CHAPTER 1

INTRODUCTION

1.1. Research Context and Motivation

The healthcare domain is recognized for its ontological complication and diversity of medical data standards and variable data quality [1], [2], [3], [4]. Making an effective and practically usable discovery in medical data is of ongoing importance in current decades due to the addition of privacy issues in patient data. Computer-aided knowledge discovery methods play an important role in transformation and understanding of concepts related to health and illness [5].

From the twentieth century, many countries have chosen e-health as a prioritized national program [6], [7]. It provides standardized aggregation of patients’ clinical information and health care services by providing instant access to this information for healthcare professionals and patients too.

For the first time in history, the research society is going to get a full set of a person’s medical history from the birth date until the decease date. This anticipated situation forecasts an incredible potential for machine learning and in particular for data mining applications in healthcare. Unequivocal methods, tools and methodology are needed for the application of Data Mining in health
care. Both structured and unstructured data is available for research as a result of progress in the computerization of data in the health care industry. Even though, there are number of algorithms are available to classify, cluster and to find hidden patterns in data, issues related to specific domains in health care is still to be resolved [8], [9].

1.2. Data Mining and Knowledge Discovery

Large volumes of data have been stored in different formats by the initiation of Information Technology in diverse fields of Health Care industry. By applying various methods and algorithms in the core functions of Data Mining (DM), the extract patterns of data can be discovered. Various methods and algorithms are applied in the core function of DM to extract the pattern of data [10]. That is, it can be used to discover the pattern of data.

The emerging trends in information technology have demonstrated its potential in different functional areas such as business, scientific research, social and environmental issues [11]. Rapid computerization of the daily transactions in these areas has resulted in the accumulation of huge amount of raw data with major organizations. It contains a wealth of knowledge about the users, which can assist in the development of marketing strategies, political campaigns, Government policies and product quality control. The repository of this historic data after some levels of pre-processing is often called a Data Warehouse [12], [13] and the process of finding useful patterns and information
from a data warehouse is often known as Knowledge Discovery in Databases or KDD. Data Mining is a particular step in this process, involving the application of specific algorithms for extracting patterns from data [14]. The actual Data Mining task is the automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns. This involves groups of similar data records, unusual records and dependencies [15]. These patterns can then be seen as a kind of summary of the input data, and these techniques find applications in a wide variety of areas like:

- Financial data analysis
- Retail industry
- Telecommunication industry
- Biological data analysis
- Web usage and Content analysis
- Multimedia and Mass communication
- Spatial analysis
- Social survey

1.3. Stages of knowledge discovery

Data mining process involves data collection, data preparation, result interpretation and reporting. To discover and to present the knowledge in the form understandable to the user, different machine learning, statistical and
visualization techniques were discovered [16]. The three technologies which support the recent development in Data Mining area are [17], [18]:

- Massive data collection
- Powerful multiprocessor computers
- Data Mining algorithms

Data Mining process involves clearly defined phases which take inputs from a preceding phase and outputs become the inputs for the next step. Figure 1.1 presents a generic model for the DM process. Depending upon the problem formation, data filtering is used to select the relevant data for the intended analysis. Data cleaning is responsible for handling missing values, smoothing noisy data, identifying or removing outliers, and resolving inconsistencies, such as to compensate for the learning algorithms’ inability to deal with such data irregularities. Data transformation activities include aggregation, normalization and solving syntactic incompatibilities, such as unit conversions or data format synchronization. Data projection translates the input space into an alternative space, of lower dimensionality. During the processing step, learning models/patterns are inferred, by applying the appropriate learning scheme on the pre-processed data. The processing activities are included in an iterative process, during which the most appropriate algorithm and associated parameter values are established. The correct choice of the learning algorithm, given the established goals and data characteristics is essential. There are situations in
which it is required to adapt existing algorithms, or to develop new methods, in order to satisfy all requirements. Subsequently, the output model is built using the results from the model tuning loop, and its expected performance is assessed.

