

Analysis study on Predictive based on Knowledge Discovery Decision Tree (KDDT- DSADS) using Machine Intelligence Learning

Dr.K.Kumaravel, Professor & Head, Dept of Computer Science, Dr.N.G.P. Arts & Science College(Autonomous), Coimbatore, India.

Mr.S.Dineshkumar, Asst.Professor, Dept of Computer Science, Dr.N.G.P. Arts & Science College (Autonomous), Coimbatore, India.

Abstract:

Astrology can be an interesting example of the application of various classification techniques of artificial intelligence. In astrology, predictions about different aspects of human life are done based on the planetary position of the stars at the time of birth of a person. In this research work, the positions of the planets and stars at the time of the birth of a person are utilized.

Astrology is the area that generates lots of curiosity among human beings. To find scientific theories and its relations with astrology, changes are to be done in the way astrology is looked upon. To find scientific basis to astrology changes are to be made from the traditional ways in which astrology is analyzed. In this paper, efforts have been made to identify the scientific basis using various classification techniques of artificial intelligence. These techniques are used to perform classification of data automatically. Various classification techniques used here have resemblance with the way astrology works. First step for prediction

Keywords: KNN, DSADN, KDDT, Bayesian, ANN etc.

INTRODUCTION:

Machine learning, by its definition, is a field of computer science that evolved from studying pattern recognition and computational learning theory in artificial intelligence. It is the learning and building of algorithms that can learn from and make predictions on data sets.

We need machine learning in the following cases:

- Human expertise is absent. E. g. Navigating on Mars.
- Humans are unable to explain their expertise. E. g. Speech Recognition.
- Solution changes with time E. g. Temperature Control.

Solution needs to be adapted to particular cases. E. g. Biometrics

- Problem size is too vast for our limited capabilities. E. g. Calculating webpage ranks.

Consider the recognition of spoken speech, where an acoustic speech signal is converted to ASCII text. The pronunciation of a word may vary from person to person due to differences in age, gender or pronunciation, so in machine learning, the approach is to collect a large collection of sample utterances from diverse people and learn to plot these to words. As another example, consider routing packets over a computer grid. The trail maximizing the quality of service from source to destination changes regularly as the system traffic changes. A learning routing procedure is able to adapt to the best path by monitoring the network traffic.

Machine learning involves two types of tasks:

- Supervised machine learning: on a pre-defined set of “training examples”, which then facilitate its ability to reach an accurate conclusion when given new data.

- Unsupervised machine learning: bunch of data and must find pat therein. Consider a situation wherein we algorithm to make predictions. Our Where and are constants.

so in machine learning, the approach is to collect a large collection of sample utterances from diverse people and learn to plot these to words. As another example, consider routing packets over a computer quality of service from source to destination changes regularly as the system traffic changes. A learning routing procedure is able to adapt to the best path Machine learning involves two types of tasks:- learning: The program is “trained” defined set of “training examples”, which then facilitate its ability to reach an accurate conclusion when learning: The program is given a bunch of data and must find patterns and relationships

OVERVIEW OF MACHINE LEARNING METHODS

In this paper we are going to concentrate on supervised machine learning methods. The basic steps used in machine learning are first of all collecting the dataset and then identifying and isolating the features which will be informative. If the dataset contains noise or missing features then preprocessing is done [3]. Next step is identification of learning algorithm to be used. During identification, different algorithms are evaluated then the algorithm with highest prediction accuracy is selected. Description of most frequently used machine learning algorithms in the field of prediction and analysis is provided

Bayesian networks this is the well-known statistical technique for machine learning. This technique provides the probability that the particular instance belongs to which class. Comprehensive book on Bayesian networks is by Jensen's in 1996[7]. The Bayesian network structure [8] is a directed acyclic graph (DAG) and the nodes of this graph defines features. The arcs represent casual influences among the features while the lack of possible arcs defines conditional independencies. The task of learning a Bayesian network is divided into the learning of the DAG structure of the network, and then determination of its parameters. Probabilistic parameters are encoded into a set of tables, in the form of local conditional distributions of a variable given its parents. The joint distribution can be reconstructed by simply multiplying these tables. In Bayesian networks if the structure of the network is given then the learning parameter in the Conditional Probability Tables (CPT) is solved by estimating a locally exponential number of parameters from the data provided [4]. Each node in the network has an associated CPT that describes the conditional probability distribution of that node given the different values of its parents. The Neural concepts of the Bayesian networks are used to detect the networks Which ode specifies some relationship of root node.

