never focus the central triad of features that is generally considered to characterize the syndrome – deficits in socialization, communication, and imagination.

Many researchers have offered insights into aspects of autism based on related neural network approaches and reported in review articles such as those by Rumsey (1992), Happe et al (1996), Cohen, (1994) and Gustafsson, (1997). In the 1980’s, powerful methods have been discovered for adjusting the connections in neural networks (e.g. Rumelhart et al (1986); Ackley et al 1985). Matthew Belmonte et al (1997) developed a two-layer and three-layer feed-forward artificial neural networks which were trained to predict behavioural performance from single trial electroencephalography (EEG) in autistic and normal subjects in a task involving response to rare stimuli and shifting of attention between vision and audition. Eye blink artifacts were removed from the data using a frequency-domain filter. Performances of the networks on separate test sets varied across subjects but were usually at least by 80%. The networks usually converged faster and attained a somewhat greater level of performance when input was presented in the frequency domain instead of in the time domain.

Batshaw et al (2001) stated that the findings suggest that autism involves the abnormal development of a distributed neural network, involving a number of regions of brain. Many researchers consider autism as a disorder

Cohen has proposed a neural network model of autism in which too many connections are established during development, perhaps owing to inadequate pruning of irrelevant connections. Thus, while too few neurons may not be able to solve a problem to begin with, having too many, may impair generalization and produce rigidity and resistance to change. Among many other predictions, Cohen’s model explains savant skills. According to Gustafsson’s model of the autistic brain, inhibitory lateral feedback synaptic connection strengths of Kohonen-type neural networks are excessive. This results in networks that attend to a restricted set of features in a stimulus configuration or too much discrimination.

Paplinski et al (2002) has modelled autism using self organizing maps. Since autism is a developmental disorder, attention shift impairment and strong familiarity preference are considered to be prime deficiencies. He modelled these two characteristics of autistic behaviour using Kohonen’s Self-Organizing Maps (KSOM). Autism is a developmental disorder in which attention shifting is known to be restricted.

Using an artificial neural network model of learning, Andrew (2003) explained the learning in narrow fields. He focused on attention shifting between different sources of stimuli which is restricted by familiarity preference. His model is based on modified Self-Organizing Maps (SOM) supported by the attention shift mechanism. The novelty seeking and the attention shifting restricted by familiarity preference learning modes are investigated for stimuli of low and high dimensionality, which requires
different techniques to visualize feature maps. The distance between a stimulus and a weight vector can now be simply measured by the post-synaptic activities. Gustafsson et al (2004) used their neural network model to test assumptions about attention shift impairments and familiarity preference in autism.

Casanova et al (2003) have reported that the minicolumns in the brains of autistic subjects are narrower, with less peripheral neurophil space. This leads to a neural network model in which autism could result from an imbalance between the process of excitation and inhibition toward too little inhibition. All the three neural network models were conceptually similar which focused on a neurodevelopment impairment in the neural networks that leads to the concept that connectivity constraints are responsible for the attention to the restricted range of details. They differ only in the types of constraints that produce these problems. In these neural network models, the brain in autism becomes impaired during early neurodevelopment. The connectivity of the networks formed, at least in part, in an unusual and aberrant fashion has been suggested by a variety of studies related to patients with high-functioning autism and Asperger syndrome that raise the question of abnormal variability and scatter of functional maps in autism (Muller et al 2001).

Gustafsson et al (2004) addressed about self-organizing maps (SOMs) which offer insights into the development of cortical feature maps. Sabina munteanu (2005) has presented a hybrid model for diagnosing multiple disorders. The first level of the hybrid model was implemented for hypothesis selection and then an original association based reasoning scheme was used, which measures the distance between the observations and the prototypical models of diseases using fuzzy decision functions.

