CHAPTER – 2

LITERATURE REVIEW

2.1 INTRODUCTION

Asthma is multi factorial disease having significant genetic and environmental components, but since its pathogenesis is not clear, much of its definition is descriptive, based on the functional consequences of airways inflammation, GINA (Global Initiative for Asthma) reported that in case of Asthma airways become narrowed and inflamed, which affect the many cells and cellular elements of the body. In vulnerable persons, inflammation of airways causes hyper responsiveness that lead to recurrent episodes of wheezing, breathlessness, chest tightness, and coughing, mostly at night or in the early morning [11]. From the different respiratory diseases Asthma is one of them which affect the Indian population. The frequently occurring respiratory diseases such as: acute respiratory infections, Asthma, COPD, TB and lung cancer that result in foremost mortality and morbidity [12].

Globally 300 million people [13] have been affected by the Asthma. Nearly 3% people of most countries are affected by Asthma [14] and WHO and Government of India Collaborative Program (2004-2005) reported that mortality of around 0.25 million per year worldwide due to Asthma [15]. Throughout the globe Asthma is a major public health problem [16]. People of all ages especially children are affected by this disease [17]. Fatal nature of Asthma and its uncontrolled conditions bring into being severe obstacle in daily life [18]. It’s an alarm for citizens in the developed world and they have recently focused on Asthma because of its rapidly increasing
frequency, affecting at least one out of four urban children [19]. Asthma is just one of the many respiratory diseases affecting us in India (Figure 2.1) and more than 15 million people have been suffering from this disease in our country [20]. The most of the diseases such as Acute Respiratory Infections, TB, Asthma, COPD and Lung Cancer cause significant mortality and morbidity [21]. In many situations the correct diagnosis of Asthma and distinguished it from other chronic diseases like TB and COPD is needed [22].

Diagnosis of Asthma beforehand can play a vital role in medical treatment [23]. It is a common knowledge that if a patient suffering with a set of certain symptoms goes to different doctors, he may get different opinions regarding the type or grade of the

Figure 2.1: Graphical distribution of Asthma and COPD burden in India. (Picture is taken from http://arachnoserver.org)
disease [24]. Also, two different persons with similar symptoms going to the same
doctor may be diagnosed differently. This indicates that there is a certain degree of
fuzziness in the thinking process of a doctor [25, 26]. To overcome from these
difficulties, the machine learning application is one of the most appropriate solutions.
Taking this solution in account and designed a decision support system for the
diagnosis of Asthma using machine learning applications fuzzy, ANN (Artificial
Neuro Network) and ANFIS (Artificial Neuro Fuzzy Inference System).

For the construction of expert system knowledge acquisition process is one of the
most difficult tasks, because some time unavailability of expert in relevant field
makes development of such system impossible [27]. Various problems come across in
drawing out knowledge from expert. Although the experts are highly skilled in
solving problems in their domain, they often have difficulty in articulating their
knowledge or sometimes even in understanding their own decision making processes.

Problems with manual knowledge acquisition have directed efforts towards
automation of the process. As a result of these numbers of automated knowledge
acquisition methods have been developed.

Alternative knowledge acquisition approaches based Fuzzy, ANN and ANFS have
recently attracted considerable attention [28]. The advantage of these methods over
other approaches is that it can produce good results in situations where data is
imprecise, noisy and inconsistent.

Evaluation of imperfect data does not require add-on mechanism as these
approaches have their inbuilt ability to deal with uncertainty. The recent popularity of
Fuzzy, ANN and ANFS expert systems, have created the need for automation of not
only the process of extraction of rules but also the process of generating the parameters of the associated sets.

This work presents the application of neural networks, fuzzy logic and combined effort neuro fuzzy to automate knowledge acquisition for diagnosis of Asthma. These techniques are capable of dealing with inexact and imprecise problem domains and have been demonstrated to be useful in the solution of classification problems. It addresses the issue of the application of appropriate evaluation criteria such as rule base accuracy and comprehensibility for new knowledge acquisition techniques. In this work neural network, ANFS and fuzzy rule induction systems has been developed.

2.2 ASTHMA

Asthma is a predisposition to chronic inflammation of the lungs in which the airways bronchi are reversibly narrowed [29]. Asthma is a serious public health problem throughout the world, affecting people of all ages. When uncontrolled, Asthma can place severe limits on daily life, and is sometimes fatal. The National Heart, Lung and Blood Institute defines Asthma as a common chronic disorder of the airways characterized by variable and recurring symptoms, airflow obstruction, bronchial hyper responsiveness (bronchospasm), and an underlying inflammation. An etiologic standpoint, Asthma is a heterogeneous disease. Location of airways & lungs in normal and during Asthma is shown in Figure 2.2.

There are two broad types of Asthma: allergic and idiosyncratic. Allergic Asthma is often associated with a personal and /or family history of allergic disease such as rhinitis urticaria, and eczema [4], with positive wheal-and flares skin reactions
to intradermal injection of extracts of airborne antigens, with increase level of IgE, and therefore has disease that cannot be classified on the basis of defined immunologic mechanisms [31]. Many patients develop a typical system complex on contracting an upper respiratory illness. The initial symptoms may be little more than a common cold, but after several days the patient begins to develop paroxysms of wheezing and dyspnea that can last for days to months. Many patients have a disease that does not fit clearly into either of the preceding categories but instead falls into a mix group with features of each. In general, bronchial Asthma that has its onset in early life tends to have strong allergic components, whereas Asthma that develops late tends to be non-allergic or to have a mixed etiology. During Asthma attacks (exacerbations of Asthma), the smooth muscle cells in the bronchi constrict, and the airways become inflamed and swollen. Breathing becomes difficult, and Asthma causes 4,000 deaths a year in the U.S [32].

Figure 2.2: (A) Location of airways & lungs (B) normal airway (C) airway during Asthma (Picture is taken from http://nhlbi.nih.gov)
2.2.1 Mechanism of Asthma

Asthma is an inflammatory disorder of the airways [33], which involves several inflammatory cells and multiple mediators that result in characteristics path physiological changes [34]. In ways are still not well understood, this pattern of inflammation is strongly associated with airway hyper responsiveness and Asthma symptoms [35].

2.2.1.1 Airway Inflammation in Asthma

The clinical spectrum of Asthma is highly variable, and different cellular patterns have been observed, but the presence of airway inflammation remains a consistent feature [36]. The airway inflammation in Asthma is persistent even tough symptoms are episodic, and the relationship between the Asthma severity and the intensity of inflammation is not clearly established. The inflammation affects all airways including in most patients the upper respiratory tract and nose but its physiological effects are most pronounced in medium sized bronchi [37]. The pattern of inflammation in the airway appears to be similar in all clinical forms of Asthma, whether allergic, non-allergic, or aspirin- induced, and at all ages release mediators that contribute to symptoms. Structural cells of the airways also produce inflammatory mediators, and contribute to the persistent of inflammation in various ways [38].

2.2.1.2 Inflammatory Mediators

Over 100 different mediators are now recognized to be involved in Asthma and mediate the complex inflammatory response in the airway [39].
Airway hyper responsiveness is the characteristics functional abnormality of Asthma, results in airway narrowing in a patient with Asthma in response to a stimulates that would be innocuous in a normal person in turn, this airway narrowing leads to variable airflow limitation and intermittent symptoms. Airway hyper responsiveness is linked to both inflammation and repairs of the airway and is partially reversible with therapy. Its mechanism is incompletely understood [40].