Knowledge presentation employs visualization methods to display the extracted knowledge in an intuitive, accessible and easy to understand manner. Decisions on how to proceed with future iterations are made based on the conclusions reached at this point. DM process modelling represents an active challenge, through their diversity and uniqueness within a certain application. All process models contain activities which can be conceptually grouped, into the three types: pre-processing, processing and post-processing. Several standard process models exist in literature, the most important being: William’s model, Reinartz’ model, CRISP-DM, I-MIN or Redpath’s model [19]. Each model specifies the same process steps and data flow; they differ in the control flow. Essentially, they all try to achieve maximum possible automation.
Figure 1.1: Knowledge representation using Data Mining Techniques
1.4. Medical Data Mining

The health care industry is one of the most information intensive industries [20]. Medical informatics plays an important role with clinical data. Health care is an intensive research field and largest consumes of public health. Human medical data are the most rewarding and difficult of all biological data to mine and analyze [21]. Following are some of the important areas of interests where data mining techniques can be of tremendous use in health care management [22], [23].

1. Data modelling for health care applications

2. Executive Information System for health care

3. Forecasting treatment costs and demand of resources

4. Anticipating patient’s future behaviour from their history

5. Public Health Informatics

6. E-governance structures in health care

7. Health Insurance

The reliance of health care on data is increasing [24]. When the database is large, it is very difficult for the medical researchers, physicians and health care providers to use the stored data more effectively. The medical database usually contains data such as patient records, physician diagnosis and
monitoring information to save lives [25]. The medical decision support system was designed to reduce medical errors and costs, earlier disease detection and to achieve preventive medicine. To improve the decision making in medical problems, the data mining has been used in medial domain. Even biomedical researchers and health care managers find very difficult to detect the association between the risk and outcomes.

1.5. Infertility

Infertility is defined as the inability to conceive naturally after one year of regular unprotected intercourse. Most of the time, infertility is some degree of sub fertility in which 1 in 7 couples need specialist help to get conceive. Sub fertility can be either primary or secondary. Primary sub fertility is a delay for a couple who have had no previous pregnancies; and, secondary sub fertility is a delay for a couple who have conceived previously, although the pregnancy may not have been successful for example, miscarriage, and ectopic pregnancy [26].

The chance to conceive depends on the length of sexual exposure, frequency of coitus and couple’s age. The normal, young aged couples have a 25% chance to conceive after 1 month of unprotected intercourse; 70% of the couple’s conceive by 6 months, and 90% of the couples have a probability to conceive by 1 year. Only 5% of the couples will conceive after one and a half year or two years [27], [28]. Both males and females are equally responsible
for the causes. Most of the infertile couples have one of these three major causes including a male factor, ovulatory dysfunction, or tubal-peritoneal disease [27]. Literature shows that vaginismus and dyspareunia are more common in 20-24 years aged females [29]. The sexual response cycle plays an important role to promote fertility, because it comprises of sequential physical and emotional changes that occur as a person becomes sexually aroused [30].

In normal physiology, the two gonadotropin hormones, Follicle Stimulating Hormone (FSH) and Luteinizing Hormone (LH) are produced in the pituitary and their secretion is controlled by Gonadotropin Releasing Hormone (GnRH) that is released by the hypothalamus. At the start of a new cycle, the hypothalamus begins to release GnRH that acts on the pituitary gland to release FSH and LH. These two hormones stimulate the ovary and cause the follicles to develop. Every month about 30-40 follicles start to grow in response to FSH, but only a single mature egg is released every month. This involves message transmission in the form of hormones from the ovary, the pituitary and the hypothalamus. When the egg is ripe, the mature follicle releases an increasing amount of estrogen that is produced by the granulosa cells lining the follicle. The estrogen produced by the dominant follicle progressively increases in quantity as the egg matures, until a surge of estrogen is released into the blood. The high level of estrogen stimulates the pituitary to release a large amount of LH, thus leading to the LH surge. This LH in turn acts on the mature follicle, causing it to rupture to release the mature egg (ovulation) in the
Males born without testes or vas deferens can become infertile. Some men have both the testes but they are not able to produce sperm or may produce very few sperms also suffer from infertility. Moreover, stress can cause decreased libido and the couple can end up in having infertility. WHO (World Health Organization) has also accredited infertility as a public health concern [31], [32]. Around 60-80% couples are infertile throughout the world among which 20-25% is from India [33]. Figure 1.2, 1.3, 1.4 and 1.5 shows increase in the infertility level throughout the world from 1955 to 2015 [34].