APPLICATIONS OF MACHINE LEARNING ALGORITHMS

Machine Learning aims to provide increasing levels of automation in the knowledge engineering process, replacing much time consuming human activity with automatic techniques that improve accuracy or efficiency by discovering and exploiting regularities in training data [8]. The ultimate test of machine learning is its ability to produce systems that are us regularly in industry, education, and elsewhere. The Machine Intelligence is key factor which applied the reality of human activities which is correlated to forecast. CART is a decision trees which classifies based on formulated decision support of each root.

C. K Nearest Neighbor K-Nearest Neighbor (KNN): is based on the concept that similar dataset will generally exist inclose proximity to each other as described by Cover and Hart in 1967[9]. If the instances are tagged with a classification label, then label of an unclassified data instance can be determined by observing the class of its K nearest neighbors. The KNN identifies the label of the instance by identifying the single most frequent class label from its k nearest neighbors. K nearest neighbors of an instance is identified by calculating relative distance between instances. This relative distance is determined by using a distance metric. Many different metrics have been presented. The most significant ones are Minkowskyi, Manhattan and Euclidean distances. D. Decision Trees Murthy in 1998[10] provided the overview of work.

K-Nearest Neighbor which two planets position which have matched with the frequency of the application with nearest planet position to be occupied with an place of ruling planet nearest to node combination two or more nodes in the root path which may be classified the planets revolving and movement of its next position or next hop.

ANNs are inspired by the architecture of biological neural networks, consisting of simple interconnected units called neurons. Neuron computes a weighted sum of its inputs and generates an excitatory (positive) output if the sum exceeds a certain threshold else generates inhibitory (negative) output. Zheng in 2000[18] provided an overview of existing work in Artificial Neural Networks (ANNs). Artificial Neural Networks is composed of an input layer (having neurons for all input variables), a hidden layer (consisting of any number of hidden neurons), and an output layer. Each neuron processes its inputs and transmits its output value to the neurons in the subsequent layer. Each such connection between neurons is assigned a weight during training. The output of hidden neuron is computed by applying an activation function to the weighted inputs and its bias term. The ANN is initialized with random weights. The network learns by adjusting the weightings when presented with a combination of inputs and

outputs that are known as the training set. There are several training algorithms that can be used to train the network each having particular area. Based on the Knowledge Discovery which can be already the node combination of two or three planets position will be calculated. It is also predictive based on search Tree method using ancestor behaviour planets.

Decision Stump Algorithm Decision Stump (KDDT- DSADS)

It's one-level decision tree. It is a decision tree with one root node connected to leaf node. A decision stump makes a prediction based on the value of just a single input feature hence are also called 1-rules. Based on input features, different types of trees are created. For nominal features, a decision tree which contains a leaf for each possible feature value is constructed or a decision tree with two leafs is created. One leaf corresponds to some chosen category and the other leaf for all other categories. For continuous features, usually, some threshold feature value is selected, and the decision tree containing two leafs are create

CONCLUSION:

An astrologer predicts different aspect of human life based on the similarity of important aspect of planetary positions, houses, zodiac signs and transition of planets. The astrologer uses the diagnosis and prediction of the previous horoscopes of the person in similar fields to determine prediction of the person in front of him/her. Hence above methods can provide a good ground for performing researches in the field of astrology. Based on our existing survey we theoretical presume that using the mentioned tools, the proposed research will give best results for astrological prediction. Practically verification of our assumptions is required and need to be thoroughly tested which will be done during the research.