2.2.1.3 Mechanism of Airway Hyper Responsiveness

- **Excessive contraction of airway smooth muscle** may result from increased volume and/or contractility of airway smooth muscle cells. **Uncoupling** of airway contraction as a result of inflammatory changes in the airway may lead to excessive narrowing of the airways and a loss of the maximum plateau of contraction found in normal airway when bronchoconstrictor substances are inhaled [41].
- **Thickening of the airway wall** by edema and structural changes amplifies airway narrowing due to contraction of airway smooth muscle for geometric reasons.
- **Sensory nerves** may be sensitized by inflammation, leading to exaggerated bronchoconstriction in response to sensory stimuli.

2.2.2 Cause of Asthma

Asthma is caused by environmental and genetic factors [42], which can influence how severe Asthma is and how well it responds to medication.

2.2.2.1 Environmental

Many environmental risk factors have been associated with Asthma development and morbidity in children, but a few stands out as well-replicated or that
has a meta-analysis of several studies to support their direct association. Environmental tobacco smoke, especially maternal cigarette smoking, is associated with high risk of Asthma prevalence and Asthma morbidity, wheeze, and respiratory infections. Low air quality, from traffic pollution or high ozone levels, has been repeatedly associated with increased Asthma morbidity and has a suggested association with Asthma development [43].

A relationship between exposure to air pollutants (e.g. from traffic) and childhood Asthma shows that both the occurrence of the disease and exacerbation of childhood Asthma are affected by outdoor air pollutants [44]. Caesarean sections have been associated with Asthma when compared with vaginal birth; a meta-analysis found a 20% increase in Asthma prevalence in children delivered by Caesarean section compared to those who are not. This is due to modified bacterial exposure during Caesarean section compared with vaginal birth, which modifies the immune system [45].

Psychological stress has long been suspected of being an Asthma trigger, but only in recent decades has convincing scientific evidence substantiated this hypothesis. Rather than stress directly causing the Asthma symptoms, it is thought that stress modulates the immune system to increase the magnitude of the airway inflammatory response to allergens and irritants [46].

Viral respiratory infections at an early age, along with siblings and day care exposure, may be protective against Asthma, although there have been controversial results, and this protection may depend on genetic context [47]. Antibiotic used early in life has been linked to development of Asthma in several examples; it is thought
that antibiotics make one susceptible to development of Asthma because they modify
gut flora, and thus the immune system [48].

2.2.2.2 Genetic

Over 100 genes have been associated with Asthma in at least one genetic association
study [49]. Through the end of 2005, 25 genes had been associated with Asthma in six
or more separate populations:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene</th>
<th>Gene</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTM1</td>
<td>CC16</td>
<td>IL13</td>
<td>FCER1B</td>
</tr>
<tr>
<td>IL10</td>
<td>GSTP1</td>
<td>CD14</td>
<td>IL4R</td>
</tr>
<tr>
<td>CTLA-4</td>
<td>STAT6</td>
<td>ADRB2 (β-2 adrenergic)</td>
<td>ADAM33</td>
</tr>
<tr>
<td>SPINK5</td>
<td>NOS1</td>
<td>CCL5</td>
<td>receptor)</td>
</tr>
<tr>
<td>LTC4S</td>
<td>TGB</td>
<td>HLA-DRB1</td>
<td></td>
</tr>
<tr>
<td>LTA</td>
<td>TBXA2R</td>
<td>HLA-DQB1</td>
<td></td>
</tr>
<tr>
<td>GRPA</td>
<td>TGFBI</td>
<td>HLA-DQB1</td>
<td></td>
</tr>
<tr>
<td>NOD1</td>
<td>IL4</td>
<td>TNF</td>
<td></td>
</tr>
</tbody>
</table>

Many of these genes are related to the immune system or to modulating
inflammation. However, even among this list of highly replicated genes associated
with Asthma, the results have not been consistent among all of the populations that
have been tested. This indicates that these genes are not associated with Asthma under
every condition, and that researchers need to do further investigation to figure out the
complex interactions that cause Asthma. One theory is that Asthma is a collection of
several diseases, and that genes might have a role in only subsets of Asthma. For
example, one group of genetic differences (single nucleotide polymorphisms in
17q21) has been associated with Asthma that develops in childhood [50].
2.2.3 **Diagnosis of Asthma**

Asthma is defined simply as reversible airway obstruction. Reversibility occurs either spontaneously or with treatment. The basic measurement is peak flow rates and the following diagnostic criteria are used by the British Thoracic Society [51]:

- $\geq 20\%$ difference on at least three days in a week for at least two weeks;
- $\geq 20\%$ improvement of peak flow following treatment, for example:
  - 10 minutes of inhaled $\beta$-agonist (e.g., salbutamol);
  - six weeks of inhaled corticosteroid (e.g., beclometasone);
  - 14 days of 30 mg prednisolone.
- $\geq 20\%$ decrease in peak flow following exposure to a trigger (e.g., exercise).

In many cases, a physician can diagnose Asthma on the basis of typical findings in a patient's clinical history and examination. Asthma is strongly suspected if a patient suffers from eczema or other allergic conditions-suggesting a general atopic constitution-or has a family history of Asthma. While measurement of airway function is possible for adults, most new cases are diagnosed in children who are unable to perform such tests [23].

In children, the key to Asthma diagnosis is the sound of wheezing or a high-pitched sound upon exhalation. Other clues are recurrent wheezing, breathing difficulty or chest tightness or a history of coughing that is worse at night. The doctor should also know if the child's symptoms are worse with exercise, colds or exposure to certain irritants such as smoke, emotional stress or changes in the weather [14].
Other information important to diagnosis is the age at which symptoms began and how they progressed, the timing and pattern of wheezing, when and how often a child had to visit a clinic or hospital emergency department because of symptoms, whether the child ever took bronchodilator medication for the symptoms and the nature of the response to medication [14]. Before diagnosing someone as having Asthma, alternative possibilities should respond to medication [14].

2.2.3.1 Differential Diagnosis

Before diagnosing someone as having Asthma, alternative possibilities should be considered. A clinician taking a history should check whether the patient is using any known bronchoconstrictors (substances that cause narrowing of the airways, e.g. certain anti-inflammatory agents or beta-blockers). Among elderly patients, the presenting symptom may be fatigue, cough or difficulty breathing, all of which may be erroneously attributed to COPD, congestive heart failure, or simple aging [13]. After a pulmonary function test has been carried out, radiological tests, such as a chest X-ray or CT scan, may be required to exclude the possibility of other lung diseases. Occasionally, a bronchial challenge test may be performed using methacholine or histamine to assess bronchial hyper responsiveness. Chronic obstructive pulmonary disease, is a syndrome which closely resembles Asthma, is correlated with more exposure to cigarette smoke, an older patient, less symptom reversibility after bronchodilator administration (as measured by spirometer), and decreased likelihood of family history of atopy [52]. Most people found to have the syndrome are subsequently found to have Asthma and many diagnosed Asthmatics go on to develop the syndrome. Pulmonary aspiration, whether direct due to dysphagia (swallowing
disorder) or indirect (due to acid reflux), can show similar symptoms to Asthma. However, with aspiration, fevers might also indicate aspiration pneumonia. In some people, Asthma-like symptoms may be triggered by gastro-esophageal reflux disease. A majority of children who are Asthma sufferers have an identifiable allergy trigger. Specifically, in a 2004 study, 71% had positive test results for more than 1 allergen, and 42% had positive test results for more than 3 allergens [53]. The majority of these triggers can often be identified from the history; for instance, patients with hay fever or pollen allergy will have seasonal symptoms, those with allergies to pets may experience an abatement of symptoms when away from home, and those with occupational Asthma may improve during leave from work. Allergy tests can help identify some avoidable symptom triggers. Asthma is categorized by the United States National Heart, Lung, and Blood Institute as falling into one of four categories: intermittent, mild persistent, moderate persistent and severe persistent.

