A MACHINE LEARNING BASED SOLUTION FOR FINDING PERFECT MARITAL PARTNER

BKTP Wickramasinghe[1], DU Vidanagama, and N Wedasinghe

Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka

[1]wickramasinghe95@gmail.com

Abstract- Marriage is a socially or ceremonially perceived joining between mates that sets up rights and commitments between those life partners. Finding a good marriage partner is one of the main reasons for the delay in marriage. Therefore, there is a need for a solution that can get user details and expected partner preferences and suggest proper matches based on their preferences. The objective of this paper is to discuss the necessity of the proposed model for Sri Lankans. The proposed solution will maintain user details and get user's preferences for their matrimonial partner. Based on the preferences, appropriate matches will be displayed to the user using a clustering algorithm along with the matching percentages. Horoscope of the user will be generated based on the planet details of the user. Furthermore, previous birth connection and 'dosha' identification will also be done. The proposed solution will also enable of sending messages to the matches and get email notifications about those matches. The final aim of this solution is to ease the matchmaking business by providing proper matches.

Keywords- Matrimonial Partner, Horoscope, Clustering, Matches

I. INTRODUCTION

Marriage is the union of two people. The definition of marriage varies around the world not only between cultures and between religions, but also throughout the history of any given culture and religion. A marriage ceremony is known as a wedding. In Sri Lanka, during the present century female age at marriage has increased almost by seven years. The delay in marriage has an enormous impact on the birth rate. Unlike in the west, where marriage is not necessarily the prosecutor of childbearing or the responsibility of running a household, in Sri Lanka procreation is almost entirely within the marriage. The United Nations World Fertility Report of 2003 reports that 89% of all people get married before the age of forty-nine. The percent of women and men who marry before the age of forty-nine drops to nearly 50% in some nations and reaches near 100% in other nations. (Un.org,2000) Finding a good marriage partner is one of the main reasons for the delay in marriage. In early times, people were not allowed to fall in love and get married. They had to find their matrimonial partner using traditional matchmaker approach ("Kapuwa") or Matrimonial advertisements. Even though those approaches were famous, there were some disadvantages. According to (Vreede-de Stuers,1969), the matrimonial advertisements that appear in English-language newspapers have attracted scholarly attention in recent years, and quite justifiably, for these items provide an abundance of information admirably adapted to statistical analysis of some of the variables determining mate selection. Yet the limitations of this material are also obvious. Verifying the accuracy of the contents is the main limitation. For instance, the beauty of a girl or the earnings of a boy can be exaggerated, and that may be misleading those who search the columns for an attractive candidate. Furthermore, the marriages happened based on these advertisements have led to divorce in most of the cases. With the advent of the Internet, a new channel in the form of matrimonial Web sites has emerged as an alternative way to find partners for marriageable members of the family. The introduction of technology in the form of matrimonial Web sites in an otherwise socially-enabled process provides the setting for a fascinating exploration of changing social mores and the interaction of technology and society. (Patnayakuni and Seth,2008)
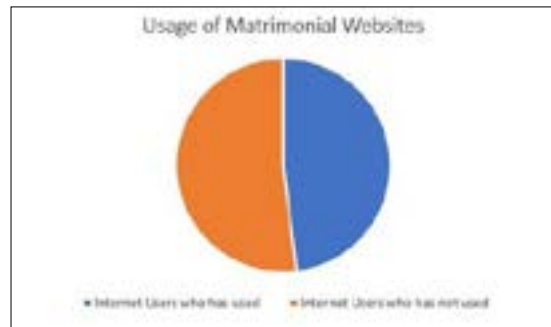
Figure 1. Usage of Matrimonial Websites

