CHAPTER - 2  
LITERATURE SURVEY

Astrology is the study of influence of planets and stars on every person and is said to play a significant role in predicting the future of person. Astrology is five thousand year old study of effects of various planets like sun, moon, etc and zodiacs on events on the earth and individuals. There were concert efforts done towards collecting of the astronomical knowledge in the field of astrology by various researchers like Galileo, Kepler, Carl Jung, Copernicus, etc.

Astrology is the search for events which takes place due to movement of various planets in the sky. Conscious attempts to record, measure and predict changes in seasons by referring astronomical cycles appears as markings on caves walls and bones. It shows that lunar cycles were observed during 25, 000 years ago. Farmers talked about agricultural needs with increasing knowledge of the constellations that appear in the different seasons and used the rising of particular star groups to identify annual floods or seasonal activities (Evelyn, 1977).

Astrology has been dated to back to 2nd century BC, and has its origin in calendrical systems used to calculate periodic changes and to predict seasonal changes to interpret celestial cycles as signs of heavenly communications (Koch, 1995). Many cultures such as Chinese, Indians and Maya have provided importance to astronomical events. They have developed detail techniques for the prediction of terrestrial events from celestial observations. Western astrology as one of the oldest astrological systems traces its roots to 19th-17th century BC. Mesopotamia, from there it spread to Ancient Greece, Rome, Arab and finally to Central and Western Europe. Western astrology is often associated with systems of horoscopes that explain different aspects of a person's personality and predict significant events in the person’s lives based on the positions of celestial objects, the majority of professional astrologers perform by such systems (Jeffrey et al., 2007).

The oldest evidence of the use of astrology as a system of knowledge is therefore assigned to the information of the first dynasty of mesopotamia during
1950–1651 BCE. This astrology had some parallels with Hellenistic greek that is western astrology. The babylonians saw celestial events as possible signs than as causes of physical events (Rochberg, 1988).

Two form of astrology are practiced more in the world. First one is Indian Astrology or Vedic Astrology and second one is Western Astrology. There is difference between the two astrological thoughts. The first difference the two astrology has is difference between birth charts. Chart in western astrology is round and that of vedic astrology is square.

Western astrology uses the moving zodiac whereas vedic astrology uses the fixed zodiac. Western astrology focuses on the psychological nature of the person, but lacks accuracy for predicting future events and cycles. Vedic astrology defines a person’s psychological nature, it focuses more on knowing when certain events are likely to take place.

Western astrologers use the tropical zodiac, where the location of the twelve signs are identified by the sun's position around the earth. Vedic astrologers use the sidereal zodiac. It is based upon the location of the constellations in the sky. The starting point is chosen as the place in the sky opposite to spica for the start point for aries, and proceed in equal angles of 30 degree for subsequent signs.

Western astrologers use an selection of aspects, in which five important ptolemaic aspects are conjunction, sextile, square, trine, and opposition. Vedic astrologers use a different strategies each planet have association with all planets in the same house, and have aspect with opposite house and planets in that opposite house. Mars, Jupiter and Saturn also have special aspects to both planets and houses.

Western astrologers also give significance to aspect patterns like grand trines, yods, T-squares etc. which are not used by vedic astrologers.

2.1 Research in Astrology
Nearly thirty three percent of the population in western countries have faith in astrology and another thirty three percent cares enough checks astrological
predictions at least some of the time (Eysenck, 1982). The scientific community questions the validity of astrology based on arguments such as it is not science but just a common sense. Many scientific test shows that astrology is not correct. Most scientists just don’t believe in astrology at all and they simply ignore and dismiss astrology (Bok et al. 1975). As a result, the scientific community has been facing the allegation of arrogance and lack of concern for astrology (Eysenck et al., 1982).

There are beliefs that astrology has two parts, predictive and non-predictive and researchers feel that predictive astrology is the actual subject to be tested that whether astrology can be utilized for prediction (Penny, 2008/2009).

Morin et al. (1974) built up an astrological framework that serves as a base for scientific astrology. Morin’s hypothesis of determination is the base of his research. In Astrologia Gallica, Morin put forward in detail his method and theory of astrological determination and laid out the standard rules and principles according to which the methods are applied for the theory. For judging the charts and performing predictions it is necessary that the details of principles, methods and rules of Morin’s system be used in application. Comprehension of whole system is necessary for one who attempted to evaluate the framework with statistical techniques.

Rudhyar (1976) was a leading non-scientific astrologer in the 20th century. Rudhyar pointed out three types of astrology. He defined first form of astrology as predictive astrology. He defined the second form of astrology as the one that tries to adopt the scientific and statistical methods of analysis and research. He defined third form of astrology as the humanistic astrology. According to him there are many astrologers who are philosophical and psychological by nature and for them the basic aim of astrology is to help every person.

Rudhyar (1967) formulates and describes the Lunation Cycle. It is a unique technical breakthrough showing the importance of the cyclic relationship of the sun and moon as part of a larger process. He describes for the first time, the eight type of lunar personalities and the importance of the new moon before birth and also described the progressed new moon charts. Rudhyar's offered understanding into
how to use events of life to identify one’s inner potentials. The occurrence of specific events in person’s life is considered as an ordered series of lunation phases within the context of person’s life from birth to death. Thus providing a clear and practical way to expect major turning events and transition periods in individuals life.

The 20th century astrologer Leo Knekt used the concept of scientific astrology to attain his best results (Smit, 1997).