2.2.3.2 Signs and Symptoms of Asthma

Because of the spectrum of severity within Asthma, some people with Asthma only rarely experience symptoms, usually in response to triggers, where as other more severe cases may have marked airflow obstruction at all times [54]. Asthma exists in two states:

The steady-state of chronic Asthma [45],

- The acute state Asthma exacerbation.
- The symptoms are different depending on what state the patient is in.

Common symptoms of Asthma in a steady-state include [46]:

1. Nighttime coughing,
2. Shortness of breath with exertion but no dyspnea at rest,

3. A chronic 'throat-clearing' type cough, and


Severity often correlates to an increase in symptoms. Symptoms can worsen gradually and rather insidiously, up to the point of an acute exacerbation of Asthma. It is a common misconception that all people with Asthma, wheeze, some never wheeze, and their disease may be confused with another chronic obstructive pulmonary disease such as emphysema or chronic bronchitis [47].

An acute exacerbation of Asthma is commonly referred to as an Asthma attack. The cardinal symptoms of an attack are [48]:

1. Shortness of breath (dyspnea),

2. Wheezing and


Steady state is often regarded as the sine qua non of Asthma, some patients present primarily with coughing, and in the late stages of an attack, air motion may be so impaired that no wheezing may be heard. When present the cough may sometimes produce clear sputum. The onset may be sudden, with a sense of constriction in the chest, breathing becomes difficult, and wheezing occurs (primarily upon expiration, but can be in both respiratory phases). It is important to note inspiratory strider without expiratory wheeze however, as an upper airway obstruction may manifest with symptoms similar to an acute exacerbation of Asthma, with strider instead of wheezing, and will remain unresponsive to bronchodilators. Signs of an Asthmatic episode include [49]:

- Wheezing,
• Prolonged expiration,
• A rapid heart rate (tachycardia), and
• Rhonchus lung sounds (audible through a stethoscope).

### 2.2.3.3 Asthma Severity

Based on the severity, Asthma can be subdivided on the level of symptoms, airflow limitation and lung function variability into four categories: Intermittent, Mild Persistent, Moderate Persistent, or Severe Persistent [50]. Classification of Asthma by severity is useful when decisions are being made about management at the initial assessment of a patient. It is important to recognize, however, that Asthma severity involves both the severity of the underlying disease and its responsiveness to treatment. Thus, Asthma can present with severe symptoms and airflow obstruction and be classified as Severe Persistent on initial presentation (Table 2.1) [1], but respond fully to treatment and then be classified as Moderate Persistent Asthma. In addition, severity is not an unvarying feature of individual patients’ Asthma, but may change over months or years.

#### Table 2.1: Classification of Asthma severity with different parameters
(As per GINA Report, 2009)

<table>
<thead>
<tr>
<th>Classification of Asthma severity</th>
<th>Symptom frequency</th>
<th>Nighttime symptoms</th>
<th>Peak expiratory flow rate or FEV1 of predicted</th>
<th>Variability of peak expiratory flow rate or FEV1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Severity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermittent</td>
<td>&lt; once a week</td>
<td>≤ twice per Month</td>
<td>≥ 80% predicted</td>
<td>&lt; 20%</td>
</tr>
<tr>
<td>Mild persistent</td>
<td>&gt; once per week but ≤ once per day</td>
<td>&gt; twice per month</td>
<td>≥ 80% predicted</td>
<td>20–30%</td>
</tr>
<tr>
<td>Moderate persistent</td>
<td>Daily</td>
<td>&gt; once per week</td>
<td>60–80% predicted</td>
<td>&gt; 30%</td>
</tr>
<tr>
<td>Severe persistent</td>
<td>Daily</td>
<td>Frequent</td>
<td>&lt; 60% predicted</td>
<td>&gt; 30%</td>
</tr>
</tbody>
</table>
2.3 **FUZZY LOGIC**

The concept of Fuzzy Logic (FL) has been conceived by Lotfi Zadeh in 1965. Fuzzy logic is a logical system, which is an extension of multivalve logic to deal with the human reasoning that ranges from ‘almost certain’ to ‘very unlikely’. In contrast to classical propositional logic (true/false), the membership value of fuzzy logic variables are not only 0 and 1 but it can be range between 0 and 1. [55]. While using linguistic variables these degrees may be managed by specific functions, as discussed below. For example, let a 100 ml glass contain 40 ml of water. Then we may consider two concepts: Empty and Full. Fuzzy set defines the meaning of both the concepts. Then one might define the glass as being 0.6 empty and 0.4 full. As the concept of emptiness would be subjective and thus would depend on the observer or designer. One might design a set membership function where the glass would be considered full for all values down to 50 ml. It is essential to realize that fuzzy logic uses truth degrees as a mathematical model of the vagueness phenomenon while probability is a mathematical model of randomness. In a probabilistic setting a Scalar variable defines the fullness of the glass, and conditional distribution describes the probability that glass is full to a given a specific fullness level. This model, however, has no sense without accepting occurrence of some event e.g. that after a few minutes, the glass will be half empty. Specific observer that randomly selects label for the glass, and achieve the condition like distribution over deterministic observers, or both. As a result probability and fuzziness are not common; these are simply different concepts. They superficially seem similar because of using the same interval of real numbers (0, 1). Still, since theorems such as De Morgan’s have dual applicability and properties of...
random variables are analogous to properties of binary logic states, one can see where the confusion might arise.

Zadeh's fuzzy inference methods have found their broad use in control systems. Fuzzy designers have taken advantage of the ability of fuzzy sets to express vague linguistic terms, and perform inference using expert-derived, intuitively phrased rules. They have exploited the capacity of fuzzy inference to create systems with a high tolerance and for uncertain or incomplete information [56]. Fuzzy inference has principally been used in medicine in diagnosis and classification engines [57, 58, 59], in control systems [60, 61], and in pattern recognition and image enhancement [62]. Though fuzzy logic has been applied to many fields, from control theory to artificial intelligence, it still remains controversial among most statisticians, who prefer Bayesian logic, and some control engineers, prefer traditional two-valued logic.

2.3.1 Linguistic Variables

While variables in mathematics usually take numerical values, in fuzzy logic applications [63], the non-numeric linguistic variables are often used to facilitate the expression of rules and facts [64]. A linguistic variable such as age may have a value such as young or its antonym old. However, the great utility of linguistic variables is that they can be modified via linguistic hedges applied to primary terms. The linguistic hedges can be associated with certain functions. Fuzzy set theory defines fuzzy operators on fuzzy sets. The problem in applying this is that the appropriate fuzzy operator may not be known. For this reason, fuzzy logic usually uses IF-THEN rules, or constructs that are equivalent, such as fuzzy associative matrices.