Ira et al (1993) proposed a non linear pattern recognition system in assisting the classification of autism. Here, Neural network technology was compared with simultaneous and stepwise linear discriminant analysis in terms of their ability to classify and predict persons (n=138) as having autism or mental retardation. The neural network methodology was superior in classifying groups and in generalizing to new cases that were not part of the training sample.

Tung-Kuang Wu et al (2006) adopted two well-known artificial intelligence techniques (Artificial Neural Network and Support Vector Machine), which have been applied successfully to the Learning Disabilities identification and diagnosis problem. Many researchers have carried out diagnosis procedure to analyze autistic patients EEG behaviour. In a related study, You et al (2007) used EEG data to compare patterns of speech and non-speech sound discrimination between non-autistic and autistic teenagers.
1.4 STATE OF THE ART IN MACHINE LEARNING ALGORITHMS

Many machine learning researchers view the task of inductive generalization as beginning after the data is collected, assuming that the useful features have been identified and that representative data has been collected. This led the researchers to focus on algorithm development and applying machine learning algorithms. Probabilistic neural network (PNN) is a kind of supervised neural network, proposed by Specht (1990) as an alternative to back-propagation neural network.

The key advantages of PNN are that, training requires only a single pass, and decision surfaces are guaranteed to approach the Bayes-optimal decision boundaries, as the number of training samples grow. Furthermore, shape of the decision surface can be made as complex as necessary, or as simple as desired, by choosing an appropriate value of the smoothing parameter; erroneous samples can be tolerated, and sparse samples are adequate for network performance. Schizas et al (1992) have used Neural networks, genetic algorithms, and the K-means clustering algorithm for the classification of quantitative electromyographic data.

The comparative study of these methods provides guidance in choosing the most appropriate method for the classification of electromyographic data. The results that have been independently obtained by the above methods were presented and the relative advantages and disadvantages were discussed. It also suggested that the problem of medical diagnosis can be handled better by combining more than one method. In the machine learning literature, there have been several recent survey articles on the application of classification algorithms (Kodratoff, 1994, Langley et al 1995, Rudstrom, 1995, Widrow et al 1994). Much of this previous works have focused on general themes which arise across many application domains, rather than on specific procedures for developing classification models.
One of the most difficult aspects of applying machine learning is defining the features that will allow a classifier to make the necessary discriminations. Solaiman et al (1994) classified the remotely sensed data using several classifiers and neural networks. The study was conducted using a test scene containing mainly agricultural areas. The main result obtained is that the application of topological map based neural networks to classify the intensity vectors issued from agricultural classes are more suited than other neural network methods, especially when the multilayer perceptron (MLP) usually employed. Obtained results are very close to those of the maximum likelihood classifier (MLC).

Drouhard et al (1995) proposed an automatic handwritten signature verification system that will eliminate random and simple forgeries rapidly. The directional probability density function (PDF) was used as a global shape factor, and its discriminatory power was enhanced by reducing its cardinality. The choice of the best pretreatment was made by means of a ‘K’ nearest neighbour classifier. This study has shown that the cardinality of the PDF can be reduced by a factor of ten while doubling its discriminatory power.

The back propagation model was retained to build the neural network classifier. An experimental protocol was used to find the best configuration of the back propagation network classifier whose performance was compared on the same database and with the same decision rule (without rejection criteria), to those of the KNN and threshold classifiers. This comparison shows that the BPN classifier is clearly better than the T classifier, and compares favorably with the KNN classifier.

Lynn Waterhouse et al (1996) compared the four systems for the diagnosis of autism with two empirically derived taxa of autism, and with three social subgroups of autism (Aloof, Passive, and Active-but-Odd) in 194 preschool children with salient social impairment. There were significant
behaviour and IQ differences between autistic and other-PDD groups for all four diagnostic systems, and a significant association was found (a) for Taxon B, diagnoses of autism, and the Aloof subgroup, and (b) for Taxon A, other-PDD, and the Active-but-Odd subgroup. Findings offer support for two major overlapping continua within idiopathic Pervasive Developmental Disorder.