Tarnas (2006) is explicit about the aim and orientation of planetary Arche typology, which is considered to be force of non-scientific or non-productive astrology. He explicitly and in practice focuses largely on universal significations. He concentrates on the meanings of planets as celestial symbols. Tarnas attends neither to the range of what Morin called celestial state nor to local determinations. He focuses his astrological interest to nature of planets, position of planets in signs and aspects between planets. The placing of planet in a sign and the aspects it has with other planets are two factors among the range of factors that celestial state comprises.

Perry (1994) feels that correlational relationship exist between person and birth charts. The astronomical body reflect experiences of human life and has to be understood. Some researchers feel that the astrology is not scientific (Ivan, 1997; John et al. 1990) while some (Ken, 2011) feel that a proper study is required in this field.

Mcgrew et al. (1990) performed experimental studies in the close collaboration with Indiana Federation of Astrologers. The experiment performed were executed only after all details were discussed and approved by the IFA as a reasonable test for astrologers capabilities for performing prediction. The trial subjects were six people one male, five female identified by the IFA as astrologers with best capabilities. In experiment six expert astrologers freely attempted to match twenty three horoscope charts to the related case files of four men and nineteen women volunteers. Case documents contained data of the volunteers’ life histories,
face and profile photographs, test profiles and personality Inventory. Six expert astrologers were not able to perform better than chance at matching birth details to the corresponding case files of twenty three persons. There was hardly any relationship between astrologer’s prediction, their confidence in their predicted results and the correctness of their predictions. Overall, mean confidence rating of 73.5% was shown by the astrologers in their predictions. Across astrologers, the Pearson correlation between number of accurate predictions and mean of self rated confidence was r=.03 which was not very good. For those astrologers who had at least one correct match, the mean confidence ratings for correct prediction was 76.4 percent and incorrect predictions was 72.8 percent that were also not significantly different from previous.

US psychologist Clark (1961) performed a number of tests between 1951 and 1970. The test performed where blind matching tests involving a total of fifty astrologers. While a control group of 20 psychologists and social workers matched 10 pairs of horoscope with corresponding professions of the person’s with an accuracy of 50% that equal to expected by chance. But the astrologers successfully matched with 65% accuracy. This result are not very significant. In further study (Clark, 1970) removed all possible cues from knowing sun sign traits, by using matched pairs with the same sun sign. The astrologers were able to match birth charts to case histories in 72% of the time. In final experiment, 59% astrologers were able to identify correctly the individual’s with a high IQ and one with cerebral palsy. Overall astrologers matched correctly 64% of the time out of 700 evaluations.

An empirical study is presented in Gauquelin et al. (1979) of relationship between the position of planets at the time of birth and the personality of the persons. The records of profile details and birth data of thousands of famous French sportsmen, scientists and actors were collected from biographies. The rising and culminating positions of planets Saturn, Mars and Jupiter were calculated and the Hans Eysenck’s scales was used to measure personality of person. It was predicted and found that introverts are born more frequently when Saturn had just risen or just passed its upper culmination and extroverts are born when Jupiter and Mars had just risen or passed their upper culminations.
French psychologist and statistician, Gauquelin (1988) performed tests focused on the correlation between the person’s birth chart and positions of the planets in the person’s horoscope at the time of birth. He also tried to identify the relation between position of planets and the person’s psychological nature, choice of profession and biographical descriptions. The data of 20,000 professional celebrities where collected from Unites states and various European countries. Gauquelin's research detected statistically anomalous positions of the Mars planet in birth chart of athletes, Jupiter planet in birth chart of actors, Saturn planet in birth chart scientists and the Moon planet in the birth chart of writers. The tests performed confirmed the claims of ancient astrologers that planets posted around the four angles are stronger and the qualities associated with the planet shows unmistakably in the person.

Ridgley (1993), performed a research on nearly a thousand persons living in California who had filed workers' compensation claims and were handicapped for not less than three months. In these research it was found that workplace accidents were found to be very frequent during the days when the sun is conjunct, opposite, or at 90 degrees to the Sun's position in the horoscope of the injured workers. This evidence confirms the traditional relation of fourth harmonic planetary aspects with difficulties and challenges in person’s life.

Johnston (2001) studied hundreds of data of earthquake from 1996 to 2000 collected by U.S. Geological survey. He found that earthquake occurs when planets occur at significant angle with each other. Hence he concluded that seismic activity is connected with major aspect of planets with each other.

A research by Kathy et al. (2006), was done on the data of 48 countries that was much larger than similar studies done previously. This study, was conducted for identifying the factors for stock market volatility and calendar-related anomalies. The results of this research had found that stock market is effected by increasing and decreasing cycle of the Moon. International stock market returns were observed to be essentially lower on days close full moon days than on the days close to new moon days.
The concept stock of market analysis, was given by 17th century astrologer Johannes Kepler, has been worked upon afterwards by researchers Addey (1976), Landscheidt (1989), and Schulz (2010) by performing computer modeling on the concept. In exploratory application by Schulz’s trend fitting application, the TAO Oscillator program, has a proved to be very successful for stock market forecasting (Snow, 2010). This application has found unexpected correlations between planetary positions and structure and frequencies of economic indexes and the concept is called as cosmic cybernetics.