Rules are usually expressed in the form:
IF variable IS property THEN action

For example, a simple temperature regulator that uses a fan might look like this:

IF temperature IS very cold THEN stop fan
IF temperature IS cold THEN turn down fan
IF temperature IS normal THEN maintain level
IF temperature IS hot THEN speed up fan

There is no "ELSE" - all of the rules are evaluated, because the temperature might be "cold" and "normal" at the same time to different degrees.

The AND, OR, and NOT operators of Boolean logic exist in fuzzy logic, usually defined as the minimum, maximum, and complement; when they are defined this way, they are called the Zadeh operators. So for the fuzzy variables x and y:

\[
\text{NOT } x = (1 - \text{truth}(x))
\]
\[
x \text{ AND } y = \text{minimum}(\text{truth}(x), \text{truth}(y))
\]
\[
x \text{ OR } y = \text{maximum}(\text{truth}(x), \text{truth}(y))
\]

There are also other operators, more linguistic in nature, called hedges that can be applied. These are generally adverbs such as "very", or "somewhat", which modify the meaning of a set using a mathematical formula.

2.3.2 Mathematical Fuzzy Logic

In mathematical logic, there are several formal systems of "fuzzy logic"; most of them belong among so-called t-norm fuzzy logics.

2.3.3 Propositional Fuzzy Logics

- The most important propositional fuzzy logics are: Monoidal t-norm-based propositional fuzzy logic is an axiomatization of logic where conjunction is defined by a left continuous t-norm, and implication is defined as the residuum
of the t-norm. Its models correspond to MTL-algebras that are prelinear commutative bounded integral resituated lattices.

- **Basic propositional fuzzy logic** is an extension of MTL logic where conjunction is defined by a continuous t-norm, and implication is also defined as the residuum of the t-norm. Its models correspond to BL-algebras.

- **Łukasiewicz fuzzy logic** is the extension of basic fuzzy logic BL where standard conjunction is the Łukasiewicz t-norm. It has the axioms of basic fuzzy logic plus an axiom of double negation, and its models correspond to MV-algebras.

- **Gödel fuzzy logic** is the extension of basic fuzzy logic BL where conjunction is Gödel t-norm. It has the axioms of BL plus an axiom of idempotence of conjunction, and its models are called G-algebras.

- **Product fuzzy logic** is the extension of basic fuzzy logic BL where conjunction is product t-norm. It has the axioms of BL plus another axiom for cancellation of conjunction, and its models are called product algebras.

- **Fuzzy logic with evaluated syntax** (sometimes also called Pavelka's logic), denoted by EVŁ, is a further generalization of mathematical fuzzy logic. While the above kinds of fuzzy logic have traditional syntax and many-value semantics, EVŁ is also evaluated the syntax. This means that each formula has an evaluation. Axiomatization of EVŁ stems from Łukasiewicz fuzzy logic. A generalization of classical Gödel completeness theorem is provable in EVŁ [64].
2.3.4 **Predicate Fuzzy Logics**

These extend the above-mentioned fuzzy logics by adding universal and existential quantifiers in a manner similar to the way that predicate logic is created from propositional logic. The semantics of the universal (resp. existential) quantifier in t-norm fuzzy logics is the infimum (resp. supremum) of the truth degrees of the instances of the quantified sub formula.

2.3.5 **Higher-Order Fuzzy Logics**

These logics, called fuzzy type theories, extend predicate fuzzy logics to be able to quantify also predicates and higher order objects.

2.3.5.1 **The Fuzzy Hypercube**

In 1992 fuzzy set is geometrically interpreted as mid- point in a hypercube by Kosko [65]. In 1998, unit hypercube has been used to represent concomitant mechanisms in stroke by Helgason and Jobe [66]. In a given set (Figure 2.3) $X = \{x_1, \ldots, x_n\}$,

A fuzzy subset is just a mapping and grade of membership has been expressed by the value of $\mu(x)$. $\mu : X \rightarrow I = [0, 1]$, the element $x \in X$ to the fuzzy subset $\mu$.

![Figure 2.3: Two dimensional hypercube I2 with the 4 non fuzzy subsets (0, 0), (0, 1), (1, 1), and the fuzzy set (0.5, 0.5).](image)
For example, let $X$ be the set of persons of some population and let the fuzzy set $\mu$ be
defined as healthy subjects. If John is a member of the population (the set $X$), then, $\mu$
(John) gives the grade of healthiness of John, or the grade of membership of John to
the set of healthy subjects. If $\lambda$ is the fuzzy set that describes the grade of depression,
then $\lambda$ (Mary) is the degree of depression of Mary.

Thus, the set of all fuzzy subsets (of $X$) is precisely the unit hypercube $I^n = [0, 1]^n$, as
any fuzzy subset $\mu$ determines a point $P \in I^n$ given by

$$ P = (\mu(x_1), \ldots, \mu(x_n)) $$

Reciprocally, any point $A = (a_1, \ldots, a_n) \in I^n$ generates a fuzzy subset $\mu$ defined
by $\mu(x_i) = a_i, i = 1, \ldots, n$. Non fuzzy or crisp subsets of $x$ are given by mappings $\mu
x : X \to \{0, 1\}$, and are located at the $2n$ corners of the $n$-dimensional unit hypercube
$I^n$. For graphical representations of the two-dimensional and three-dimensional
hypercube [64], Given,

$$ p = (p_1, p_2, \ldots, p_n), \quad q = (q_1, q_2, \ldots, q_n) \in I^n $$

both are not equal to empty $\emptyset = (0, 0, \ldots, 0)$, we define the difference between $p$
and $q$ as

$$ d(p, q) = \frac{\sum_{i=1}^{n} |p_i - q_i|}{\sum_{i=1}^{n} \max\{p_i, q_i\}}. \quad (1) $$

Of course $d(\emptyset, \emptyset) = 0$. We know that $d$ is indeed a metric [67]. Hyper-cubical
calculus has been described in [68], while some biomedical applications of the fuzzy
unit hypercube are given in [66, 69]. Recently, the fuzzy hypercube has been utilized
to study differences between polynucleotide [69] and to compare genomes.
2.3.6 Time Line of Fuzzy Logic Publication

The concept of fuzzy logic has been given by Lofti Zedah in 1965 after his application of fuzzy set theory. Fuzzy logic has been applied in various fields for designing of artificial intelligence system. Generally, statisticians and control engineers use this concept but some of them prefer Bayesian logic [70]. By the help of Nicholas Sheble the concept of fuzzy logic comes in to the life [71]. Prof. R Russell Rhineheart, Head, Chemical Engineering School, Oklahoma State University, first declared that the fuzzy logic term is originally nonsensical [72]. The fuzzy logic concept is simple but the jargon obscure that because fuzzy logic concept is totally depends on logic. In fact the fuzzy is absurdly simple and it provides both command line function and graphical user interface. Many researchers and also many chemical industries used the fuzzy logic concepts [70]. One of the most important uses of fuzzy logic is analysis of pollutants and guide line for water quality improvement [73].

Fuzzy logic is type of computer software that recognizes human emotions from conversion analysis [74]. Many researchers have developed the computer softwares with help of fuzzy logic, which are capable to recognize the human emotions from programmed voice analysis. This software is also useful for the programming of robot for the performance of different activities in the world. Lot of problems related to the biological fields also been designed and analyzed by the help of fuzzy logic.

Description of fuzzy logic applications in different years-

1965: In 1965, Lotfi A Zadeh of the University of California at Berkeleyly published fuzzy sets. In 1965 the fuzzy logic concepts are started from fuzzy set theory.