Lorenz et al (1997) selected five trainable Neuro-Fuzzy classification algorithms in order to investigate their ability to differentiate areas of malign tissue in ultrasonic prostate images. The algorithms were compared with results from two commonly used classifiers, the K-nearest neighbor (KNN) classifier and the Bayes classifier. The best Neuro-Fuzzy classification system, which is based on a mountain clustering algorithm published by Yager et al (1992) and refined by Chiu (1996) reached recognition rates above 86 % in comparison to the Bayes classifier (79 %) and the KNN classifier (78%). Finally results suggest that Neuro-Fuzzy classification algorithms have the potential to significantly improve common classification methods for the use in ultrasonic tissue characterization.

Leekam et al (2002) proposed an interviewer-based schedule called Diagnostic Interview for Social and Communication Disorders (DISCO) which has been used with parents and carers where the authors compared two algorithms based on the ninth revision of the schedule (DISCO 9). The algorithm for International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) childhood autism was comprised of 91 individuals having operationally defined items covering the behaviour outlined in the ICD-10 research criteria.

The algorithm for the autistic spectrum disorder, as defined by Wing et al (1979), was based on 5 DISCO items that represented overarching categories of behaviour crucial for the diagnosis of autistic disorders. Authors
examined the implications for clinical diagnosis of these two different approaches. Parents of 36 children with clinical diagnoses of autistic disorder, 17 children with learning disability and 14 children with language disorders were interviewed by the two interviewers. Algorithm diagnoses were applied to interview items in order to analyze the relationship between clinical and algorithm diagnoses and the inter-rater reliability between interviewers.

Clinical diagnosis was significantly related to the diagnostic outputs for both algorithms. Inter-rater reliability was also high for both algorithms. The ICD childhood disorder algorithm produced more discrepant diagnoses than the Wing and Gould autistic spectrum algorithm. Analysis of the ICD-10 algorithm and combination of other items, helped to explain the reason for these discrepancies. The results indicate that the DISCO is a reliable instrument for diagnosis when sources of information are used from the whole interview. It is particularly effective for diagnosing disorders of the broader autistic spectrum.

Amendolia et al (2003) investigated the feasibility of two typical techniques of Pattern Recognition in the classification for Thalassemia screening. They were Support Vector Machine (SVM) and the K-Nearest Neighbour (KNN) techniques. He compared SVM and KNN with a Multi-Layer Perceptron (MLP) classifier and proposed a two-classifier system based on SVM. The first layer is used to differentiate between pathological and non-pathological cases while the second layer is used to discriminate between two different pathologies (α-thalassemia carrier against β-thalassemia carrier) from the first output layer (pathological cases). Using the parameters sensitivity (percentage of pathologic cases correctly classified) and specificity (percentage of non-pathologic cases correctly classified), the results obtained with this analysis show that the MLP classifier gives slightly better results
than SVM although the amount of data available is limited. Both techniques enable thalassemia carriers to be discriminated from healthy subjects with 95% specificity, although the sensitivity of MLP is 92% while that of SVM is 83%.

Dong Wang et al (2006) analyzed DNA microarray data for classification of a disease. He demonstrated three kinds of classifiers like support vector machine (SVM) classifiers, K nearest neighbour (KNN) imputation algorithm and classification and regression tree classifiers. The performance of the KNN classifiers based on all measured genes can be significantly deteriorated with larger missing rate (MR) (e.g. $MR > 5\%$). Another classification technique based on radial basis function neural network (RBFNN) produces successful results and outperforms a number of classifiers which are based on the feed forward neural network architecture (Haralambos et al 2006).