2.2 Case Based Reasoning and its Application Area

In the research work provided, we are using case based reasoning method along with various classification techniques to perform astrological predictions of future of the people. Till now hardly any work has been done in the field of astrological prediction using Case Based Reasoning along with various classification techniques of artificial intelligence. Rishi et al. (2010) has done some initial research using case base reasoning along with nearest neighbor algorithm. Hence only those papers are available on the topic.

Case Based Reasoning (CBR) is a technique used for problem solving. It replicates human thinking by making a decision based on previous experiences. This concept was first introduced around the year 1977 and originated in the field of cognitive science. The idea of CBR can be described in one sentence that is a Case Based Reasoning solves new problems by using/reusing solutions that were used previously to solve similar problems.

Kolonder et al. (1993) defined the case-based reasoning as “adapting old solutions to meet new demands, using old cases to explain new situations, using old cases to critique new solutions, or reasoning from precedents to interpret a new situation”. Case based reasoning provides a model of reasoning that handles problem solving, learning, and combine everything with memory processes.

In this section, the aim is to provide a brief review of work done in the application of case-based reasoning (CBR). CBR systems have pulled a lot of attraction
in the medical and legal fields, especially as diagnostic and care systems. It is also used in insurance and finance field for customer support and credit evaluation (Allen, 1994). In addition CBR has been used in diverse applications such as in intelligent Web-based sales services (Wilke et al., 1998; Watson et al., 1999), software interface design (Tsatsoulis et al., 1997), in building and mechanical design (Mileman et al., 2002; Rivard et al., 2000) and architecture design (Domeshek et al., 1992; Flemming et al., 1997).

CBR is used in chemical engineering field for quality design (Suh et al., 1998). It is also used thermal analysis support (Nakayama et al., 1999). The attempts were made in developing intelligent systems in the domain of law. This attempt were boosted by HYPO, a CBR system that was used to creates legal arguments in the domain of trade secret law using case base(Ashley et al., 1988). In this program, dynamically indexing and relevancy analysis of past cases are performed.

Rissland et al. (1995) built a hybrid CBR and Information Retrieval (IR) system in that Case Based Reasoning technology performs the task of improving the query send to the Information Retrieval System thus improving the results. The CBR system is also explored in music applications by Macedo et al., (1996). The work is done by using CBR techniques for music composition. The tree like structure is used for representation of cases. Previous work was done in the field of learning to generate expressive performances based on various sample data (Widmer, 2002). It uses explanation based techniques to learn rules for dynamics and rubato for MIDI electronic piano.

The Researcher (Aleven, 2003) presented a model of case-based legal argumentation (CATO model). It uses the concept used by attorneys to use previous cases in arguments. It concentrates on the representation of middle level background knowledge. The knowledge generated is used for purposes of case selection and argumentation. Factor Hierarchy, is a new knowledge source implemented in the CATO program was used which deals with representation of knowledge about the meaning of the factors that are used to represent cases. Background knowledge was used to organize arguments related multi case discussion by issues. Then arguments are
generated about the similarity of cases, which focus on alternative characterizations. Then assessing of the best case is done for citing it in an argument based on the similarity of cases. The preliminary evaluation presented in this research provides evidence that CATO’s provide sound arguments about the significance of distinctions. The results provided by the research shows that, Case Based Reasoning methods applied for the background knowledge in reasoning help to make better predictions of the outcome.

Mota et al. (2008) apply the Case Based Reasoning (CBR) technique to establish rules and detect object deforestation evolution in an Amazonia region. The analysis of objects is done by using a sequence of images from different durations for creating the base and rules to apply CBR. CBR interact with an already existent structural classifier. This classifier classifies deforestation by identifying the patterns that agree with a land use and land cover typology. First of all the classified images are used to create the new case then the database cases is searched to identify previous cases having similar description to the current situation. The application of Case Based Reasoning helps the expert to understand how different objects in the nature evolve, to establish and to represent that evolution of the objects in the nature.

Expression CBR (De Paz et al., 2008) is a decision support system for cancer diagnosis. The system by classifies the Leukemia patients automatically from Exon array data. It helps in diagnosis of patients with various cancer types data filtering algorithm is used that deals with the dimensionality problem in data sets. Clustering algorithm for classification approach is used to speeds up the process of diagnosis.

Ahmed et al. (2008) proposes a three phase computer assisted sensor based biofeedback decision support system to provide treatment for stress related disorders. A CBR framework is created to help to classify a patient, estimate initial parameters and to advice the recommendations for biofeedback training. Fuzzy techniques are used to handle uncertainty in clinical reasoning as well as help in decision analysis. The intention of the system is to enable a patient for treatment without any particular supervision.
Jianhua Chen et al. (2010), this paper carried out research work on prediction of regional metallogenic using Case Based Reasoning and GIS approach. In the metallogenic prediction Case Based Reasoning model along with metallogenic case expression model is used. The reasoning flow was given, experiments were implemented to perform prediction for Gold mine. The experiment results show the prediction of metallogenic is effective using Case-Based Reasoning.

Martin et al. (2011) provided an investigation on Impact of Feature Selection in Case based Reasoning for Predicting Bankruptcy. Forward feature selection and backward feature elimination method is used. For case retrieval the best features where obtained using K-Nearest Neighbor algorithm. In this investigation comparative study on above two feature selection techniques was done and influencing features where selected by genetic algorithm. Effectiveness of Case Based Reasoning bankruptcy model depends on the feature selection method and case retrieval algorithm used. The 82 % accuracy results were generated by forward feature selection method. This method was better than other feature selection techniques.