1990: In Nov 1990 Zadeh delineate the novel ideas for analysis of complex systems and decision processes [75]. The fuzzy logic concept mostly used in the field of
decision making process and analysis of intricate system during 1990. Total 03 [76] research papers have been published on different topics. Some of them related to data mining techniques based on fuzzy set model.

1992: Main emphasis in the field of design of fuzzy logic controllers and design of machines [77]. This year entirely 20 (Pubmed) research papers have been published.

1993: Fuzzy logic concepts have been used in the area of artificial intelligence this year. The concept of artificial intelligence has been applied to construct various types of systems, robots and machines etc. Nearly 30 research papers [76] have been published on different topics like fuzzy logic and neural networks [61].

1994: In this year fuzzy logic helps to develop different expert systems (software developments). Total 46 publications [76] came from different topics like optical implementation of fuzzy logic controllers [78].

1995: Work related to audio command recognition and prediction of protein structure classes have been introduced and applied [79]. Total 73 research papers [76] have been published.

1996: Concept of fuzzy optimization techniques in the area of drug therapy being used. Total 67 research papers [76] have been published on different topics like algorithms for optimizing drug therapy.

1997: This was the year of inventions, related to data processing with the help of fuzzy logic operations. Total 95 research papers [76] have been published.

1998: In 1998 the fuzzy logic used the controllers, design of fuzzy logic controllers, frameworks of implementing etc. Total 84 [76] research papers are published in different topics like steady state error of a system with fuzzy controller.
1999: Fuzzy logic activity concerned with the design, development and testing of model of a spacecraft control have been initiated this year [80]. Total 90 research papers [74, 76] have been published.

2000: In 2000 total 96 research papers have been published on different topics such as controlled genetic algorithm using fuzzy logic, belief functions for job-shop scheduling, design and stability analysis of single-input fuzzy logic controller etc. [81].

2001: Total 151 research papers have been published on self-learning fuzzy discrete system [25, 82].

2002: In pubmed total 119 results have been found. Researches conceded onward on designing of decision support system in the medical field [83].

2003: In 2003 total 141 research papers are published in different topics. Research work on fuzzy logic with support vector machine has been done this year [60, 84].

2004: Total 182 results have been found in the form published research articles. Fuzzy vector median based researches have been started this year [85, 86].

2005: Total 194 results are found as research articles. Researchers have been carried out on identification of uncertain nonlinear systems for robust fuzzy control and brain segmentation with competitive level sets and fuzzy control etc [87].

2006: Research articles related to the adaptive fuzzy associations are coming; and total 277 research papers have been published.

2007: Total 253 research papers have been published on different topic such as data mining techniques and fuzzy neural based systems.

2008: A neuro-fuzzy approach has been applied by the most of the researchers this year and in totality 290 research papers came out as result.
2009: In 2009 major research contributions in fuzzy logic came from the clinical area and total 312 research papers have been published on different topics.

2010: Total 306 research papers have been published on different topics.

2011: During this year different articles related to the diverse field have been published. Nearly 146 research article found till now.

2.3.7 Fuzzy Logic Article in Pubmed

Pubmed source have been used to identify the number of publications related to fuzzy in the area of Bioinformatics, Life Sciences and all area. The total number of articles per year from 2000 to 2010 appears in Table 2.2 indicates a comparison in the number of publications per year indexed in PUBMED [88] based on fuzzy logic

Table 2.2: Number of papers published by Pubmed using fuzzy logic

<table>
<thead>
<tr>
<th>Year</th>
<th>Life Sciences</th>
<th>Bio-Informatics</th>
<th>All Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>5</td>
<td>2</td>
<td>96</td>
</tr>
<tr>
<td>2001</td>
<td>7</td>
<td>2</td>
<td>151</td>
</tr>
<tr>
<td>2002</td>
<td>11</td>
<td>5</td>
<td>119</td>
</tr>
<tr>
<td>2003</td>
<td>16</td>
<td>14</td>
<td>141</td>
</tr>
<tr>
<td>2004</td>
<td>18</td>
<td>12</td>
<td>182</td>
</tr>
<tr>
<td>2005</td>
<td>10</td>
<td>13</td>
<td>194</td>
</tr>
<tr>
<td>2006</td>
<td>27</td>
<td>21</td>
<td>277</td>
</tr>
<tr>
<td>2007</td>
<td>28</td>
<td>16</td>
<td>253</td>
</tr>
<tr>
<td>2008</td>
<td>31</td>
<td>16</td>
<td>290</td>
</tr>
<tr>
<td>2009</td>
<td>32</td>
<td>25</td>
<td>312</td>
</tr>
<tr>
<td>2010</td>
<td>33</td>
<td>28</td>
<td>306</td>
</tr>
</tbody>
</table>
2.3.8 **Applications of Fuzzy Logic in Bioinformatics**

Bioinformatics combines the multi disciplinary area such as computer science, biology, physical and chemical principles, designing of tools utilized for the analysis and modeling of large biological data sets, chronic diseases management, learning of molecular computing and cloning etc.[89]. The field of bioinformatics is intensifying for research and development of new technology [90]. Now fuzzy inference technologies are repeatedly applied in bioinformatics. For example, increase the suppleness of protein motifs and learn about the distinction among polynucleotide, utilizing the fuzzy adaptive resonance theory for the analysis of experimental expression data, applying the dynamic programming algorithm for the alignment of the sequences based on fuzzy recast, fuzzy k-nearest neighbors algorithm used to identify the proteins sub-cellular locations from their dipeptide composition, applying fuzzy c-means and partitioning method for characteristic cluster relationship values of genes, analysis of gene appearance data, functional and ancestral relationships between amino acids with the help of fuzzy alignment method, fuzzy classification rules generated by neural network architecture for the analysis of affairs between genes and decipher of a genetic set-up to process micro-array images, use of fuzzy vector filtering framework in the classification of amino acid sequences in to different super families etc.
2.4 **ARTIFICIAL NEURAL NETWORK**

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information [91]. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems [92]. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons [93]. This is true of ANNs as well. Another aspect of the artificial neural networks is that there are different architectures, which consequently requires different types of algorithms, but despite to be an apparently complex system, a neural network is relatively simple [94]. An Artificial Neural Network is an adaptive, most often nonlinear system that learns to perform a function (an input/output map) from data [95]. Adaptive means that the system parameters are changed during operation, normally called the training phase [96]. After the training phase the Artificial Neural Network parameters are fixed and the system is deployed to solve the problem at hand (the testing phase) [97]. The Artificial Neural Network is built with a systematic step-by-step procedure to optimize a performance criterion or to follow some implicit internal constraint, which is commonly referred to as the learning rule [98, 99,100]. The input/output training data are fundamental in neural network technology, because they convey the necessary information to "discover" the optimal operating point [101]. The nonlinear nature of the neural network processing elements (PEs) provides the system with lots of flexibility to achieve practically any desired input/output map...
[102], i.e., some Artificial Neural Networks are universal mappers. There is a style in neural computation that is worth describing (Figure 2.4).

