Recommendation of food tourism using Artificial Neural Network – A survey

Suruthi P.
psruthi557@gmail.com
KPR Institute of Engineering and Technology, Coimbatore, Tamil Nadu

Varsa Sukirthana S.
sukirdhanasubramaniam13@gmail.com
KPR Institute of Engineering and Technology, Coimbatore, Tamil Nadu

Raguvaran S.
s.raguvaran@kpriet.ac.in
KPR Institute of Engineering and Technology, Coimbatore, Tamil Nadu

ABSTRACT

According to "Global Report on Food Tourism" of the World Tourism Organization (UNWTO), food tourism is commented as a fast-growing segment of the tourism industry. Food tourism means travelling to seek enjoyment via eating and drinking experiences at the destination movies. Such a system can suggest a set of recipes to users based on their interest, or the popularities of the dishes. For tourism services like "Trip Advisor", the most successful online forum in the travel and tourism industry which recommends the food items based on the search history. Here, we using the ANN algorithm to predict which algorithm is better to analyses the datasets of food recipes. Finally, it can make some recommendations for the tourist with better-personalized travelling experience and food services.

Keywords— Analysis, ANN Algorithm, Datasets, Neural Networks

1. INTRODUCTION

Although frequently referred to as "black boxes", artificial neural networks (ANN) find increasing application in intelligent- and recommender systems in a wide range of industries. In travel and tourism research ANNs have, however, not been extensively used so far.

This is despite the fact that first empirical studies in peer-reviewed tourism journals have already been published in the late nineties, introducing ANNs as a valid alternative to traditional regression-based approaches, mostly with regard to demand forecasting purposes.

This study aims to demonstrate the particular advantages and shortcomings of ANN-based applications using an empirical case example. For this purpose, this study uses data from a survey on attitudes and expenditures of tourists in Sarajevo, Bosnia and Herzegovina, conducted by the Institute for Tourism, Zagreb during summer2010. In particular, a multilayer perception-based key-driver analysis is performed on the data to obtain insight into those destination attributes that have a predominant influence on the overall tourist experience in Sarajevo. Finally, results from the ANN-based analysis are opposed to results from a (traditional) regression-based key-driver analysis in order to identify possible significant differences between the approaches.

2. RELATED WORK

Tseng and Lin [1] thought that the service and location are the influence factors of customer behavior in a mobile service environment; they proposed a method named SMAP-Mine to mine customer behaviors [15]. Ma et al. took the time context into consideration and constructed a temporal sequence of mobile access patterns mining model based on context awareness [16]. They have used the Apriori algorithm for this project. The algorithm scans database, accumulates each item count, collects the items which meet the minimum support (min_sup), finds out the frequent 1-itemsets, and named it. Then, the algorithm uses to find out the frequent 2-item sets and uses to find out the frequent 2-item sets and so on and keeps doing these until it cannot find out the frequent -itemsets. In these frequent itemsets, it will be defined as a strong-association rule if it reaches minimum confidence [12].

Dong-sheng Liu and Shu-jiang Fan [2] analyze the context which influences the tourist behavior patterns, select the main context factors, and construct the tourist behavior pattern model based on it; then, calculate the interest degree of the tourist behavior pattern and mine out the rules with high interest degree with the association rule algorithm; It can make some recommendations for the tourist with better-personalized travelling experience and services.

Lori Cameron and Michael Martinez [3] was created the "Trip Advisor", the most successful online forum in the travel and tourism industry. Analysts mine social media and other outlets for "sentiment analysis" and use that information to study what people like and don't like and, more importantly, why so that they can develop better marketing campaigns.

Fonseka P.D.K.P, Kaluarachchi S. J., Jayaneththi D.D.M, Thennakoon D.B.S.P. N Vithana[4] research is focusing to developing a web-based solution to find out most frequent and potential customers who arrive at the tourism companies and important details about them. The data mining techniques will

be applied to find out this information using the data collected from tourism companies. And discovered knowledge will be presented in graphs and send this information Tourism Company management. So this information is more valuable to change their businesses into effective ways by using Time series algorithm.

Paula Odette Fernandes, João Paulo Teixeira, João Matos Ferreira and Susana Garrido Azevedo[5] was developed the methodology has been considered interesting in economics and business areas since it is viewed as a valid alternative to the classical forecast approaches event in complex situations. The artificial neural network methodology suggested that this method should be used in time series with non-linear behavior.

3. FOOD TOURISM RECOMMENDATION SYSTEM

3.1 ANN Algothirm
The network to represent probabilistic representation. This represents a set of random variables. Artificial neural networks are used in sequence and pattern recognition systems, data processing, robotics, modelling, etc. ANN acquires knowledge from their surroundings by adapting to internal and external parameters and they solve complex problems which are difficult to manage. It generalizes knowledge to produce adequate responses to unknown situations.