In 2015, India is estimated to possess 7% share of the total Internet users of the world and now is ranked 4th in the world in terms of Internet usage. 11% of the share is estimated to come from the Internet users who use matrimony sites (Figure 1). The total increase in growth of matrimonial sites is pegged to be 1500%. (Dasgupta, 2016) Interest in Online matrimonial reflects a new mindset among the urban youth and a major reason for that is the Internet and the large number of matrimonial sites offering a new landscape for romance and the search for a life partner without the traditional trappings or restrictions. Even though there are vast number of matrimonial websites, they do not support key features required in a Sri-Lankan marriage. By considering the facts from research papers and studying about existing systems, it could be understood that there is no proper system in Sri Lanka to find a matrimonial partner and fulfil the other requirements such as horoscope matching, wedding services etc. The main objectives of this research are: (1) Study the problems of existing matrimonial websites (2) Study the techniques and tools that can be used to develop the model (3) Present the developed model to solve the problem. The paper is organized as follows. Section 2 describes the background study. Section 3 describes the methodology. Section 4 presents the evaluation criteria and Section 5 presents conclusion about the framework together with directions in the future.

II. LITERATURE REVIEW

A. Existing Systems

There are vast number of matrimonial websites available in the market. Out of those websites huge number of websites belongs to Indian authors and dedicated to Indian users.

1) Shaadi.com: Shaadi.com is an online wedding service founded by Anupam Mittal in 1997. Shaadi.com began as Sagaai.com in 1997. Shaadi.com saw success over the next fifteen years as Internet adoption increased and people became more receptive to online matchmaking. (Shaadi.com, n.d.) Even though, shaadi.com is quite famous in India, it is not famous in Sri Lanka. It is basically created for India. If a Sri Lankan user is creating a profile, he cannot state the exact location. Location needs to be selected from the given set of values. Since it lacks Sri Lankan users, all the suggestions are for Indian candidates. It can be a huge limitation for those who strictly consider the religion, cast and nationality of their partner.

2) Jeevansathi.com: Jeevansathi.com is an Indian matrimonial portal owned by Info Edge. The website was bought in 2004 by Info Edge Ltd. The website has more than 6.1 million registered members as of March 31, 2014. The organization has around 230 employees working in 54 offices in 37 cities spread across the country, headquartered in Noida. It is listed in Bombay Stock Exchange.

Even-though it is quite famous, it does not provide you with perfect matches. Furthermore, Jeevansathi.com does not support Sri Lankan Candidates. Sri Lankan candidates cannot be searched using this website. (Figure 2) Horoscope matching service is also lacking in this website.

Figure 2. Browse Section of Jeevansathi.com

3) SimplyMarry.com: It is India's premier matrimonial service portal, promoted by the Time of India group. SimplyMarry.com offers one stop platform for online matchmaking that allows prospective Indian brides and grooms to meet and communicate regarding matrimony. Only, limitation in this site is lack of Sri Lankan candidates. Since it is an Indian site, it mainly focuses on Indian candidates. Furthermore, it lacks horoscope matching service. Out of these Indian websites, shaadi.com has been identified as the most user friendly, most used websites. The daily traffic rank trends (a measure of website's popularity) of these matrimony sites over a period of six months, are shown in Fig 3. Traffic rank is calculated using a combination of average daily visitors and page-views. The lesser the rank is of course better for a site in terms of popularity. Trend graph indicates daily traffic rank (charted over time), where shaadi.com leads currently. (Pal,2011)
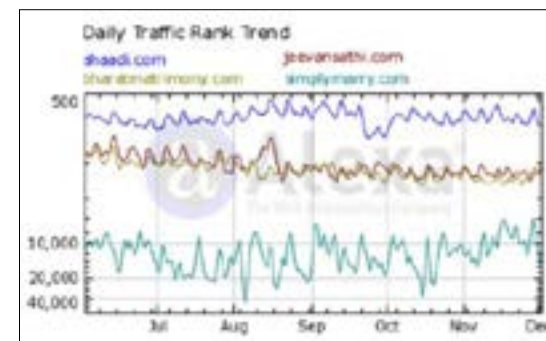

Figure 3. Indian Matrimony Traffic Comparison

Even though shaadi.com is quite famous and trending, Sri Lankan candidates face a great deal in finding a perfect matrimonial partner through these Indian Websites. In the matrimonial market there are Sri Lankan matrimonial sites as well. They lack many important features.

