

Best Match Selection with Preference Priority using Hunt's algorithm

A.Arjuna Rao¹, K. Sujatha², M. Sankara Sastry², V.Venkata Rao³

¹Professor and Principal, Miracle Educational Society Group of Institutions, Bhogapuram, Vizianagaram.

² Associate Professor, Miracle Educational Society Group of Institutions, Bhogapuram, Vizianagaram

³MCA Student, Miracle Educational Society Group of Institutions, Bhogapuram, Vizianagaram.

Abstract— Now-a-days there are several marital status websites that don't seem to be used properly because of the issues like they're business oriented and not reliable. The proposed system, Best Match Selection with Preference Priority, that is a marital website developed to boost responsibility with less issues. This helps in checking out data reality of brides and grooms related to concerned registration done with reference and is termed Lifemate. The system permits to make the mind upon to stay honest in efforts of married and deliver the foremost effective results whereas not commercial. The seriousness concerning the priority is finding the acceptable right life partner basing on preferences. Finally, marital status matchmaking isn't simply a business, however additionally a social cause. This perspective makes Lifemate distinctive amongst many alternative Matrimony and marital internet sites. Reliable matches are electing by drawing decision tree by victimization Hunt's algorithm. A decision tree is considered to be a structure that encompasses a root node, branches, and leaf nodes. Each internal node denotes a check on degree attribute, each branch denotes the tip results of a check, and every leaf node holds a class label. The highest node at intervals the tree is that the basis node. The choice creating victimization Hunt's algorithm is reliable and therefore the preferences don't seem to be hindered.

Keywords— Best Match Selection with Preference priority(BMSPP), Data Reality, Lifemate, Reliable matches, decision tree, node, Hunt's algorithm

1. INTRODUCTION

Match creating forms in lead role in choosing the bride or groom as per the preferences. Earlier days there were middlemen who rapidly worked on helping people choose best matches as per their status and requirements. The middlemen act consequently in showing the main points and pictures of the bride or groom as outlined. Their roles are listed as follows.

- (1) Assortment of profiles.
- (2) Reveal profiles to acceptable match.
- (3) Easy search for match.
- (4) Horoscope checking if desired.
- (5) Convince both sides
- (6) Help them in searching better match

However over and over the world is confined to geographical and that they physically fail to point out the appropriate match. This on-line wedding web site provides the means that to pick best match for brides and grooms. It provides decision tree implementation mistreatment Hunt's formula, formula to decide on right person. Manually choosing best match take longer and that they search solely in restricted area. Currently there are several marital status internet sites however they're industrial not reliable in nature. It should contain a lot of pretend profiles. Projected system provides best match choice and it will hold vast quantity of information and solely the reference persons will register during this system and it involves certificate cross validation what the user transfer.

2.LITERATURE REVIEW

Esposito mentioned comparative analysis of ways for pruning decision tree[1]. Niuniu mentioned notice of retraction review of decision trees[5]. The choice tree formula could be a hot purpose within the field of information mining, that is typically accustomed type classifiers and prediction models. Quinlan, J.R mentioned decision trees and decision-making[6]. Sethi, I.K mentioned Entropy nets From decision trees to neural networks[8]. A multiple-layer artificial network (ANN) structure is capable of implementing discretionary input-output mappings. Tsang mentioned decision tree for unsure knowledge[12]. Ancient decision tree classifiers work with knowledge whose values are a unit identified and precise.

Thangaparvathi,B.,Anandhavalli,D.andShalini,S.M. mentioned ahighspeed decision treeclassifier formula forvast knowledge[11].

Information discovery is a crucial tool for the intelligent business to rework knowledge into helpful data which will increase the business revenue. Takahashi, F mentioned decision tree based multicategory support vector machines that area unit termed as Decision-tree-based multiclass support vector machines[10]. In work, at the very best node, we tend to confirm the hyper plane that separates a class (or some classes) from the others. Sheng, Y mentioned decision tree based mostly methodology for top electrical phenomenon fault detection[9]. Paper presents a high electrical phenomenon fault (HIF) detection technique supported decision trees (DTs).