REFERENCES:

- [1].Zhang, S., Zhang, C., Yang, Q, Data Preparation for Data Mining. Applied Artificial Intelligence, 2002, Volume 17, pp. 375 – 381.
- [2] M. Zickus, A. J. Greig and M. Niranjana, Comparison Of four machine learning methods for predicting PM10 concentration , Helsinki finland Water, Air, and Soil Pollution , Kluwer Academic Publisher , 2002, Focus 2, pp. 717–729.
- [3] Agresti. A., “Introduction to Categorical Data Analysis”, in journal Computational Statistics & Data Analysis ,New York, Wiley & S Sons, ISBN 0-471-11338-7 , 1996, pp. 565-563.
- [4] Van Stiphout, E.O. Postma , V. Valentini , P. Lambin The contribution of machine learning to predicting cancer outcome , R.G.P.M.
- [5] Jensen F., An introduction to Bayesian networks, Springer,1996. [8] S. B. Kotsiantis,” Supervised Machine Learning: A Review of Classification Techniques”, in Informatics, 2007 pp. 249-268. [9] Cover T., Hart P., Nearest neighbor pattern classification, IEEE Transactions on Information Theory, vol. 13(1), pp. 21–7.
- [6] Ken McRitchie. August 2011. Support for Astrology from the Carlson Double-blind Experiment. ISAR International Astrologer.
- [7] Penny Seator. 2008/2009 Astrological Prediction and Statistical Tests. The British Astrological Association.
- [8] S. B. Kotsiantis, I. D. Zaharakis, P. E. Pintelas. 2007. Machine learning: a review of classification and combining techniques. Springer Science Business Media
- [9] B Folorunsho Olaiya, Adesesan Barnabas Adeyemo. 2012. Application of Data Mining Techniques in Weather Prediction and Climate Change Studies. I. J. Information Engineering and Electronic Business.
- [10] Joseph A. Cruz and David S. Wishart. 2006. Applications of Machine Learning in cancer prediction and prognosis cancer.Informatica

- [11] Harold Somers. June 1999. Review Article: Example based Machine Translation. Machine Translation. [10] Neelam Chaplot, Praveen Dhyani, O.P.Rishi. March 2012.
- [12] V.S. Desai, J.N. Crook & G.A, Overstreet,” A comparison of neural networks and linear scoring models in the credit union environment”, in European Journal of Operational Research, 1996, vol. 95(1), pp. 24– 37.
- [13] Ynawen Dong, “A Case Based Reasoning System for Customer Credit Scoring: Comparative Study of Similarity Measure”, in Proceedings of 51st Annual Meeting of the International Society for the Systems Sciences, Tokyo, Japan, 2007, pp. 5-10.
- [14] N.C. Hsieh, Hybrid mining approach in the design of credit scoring models in Expert Systems with Applications, 2005, vol. 28(4), pp. 655–665.
- [15] K.S. Shin, and I. Han, A case-based approach using inductive indexing for corporate bond rating, Decision Support Systems, 2001, vol. 32, pp. 41–52.
- [16] Simes R.J, Treatment selection for cancer patients: application of statistical decision theory to the treatment of advanced ovarian cancer, J. Chronic disease, 1985, vol38, pp. 171–86.
- [17] Maclin P.S., Dempsey J, Brooks J, Using neural networks to diagnose cancer, Medical System, 1991 , vol. 5 pp. 11–9.

Authors



Dr.K.Kumaravel Working as a professor & Head, Department of Computer Science, he has more experience in teaching and Research. He has contributed more 40 research publications with UGC, Scopus, Sci, Thomson Reuters, Google scholar with high impact factor and presented paper in various international conference seminar in overseas and around india and also delivered various keynote and guest address . Organized and conducted various programmes related to his academic and research recently he got Dr.Sarvepillai Radhakrishna life time achievements award for his dedication in research and consultancy with social impact values.



Mr.S.Dineshkumar working as a Asst.Professor, Computer Science he has more than 7 years experience in teaching and 3 years in Industry. He is Oracle Certifier and sound Technical knowledge in curriculum . He is participated in various academic activities currently pursuing his Doctorate degree.