An input is presented to the neural network and a corresponding desired or target response set at the output (when this is the case the training is called supervised). An error is composed from the difference between the desired response and the system output [103]. This error information is fed back to the system and adjusts the system parameters in a systematic fashion (the learning rule). The process is repeated until the performance is acceptable [104, 105]. It is clear from this description that the performance hinges heavily on the data [106]. If one does not have data that cover a significant portion of the operating conditions or if they are noisy, then neural network technology is probably not the right solution [107]. On the other hand, if there is plenty of data and the problem is poorly understood to derive an approximate model then neural network technology is a good choice [108]. In artificial neural networks, the designer chooses the network topology, the

![Figure 2.4: The Style of Neural Computation](www.rajalakshmi.org/dept/BME/BM2401-NOL.doc)..
performance function, the learning rule, and the criterion to stop the training phase [109], but the system automatically adjusts the parameters. So, it is difficult to bring a priori information into the design, and when the system does not work properly it is also hard to incrementally refine the solution [105]. But ANN-based solutions are extremely efficient in terms of development time and resources, and in many difficult problems artificial neural networks provide performance that is difficult to match with other technologies [99, 110].

2.4.1 The Biological Model

Artificial neural networks emerged after the introduction of simplified neurons by McCulloch and Pitts (1943). These neurons are presented as models of biological neurons and as conceptual components for circuits that could perform computational tasks. The basic model of the neuron is founded upon the functionality of a biological neuron [111]. “Neurons are the basic signaling units of the nervous system” and “each neuron is a discrete cell whose several processes arise from its cell body” [112].

![Biological Neuron](www.rajalakshmi.org/dept/BME/BM2401-NOL.doc)

Figure 2.5: Biological Neuron (www.rajalakshmi.org/dept/BME/BM2401-NOL.doc).
The neuron has four main regions to its structure (Figure 2.5). The cell body, or soma, has two offshoots from it, the dendrites and the axon, which end in presynaptic terminals. The cell body is the heart of the cell, containing the nucleus and maintaining protein synthesis. A neuron may have many dendrites, which branch out in a treelike structure, and receive signals from other neurons. A neuron usually only has one axon which grows out from a part of the cell body called the axon hillock. The axon conducts electric signals generated at the axon hillock down its length. These electric signals are called action potentials. The other end of the axon may split into several branches, which end in a presynaptic terminal. Action potentials are the electric signals that neurons use to convey information to the brain. All these signals are identical. Therefore, the brain determines what type of information is being received based on the path that the signal took. The brain analyzes the patterns of signals being sent and from that information it can interpret the type of information being received. At Nodes of Ranvier, the signal traveling down the axon is regenerated [113]. This ensures that the signal traveling down the axon travels fast and remains constant [114] (i.e. very short propagation delay and no weakening of the signal). The synapse is the area of contact between two neurons. The neurons do not actually physically touch. They are separated by the synaptic cleft, and electric signals are sent through chemical 13 interactions [115]. The neuron sending the signal is called the presynaptic cell and the neuron receiving the signal is called the postsynaptic cell [116]. The signals are generated by the membrane potential, which is based on the differences in concentration of sodium and potassium ions inside and outside the cell membrane. Neurons can be classified by their number of processes (or appendages), or by their function [117].
If they are classified by the number of processes, they fall into three categories.

- **Unipolar** neurons have a single process (dendrites and axon are located on the same stem), and are most common in invertebrates [117].

- In **bipolar** neurons, the dendrite and axon are the neuron's two separate processes. Bipolar neurons have a subclass called pseudo-bipolar neurons, which are used to send sensory information to the spinal cord [118].

- **Multipolar** neurons are most common in mammals. Examples of these neurons are spinal motor neurons, pyramidal cells and Purkinje cells (in the cerebellum) [119].

If classified by function, neurons again fall into three separate categories.

- The first group is sensory, or afferent, neurons, which provide information for perception and motor coordination [120].

- The second group provides information (or instructions) to muscles and glands and is therefore called motor neurons [121].

- The last group, interneuronal, contains all other neurons and has two subclasses [122]. One group called relay or projection interneurons have long axons and connect different parts of the brain. The other group called local interneurons are only used in local circuits.

### 2.4.2 The Mathematical Model

When creating a functional model of the biological neuron, there are three basic components of importance. First, the synapses of the neuron are modeled as weights [111]. The strength of the connection between an input and a neuron is noted
by the value of the weight. Negative weight values reflect inhibitory connections, while positive values designate excitatory connections [123,124,125]. The next two components model the actual activity within the neuron cell. An adder sums up all the inputs modified by their respective weights. This activity is referred to as linear combination [126]. Finally, an activation function controls the amplitude of the output of the neuron. The activation function dampens or bounds the neuron's output, Figures 6 and 7 represent two common activation functions which also happened to be used by the network tested during this research [127]. The first is the logistic or sigmoid function \( f(x) = \frac{1}{1+\exp(-X)} \). The second is the gaussian function \( f(x) = \exp (- 2 X ) \) [128].

The final output of the neuron represents the output of the activation function (Figure 2.6 and 2.7). Large numbers of interconnected artificial neurons have the ability to learn or store knowledge in their synaptic weights [129]. The learning ability of an artificial neural network makes it ideal for trying to identify signals within a noisy data flow [130].
Figure 2.6: Logistic Activation Function.

Figure 2.7: Gaussian Activation Function.

An acceptable range of output is usually between 0 and 1, or -1 and 1.

Mathematically, this process is described in the Figure 2.8.
From this model the interval activity of the neuron can be shown to be:

$$v_k = \sum_{j=1}^{p} W_{kj} \cdot x_j$$

The output of the neuron, $y_k$, would therefore be the outcome of some activation function on the value of $v_k$ [131].

### 2.4.2.1 Activation Functions

As mentioned previously, the activation function acts as a squashing function, such that the output of a neuron in a neural network is between certain values (usually 0 and 1, or -1 and 1). In general, there are three types of activation functions, denoted by $\Phi(.)$ [132].
First, there is the Threshold Function which takes on a value of 0 if the summed input is less than a certain threshold value \( v \), and the value 1 if the summed input is greater than or equal to the threshold value \[133\].

\[
\varphi(v) = \begin{cases} 
1 & \text{if } v \geq 0 \\
0 & \text{if } v < 0
\end{cases}
\]

Secondly, there is the Piecewise-Linear function. This function again can take on the values of 0 or 1, but can also take on values between that depending on the amplification factor in a certain region of linear operation \[134\].

\[
\varphi(v) = \begin{cases} 
1 & v \geq \frac{1}{2} \\
\frac{1}{2} & -\frac{1}{2} < v < \frac{1}{2} \\
0 & v \leq -\frac{1}{2}
\end{cases}
\]

Thirdly, there is the sigmoid function. This function can range between 0 and 1, but it is also sometimes useful to use the -1 to 1 range. An example of the sigmoid function is the hyperbolic tangent function.

\[
\varphi(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - \exp(-v)}{1 + \exp(-v)}
\]

An artificial neural network consists of a pool of simple processing units which communicate by sending signals to each other over a large number of weighted connections \[135\]. A set of major aspects of a parallel distributed model can be distinguished:

- a set of processing units ('neurons,' 'cells');
• a state of activation \( y_k \) for every unit, which equivalent to the output of the unit;

• Connections between the units. Generally each connection is defined by a weight \( w_{jk} \) which determines the effect which the signal of unit \( j \) has on unit \( k \);

• a propagation rule, which determines the effective input \( s_k \) of a unit from its external inputs;

• an activation function \( F_k \), which determines the new level of activation based on the effective input \( s_k(t) \) and the current activation \( y_k(t) \) (i.e., the update);

• an external input (aka bias, offset) \( \theta_k \) for each unit;

• a method for information gathering (the learning rule);

• An environment within which the system must operate, providing input signals and if necessary error signals [136].