Artificial neural networks are flexible and have the ability to learn, generalize and adapts to situations based on its findings.
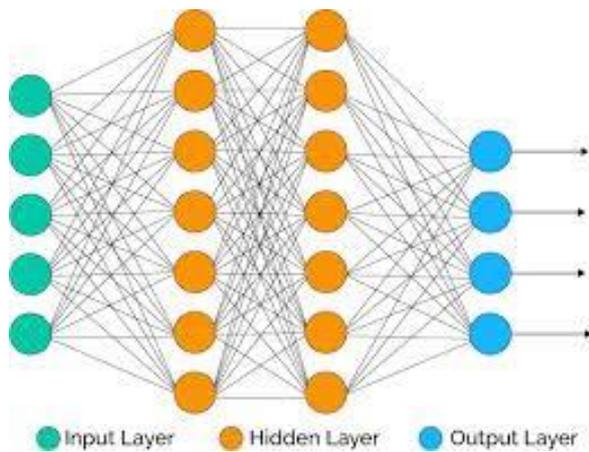

Fig. 1: Structure of ANN

Considering that the above-mentioned information may be difficult to obtain for the most recommendation system, recommendation model based on ANNs that does not need any extra information aside from the interaction between users and food items.

3.2 K-Means Algorithm
The method of K-means algorithm as follows:
(a) Determine the number of clusters k as in shape. To determine the number of clusters K was done with some consideration as theoretical and conceptual considerations that may be proposed to determine how many clusters.

(b) Generate K centroid (the centre point of the cluster) beginning at random. Determination of initial centroid done at random from objects provided as K cluster, then to calculate the in cluster centroid next, use the following formula:

$$v = \frac{\sum_{i=1}^{n} x_i ; i = 1,2,\dots,n}{n}$$

v: Cluster centroid
n: The number of objects to members of the cluster
x_i: The object to-*i*

(c) Calculate the distance of each object to each centroid of each cluster. To calculate the distance between the object with the centroid author using Euclidian Distance.

$$d(x,y) = ||x - y|| = \sqrt{\sum_{j=1}^{n} (x_i - y_i)^2}$$

x_i: object x to-*i*
y_i: object y to-*i*
n: the number of objects

(d) Allocate each object into the nearest centroid. To perform the allocation of objects into each cluster during the iteration can generally be done in two ways, with a hard K-means, where it is explicitly every object is declared as a member of the cluster by measuring the distance of the proximity of nature towards the centre point of the cluster, another way to do with fuzzy C-Means.
(e) Do iteration, then specify a new centroid position using equation (1). Through the clustering process, it can analyze the dataset with those requirements to find out the better results.

3.3 Decision tree
A decision tree is a flowchart-like tree structure, where each internal node represents a test on an attribute, each branch represents an outcome of the test, class label is represented by each leaf node (or terminal node). Given a tuple X, the attribute values of the tuple are tested against the decision tree. A path is traced from the root to a leaf node which holds the class prediction for the tuple. It is easy to convert decision trees into classification rules. Decision tree learning uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. It is one of the predictive modelling approaches used in statistics, data mining and machine learning. Tree models where the target variable can take a finite set of values are called classification trees, in this tree structure, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. A decision tree can be constructed relatively fast compared to A decision tree is a flowchart-like tree structure, where each internal node represents a test on an attribute, each branch represents an outcome of the test, class label is represented by each leaf node (or terminal node). Given a tuple X, the attribute values of the tuple are tested against the decision tree. A path is traced from the root to a leaf node which holds the class prediction for the tuple. It is easy to convert decision trees into classification rules. Decision tree learning uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. It is one of the predictive modelling approaches used in statistics, data mining and machine learning. Tree models where the target variable can take a finite set of values are called classification trees, in this tree structure, leaves represent class labels and branches represent conjunctions of features that lead to those class labels.

A decision tree can be constructed relatively fast compared to a decision tree may be a flowchart-like tree structure, where each internal node represents a take a look at on associate attribute, every branch represents the associated outcome of the take a look at, category label is depicted by every leaf node (or terminal node). Given a tuple X, the attribute values of the tuple square measure tested against the choice tree. A path

is derived from the foundation to a leaf node that holds the class prediction for the tuple. It's simple to convert decision trees into classification rules. Call tree learning uses a choice tree as a prognosticative model that maps observations regarding the associated item to conclusions regarding the item's target worth. It's one amongst the prognosticative modelling approaches used in statistics, data processing and machine learning. Tree models wherever the target variable will take a finite set of values square measure known as classification trees, during this tree structure, leaves represent category labels and branches represent conjunctions of options that cause those category labels. A decision tree is created comparatively quick compared to other strategies of classification. SQL statements are constructed from a tree that may be accustomed to accessing databases efficiently. Call tree classifiers get similar or higher accuracy compared with alternative classification strategies.