Salem et al (2007) designed and implemented a clustering algorithm called (NNCA), Nearest Neighbor Clustering Algorithm, which was proposed in conjunction with a Fast $K$ Nearest Neighbor (FKNN) strategy for further reduction in processing time. The parallel algorithm (PN NCA) has the ability to cluster pixels of retinal images into those belonging to blood vessels and others not belonging to blood vessels in a reasonable time. Mohamed et al (2007) introduced an aid to radiologists, a computer diagnosis system, which could be helpful in diagnosing abnormalities faster than traditional screening program without the drawback attribute to human factors.

The techniques used for feature extraction is based on the invariant features and fractal dimensions of locally processed image. Two statistical classifiers (The minimum distance classifier and the voting K-Nearest
Neighbor classifier) were used and compared through the system to reach a better classification decision. Feyzullah (2009) proposed a comparative study on thyroid disease diagnosis using neural networks. In his study, a comparative thyroid disease diagnosis was realized by using multilayer, probabilistic, and learning vector quantization neural networks. A thyroid disease dataset which was taken from UCI machine learning database has been used.

Katherine Gotham et al (2007) proposed an Autism Diagnostic Observation Schedule (ADOS) which contains Modules of 1–3 item and domain total distributions were reviewed for 1,630 assessments of children aged 14 months to 16 years with an autism spectrum disorder (ASD) or with heterogeneous non-spectrum disorders. Children were divided by language level and age to yield more homogeneous cells. Items were chosen that are best differentiated between diagnoses and were arranged into domains on the basis of multi-factor item-response analysis. The revised algorithm consists of two new domains, Social Affect and Restricted, Repetitive Behaviours (RRB), combined to one score to which thresholds are applied, resulting in generally improved predictive value.

Rizyan Erol et al (2008) have studied Radial Basis Function network for structural classification of thyroid diseases. In that work, experimental results shows the trained RBFNN model outperforms the corresponding MLP neural network model. Ramakrishnan et al (2008) reviewed the PNN, modified PNN and various learning approaches employed to train the PNN and some comparisons of various types of PNN. Experimental results have been carried out to verify the ability of modified PNN in achieving good classification rate over traditional PNN, BPNN and KNN.
1.5 STATE OF THE ART IN HYBRID LEARNING ALGORITHMS

Lin et al (1991) proposed a general neural-network (connectionist) model for fuzzy logic control and decision systems. This connectionist model, in the form of feed forward multilayer net, combines the idea of fuzzy logic controller and neural-network structure and learning abilities into an integrated neural-network-based fuzzy logic control and decision system. A fuzzy logic control decision network is constructed automatically by learning the training examples itself. By combining both unsupervised (self-organized) and supervised learning schemes, the learning speed converges much faster than the original back propagation learning algorithm. The connectionist structure avoids the rule-matching time of the inference engine in the traditional fuzzy logic system. Jude et al (1991) experimented comparing the ID3 symbolic learning algorithm with the perceptron and back propagation neural learning algorithms using five large, real-world data sets. Overall, back propagation performs slightly better than the other two algorithms in terms of classification accuracy on new samples, but takes much longer to train. Experimental results suggest that back propagation can work significantly better on data sets containing numerical data. Back propagation occasionally outperforms the other two systems when given relatively small amounts of training data.

Hudson et al (1994) proposed a hybrid system approach which combines a knowledge-based system, neural network model, and time series analysis into one system for the diagnosis and treatment of heart disease. A system made of Self-Organizing Maps and back propagation has been built as an aiding tool in the analysis of mammograms for the diagnosis of breast cancer (Santos Andre et al 1999). It is a hybrid neural network system, consisting of a self-organizing map followed by back propagation network, to
restrict the number of spatial grey level dependence matrices that need to be computed (Arrowsmith et al 1999).

Brasil et al (2001) proposed a hybrid expert system (HES) to minimize some complexity problems pervasive to the artificial intelligence such as the knowledge elicitation process, known as the bottleneck of expert systems, the model choice for knowledge representation to code human reasoning, the number of neurons in the hidden layer and the topology used in the connectionist approach and the difficulty to obtain the explanation on how the network arrived to a conclusion. Two algorithms applied for developing HES were also suggested where one of them is used to train the fuzzy neural network and the other to obtain explanations on how the fuzzy neural network attained a conclusion.