Figure 1.2: Infertility level around the world in 1955-Source: United Nations-Population Division (2017 Revision)
Figure 1.3: Infertility level around the world in 1995
Source: United Nations-Population Division (2017 Revision)
Figure 1.4: Infertility level around the world in 2010 - Source: United Nations-Population Division (2017 Revision)
Figure 1.5: Infertility level around the world in 2015-Source: United Nations-Population Division (2017 Revision)
1.5.1. Causes of Female Infertility

According to the Center of Disease Control (CDC 2013) [35], the causes of female infertility can be divided into three broad categories including defective ovulation, transport and implantation. These categories are further discussed below in detail.

Figure 1.6: Causes of Female Infertility

1.5.1.1. Defective Ovulation

Defective ovulation occurs because of the following causes:

**Endocrine disorders:** The disfunction of hypothalamus and pituitary gland can lead to the secretion of excess amount of prolactin, which prevents ovulation. Moreover, other endocrine glands including adrenals and thyroid may also delay ovulation. When the corpus luteum, fails to produce enough
progesterone required to thicken the uterine lining, the fertilized egg may not be able to implant, thus leading to infertility.

**Physical disorders:** Certain physical disorders such as obesity, anorexia nervosa, and excessive exercise may lead to overweight or malnutrition, and later the menstrual cycle, thus make the couple infertile.

**Ovarian disorders:** Polycystic Ovarian Disease (PCO) can lead to infertility due to increased amount of testosterone and LH and decreased uptake of glucose by muscle, fat and liver cells resulting in the production of large amount of insulin by the pancreas. Low FSH levels also hinder the production of eggs from the ovarian follicles, and results in fluid-filled ovarian cysts that eventually cover the whole ovaries and prevent conception. Figure 1.7 depicts the ovarian disorder.

**Endometriosis:** It is a disorder where the tissue that belongs to the uterus grows in other locations like ovaries, fallopian tubes, tissues surrounding uterus and pelvic cavity [36]. If it affects ovaries, cysts develop and anovulation occurs. It can also form scars and adhesion that block fallopian tubes or disrupt implantation of fertilized egg. It is also accompanied by painful and heavy periods, pain at the ovulation. Figure 1.8 shows the endometriosis inner layer of the uterus growing in abnormal locations.

**Tubal blockage [37]:** Fallopian tubes are the place where the sperm fertilizes the egg; hence it should be healthy and open. When the tubes are blocked or
damaged, the sperm cannot reach the egg to fertilize it [38]. Tubal blockage can be caused by anatomical abnormalities, previous surgeries in uterus or tubes and infections due to sexually transmitted diseases, especially Chlamydia or gonorrhoea. Figure 1.9 shows the blockage in tube.

Figure 1.7: Ovarian disorder
Figure 1.8: Endometriosis—inner layer of the uterus growing in abnormal locations

Figure 1.9: Tubal blockage
1.5.1.2. Defective Transport

The following can lead to defective transport of ovum and sperm:

**Ovum:** Occurrence of Pelvic Inflammatory Disease (PID), gonorrhea, peritonitis, previous tubal surgery, and fimbrial adhesions can cause tubal obstruction; as a result the egg is not released or trapped, therefore, delaying conception.