Patel S.B mentioned a totally distinctive approach victimization transformation techniques and decision tree formula on pictures for acting Digital Water Marking[7]. Digital Watermarking is related to rising copyright protection technology. The paper presents a replacement durable watermarking technique supported combining the power of transform domain technique, the separate trigonometric function transform (DCT) and thus the information process technique like decision Tree Induction (ID3). we tend to train the image blocks for account the classification. Friedl M.A mentioned decision tree classification of land cowl from remotely detected knowledge [4].

M.A.Friedl aforesaid his views concerning decision tree classification algorithms have important potential for land cowl mapping issues and haven't been tested thoroughly by the remote sensing community relative to additional typical pattern recognition techniques like most probability classification. Janikow C.Z mentioned fuzzy call trees: problems and ways[3]. Sun, J. mentioned Application of information Mining for decision Tree Model of Multi choice separate Production and Manufacture[2]. During this paper, taking the choice attributes, along side order selection, product selection, person hour and comprehensive analysis into thought, the choice tree model of separate production and manufacture has been bestowed.

3. BEST MATCH SELECTION WITH PREFERENCE PRIORITY USING HUNT'S ALGORITHM

First, preferences are gathered at the time of registration and basing on this next best match selection is done. Match creating exploitation Hunt's algorithm. Match creating is finished by exploitation decision trees supported Hunt's algorithm. The system has the following advantages.

1. Less time overwhelming
2. Covers massive space
3. Suitable match suggestions

3.1 Decision tree Induction

A decision tree could be a structure that features a root node, branches, and leaf nodes. Every internal node denotes a look at on an attribute, every branch denotes the result of a take a look at, and every leaf node holds a category label. The uppermost node within the tree is that the root node.

1. A root node, that has no incoming edges and zero or further outgoing edges
2. Internal nodes maintain exactly one incoming edge and a pair of or further outgoing edges
3. Leaf nodes maintain exactly one incoming edge and no outgoing edges. Each leaf node conjointly contains a class label attached thereto.

A decision tree with preferences is drawn as shown in Figure 1.

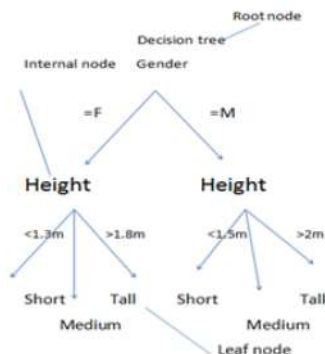


Figure 1 : Decision tree on a chosen sample set

The benefits of obtaining a selection tree area unit as follows:

- This does not would like any domain info.
- This is easy to understand.
- The classification steps of a selection tree area are straightforward and fast.

3.2 Decision Tree Induction Algorithm

Decision tree Induction algorithm is a machine investigator named J. Ross Quinlan in 1980 developed a selection tree rule known as ID3 (Iterative Dichotomies). Later, he bestowed C4.5 that was the successor of ID3. ID3 and C4.5 adopted a greedy approach. During this rule, there's no back tracking; the trees area unit created during a top-down algorithmic divide-and-conquer manner. Tree pruning is performed therefore on get eliminate anomalies inside the employment info because of noise or outliers. The cropped trees area unit smaller and fewer advanced. Tree Pruning Approaches

Here is that the Tree Pruning Approaches listed below :

- Pre-pruning – The tree is cropped by halting its construction early.
- Post-pruning - This approach removes a sub-tree from a completely mature tree.

Cost quality The price quality is measured by the following two parameters :

- vary of leaves inside the tree and
- Error rate of the tree.

3.3. Hunt's Rule

Most of the selection tree induction algorithms are supported original ideas planned in Hunt's algorithm. Let D_t be the coaching job set and y be the set of sophistication labels.

1. If D_t contains records that belong to an analogous class y_k , then its decision tree consists of leaf node tagged as y_k .
2. If D_t is degree empty set, then its decision tree may well be a leaf node whose class label is ready from various data such as a result of the bulk class of the records.
3. If D_t contains records that belong to several classes, then a take a glance at condition, supported one of the attributes of D_t , is applied to separate the information in to plenty of consistent subsets.