Stephen Grossberg et al (2006) proposed an Imbalanced Spectrally Timed Adaptive Resonance Theory (iSTART) to explain the cognitive, emotional and motor symptoms of autism. The iSTART model is actually a combination of three models. First component is the Adaptive Resonance Theory (ART), where object recognition occurs. The iSTART model proposes that autistics have their vigilance set at very high levels and causes the problems in cognition, attention and learning abilities. The second component of iSTART is the Cognitive-Emotional-Motor (CogEM) model, which extends the ART model to the learning of associations between external objects and the emotional states which give them value. The iSTART model proposes that individuals with autism experience under-aroused emotional depression that helps to explain symptoms like reduced emotional expression as well as emotional outbursts. iSTART depicts how autistic symptoms may arise from breakdowns in normal brain processes and opens up a wide range of possible new experiments which make it easier for scientists studying normal behaviour to connect their work to autism research.
Recently, a hybrid artificial neural network model has been developed which consist of a self-organizing map followed by a back propagation neural network that can automate between the normal subjects and the subjects with Parkinson’s disease or the spino cerebeller degeneration patients (Yamaguchi Tsuyoshi et al 2006). Cunjie Wang et al (2007) proposed a new fault diagnosis method, which combines negative selection algorithm and conventional classification algorithm. All the available training samples including normal samples and known anomaly are treated as positive samples (self). The real-valued negative selection algorithm is adopted to generate the negative samples (non-self), which distribute among the rest of the detected space. Both the positive and negative samples are used to train a feed forward artificial neural network classifier, whose output neurons are assigned to fault indicators. The simulation result demonstrates that the performance looks satisfactory.

Machine learning could provide invaluable support for automatically inferring diagnostic rules from descriptions of past cases, making the diagnosis process more objective and reliable. Since the problem involves both test and misclassification costs, Bratu et al (2007) proposed a hybrid algorithm which tries to avoid the pitfalls of traditional greedy induction by performing a heuristic search in the space of possible decision trees through evolutionary mechanisms.

Kemal et al (2007) presented a hybrid approach based on feature selection, fuzzy weighted pre-processing and artificial immune recognition system (AIRS) to medical decision support systems. Babita et al (2009) designed a system where a Knowledge-based systems (KBS) and intelligent computing systems have been used in the medical planning, diagnosis and treatment. The KBS consists of rule-based reasoning (RBR), case-based reasoning (CBR) and model-based reasoning (MBR) whereas intelligent computing method (ICM) encompasses genetic algorithm (GA), artificial
neural network (ANN), fuzzy logic (FL) and others. The combination of methods in KBS such as CBR-RBR, CBR-MBR and RBR-CBR-MBR and the combination of methods in ICM is ANN-GA, fuzzy-ANN, fuzzy-GA and fuzzy-ANN-GA. The combination of methods from KBS to ICM is RBR-ANN, CBR-ANN, RBR-CBR-ANN, fuzzy-RBR, fuzzy-CBR and fuzzy-CBR-ANN. Yamaguchi Tsuyaoshi et al (2006) developed a hybrid artificial neural networks consisting of a self-organized maps (SOM) followed by a back propagation neural networks which can classify Parkinson’s disease and the Spino cerebellar degeneration patients.

1.6 STATE OF THE ART IN FUZZY COGNITIVE MAPS

Fuzzy Cognitive Maps (FCM) is better than traditional rule-based reasoning since it uses stronger mathematical analysis. FCM is an efficient modeling method, which is based on human knowledge and experience. Fuzzy cognitive maps accommodate the knowledge–base building property. It can handle with uncertainty and it is constructed by extracted knowledge in the form of fuzzy rules. Researchers have used FCMs’ for many tasks in several domains like disease diagnosis in the medical domain and fault management in distributed network environment (Ndousse, et al 1996) (Taber, 1991). Kim et al (1996) proposed a model based on time-dependent influences, which appropriately describes the progress of multiple disorders, and gives some modifications for applying this model to medical domains. Here the algorithm is used for forecasting the state of each disease on the time horizon and for evaluation of therapy alternatives.