4) SriLankanMatrimony.com: It is an exclusive matrimony portal catering to the special matrimonial needs of Sri Lankans across the globe. The portal offers several benefits to its members, primary being the pleasure of searching for a life partner within one's own community from across the globe at the click of a mouse. Even though there are some interesting features. There are limitations as well. To maintain a profile at SriLankanMatrimony.com, user needs to select a type of membership and needs to make the payment. Limitations in existing systems are summarized as follows. (Table 1)

B. Technology Used

1) Statistical Analysis Approach: The general approach of a statistical Matchmaker is to exploit that users

Table 1. Limitations of Existing Systems

Existing Systems	Limitation
Shaadi.com	• Lack of Sri-Lankan candidates. • Incapability of stating Sri-Lankan towns as location,
Jeevansathi.com	• Lack of Sri-Lankan candidates. • Lack of Perfect match feature. • Lack of Horoscope Matching Feature.
SimpleMarry.com	• Lack of Sri-Lankan candidates
SriLankaMatrimony.com	• Lack of free services
LankaMatrimony.com	• Lack of Perfect match feature. • Lack of Horoscope Matching Feature.

are confirming preferences that they see fitting for themselves, without the need to call for an external expert or framework. A statistical Matchmaker tries to identify similarities between preference sets to come up with expectations or recommendations for a user, based on what user, identified as 'similar' by a distance function - expressed as their preference in the target context.

Dating agencies are using statistical methods to find good matches between people. They analyse the user profiles and their collected data about successful meetings to continuously improve their matching. (Segaran,2007)

In the statistical approach, appropriate algorithms from the field of machine learning and statistical analysis will be used for the matching of user profiles. A classical candidate approach is based on Support Vector Machines (Wassermann and Zimmermann,2011). Online dating services use statistics, data mining, and activity monitoring to provide appropriate matches; thus, differentiating their services and understanding the success of their product offering. (Smith,2005)

2) Clustering: Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters). It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, bioinformatics, data compression, and computer graphics. (En.wikipedia. org,208)

(Gal, n.d.) has conducted a research on how k-means clustering allows to give the perfect match in dating apps. K-means algorithm was used to produce 15 clusters. A relatively high number of clusters was specified since the study deals with human behaviour, and more clusters are generally needed to see trends. Clustering algorithms find groups of observations that are similar to each other and different from the observations in other groups. For the study, both the k-means and EM clustering algorithms were run on the online dating data set to see if patterns could be found in the usage of and attitudes about online dating. In this case, the k-means algorithm was chosen as it had the most interesting results.

(Tausch et al., n.d.) has introduced a tree structure for efficient service matchmaking which was created by using the hierarchical clustering algorithm. Tree nodes represent a superset of all service descriptions in the leaves below. During query processing matchmaking can be restricted to the branches of the tree where tree nodes indicate overlapping between user requests and service descriptions. Good clustering of n service descriptions may improve retrieval time from O(n) to O (log n) for concise queries.

(Pei,2008) states that their experimental study reveals that the Probabilistic-based approach, employing the EM algorithm, yields better results than the Rule-based approach when the requirements to generate the probabilistic parameters are satisfied.

The results are summarized in Table 2. As per the literature review it was revealed that K-means clustering, and Statistical Analysis approach is suitable for the match-making process.

Table 2. Summary on Used Technologies

Machine Learning	Performance	Applicability
Statistical Analysis Approach	Good	Good
K-means Clustering	Good	Good
Hierarchical Clustering	Weak	Weak
EM Algorithm	Good	Weak
Canopy Clustering	Good	Weak
K-Nearest Neighbour	Good	Weak
Artificial Neural Network	Good	Weak

III. METHODOLOGY

A. Data Gathering

Qualitative and quantitative data required for designing the requirement specification for the new system were gathered by conducting a survey.