Variety of knowledge mining techniques have already been done on academic data processing to enhance the performance of students like Regression, Genetic formula, Bayes classification, k-means clump, associate rules, prediction etc. data processing techniques are utilized in academic field to enhance our understanding of learning method to specialize in identifying, extracting and evaluating variables associated with the learning method of scholars. Classification is one amongst the foremost frequently. The C4.5, ID3, CART call tree square measure applied on the info of scholars to predict their performance. These algorithms square measure explained below

3.4 ID3 rule
Iterative Dichotomize three may be an easy call tree learning algorithm introduced in 1986 by Quinlan Ross. It's serially implemented and supported Hunt's rule. The fundamental plan of ID3 rule is to construct the choice tree by employing a top-down, greedy search through the given sets to test every attribute at each tree node. Within the call tree method, the info gain approach is mostly wont to determine appropriate property for every node of a generated decision tree. Therefore, we will choose the attribute with the highest info gain (entropy reduction within the level of maximum) because they take a look at the attribute of the current node. During this manner, the information required to classify the coaching sample set obtained from short partitioning is the tiniest. So, the use of this property for partitioning the sample set contained in the current node can create the mixture degree of different types for all generated sample subsets reduced to a minimum. Hence, the employment of AN scientific theory approach will effectively cut back the specified dividing variety of object classification.

4. PRINCIPAL COMPONENT ANALYSIS
4.1 Definition of PCA
The goal of the PCA technique is to find a lower dimensional area or PCA area (W) that is employed to rework the information (X=)from the next dimensional area (RM) to a lower dimensional space (Rk), wherever N represents the overall range of samples or observations and xi represents Ith sample, pattern, or observation. All samples have the same dimension (xi∈ RM). In different words, every sample is pictured by M variables, i.e. every sample is pictured as a degree in M-dimensional area. The

direction of the PCA area represents the direction of the utmost variance of the given data as shown in Figure one. As shown within the PCA area is consists of a number of PCs. every principal part incorporates a completed different lustiness in keeping with the number of the variance in its direction.

4.2 Principal elements (PCs)
The PCA house consists of k principal elements. The principal elements are orthonormal, un-correlated b, and it represents the direction of the most variance. The first principal element ((P C1 or v1) ∈ RM × 1) of the PCA house represents the direction of the most variance of the info, the second principal element has the second largest variance, and so on. Figure one shows however the first information area unit transformed from the first house (RM) to the PCA house (Rk). Thus, the PCA technique is considered an orthogonal transformation because of its orthogonal principal elements or axes rotation due to the rotation of the first axes. There area unit two methods to calculate the principal elements. The first technique depends on calculative the covariance matrix, while, the second uses the SVD technique.

4.3 Support Vector Machines
Let E= f(x1; y1); (x2; y2); (xl; yl)g, wherever xi a pair of RN and Loloish a pair of f1; 1g be a collection of coaching examples for a 2-category classier. Suppose the coaching knowledge is linearly separable. Then it's attainable to and a hyper plane that partitions the N-dimensional pattern area into 2 half spaces R+ and R. The set of such hyper planes is given by fw;b(x) = sign(wx+b). SVM selects among the hyper planes that properly classify the coaching set, the one that minimizes kwk 2, which is that the same as hyper plane that the margin of separation between the two categories, measured on a line perpendicular to the hyper plane, is maximized. If the goal of the classification drawback is to and a linear classier for a non-separable coaching set, a replacement set of weights, referred to as slack weights (measuring the extent to that the constraints are violated) is introduced During this case the margin is maximized, paying a penalty proportional to the cost of constraint violation. The choice perform is similar to the one for the linearly dissociable drawback. If the coaching examples aren't linearly dissociable, the SVM works by mapping the coaching set into a better dimensional feature area victimization AN acceptable kernel perform. Therefore, the matter is resolved victimization linear call surfaces within the higher dimensional area.

Any consistent training set (i.e., one within which no instance is appointed a lot of that one category label is created dissociable with AN acceptable selection of a feature area of a sufficiently high spatiality. However, in general, this will cause the training algorithm to overt the coaching knowledge leading to poor generalization. During this paper, though, we tend to consider linear classifiers, however, discuss extensions to non-linear cases within the final section. For the experiments during this paper, we tend to used SVMlight3:50 implementations of SVM formula, which will handle large knowledge sets, as against the standard quadratic improvement programs that have tight limitations relating to memory and time. SVM light is presently one among the foremost widely used implementations of SVM formula.
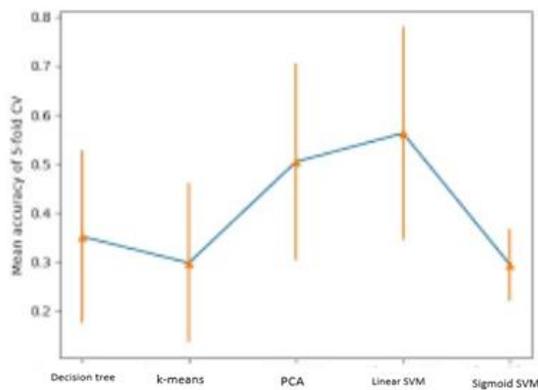
The graph shows that SVM predicts the highest accuracy when compared with other algorithms.