Gopal et al. (2012) performed studies on project estimation based on case based reasoning methods. In his work he combined two available techniques. Prototype tool was created that permits addition of parameters dynamically, based on the context of the project using case based reasoning. Estimation of the project is done by computing similarity index of the parameters. After that revision in estimation can also be done based on delay causing incidents that are present during the execution of the project. Data from about ten projects implemented in industry previously shows that estimates calculated using the given approach can converge to actual values over a period of time. Illustrating the variations between estimated efforts against the actual effort for all the ten projects. Then graph plots estimated effort, revised effort and actual effort as it existed in the data source for the task. The results generated from the graph shows clearly that the actual and the revised estimates vary significantly for the first few projects, but the revised estimates converges for the last few projects.
A new storytelling system, Say Anything (Swanson et al., 2012), writes textual stories with human collaboration. Interactive tale is a category of storytelling that allows active contribution by the audience to affect the development and results of the story. It is a powerful method because of the essential connection between humans experiences and individual instinct. The computer carries on the storyline with narration that is both consistent and interesting. This system for open-domain interactive storytelling is supported by a very large database of nonfiction stories by persons, which is used as a knowledge base in a case based reasoning system. It is composed of three primary components of case based reasoning method is million item corpus of personal stories taken from internet weblogs. A case retrieval strategy that is optimized for story coherence, and a strategy that make sure that the adapted sentences from the case base are correct for the story created by user’s. It also have a series of evaluations for the system’s ability to produce consistent and entertaining stories, and these stories are compared with single-author stories posted to internet weblogs. This system is reasonable, user friendly and found to be effective for limited domain of use cases. However, major part of its operation and success continues to remain a roadblock for judgement of common sense reasoning about real-world stories.

In Wang et al. (2012), a Triple Artificial Neural Network Model using Case Based Reasoning for Credit Risk Assessment is created and tested. For this system the history data are preprocessed by Case Based Reasoning, only similar cases are chosen for credit risk assessment. Various attributes such as numerical and categorical are used independently in different stages of the model, and their influence is overcome. This model utilizes the advantages of CBR and ANN. Practical implementation of model show that it is feasible and effective. Comparison of results generated by different models shows that proposed model is superior to individual SVM and ANN models for prediction precision.

Zubair et al. (2012) proposed a system that apply CBR to predict, estimate and evaluate air accidents and crashes. Detailed algorithm has been defined and coded for retrieval of the cases from the case base of air crashes having solution algorithms and revision mechanism. Five different type of distance functions are used for
retrieval of nearest neighbors of the new case encountered. Solution and adaption system has been applied and detailed evaluation of the suggested method has been presented. It was found that upto 87% accuracy can be achieved using the suggested algorithm.

This paper (Wah et al., 2013) suggest to apply case based reasoning in solving legal cases due to its resemblance with the human cognition in solving problems. Courtroom decision support system using case based reasoning is found to be productive and constructive. Testing of initial implementations have shown that the concept of reusing solutions of past cases for current problem is well accepted by both litigants and offenders. Police officers also support it as a step to reduce the number of instances of the cases for hearing.

Elhalwany et al. (2014), proposed an enhanced version of SOPHIA (SOPHisticated Information Analysis)-TCBR (Textual Case Based Reasoning), which provides higher precision and better time efficiency. The suggested application is assessed in the area of Arabic Islamic Jurisprudence, which is a challenging case study with its huge case base and large number of user requesting questions daily. This process needs a intelligent system that can help in satisfying people's needs for solutions by implementing the suggested strategy in this area and conquering difficulties related to the terminology format and semantics.

Kawabe et al. (2014) introduces a Case Based approximation techniques to solve Traveling Salesman Problems in a short duration of around three seconds with less than 3% error rate. This approach is in accordance with the understanding, that a majority of real life problems are very often identical to previous ones. Thus, a solution of the problem can be derived from past solutions by selecting most identical Traveling Salesman Problems from the records of Case Base of previously solved TSP solutions then removing the places that are not included in the newly given TSP problem and adding the new locations by nearest Insertion and adjusting the nearest Insertion using Genetic Algorithm. This way of creating solutions by Case Based Reasoning (CBR) prevents the computational costs to create new solutions from scratch. The assessment of this technique showed amazing results.
Till now the fastest optimal approximate TSP solving techniques needed more than 3 seconds. The worst error rate by the technique suggested is less than 1 %. This is about 10 to 100 times more faster than the previous approach.

The author (Khan, 2014) has provided a wide and spectacular methodical applying research of various CBR programs in SE domain and identified some significant research directions for future. Application areas using CBR are identified as cost and effort estimation, software quality, software reuse, software project management, software design and software self-management. The results of the study reflect that CBR has been used more substantially for software cost and effort estimation.

Kim et al. (2014) designed a process in which the most appropriate technique will be used in the existing battle situation depending on the similar past situations, by using case based reasoning to design fighting strategies. In such system, when the squad persons gather to perform an army operation on a squad, the squad persons will act as per the specific plan, to achieve the final goal. This process should involve army techniques for activities such as attack, trap, and strategic plans. In other words, new methods are necessary to be used based on the members circumstances, such as the location of enemies and shelters, and such techniques can be produced depending on the learning obtained by the squad or platoon commanders from army manuals, and their collected experiences. In this research, a combat strategy generation system using Case Based Reasoning is proposed, in which the most identical cases saved in a case base were extracted and used after appropriate modification to the current conditions.