3.4 Match creating exploitation Hunt's rule

- Consider preferences.
- Generate decision tree.
- Apply Hunt's rule for choosing best match

4. RESULTS

The projected methodology is developed and tested exploitation the coaching and testing sets as input.

4.1 Training and Test Set

The training set may consists of Name, Marital status, age, caste, job, income, property, horoscope, family background, class label as attributes.

Table 1: Training Set

Name	Marital Status	Age	Caste	Job	Income	Property	Horoscope	Family Background	Class Label
Venky	Unmarried	24	B	Yes	>5.5	Yes	Matched	Good	No
Murali	Unmarried	22	B	Yes	<5.5	Yes	Unmatched	Good	No
Nani	Unmarried	26	F	No	No	Yes	Matched	Good	No

Indu	Divorced	22	F	No	No	No	Unmatched	Good	No
Keerthana	Unmarried	21	B	No	No	No	Unmatched	Good	Yes
Niharika	Unmarried	25	F	No	No	No	Matched	Bad	No
Suneetha	Divorced	22	B	Yes	>5.5	Yes	Matched	Bad	
Gayatri	Unmarried	20	F	No	No	No	Unmatched	Bad	Yes

The sample test set is as given in Table 2.

Table 2 : Test Set

Name	Marital Status	Age	Caste	Job	Income	Property	Horoscope	Family Background	Class Label
Ram	Unmarried	25	B	yes	7lakhs	Yes	Matched	Good	Yes

4.2 Performance Analysis

The performance analysis of the proposed algorithm is done by mistreatment the False Acceptance Rate (FAR) and False Rejection Rate(FRR). For a given threshold, FAR is the consolidated number with in the classes in the training set whose distance with completely different class test sample is larger than the threshold and FRR is the number of classes in the training set whose distance with the equivalent class test sample is larger than the threshold.

Table 3 : Performanse Analysis based on FAR and FRR

Set	No. of Samples	FAR%	FRR%
1	50	96	98
2	150	99	99
3	30	100	97
4	10	90	100
5	60	97	99
6	40	98	95

5. CONCLUSION

The proposed match making website is Reliable and not commercial and eliminates all the problems in existing system. This will generate 100% genuine profile matching. Finally, matrimonial matchmaking is not just a business, but also a social cause. Decision trees are rapidly used in decision making preferences and are always secure and reliable. These methods are used for both prediction and classification as they are easily converted to classification rules. Best matches are selected by drawing decision tree by using Hunt's algorithm using preference parameters. This system is developed and tested on sample dataset and is found to be reliable and assists in making better decisions.

REFERENCES:

- [1] Esposito, F. "A comparative analysis of methods for pruning decision trees". IEEE, 476 – 491, 1997.
- [2] Sun, J. Application of knowledge Mining for decision tree Model of Multi-variety distinct Production and Manufacture. 724-728, 2010
- [3] Janikow, C. Z., Fuzzy decision trees: issues and methods. IEEE Systems, Man, and Cybernetics Society, pp 1-14, 1998.

- [4] M.A. Friedl, C. B., Decision tree classification of land cover from remotely sensed data. pp 399-409, 1997.
- [5] Niuniu, X. Notice of Retraction Review of decision trees. IEEE, pp 105-109, 2010.
- [6] Quinlan, J. R, Decision trees and decision-making. IEEE, pp 339 – 346, 1990.
- [7] S.B.patel, A novel approach exploitation transformation techniques and decision tree formula on pictures for performing Digital Watermarking. IEEE, pp 1-6, 2009.
- [8] Sethi, I. K, Entropy nets: from decision trees to neural networks. IEEE, pp 1605 – 1613, 1990.
- [9] Sheng, Y, Decision tree-based methodology for high impedance fault detection. IEEE, pp 533 – 536,2004.
- [10] Takahashi, F, Decision-tree-based multiclass support vector machines. IEEE, pp 1418 – 1422, vol.3,2002.
- [10] Thangaparvathi, B., Anandhavalli, D., & Shalinie, S. M).,A high speed decision tree classifier algorithm for huge dataset. IEEE, pp 695-700, 2011.
- [11] Tsang, S., Decision Trees for Uncertain Data, pp 441 – 444, 2009.