B. Data Analysis

The data which were gathered during the data collection phase described were analysed by using charts and presented in this section. Through analysing the data, was able to reveal the need of a proper matrimonial web solution for Sri-Lanka.

C. Approach

The users of this system are Candidates and Admin. There are five major different types of inputs from the candidates including Personal Details, Account Details, Education, Expected Partner Details and Horoscope Details. The outputs from the system are basically appropriate matches with horoscope matching details, previous birth connection details and dosha. Admin will generate user reports based on the candidate details.

D. Technology adopted

The input processing has been implemented using selected programming languages Html, CSS, JavaScript and Python. The technology that suits for the system development should be decided by considering the domain and the requirements for the system. It is important to identify most appropriate technological methodologies to satisfy the functional requirements and the non-functional requirements of the system in the system development procedure.

Proposed Matrimonial Solution is a web-based system. Time for the production, Efficiency and the performance of the system, Usability and Flexibility of the system and functionality of the system should be considered in developing the system. Programming language is the most important technical factor that should be focused in implementing the system. The programming language should in compatible with the development tools that are used in implementing the system. HTML, CSS, JavaScript were employed to develop the front end and Python was used to develop the backend. Flask, a python web framework was used and Pycharm IDE has been used.

E. Proposed Design

The architectural perspective of the developed system breaks into four main components as; software design, database design, module design and the interfaces. The gathered data from the analysis phase was used in creating the system's design.

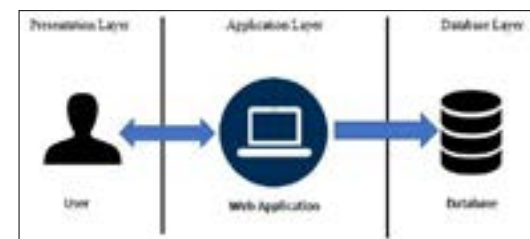#### 1) Overall System Architecture

Figure 4. Overall System Architecture

Client layer: The user accesses the application through the client layer. The developed system provides access for the users in different user levels and each hold the interfaces which satisfy the requirements of each user type.

Application layer: Web Application belongs to the application layer. All the algorithms including porondam matching, clustering, dosha identification and previous birth connection identification are included in this layer.

Database layer: Database layer is responsible to manage the entire database of the system. The database will store the data enters to the system by the web application after the manipulation of the application layer.

2) Modular Architecture

The discussed modules in the software architecture will be elaborated in this section to show how the modules will be dealing with the users' actions in each layer. The main modules and their sub modules are given below.

i. **Login and Authentication Module**

Only the authenticated users can access the system. There are two user levels namely candidates and admin.

ii. **User Registration Module**

Candidates can register through the User Registration Module. Registration Module will be divided into five sub-forms namely Account Details, Personal Details, Education, Expected Partner Details and Horoscope Details.

iii. **Matching Process**

This module includes a clustering algorithm. Expected Partner preferences will be recorded during registration. Based on the preferences, appropriate matches will be displayed to the user using K-means Clustering technique. K-means clustering is a type of unsupervised learning, which is used when you have unlabelled data. The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K. The algorithm works iteratively to assign each data point to one of K groups based on the features that are provided.

iv. **Porondam Match Module**

Horoscope Details of a candidate will be stored in the database during registration. For the matches derived from the clustering algorithm, porondam will be matched. In this process five porondams are being matched.

- Nakath Porondam

   The most important aspect and crucial for a successful relationship between two persons. The Nakath porondam is the checking of whether the natures of the two marriage partners minds or in Sinhalese Sith are compatible.

- Rajju Porondam

   Physical bond of the marriage is through the certain parts of the body starting from the head down to toe. This porondam is about the compatibility of those parts which are producing sensations to the partners.