Leite et al. (2014) suggests a hybrid software architecture that used case based reasoning, reactive behavior and learning. The main feature of the proposal is the learning of reactive behavior, very quickly and efficiently. This learning is done through continuous interactions of the agent with its environment. Currently, a software agent based on this architecture is being developed in order to evaluate it.
Li et al. (2014) reports the definition of the context model for a small home and how to reason on that model to provide suitable services. The prototype system integrates a context aware system with an OSGI platform to allow the context-aware system to operate home devices easily. Case-based reasoning is used determining use, behaviors and suggests a service fitting to the needs. Personal ontology is kept as cases that represent user preferences, and the system will use this information to choose the suitable services of users.

Lin et al. (2014), the techniques of over sampling and under sampling are used for managing the problems of data imbalanced. The case based reasoning is used for creating classification techniques to predict possibility of repetition of liver cancer in patients. Classification results of over sampling and under sampling methods are compared with original set of imbalanced dataset by conventional indicators, such as specificity, sensitivity, accuracy, balanced accuracy, positive predictive value and negative predictive value. According to the initial results of classification techniques, average balanced accuracy of methods of the under-sampling was 66.07% and the over-sampling was 54.24% that was significant enhancement in comparison with the imbalanced group of dataset that is 48.33%. Most importantly, the highest mean accuracy of the under-sampling is 66.76%, over-sampling is 53.47% and imbalanced is 48.58% the highest value is of under sampling. In under sampling method mean PPV is 65.44%, NPV is 69.44% and accuracy is 66.76. The balanced datasets can provide benefits for classification techniques and efficiently reduce biased interpretations.

Moreno et al. (2014) presents a system designed and built from a focus on Case Based Reasoning (CBR), PeCoS-CBR, for generating personalized courses, taking into account the learning style and preferences of the student. The adaptation of the teaching process is a subject of great interest in the community and the ability to leverage the same educational activities with successful results obtained by students in past sessions is emerging as a promising approach. To resolve the problem by not having RBC in the early cases, mapping metadata for learning objects with some values of variables in the student profile is performed. The tool allowed validates the proposal in a specific topic area but allows it to be used in any course.
Muangprathub et al. (2014) suggested a case based reasoning medical diagnostic system assisting enquiry in Thai language. The system is created to reduce both cost and time required by doctor’s trip by giving a prior medical diagnosis. It is a web based application created using PHP language and MySQL relational data base management system. Previous diagnoses of cases are stored by doctors using MySQL. The persons can enquire by providing their symptoms in Thai. The inputs are used to identify the best similar case among pervious cases in the case base. The proposed system was evaluated with 33 queries for all 50 diseases. The accuracy of diagnoses is 67.87%, with a 99% confidence interval from 53.98 to 81.74%.

An adaptive learning based C programming e-learning system is proposed (Pandey et al., 2014). It considers characteristics of individual students. Case based reasoning is used to streamline and improve learning efficiency of learners. Different types of mistakes committed by learners are categorized like logical, syntax, and application usage feasibility while designing programs. Comparison is done between the students in control group with adaptive learning experimental group. Results showed that performance of the experimental group using e-learning system was better than control group who relied on c programming regular course.

Rueangsirarar et al. (2014) designed a Falling Risk Screening System to identify falling patterns in older people using Motion Capture Technology. The idea is to incorporate an appropriate method including case based reasoning and motion capture to create a good decision support system. The analysis information produced from the process of case based reasoning helps support the physiotherapist to determine serious falling risks in the elderly and recommend guidelines for medical treatment. In this study, the limited sample data leads to use stratified 10-fold cross-validation method for performance evaluation of the CBR’s retrieval mechanism. It shows the very high performance of 81.67% of accuracy.

Widyantoro et al. (2014) proposes a system to support the design process by utilizing case based reasoning approach to help the cognitive procedure in form based user interface. The cases consist of problems in the form of form specifications and
their solutions in the form of form designs. In addition to case base, additional information collected from previous design cases are also included in knowledge base so that the efficiency during case retrieval is improved. This additional knowledge comprises of details about component grouping, component ordering as well as group layout. The dictionary of various words that are used commonly to denote form fields are maintained by system. The user’s form specification are taken as input by the system. It then perform normalization to the user’s requirement to ensure that all field requirements are identified precisely. After normalization the resulting form specification is matched with the stored form specifications in case base. The best matching case in case base is retrieved and its form design is used as the initial form design. It then executes form design revision process to meet with the user form specification. Revision processes consists of transformation and composition. The composition requires form fields grouping, field sorting and fields group lay outing.

Based on the investigation of practical typhoon cases, a spatial awareness CBR approach aiming to solve matching problems of the typhoon disaster cases is proposed by Zhou et al. (2014). It is a novel approach to define the case with explicit and comprehensive spatial multi-elements. Knowledge learning from similar historical cases is critical for emergency decision making against typhoons. A comprehensive case model with spatial elements is introduced, and based on that, a matching method for finding high-quality results in the repository is explained. Typhoon disaster spatial elements such as regional geographic information, population distribution, and secondary derivative disasters are added to this method. In order to manage the spatial elements, this method is integrated with GIS technologies. With the spatial analysis function and the data management capabilities of GIS, the spatial awareness CBR method is discussed. Lastly, a framework of applying the proposed method to Guangdong Emergency Platform System for testifying is introduced. Case based reasoning technology is increasingly adopted to find the appropriate disaster cases matched to current typhoons.