2.4.3. Processing Units

Each unit performs a relatively simple job: receive input from neighbors or external sources and use this to compute an output signal which is propagated to other units [137]. Apart from this processing, a second task is the adjustment of the weights. The system is inherently parallel in the sense that many units can carry out their computations at the same time. Within neural systems it is useful to distinguish three types of units [138]:

• input units (indicated by an index \( i \)) which receive data from outside the neural network,
• output units (indicated by an index o) which send data out of the neural network, and

• Hidden units (indicated by an index h), whose input and output signals remain within the neural network.

During operation, units can be updated either synchronously or asynchronously. With synchronous updating, all units update their activation simultaneously; with asynchronous updating, each unit has a (usually fixed) probability of updating its activation at a time t, and usually only one unit will be able to do this at a time. In some cases the latter model has some advantages [139,140].

2.4.4 Neural Network Topologies

This section focuses on the pattern of connections between the units and the propagation of data.

• Feed-forward neural networks, where the data flow from input to output units is strictly feed forward. The data processing can extend over multiple (layers of) units, but no feedback connections are present, that is, connections extending from outputs of units to inputs of units in the same layer or previous layers [141].

• Recurrent neural networks contain feedback connections. Contrary to feed-forward networks, the dynamical properties of the network are important [142, 143]. In some cases, the activation values of the units undergo a relaxation process such that the neural network will evolve to a stable state in which these activations do not change anymore [144]. In other applications, the
change of the activation values of the output neurons is significant, such that the dynamical behavior constitutes the output of the neural network [145]

2.4.5 Training of Artificial Neural Networks

A neural network has to be configured such that the application of a set of inputs produces (either 'direct' or via a relaxation process) the desired set of outputs. Various methods to set the strengths of the connections exist [146]. One way is to set the weights explicitly, using a priori knowledge. Another way is to 'train' the neural network by feeding it teaching patterns and letting it change its weights according to some learning rule [147].

We can categories the learning situations in two distinct sorts. These are:

- Supervised learning or Associative learning in which the network is trained by providing it with input and matching output patterns [148]. These input-output pairs can be provided by an external teacher, or by the system which contains the neural network [149] (self-supervised) (Figure 2.9).

![Figure 2.9: Supervised Learning.](image-url)
• Unsupervised learning or Self-organization in which an (output) unit is trained to respond to clusters of pattern within the input. In this paradigm the system is supposed to discover statistically salient features of the input population [150]. Unlike the supervised learning paradigm, there is no a priori set of categories into which the patterns are to be classified; rather the system must develop its own representation of the input stimuli [151].

• Reinforcement learning this type of learning may be considered as an intermediate form of the above two types of learning. Here the learning machine does some action on the environment and gets a feedback response from the environment [152]. The learning system grades its action good (rewarding) or bad (punishable) based on the environmental response and accordingly adjusts its parameters. Generally, parameter adjustment is continued until an equilibrium state occurs, following which there will be no more changes in its parameters [153].

2.4.6 Modifying Patterns of Connectivity of Neural Networks

Both learning paradigms supervised learning and unsupervised learning result in an adjustment of the weights of the connections between units, according to some modification rule. Virtually all learning rules for models of this type can be considered as a variant of the Hebbian learning rule suggested by Hebb in his classic book Organization of Behavior (1949). The basic idea is that if two units’ j and k are active simultaneously, their interconnection must be strengthened. If j receives input from k, the simplest version of Hebbian learning prescribes to modify the weight $w_{jk}$ with
where $\gamma$ is a positive constant of proportionality representing the learning rate. Another common rule uses not the actual activation of unit $k$ but the difference between the actual and desired activation for adjusting the weights:

$$\Delta w_{jk} = \gamma y_j (d_k - y_k),$$

in which $d_k$ is the desired activation provided by a teacher. This is often called the Widrow-Hoff rule or the delta rule.

### 2.4.7 Historical Sketch of Neural Networks

The first neural network has been presented by Alexander Bain (1818 - 1903) of the United Kingdom in his 1873 book entitled "Mind and Body. But an idea came in to the existence in 1940s that “Natural components of mind-like machines are simple abstractions based on the behavior of biological nerve cells, and such machines can be built by interconnecting such elements” and starts the new era of neural computing. 1938- N. Rashevsky proposed that the brain could be organized around binary logic operations since action potentials could be viewed as binary 1 (true) value.

1943- W. McCulloch & W. Pitts gives the first theory on fundamental neural computing.

1947- McCulloch-Pitts build the neuron model having capability to recognize the spatial patterns invariant of geometric transformations. This year Cybernetics
combines the concepts from biology, psychology, mathematics, and engineering.

1949- D.O. Hebb conceived the concept of “The Organization of Behavior” the first theory that describes the idea of internal representation in the form of cell assembly. Hebb established the learning rule based on the supportive nature of neurons to another in their activities.

1951- M. Minsky built a reinforcement-based network learning system.

1955- IRE Symposium designs the machine which simulates the behavior of human brain.

1958- The first practical Artificial Neural Network (ANN) model came in to existence with the effort of F. Rosenblatt. At the end of 50s, the NN field became undeveloped because of the new AI advances based on serial processing of symbolic expressions.

1960- In this year B. Widrow & M.E. Hoff introduces the “Adaptive Switching Circuits” in which the weights are adjusted so to minimize the mean square error between the actual and desired output.

1961- Widrow and his students developed the “Generalization and Information Storage in Networks of Adeline “Neurons.

1969- M. Minsky & S. Papert gave the concept of relationship between the perceptron’s architecture and what it can learn.

1972- “Correlation Matrix Memories” a mathematical oriented paper of T. Kohonen, proposing a correlation matrix model for associative memory which is trained,
using Hebb’s rule, to learn associations between input and output vectors. In the same year J.A. Anderson gave his ideas on memories aspects, “A Simple Neural Network Generating an Interactive Memory” a physiological oriented paper proposing a “linear associator” model for associative memory, using Hebb’s rule, to learn associations between input and output vectors.

1976- S. Grossberg developed the continuous-time competitive network that forms a basis for the Adaptive Resonance Theory (ART) networks.


1988- “Neurocomputing: Foundations of Research” contains over forty seminal papers in the field of ANN. In the same year DARPA Neural Network Study a comprehensive review of the theory and applications of the Neural Networks has been started. International Neural Network Society came in to existence.

1990- After 90s the ANN established its status in field of research and development.
2.5 **ARTIFICIAL NEURO FUZZY INFERENCE SYSTEM**

In real world our aim could not to acquire the knowledge from various sources, but also to combine different intelligent technologies. The combination of fuzzy logic and neural networks constitutes a powerful means of designing intelligent systems [154]. A hybrid intelligent system is one that combines at least two intelligent technologies. For example: combining a neural network with a fuzzy system results in a hybrid neuro-fuzzy system [155].

Fuzzy logic and neural networks are natural complementary tools in building intelligent systems. While neural networks possess low-level computational structures that perform well when dealing with raw data, fuzzy logic deals with reasoning on a higher level, using linguistic information acquired from domain experts [156]. However, fuzzy systems lack the ability to learn and cannot adjust themselves to a new environment. The merger of a neural network with a fuzzy system into one integrated system therefore offers promising approach to building intelligent systems [157]. Integrated neuro-fuzzy systems can combine the parallel computation and learning abilities of neural networks with the human like knowledge representation and explanation ability of fuzzy systems [158,159,160]. This gives rise to neuro-fuzzy systems, which have recently been investigated by many researchers; see, for instance, [161] and references therein.