**Scar tissue after abdominal surgery:** After abdominal surgeries, presence of scar tissue may alter the movement of the ovaries, fallopian tubes, and uterus, resulting in infertility.

**Sperm:** Presence of psychosexual problem such as vaginismus, or dyspareunia may hinder fertilization and make the couple infertile.

**Cervix:** Trauma, surgery, infection, anti-sperm antibodies in the cervical mucus may also delay pregnancy.

1.5.1.3. Defective Implantation

Defective implantation can occur because of the following causes:

**Congenital anomaly and fibroids:** Congenital uterine anomaly such as bicornuate uterus and uterine fibroids near the fallopian tubes or cervix may alter implantation of the zygote and cause infertility. The causes of female infertility are discussed. The reviews for causes of male infertility are discussed further.
1.5.2. Causes of Male Infertility

According to the CDC (2013), male causes of infertility are divided into the following four main categories:

**Defective Spermatogenesis:** Presence of endocrine disorders such as diabetes mellitus and hyperthyroidism lead to azospermia or the formation of faulty sperms that are not capable to fertilize the ovum [39]. Moreover, testicular disorder such as undescended testis can also affect fertility.
**Defective Transport:** Obstruction of the seminal vesicles or absence of the seminal ducts may affect the mobility of the sperms, and thus end up in infertility.

**Ineffective Delivery:** The psychosexual problems like impotence, ejaculatory dysfunction, physical disability, hypospadias, and epispadias can affect fertility of males.

After discussing both male and female causes of infertility, further discussed to states the history and physical examination component of infertility.

**1.5.3. Treatment Modalities**

Infertility treatment depends on the cause, duration, both partners age, and personal preferences. Financial, physical, and time commitment is required for infertility treatment. The following treatment modalities can be explained to the couple after assessing and evaluating the couple’s health:

**Intra-Uterine Insemination (IUI):** This could be used for unexplained infertility of female cases with minimal endometriosis and mild male factor infertility problems [40], [41]. In this, healthy sperms that are collected and concentrated are placed directly in the uterus at the time of ovulation. The timing of IUI can be coordinated with the normal cycle or by using fertility medications.
**In-Vitro Fertilization (IVF):** In IVF, multiple mature eggs from a woman are retrieved, and fertilized with a man’s sperm outside the womb and inside a laboratory [42]. Then, the fertilized embryos are implanted in the uterus after three to five days of fertilization.

**Zygote Intra-Fallopian Transfer (ZIFT) [43] and Gamete Intrafallopian Transfer (GIFT):** In ZIFT, the fertilized egg is directly transferred into the fallopian tube; whereas, in GIFT a mixture of sperms and eggs is placed in the fallopian tube and fertilization occurs there [44].

**Intra Cytoplasmic Sperm Injection (ICSI):** In ICSI, a single healthy sperm is injected directly into a mature egg [45] [46]. ICSI is used when there is a problem with the quality of the semen, or there are few sperms, or prior IVF cycles have failed.

**Assisted Hatching:** Through this technique, implantation of the embryo into the uterus is assisted by breaking the outer covering of the embryo. This helps the embryo to smoothly implant

**Donor Eggs and Sperms:** Assisted reproductive technology mostly uses the married couple’s eggs and sperms, but when there are severe issues with the eggs and sperms, then donor sperms or even embryo is taken to enhance fertility.
**Gestational Carrier:** This is sometimes called as surrogate pregnancy. When a woman who does not have a uterus or if the uterus is not functional and to whom the pregnancy can endanger health, the couple can decide to have a gestational carrier, who carries the couple’s embryo in the uterus.

The couple should be explained that some of the causes of infertility cannot be corrected. Counselling can be given for adoption of child for couples who have multiple unexplained IVF failure cycles.