A decision tree is a flowchart-like tree structure, where each internal node represents a test on an attribute, each branch represents an outcome of the test, class label is represented by each leaf node (or terminal node). Given a tuple X, the attribute values of the tuple are tested against the decision tree. A path is traced from the root to a leaf node which holds the class prediction for the tuple. It is easy to convert decision trees into classification rules. Decision tree learning uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. It is one of the predictive modelling approaches used in statistics, data mining and machine learning. Tree models where the target variable can take a finite set of values are called classification trees, in this tree structure, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. A decision tree can be constructed relatively fast compared to other methods of classification. SQL statements can be constructed from a tree that can be used to access databases efficiently. Decision tree classifiers obtain similar or better accuracy when compared with other classification methods.

A number of data mining techniques have already been done on educational data mining to improve the performance of students like Regression, Genetic algorithm, Bays classification, k-means clustering, associate rules, prediction etc. Data mining techniques can be used in the educational field to enhance our understanding of the learning process to focus on identifying, extracting and evaluating variables related to the learning process of students. Classification is one of the most frequently. The C4.5, ID3, CART decision tree are applied to the data of students to predict their performance. These algorithms are explained below-

5. CONCLUSION
The comparison of different algorithms to find the better results to food destination for the people using the neural networks. Based on the conclusions obtained by the researcher wants the results especially in the field of tourism in order to be used as a reference to improve quality and publications related to food tourism that has been considered popular. This is certainly in order to bring in more tourists-both from within the country and abroad.

6. REFERENCES
[1] Burger, C. J. S. C., Dohnal, M., Kathrada, M., Law, R. (2001) A practitioners guide to time-series methods for tourism demand forecasting – a case study of Durban, South Africa, Tourism Management, 22 (4), pp. 403-

[2] Chen, C. F., Lai, M. C., Yeh, C. C. (2012) Forecasting tourism demand based on empirical mode decomposition and neural network, Knowledge-Based Systems, 26 (February), pp.281-28

[3] Cho, V. (2003) A comparison of three different approaches tor tourist arrival forecasting, Tourism Management, 24 (3), pp. 323-33

[4] Hu, H., Lee. Y., Yen, T., Tsai, C. (2009) Using BPNN and DEMATEL to modify importance-performance analysis model – A study of the computer industry, Expert Systems with Applications, 36 (6), pp. 9969-9979

[5] Kon, S. C., Turner, W. L. (2005) Neural network forecasting of tourism demand, Tourism Economics, 11 (3), pp. 301-328

[6] Law, R. (2000) Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting, Tourism Management, 21 (4), pp. 331-340.

[7] Law, R. (2001) The impact of the Asian financial crisis on Japanese demand for travel toHong Kong: a study of various forecasting techniques, Journal of Travel & Tourism Marketing, 10 (2-3), pp. 47-65

[8] Mikulić, J, Paunović, Z. I Prebežac, D. (2012) An extended neural network-based importance-performance analysis for enhancing wine fair experience, Journal of Travel &Tourism Marketing 29 (8), pp. 744-749

[9] M. Ge, M. Elahi, I. Fernández-Tobías, F. Ricci, D. Massimo, 2015, "Using Tags and Latent Factors in a Food Recommender System", The 5th International Conference on Digital Health, 18-20 May 2015, Florence, pp. 105-112.

[10] Chun-Yuen Teng, Yu-Ru Lin, Lada A. Adamic, 2012, "Recipe recommendation using ingredient networks", 4th Annual ACM Web Science Conference, 22-24 June 2012, Illinois, pp. 298-307.

[11] F. Chollet, 2015, "Keras", https://keras.io

[12] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. J. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Józefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. G. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. A. Tucker, V. Vanhoucke, V. Vasudevan, F. B. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, 2016, "TensorFlow: a system for large-scale machine learning". Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation, Berkeley, pp. 265-283.

[13] Y. Bengio, 2009, "Learning Deep Architectures for AI", Foundations and Trends® in Machine Learning archive, Vol. 2 Issue 1, January, pp. 1-127.

[14] V. Nair, G. E. Hinton, 2010, "Rectified Linear Units Improve Restricted Boltzmann Machines", Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, pp. 807-814.