Researcher (Ahmad, 2014) developed Computer Fault Diagnosis (CFD) system, it can detect any corruption or problems that happened to the computer and
can help user to maintain their computer by themselves. The approach of Case-Based Reasoning (CBR) has been used to implement the system by using the previous experience to solve the current problems. In this paper, Computer Fault Diagnosis System based on CBR (CFD) is briefly explained. It can also serve as a base for development of self-diagnostic systems in various domains.

Ayoung et al. (2015), uses the well-established Case Based Reasoning (CBR) methodology to predict failure of telecentres. It is demonstrated through experiments on real life dataset that CBR can be used to predict sustainability of their telecentre. This gives a refreshing indication suggesting that it may be feasible to use CBR to evaluate Information and Communication Technology initiatives and to predict adequately outcome of an initiative. Through this mechanism, it may be possible for managers and owners of telecentres to pre-empt an outcome and have the advantage to take mitigating steps. This affords managers an opportunity for remedial action for sustainability.

Shweta et al. (2015) developed a multiagent system named ABDITS (Agent Based Distributed Intelligent Tutoring System) is an integration of adaptive web-based learning with expert systems. ABDITS uses case based reasoning approach for student modeling. The system will categorize students in step with their skills in processing, perceiving, entering, organizing and understanding the knowledge. The proposed multi-agent system relies on two level agent architecture that Personal agent, Teaching Agent and Course Agent at higher level and Profile Agent, Evaluation Agent, Test Agent and Exercise Agent at lower level. Higher level agents realize continuous student help and tutoring throughout learning sessions.

The design of the proposed e-learning system by Hernandez et al. (2016), uses two methods of artificial intelligence i.e. Intelligent tutoring systems and case based reasoning. The results generated showed notable benefits by using these two techniques. Two surveys were performed on the users for cognitive follow up. one on group of learners using a traditional e-learning environment, and the other one on group of learners using an e-learning environment with artificial intelligence. The total student participants were 222 for four courses i.e. Finances for non-financiers,
SARLAF, Corporate Introduction and Document Identification. The age of the learners was between 22 to 60 years. Every course had duration of one month. At the end of the course learners were asked to answer the survey. It was found that the learners using e-learning with artificial intelligence performed better.

In this paper, an e-learning design using artificial intelligence techniques was introduced. This design was verified using Edupymes e-learning with 2, 400 learners in profession of IT engineering projects in the financial sector. The results show that the e-learning environment suggested is an excellent substitute for the education in subjects of IT engineering projects. Other benefits such as time and resource saving are taken into account, and have a permanent tutor guiding each student.

Hogfa et al. (2016) proposed a four-step method which combining case based reasoning and grey relational analysis. By utilizing prior information, a case library, in which every case is described in details representing for certain kind of technical risk, is built. With a new kind of risk symptom rising, the previous library is retrieved to find the case with largest conditional degree, and then a recommendation solution for the new risk symptom is given according to the solution which is taken to solve the risk with largest conditional degree mentioned above. The knowledge comes into being in the process of dealing with the new risk symptom, is combined into a new case, by creating or modifying the existing case. At last, the new case is stored in the case library. A practical example is provided. The results show that the CBR-based grey relational technical risk identification method is sensible and feasible.

Li et al. (2016) proposed a new remote sensing image discovery method for discovering observational data based on task, location, and time. It lowers the threshold of user expertise required when searching for RS data. Our method focuses on the application semantics of RS rather than detailed complex image parameters. Spatial and temporal constraints are coupled with tasks to discover RS data more precisely. This method approaches location and time not only as filters but also as spatial–temporal constraints in the discovery process, and exploits the relationships between tasks and RS data sources under spatial–temporal constraints through case-
based reasoning (CBR). In this method, cases are past experiences which comprise task, time, location, and image parameters, describing what images were used to satisfy a particular task, and at what time and place each discovery was made. CBR, given its similarity to assessment and result reasoning models, finds past cases that satisfy a user’s query and generates image parameters for specific RS data needed to satisfy that request. A prototype called iGeoportal was developed to evaluate the effectiveness of the proposed method. Experiments show that it performs efficiently when discovering RS images and can be easily integrated into current SDIs through a service-oriented architecture.

Lim et al. (2016) presents a scheduling method for semiconductor manufacturing systems through utilizing a case based reasoning approach that consists of modelling, case base building by aiming to address the need for fast computation of schedules while achieving significantly higher performances than traditional dispatching rules, and considered a scheduling problem for the die attach and wire bonding stages of a semiconductor packaging facility.

In order to accommodate the constraints that are considered in a real-world semiconductor manufacturing system and to model operation of the system, they employed Petri nets by which a sequence of operation types with their respective machine assignments is decoded into cases for the purpose of building a case base. Furthermore, they also proposed a method for generating a solution that specifies a schedule through sequentially retrieving and reusing the cases stored in the case base. The proposed framework requires that the Petri net structure for the scheduling problem to solve is same as that employed at the time of case base construction, posing limitation in its applicability to certain manufacturing environments in which job types, process plans, precedence constraints, or alternative routes vary dynamically.