### 1.6. Artificial Insemination

Artificial insemination is an assisted conception method, which can be used to alleviate infertility in selected couples. The rationale behind the use of artificial insemination is to increase the gamete density near the site of fertilization. The effectiveness of artificial insemination has been clearly established in specific subsets of infertile patients such as those with idiopathic infertility, infertility related to a cervical factor, or a mild male factor infertility [47], [48]. An accepted advantage of artificial insemination is that it is generally less expensive and invasive than other Assisted Reproductive Technology (ART) procedures.

#### 1.6.1. History

Artificial insemination has been used in clinical medicine for more than 200 years for the treatment of infertile couples [49]. In 1785, John Hunter, a
Scottish surgeon from London, advised a man with hypospadias to collect his semen and have his wife inject it into her vagina. This was the first documented case of successful artificial insemination in a human.

In the second half of the nineteenth century, numerous reports were published of human artificial insemination in France, England, Germany, and the United States. In 1909, the first account of successful donor artificial insemination was published in the United States. By 1949, improved methods of freezing and thawing sperm were being reported.

Today, artificial insemination is frequently used in the treatment of couples with various causes of infertility, including ovulatory dysfunction, cervical factor infertility, and unexplained infertility as well as those with infertility caused by endometriosis, male, and immunologic factors. Artificial insemination with donor semen has become a well-accepted method of conception.

1.6.2. General Considerations

The source of semen for artificial insemination can be either from the woman’s male partner or from a donor, who usually remains anonymous. When the concept of donor insemination become to gain momentum the terms homologous artificial insemination and heterologous artificial insemination were used to differentiate these two alternative sources. However, the usage of these biomedical terms here is at variance with their scientific meaning,
where they denote different species or organisms (as in, e.g., homologous and heterologous tissue grafts).

In the latter half of the 20th century, the terms Artificial Insemination, Donor (AID) and Artificial Insemination Husband (AIH) found common use. However, the widespread use of the acronym AIDS for Acquired Immuno Deficiency Syndrome resulted in the replacement of AID with Therapeutic Donor Insemination (TDI). An analogous alternative term for AIH has not evolved, probably in part because of the increasingly common situation where the woman’s partner is not her legal husband.

1.7. Ant Colony Optimization Algorithm

Ant Colony Optimization (ACO) Algorithm is a population-based metaheuristic algorithm. It is used to find approximate solutions to difficult optimization problems. ACO uses a set of software agents called artificial ants which search for good solutions to a given optimization problems. The ACO is applied by transforming the optimization problem into the problem of finding the best path on a weighted graph. The solution for the problem is build by the ants by moving on the graph incrementally. The solution construction is biased by pheromone model, in which a set of parameters associated with graph components whose values are modified at runtime by the ants.

The ACO algorithm is an algorithm for finding optimal paths that is based on the behaviour of ants searching for food [50]. At first, the ants wander
randomly. When an ant finds a source of food, it walks back to the colony leaving "markers" (pheromones) that show the path has food. When other ants come across the markers, they are likely to follow the path with a certain probability. If they do, they then populate the path with their own markers as they bring the food back. As more ants find the path, it gets stronger until there are a couple streams of ants travelling to various food sources near the colony.

Because the ants drop pheromones every time they bring food, shorter paths are more likely to be stronger, hence optimizing the "solution." In the meantime, some ants are still randomly scouting for closer food sources. Once the food source is depleted, the route is no longer populated with pheromones and slowly decays. Because the ant-colony works on a very dynamic system, the ant colony algorithm works very well in graphs with changing topologies. Examples of such systems include computer networks, and artificial intelligence simulations of workers.

The ACO meta heuristic is as

Set parameters, initialize pheromone trails

SCHEDULE _ACTIVITIES

ConstructAntSolutions

DaemonActions

UpdatePheromones

END_SCHEDULE_ACTIVITIES
The metaheuristic consists of an initialization step and three algorithmic components regulated by SCHEDULE_ACTIVITIES construct [51]. This is repeated until the termination criterion like maximum number of iteration or a maximum CPU time. The scheduling and synchronizing of the algorithmic components is not constructed in SCHEDULE_ACTIVITIES. The three algorithmic components undergo a loop that consists in

(i) The construction of solution by all ants

(ii) The improvement of the solution via the use of local search algorithm

(iii) The update of the pheromones.