- Rashi Porondam

   Rashi means constellation. Rashi porondam is for checking and comparison of the status of the subconscious minds for compatibility.

- Nadi Porondam

   Nadi porondam considers the compatibility of the nerves of the two people, for an example if both partners get easily exited or angered then it is very difficult to maintain a long-term relationship.

- Graha Porondam

   This is about the influences of the destiny or in other words karmic factors. In other words, a person has certain control of his or her own destiny and others in the relationship and is checking to see whether this relationship is feasible to produce mutually beneficial outcomes.

\v. **Previous Birth Connection Identification Module**

   To identify the previous birth connection, rashi should be generated separately, for both the horoscopes. Rashi of the horoscopes should be tallied. There are twelve rashis available. They are Aries, Taurus, Gemini, Cancer, Leo, Virgo, Libra, Scorpio, Sagittarius, Capricorn, Aquarius and Pisces.

vi. **Dosha Identification Module**

   Apart from Porondam Match, dosha needs to be identified as well. Dosha will be identified individually for each of the candidate. If a certain horoscope contains a dosha, horoscope of the matching candidate should be able to cope up with the dosha.

vii. **Message Sending Module**

   For the derived matches user should be able to send messages and interact with the matches.

viii. **Feedback & Success Stories Module**

   User can submit feedback. Once a candidate meets someone to marry through this solution they can update the Status and once it happens the couple details will be displayed as a success story. Recent feedback and success stories of the couples will be displayed to the user.

ix. **Email Notification Module**

   A trigger will be fired at the end of the day and the matches derived from the clustering technique will be sent to the user as email notifications.

x. **Reporting Module**

   Admin will generate reports based on the candidate data. A dashboard will be generated.

The developed Machine Learning Based Web Solution is comprised of two main interface categorizations (user-levels) namely interfaces of Candidates and Admin.
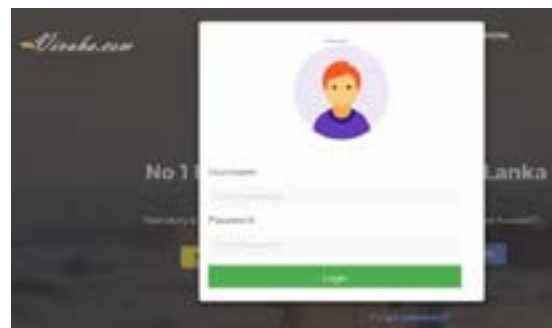

Figure 5. Interface of the Home Page
Figure 6. Login Interface

IV. EVALUATION

Summative evaluation was used as the evaluation method to find how the system functions and whether it is up to the expected level to fulfil the clients' requirements. At the system finalizing stage this evaluation is done to evaluate the product's stability. In summative evaluation a prototype with most stable build is shown to the client and the feedback is taken to find how far the system is success. In here the using prototype must be very much alike to the final product's functions and features.

The overall Evaluation of the product was carried to verify whether the system's final outcome meets the functional requirements of the users and the successfulness of the system tasks and the functions of each component are also evaluated here. This was done by comparing the functions of the new system against the problems and limitations addressed during the Requirements analysis process, also considering the functional requirements specified by the system specifications.

The prototype was given to ten people and they were asked to rate the system based on the following attributes their responses have been recorded and summarized as below (Table 3).

Table 3. Summary on Used Technologies

Parameter	Developed System
Efficiency	95%
Accuracy	75%
Availability	80%
Cost Reduction	98%
User Friendliness	97%

Furthermore, the prototype was given to five astrologers to check the accuracy of the horoscope algorithms. The results are summarized below (Table 4)

Table 4. Summary on Used Technologies

Algorithm	Accuracy
Porondam Match	64%
Previous Birth Connection Identification	96%
Dosha Identification	96%

According to the results, more than 95% of candidates have been satisfied by the mentioned parameters taken by the group for analysing the developed system and most of the astrologers state that the accuracy of horoscope algorithms are high.