Some advantages of ANFIS [163] are:

- Refines fuzzy if-then rules to describe the behaviour of a complex system.
- Does not require prior human expertise
uses membership functions plus

desired dataset to approximate

- Greater choice of membership functions to use.
- Very fast convergence time.

Another advantage of these combined systems is that they often are more precise than pure Fuzzy systems because the learning algorithm may find a better solution. One of the best known of these combinations is the Neuro Fuzzy system called ANFIS [163]. These combined systems, such as ANFIS have been used in many different cases, for instance [164] [165] or [166]. The usage of the ANFIS methods in disease, impairments and disabilities diagnosis has been increasing gradually [167]. Diagnosis of lymph diseases which is a very common and important disease has been conducted with an ANFIS system [168,169,170]. Other studies in this field use the ANFIS method to evaluate fluctuations in Parkinson’s disease [171], to predict respiratory motion in breast cancer patients [172]. These applications give clinicians a valuable tool to explore the importance of different variables and their relations in a disease and could aid treatment selection [173].

2.5.1 Application of ANFIS in Diverse Area

Jang (1992) proposed to use the ANFIS architecture to improve the performance of the fuzzy controllers. The performance of the fuzzy controller relies on two important factors: knowledge acquisition and the availability of human experts [167]. For the first problem, Jang proposed the ANFIS to solve the automatic elicitation of the knowledge in the form of fuzzy if then rules. For the second problem, that is how
the fuzzy controller is constructed without using human experts; a learning method based on a special form of gradient descent (back propagation) has been used. The proposed architecture identified the near optimal membership functions and the other parameters of a controller rule base for achieving a desired input-output mapping. The back propagation type gradient descent method has been applied to propagate the error signals through different time stages to control the plant trajectory. The inverted pendulum system has been employed to show the effectiveness and robustness of the proposed controller.

In 1992, Uchikawa et. al. presented a fuzzy modeling method using fuzzy neural networks, FNNs, with the back propagation algorithms. They proposed three types of NN structures of which the connections weights have particular meanings for getting fuzzy inference rules for tuning membership functions. These structures are categorized into FNNs and these different types FNNs realize three different types of reasoning [174].

Rao and Gupta (1994) described the basic notions of biological and computational neuronal morphologies and the principles and architectures of FNNs. Two possible models of FNN are given. In first one, the fuzzy interface provides an input vector to a multilayered network in response to linguistic statements [175]. Then the NN can be trained to yield desired output. In the second scheme, a multilayered NN drives the fuzzy inference mechanism. It has been pointed out that using FNN approaches having the potential for parallel computation could eliminate the amount of computation required.
In another paper, Uchikawa et. al. (1995) presented a new design method of adaptive fuzzy controller using linguistic rules of fuzzy models of the controlled objects [176]. FNNs identify fuzzy models of nonlinear systems automatically with the back propagation algorithm in this method [177]. Authors also presented a rule-to-rulc mapping method for describing the behavior of fuzzy dynamical systems. Using this methodology, first, the control rules are modified by considering rule-to-rule transitions. After that, designed controller has been implemented with another FNN. The adaptive tuning of the control rules has been done using the fuzzy model of the controlled object by utilizing the derivative value from the fuzzy model. A second order system has been simulated to show the feasibility of the proposed design method [178].

Luigi Benecchi (2006) proposed the ANFIS to predict the presence of prostate cancer. And describe the comparative analysis of the predictive accuracy of neuro-fuzzy system with that obtained by total prostate-specific antigen (tPSA) and percent free PSA (%fPSA) [179].

Mohd Ariffanan Bin Mohd Basri (2008) developed the conventional method in medicine for brain MR images classification and tumor detection in human brain. The use of artificial intelligent techniques, for instance, neural networks, fuzzy logic, neuro fuzzy have shown great potential in this area. Hence, the neuro fuzzy system or ANFIS has been applied for classification and detection purposes.

Ubeyli ED. (2009) in his research article proposed an integrated view of implementing adaptive neuro-fuzzy inference system (ANFIS) for breast cancer detection [180].
In 2010 Akinyokun C O et. al. designs a neuro fuzzy decision support system for the heart failure [181] and in the same year Akgundogdu A et. al. provides the solution for the diagnosis of renal failure disease with the help of ANFIS [182].

Jalali A et. al. (2011) proposed nonlinear model for cardiovascular system regulation by utilizing the neuro fuzzy system [183,184]. In 2011 a novel neuro-fuzzy computing system is proposed where its learning is based on the creation of fuzzy relations by using a new implication method without utilizing any exact mathematical techniques, And implementing the neuro-fuzzy system on memristor crossbar-based analog circuit for creating artificial brain [185].

In 2012, Babazadeh Khameneh N et. al. proposed a method to detect abnormality in red blood cells using cell microscopic images. And for blood sample classification An adaptive network-based fuzzy inference system (ANFIS) is being used [186, 187].

2.5.2 **ANFIS Article in Pubmed**

Pubmed is the reach source of research publications in the diverse area of sciences. The total numbers of articles related to artificial neuro fuzzy inference system have been indexed in Pubmed per year from 2000 to 2012 appeared in Figure 2.10.
2.5.3 Structure of Neuro-Fuzzy System

The structure of neuro fuzzy system is most likely to the neural network. According to the Mamdani [188] fuzzy inference model, in neuro fuzzy system three hidden layers are exist in between input and output layers, which represents the membership function and fuzzy rules. Finally, ANFIS (Adaptive Neuro-Fuzzy Inference System) a neural network that is functionally equal to a Sugeno fuzzy inference model has been proposed by Jang (1993).

The ANFIS architecture is shown in Figure 2.11. The nodes of the same layer have the same functions.
Figure 2.11: Neuro-Fuzzy equivalent system [188].

The ANFIS consists of five layers. The explanations of these layers are below:

**Layer 0:** It consists of plain input variable set.

**Layer 1:** In this layer every node is adaptive. $x$ (or $y$) is the input to node $i$, (or ) is linguistic label, $O_{i1}$ is the membership function of $A_i$. $\{a_i, b_i, c_i\}$ is the parameter set which called premise parameters, values effects the membership function as you can see at the equations below

$$O_{i1}(x) = \mu A_i(x)$$

$$\mu A_i(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i}\right)^2\right]^{b_i}}$$

**Layer 2:** Every node in this layer is a fixed node labeled $\Pi$ which calculates the incoming signals and sends the product out.
Layer 3: Every node in this layer is a fixed node and labeled N. The node calculates the ratio of the rules firing strength to the sum of all rules’ firing strengths. The outputs are called normalized firing strengths.

\[ w_i = \frac{w_i}{w_1 + w_2}, i = 1,2, \]

Layer 4: Every node in this layer is adaptive node with a node function:

\[ w_if_i = w_i (p_0x + q_iy + r_i), i = 1,2. \]

The consequent parameter set is \{p, q, r\} \(w_i\) is a normalized firing strength from the result of layer 3.

Layer 5: The single node in this layer labeled \(\Sigma\) which computes the overall output as the summation of all incoming signals [189].

\[ \sum_i w_if_i = \frac{\sum_i w_if_i}{\sum_i w_i} \]

(Jang, 1993)

The next chapter to follow gives the description of materials and methodology for designing of three inference system (Fuzzy, ANN and ANFIS). This chapter includes the scope of work, background, study area and availability of data from two different locations.