The most successful ACO algorithms are

(a) Ant system
(b) Ant colony system
(c) MAX-MIN system

Initially, ACO was used in the domain of NP-hard combinatorial optimization problems. It was also used in routing the telecommunication network. The successful example of ACO used in telecommunication network is AntNet. The uses of ACO for the solution of dynamic, multi objective, stochastic, continuous and mixed-variable optimization problem are a current topic. The flowchart of ACO Algorithm is illustrated in Figure 1.11.
Figure 1.11: Flowchart of Ant Colony Optimization Algorithm
1.8. Rough Set Theory

The concept of Rough Set was first introduced by Z.I. Pawlak in 1982 [52]. Any external parameter is not required by Rough Set Theory to analyze and conclude about data set. It is used to find the indispensable features. The discernibility matrix is used to recognize dispensable and indispensable Features [53], [54].

1.8.1. Basics

Let I=(U,A) be an Information System. Here, U is an non-empty set of objects called universe and A is non-empty finite set of attributes such that 
\[ a: U \rightarrow V_a \] for every \( a \in A \). \( V_a \) is a set of values that attribute \( a \) may take. If \( P \subseteq A \), there is an associated equivalence relation:

\[ IND(P) = \{ (x, y) \in U \times U : \forall a \in P, f_a(x) = f_a(y) \} \] 

The partition of U generated by IND (P) is denoted by \( U/P \). If \( (x, y) \in IND(P) \), then \( x \) and \( y \) are indiscernible by attributes from \( P \). The equivalence classes of the P-indiscernibility relation are denoted by \( [x]_P \). Let \( X \subseteq U \), the P-lower Approximation \( _P X \) and P-upper approximation \( _P X \) of set \( X \) can be defined as:

\[ _P X = \{ x [x]_P \subseteq X \} \] 

\[ _P X = \{ x [x]_P \subseteq X \} \]
Let P, Q ⊆ A be equivalence relations over U, then the positive, negative and boundary regions can be defined as:

\[ POS_p (Q) = \bigcup_{x \in Q} PX \]  
\[ NEG_p (Q) = U - \bigcup_{x \in Q} \bar{PX} \]  
\[ BND_p (Q) = \bigcup_{x \in Q} \bar{PX} - \bigcup_{x \in U \setminus Q} P.X \]

The goal of attribute reduction is to confiscate redundant attributes so that the abridged set provides the same superiority of cataloguing as the original. The set of all reduct is defined as:

\[ \text{Red} (C) = \{ R \subseteq C \mid \gamma_R (D) - \gamma_C (D), \forall B \subseteq R, \gamma_B (D) \neq \gamma_C (D) \} \]

A dataset may have many attribute reduct. The set of all optimal reduct is:

\[ \text{Red} (C)_{\text{min}} = \{ R \in \text{Red} \mid \forall R' \in \text{Red}, |R| \leq |R'| \} \]
1.8.2. Relative Reduct Algorithm

The degree of relative dependency is calculated after removing the attributes from the set. If the value of the relative dependency for the attribute which is removed is found to be one, it is eliminated. Otherwise, it is added to the core reduct. The process is repeated until the value becomes one. Figure 1.12 explains the Relative Reduct Algorithm. The algorithm for relative reduct is as follows:

\[
\text{R} \leftarrow \{\text{list of conditional attributes}\}
\]

Select \(x\) \(\leftarrow\) conditional attribute from R

If dependency = 1

Then “eliminate the attribute”

Else

\[\text{R} \leftarrow \{\text{list of reduct}\}\]

End
Figure 1.12: Flowchart of Relative Reduct (RR) Algorithm
1.9. Artificial Neural Network

Artificial Neural Network (ANN) is developed to mimic human brain technique. ANN is a Data Mining which is defined as a closely interconnected network or mathematical model or computational model [55] with some of the features of biological neural network. The complexity behind studying a system defines the number of hidden layers and neurons in each layer [56]. The learning methods adopted in the Artificial Neural Network are

(i) Supervised Learning.
(ii) Unsupervised Learning.
(iii) Reinforcement Learning.