Smith et al (2000) reported their work on enhanced risk assessment in a health care institution using cognitive fuzzy modeling which is a combination of fuzzy cognitive models and fuzzy rule based techniques. The performance of FCM can be enhanced by incorporating the advantages of neural networks.
Papageorgiou et al (2004a, 2007) proposed learning procedure which is a promising approach for exploiting experts’ involvement with their subjective reasoning and at the same time improving the effectiveness of the FCM operation mode and thus it broadens the applicability of FCMs modeling for complex systems. Papageorgiou (2004b) developed an advanced diagnostic method for urinary bladder tumour grading. A novel soft computing modelling methodology based on the augmentation of fuzzy cognitive maps (FCMs) with the unsupervised active Hebbian learning (AHL) algorithm is applied. Here the FCM grading model achieved a classification accuracy of 72.5%, 74.42% and 95.55% for tumours of grades I, II and III, respectively. An advanced computerized method has been developed to support tumour grade diagnosis. The novelty of the method is based on employing the soft computing method of FCMs to represent specialized knowledge on histopathology and on augmenting FCMs’ ability using an unsupervised learning algorithm, the AHL. This method performs with reasonably high accuracy compared to other existing methods and at the same time meets the physicians’ requirements for transparency and explicability.

Fuzzy Cognitive Maps (FCMs) constitute an attractive knowledge-based methodology, combining the robust properties of fuzzy logic and neural networks. FCMs represent causal knowledge as a signed directed graph with feedback and provide an intuitive framework which incorporates the experts’ knowledge. FCMs handle available information and knowledge from an abstract point of view. They develop behavioural model of the system exploiting the experience and knowledge of experts. The construction of FCMs is based mainly on experts who determine the structure of FCM, i.e. concepts and weighted interconnections among concepts. But this methodology may not be a sufficient model of the system because the human factor is not always reliable. Thus the FCM model of the system may requires restructuring which is achieved through adjusting the weights of FCM.
interconnections using specific learning algorithms. Elpiniki et al (2003) compared two unsupervised learning algorithms for training FCMs. The simulation results of training the industrial process system verify the effectiveness, validity and advantageous characteristics of those learning techniques for FCMs.

Elpiniki et al (2003) introduced a novel approach that uses fuzzy cognitive maps (FCMs) as the computational modelling method, which tackles the complexity and allows the analysis and simulation of the clinical radiation procedure. Specifically, this approach is used to determine the success of radiation therapy process based on the soft computing technique of FCMs. Furthermore, a two-level integrated hierarchical structure was proposed to supervise and evaluate the radiotherapy process prior to treatment. The supervisor determines the treatment variables of cancer therapy and the acceptance level of final radiation dose to the target volume. Two clinical case studies were used to test the methodology and to evaluate the simulation results.

Elpiniki et al (2005) proposed a novel hybrid method based on evolutionary computation techniques for training Fuzzy Cognitive Maps. The methodology of developing Fuzzy Cognitive Maps relies on human expert experience and knowledge, but still exhibits weaknesses in utilization of learning methods and algorithmic background. She investigated a coupling of differential evolution algorithm and unsupervised Hebbian learning algorithm, using both the global search capabilities of evolutionary strategies and the effectiveness of the nonlinear Hebbian learning rule. The use of differential evolution algorithm is related to the concept of evolution of a number of individuals from generation to generation and that of nonlinear Hebbian rule to the concept of adaptation to the environment by learning. The hybrid algorithm is introduced, presented and applied successfully in real-world
problems, from chemical industry to medicine. Experimental results suggest that the hybrid strategy is capable to train FCM effectively leading the system to desired states and determining an appropriate weight matrix for each specific problem.