2.3 CBR and AI Techniques Used for Classification and Prediction

Learning, analysis and prediction of data plays important role in person’s life. Now days classification and prediction tasks are performed by many sophisticated applications so that quick decisions can be taken for optimization of the
resources, efforts and financial benefits. Few application and classification techniques and prediction techniques used in that application are described below


In the field of financial analysis a lot of research have been done and number of studies have been done Desai et al. (1996) investigated various techniques to identify loan application for good and bad credits and found that neural networks, logistic regression performs much better than other statistical methods.

Gregory et al. (1997) predicted pneumonia mortality using eight type of machine learning techniques. For this study 9847 training sets and 4352 test sets were used for the research. 63 attributes were recorded for each patient case and a variable called vital status which stored value alive if the patient was discharged or shifted or was alive after 60 days of his arrival. He was given status expired if he died within 60 days of hospitalization. All model except k2MB predicted out of 4352 cases 435 patient will survive out of which actually 433 or 434 actually survived.

Shin et al. (2001) proposed a case-based reasoning approach to predict bond rating of firms. It was a hybrid method that used inductive learning for case indexing, and best results were produced by nearest-neighbor matching algorithms to predict data using similar past cases with 75.5% accuracy.

Joseph et al. (2006) used and analyzed different methods for cancer prediction and prognosis. Firstly the Cancer Risk or Susceptibility Prediction was done in which (steroid metabolism gene) SNP data were collected 98 SNPS from 45 different cancer associated genes for 63 patients with breast cancer and 74 cases without breast cancer. On this data several machine learning techniques such as naïve Bayes model, several decision tree models and support vector machine (SVM) were applied. The SVM classifier performed the best with an accuracy of 69%. Second study done by author was for prediction of cancer survivability in which author collected the data of microarray expression and clinical data of 56 DLBCL patients. The clinical information was obtained from the International Prediction
Index (IPI). It consists of set of risk factors which divides the patient into groups of low risk to high risk. This data was then passed to Bayesian classifier which predicted the mortality of DLBCL patient with 73.2% accuracy. Last study done by author did prediction of cancer recurrence for this author collected data from 2441 breast cancer patient yielding a data set with more than 17000 data points. The artificial neural network model performed 72.6%

Hsieh et al. (2005) proposed a credit scoring model based on clustering and neural network techniques.

Yanwen et al. (2007) created a model that can be used in small companies to calculate the credit rating of the customers. This paper does credit assessment for a small company whose main business is selling school uniform. Here case based reasoning along with similarity matrix was used to calculate credit rating. The system gave best results that are 98.2% hit rate with Euclidean distance.

Elia et al. (2009) used Classification and Regression Trees (CART) for weather prediction. The meteorological data registered during the last years in the capital of China was used to forecast the future temperature values in Hong Kong. The weather data was collected between 2002 and 2005. Total of 48 instances were checked and out of that only 40 were correctly classified thus giving the results 83.3333 % positive instances were classified as positive.

Van et al. (2010) studied the long term outcome of cancer treatment. The dataset of 1552 patients with clinical and pathological features was collected. Logistic regression and support vector machine were used to train the system to predict three outcomes after five years first one was recurrence of cancer, second one metastases and last overall survival of the patient the data was collected within a week of high dose of radio therapy followed by immediate surgery.

Ten features used were Age, Gender, Distance, Surgery type, surgery group, residual, Pathological tumor stage (PT), Pathological nodal stage(PN), overall pathological stage(PA stage), post-surgical complications (psurgcom). The features
were normalized using z score. For Local recurrence mean was 0.7703 and 0.7473 and SD was 0.0853 and 0.1049. For Metastases mean was 0.7777 and 0.7757 and SD was 0.039 and 0.0562. For Survival mean was 0.7272 and 0.7391 and SD was 0.0520 and 0.029. For PSVM and LR on test data. Thus both methods perform equally well on the data.

In Alonso et al. (2011) a set of prediction algorithms has been analyzed for predicting system crashes due to the resource exhaustion caused by software anomalies. Extensive experimental studies, in three different and complex scenarios, have been conducted to show the level of adaptability and prediction accuracy of the algorithms. The results show clearly the effectiveness of predictive techniques for designing monitoring frameworks for improving proactive rejuvenation techniques.

Folorunsho et al. (2012) collected meteorological data from Ibadan Synoptic Airport through the Nigerian Meteorological Agency, Oyo State office from Jan 2000 - Dec2009. C5 Decision tree classifier was used using 10 fold classification to do analysis. 12 rules were generated for above data set and 7th rule gave least error with 25%. Fully recurrent and partial recurrent network types were used, best results were produced by A fully recurrent network that implemented the TLFD network using the TDNN memory component, one hidden layer with eight neurons and which used the Lavenberg –Marquet learning algorithm.

In Dong et al. (2012) used SVM method of machine learning for prediction of College Entrance Examination (CEE) aspiration. Prediction of college entrance Exam is performed using support vector machine. For training the system feature extraction technique is implemented on data of two years obtained from the training staff. Then machine learning algorithm called Support Vector Machine is used to train the decision-making process of CEE aspiration and the systematic model for prediction is established. The result reveals that the suggested technique performs a very good effect. The prediction accuracy reaches 90%, giving very favorable guidance to examinees for aspiration filling.
In Kandias et al. (2013), a methodology to detect users holding negative attitude towards authorities is proposed. For doing so, a brief analysis of the medium (YouTube), machine learning techniques and a dictionary-based approaches facilitated in order to detect comments expressing negative attitude. Thus, conclusions over a user behavior and beliefs via the content the user generated within the limits a social medium were drawn. The insight offered by these classification techniques over the psychological aspects of the users could be proved useful.