There are different ANN types available. The most popular and more frequently used types are as

(i) Feed Forward Neural Network (FFNN)
(ii) Recurrent Neural Network (RNN)

In FFNN, there will be no feedback loops. The flow of information is always in one direction. In RNN, output is fed back to the input. An ANN is typically defined by three types of parameters:

1. The interconnection pattern between the different layers of neurons
2. The learning process for updating the weights of the interconnections
3. The activation function that converts a neuron's weighted input to its output activation.

The tasks Artificial Neural Networks are applied to tend to fall within the following broad categories:

- Function approximation or regression analysis
- Classification, which includes pattern and sequence recognition.
- Data processing which includes filtering, clustering and compression.
- Robotics.
- Computer numerical control

Figure 1.13 gives the sample of ANN which input layer, hidden layer and an output.

![Figure 1.13: A sample Artificial Neural Network](image-url)

---

33 | HYBRID SYSTEM FOR PREDICTING IVF SUCCESS RATE
1.10. Scope of the work

Infertility is defined as the incapability to attain pregnancy after one year of unprotected intercourse. It is distressing a growing number of married couples around the planet. It is considered as the most important health problems in developing countries. WHO has also accredited infertility as a public health concern. The effective ways like Assisted Reproductive Technologies (ART) and In Vitro Fertilization (IVF) have evolved in medical field to deal with this cause. The success rate achieved by this treatment has been increased recently up to 10%. But it still fits only to the 40% of the people [56]. It is a challenging task for an embryologist to analyze and associate the large number of features. Technology is still lagging behind to improve the success rate in infertility treatment. Hence, this work aims at developing a hybrid success rate prediction tool for IVF treatment. This hybrid success rate prediction tool combines the features of Ant Colony Optimization Algorithm, Relative Reduct Algorithm and Artificial Neural Network with Back Propagation Learning.

1.11. Objectives of the work

The main objective behind this research work is to develop a hybrid model which combines different Data Mining techniques for the efficient fertility data analysis, processing and prediction. This work involves in hybridizing Ant Colony Algorithm, Relative Reduct Algorithm and Artificial
Neural Network together to produce an effective hybrid technique to predict the success rate of IVF outcome. Further this tool can also be suggested as a tool for assisting IVF treatment. To achieve this, the work is divided into following contributions as:

(i) To propose a hybrid Algorithm for the purpose of Feature Reduction.

(ii) To train the Neural Network by using Back Propagation Learning Algorithm to classify the reduced data set.

(iii) To develop a hybrid technique for success rate prediction.

1.12. Organization of the Thesis

The research works carried out so far using the Data Mining Techniques in medical field is discussed in Chapter 2. The proposed model in this work and techniques used for constructing model are discussed in Chapter 3.

The chapter 4 presents the implementation of proposed hybrid algorithm for reducing the features of the data set without affecting the original knowledge of the data set.

Neural Network Architecture - Multi-Layered Perceptron Network MLP is applied for classifying the extracted features is discussed in Chapter 5. The classified results are interpreted with different parameters for performance measure. Chapter 6 explains the hybrid model used for predicting success rate
of IVF treatment. Chapter 7 compares the results obtained from the proposed model with different existing model/methodologies. Chapter 8 presents the conclusion of this research work and future works to be carried out in this direction.

1.13. Chapter Summary

This chapter summarizes the brief introduction about Data Mining and its various techniques. It also briefly describes about the Medical Data Mining. The causes of Infertility, treatments and various process of Assisted Reproductive Technology applied for treatments and Artificial insemination process are also explained. The objective of this work and thesis organization also explained in this chapter.