Papageorgiou et al (2008) proposed a new method for characterizing brain tumors which models the human thinking approach and the classification results were compared with other computational intelligent techniques proving the efficiency. The novelty of the method is based on the use of the soft computing method of fuzzy cognitive maps (FCMs) to represent and model experts’ knowledge (experience, expertise, heuristic). The FCM grading model classification ability was enhanced introducing a computational intelligent training technique, the Activation Hebbian Algorithm. The method was validated for clinical material, comprising of 100 cases. FCM grading model achieved a diagnostic output of accuracy of 90.26% (37/41) and 93.22% (55/59) for brain tumors of low-grade and high-grade, respectively. The results of the proposed grading model present reasonably high accuracy, and are comparable with existing algorithms, such as decision trees and fuzzy decision trees which were tested at the same type of initial data. The main advantage of the proposed FCM grading model is the sufficient interpretability and transparency in decision process, which make it a convenient consulting tool in characterizing tumor aggressiveness for every day clinical practice.

FCMs can successfully represent knowledge and experience, introducing concepts for the essential elements and through the use of cause and effect relationships among the concepts. It follows an approach similar to human reasoning and human decision-making process, considering them a valuable modeling and simulation methodology. Medical Decision Systems are complex systems, consisting of irrelevant and relevant subsystems and
elements, taking into consideration many factors that may be complementary, contradictory, and competitive. All the factors influence each other and determine the overall diagnosis with a different degree. Chrysostomos et al. (2008) developed FCMs model “Medical Decision Support Systems” and FCM structures for medical disciplines like speech and language pathology and obstetrics for differential diagnosis of specific language impairment.

1.7 SUMMARY

All the previous works of supervised and unsupervised algorithm in medical decision making were focused on psychological aspects of autism like speech discrimination, EEG behaviour, sleeping disorders, language disabilities and attention shift impairments. But they failed to characterize the deficits in socialization, communication and imagination features which are considered as central triad of autism. Other researchers have applied the algorithm for identification and classification of many diseases but not for identification of autistic disorder using neural networks.

Even the researchers have developed a hybrid model of combinational neural network algorithms to classify Parkinson’s disease and speech disorders. Many others used pattern recognition system and statistical techniques in classification of autism. Even Fuzzy cognitive maps have been used only for diagnosing diseases, language impairments, grading urinary bladder tumor, simulation of clinical radiation and for identifying multiple disorders using time-dependent models. But neuro fuzzy hybridization and fuzzy cognitive maps proposed in this thesis work are new approaches in identifying autistic disorder.
1.8 OVERVIEW OF THE THESIS

Chapter one has reviewed the literature in supervised, unsupervised, machine learning and hybrid learning algorithms.

Chapter two introduces and discusses about neuro fuzzy system, fuzzy cognitive maps, supervised and unsupervised learning algorithms.

Chapter three proposes a new model to identify Autism using neuro fuzzy system. Here the performance of Back propagation training algorithm is analyzed. A new approach to calculate the initial weights using Nguyen – Widrow initialization for faster learning of Back Propagation networks is proposed.

Chapter four provides a comparative study between neural networks and K-nearest neighbour algorithm. When fuzzy membership values are given as input, neural networks gives better result than KNN.

Chapter five proposes a new Hybrid model which combines supervised Radial Basis function network with unsupervised Self-Organizing maps for better identification of ADHD. Here the resultant weight of first level of network is given as initial weights to second level of network and the network is trained.

Chapter six classifies autistic datasets using fuzzy cognitive maps trained using non linear Hebbian Learning algorithm.

Chapters seven concludes with a summary of the thesis and its success. Also it discusses the possible future enhancement in identification of Autism and other neurological disorder.