Many companies like insurance, credit card, retail industry, bank require direct marketing. Researchers (Karim et al., 2013) had investigated two techniques, Naïve Bayes and C4.5 decision tree algorithms. The goal of this work was to predict whether a client will subscribe a term deposit or not for this purpose publicly available UCI data is used for training and testing the performance of algorithms. Besides, that actionable knowledge was also extracted from decision tree so that effective and important decision can be taken for the business.

The Research (Mythili et al., 2013) has attempted to pinpoint the most influential factors of heart disease as well as accurately predict the overall risk using homogenous data mining techniques. This paper proposes a framework using combinations of support vector machines, logistic regression, and decision trees to arrive at an accurate prediction of heart disease by using the Cleveland Heart Disease database. Guidelines are provided to train the system to attain the most efficient model for multiple rule based combinations.

Huge amount of data is stored in educational database that can be used to generate useful information for prediction of students performance. In this paper (Ahmed et al., 2014) the classification task is used to predict the final grade of students. Decision tree (ID3) classification technique is used here for prediction. Various attributes were identified to be used for predicting the final grade of student's. Results generated can be used to identify the special needs of the students and thus initiating appropriate action required.
This research (Thakkar et al., 2014) is based on decision tree which generates classification rules that will help in identifying the trend of stocks for next day. The research paper provides a glimpse of the market and trading tips. This research paper predict the next day trend of stock based on daily price movement of the stock (Open_price, High_price, Low_price, Close_price) as compare to that of with previous day price movement of the stock (Open_price, High_price, Low_price, Close_price). Here only those classification rules are considered which have accuracy more than ninety percent.

In Kirchner et al. (2015), used analysis of two machine learning methods to predict the best service within the given selection domain. The first strategy method concentrates on classification thus predicting the best performing service, while the second approach focuses on regression thus predicting service performances that can be used for identifying the best candidate. Both the methods are assessed and compared for service recommendation based on the performance gain. The evaluation of services is based on data measured on actual web services as well as on simulated environment. Both methods can be used for the recommendation of services based on consumer perspective and choices. Both methods have different pros and cons, which were recognized. Also guidelines are described to identify which approach is better under which circumstances. Classification achieved good results with fewer efforts, especially considering periodic behavior of services. However, classification lacks in identifying non best fit services, which may be worth recommending and contributes to performance gain. Regression is not focused on the direct determination of best fit services.

This paper (Kaur et al., 2015) focus on determining the slow learners among the participants by using predictive techniques of data mining. For that it executes classification algorithms on WEKA tool. The classification algorithms are applied on dataset of student academic records and then tested. Various classification algorithms used for the purpose is Multilayer Perception, Naïve Bayes, SMO, J48 and REP Tree. The Classification techniques are used on the dataset of 152 students, for prediction and analysis of student’s performance as also to identify slow learners.
among them. Among all classification algorithm Multi Layer Perception performs best with 75% accuracy and therefore MLP proves to be potentially effective and efficient classifier algorithm. Also comparison of all 5 classifiers with the help of WEKA experimenter is also done, in this case also MLP truns out to be best with F-measure of 82%. This research help the institutions to identify slow learner students so that special classes or aid can be designed for them.

The Erythemato-Squamous Diseases (ESDs) require huge computational efforts to predict the diseases because all the six diseases studied in this group have more than 90% common features. Kolkur et al. (2016) studied the use of machine learning algorithms in R software for prediction of ESD Diseases. Various methods were implemented from Machine learning algorithms in R software which can serve as a primary tool for naive user for prediction of ESD diseases. Accuracy of these algorithms was as provided. K- Nearest Neighbour Classifier gave 97.29% results, Support Vector Machine gave 94.44 % results, Navie Bayes gave 91.58%, LDA gave 94.44%, Neural Network (Feed FW) 94.39% Net, Neural Network (BPA) gave 96.26% Neural net, SOM gave 91.66%, SOM (Supervised) gave 93.55%, Classification Tree gave 92.59%, Random Forest gave 87.96%.

2.4 Scope of the CBR and AI Techniques Used for Astrological Prediction

Astrology can be a good application for artificial intelligence based prediction system because as described by Kelly (1997), there is little unity between various astrologers on issues in astrology and they also don’t agree on the ways to settle differences between astrological concepts and theories. So by making used of different learning methods on huge amount of data available for the planetary position of the people and their biographies a system can be developed for prediction and thus basic issues and differences among various astrological techniques can solved. Thus, it creates a set of predictive methods for various astrological predictions.

Astrological domain requires a lot of knowledge and research to prove the validity. We are planning to perform astrological prediction by using case based reasoning method along with various artificial intelligent techniques so that the
system can be taught to perform prediction based on the learning it does from AI techniques. The rules will be generated based on the data set and the type of prediction to be performed. Various algorithms for each and every type of prediction will be tested and the algorithm giving the best result for a given type of prediction will be used to perform the prediction.

2.5 **Summary of Chapter**

This chapter reviews the literature of the astrology that is our research area. It also provides glimpses of successful researches done to support scientific base for astrology. It also discusses the review of various successful applications implemented using case based reasoning. Various classification techniques are used for learning and prediction hence review of successful implementation of classification techniques is done in diverse fields such as weather prediction, financial analysis, medical prediction, identification of system crashes and so on.