V. CONCLUSION

In Sri Lanka, during the present century female age at marriage has increased almost by seven years. (Patnayakuni,2008) Finding a good matrimonial partner has become the major issue. Even-though there are traditional methods, they are not cost effective. Based on the results it is identified that the developed system performs well in finding a perfect marital partner. There are vast number of matrimonial websites available in the market. Many websites belong to Indian market and lacks Sri-Lankan users. The existing Sri-Lankan matrimonial websites are not up to the standard and lack many key features and do not support fee services. This paper proposes a web based matrimonial solution, which will generate appropriate matches based on the preferences along with the horoscope matching percentage. The authors will plan to further improve and refine the techniques, and to deal with the outstanding problems identified. The authors have also planned to generate an algorithm for fake detection of the feedback and the comments. These will be addressed as a further work.

ACKNOWLEDGEMENT

REFERENCES

Un.org. (2000). *World Marriage*. [online] Available at: http://www.un.org/esa/population/publications/worldmarriage/worldmarriage.htm [Accessed 10 Feb 2018].

Vreede-de Stuers, C. (1969). The relevance of matrimonial advertisements for the study of mate selection in India. Bijdragen tot de taal-, land- en volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia, 125(1), pp.103-117.

Patnayakuni, R. and Seth, N. (2008). *Online Matrimonial Sites and the Transformation of Arranged Marriage in India*. [online] Available at: https://www.researchgate.net/publication/266908028 [Accessed 10 Feb 2018].

Dasgupta, S. (2016). "*A Suitable Bride: A Comparative Analysis of Matrimonial Advertisements of Indian Women Published in*

Newspapers and Matrimonial Websites". [online] Bing.com.

shaadi.com. (n.d.). [online] Available at: http://www.shaadi.com [Accessed 2 March 2018].

Pal, J. (2011). [online] Available at: https://www.researchgate.net/publication/256484632 [Accessed 2 March 2018].

Smith, A. (2005). Exploring online dating and customer relationship management. *Online Information Review*, 29(1), pp.18-33.

En.wikipedia.org. (n.d.). Cluster analysis. [online] Available at: https://en.wikipedia.org/wiki/Cluster_analysis [Accessed 15 April 2018].

Gal, M. (n.d.). *Online Dating and Marketing Using Machine Learning*. [ebook]

Tausch, B., d'Amato, C., Staab, S. and Fanizzi, N. (n.d.). Efficient *service matchmaking using tree-structured clustering*. [online] Academia.edu. Available at: http://www.academia.edu/2884228/Efficient_service_matchmaking_using_tree-structured_clustering [Accessed 18 April 2018].

Pei, S. (2008). *A Comparative Study of Record Matching Algorithms*. Edinburgh.

Segaran, T. (2007) *Programming Collective Intelligence*, O'Reilly, pp. 197-224 [10]

Wassermann,B and Zimmermann,G. (2011) "*User Profile Matching: A Statistical Approach*", in The Fourth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services, p. 4.

CLOUD BASED POWER CONSUMPTION ESTIMATION FOR ELECTRIC VEHICLES

P.G.H.L.R. Fernando[1] and P.P.N.V. Kumara

Department of Computer Science, General Sir John Kotelawala Defence University, Sri Lanka
[1]ruk.fernando93@gmail.com

Abstract- Inaccurate range estimation is a major problem which comes with Electric Vehicles. Because of this many people face issues when planning long trips and short trips with limited battery capacity. To overcome this issue, it is necessary to have a better power consumption prediction algorithm which uses vehicle data and other dynamic environmental conditions. This paper is based on cloud based power consumption estimation system which uses linear regression in machine learning to obtain a better estimation based on above mentioned areas.

Keywords- Electric Vehicle, Range, Power Consumption, Estimation

I. INTRODUCTION

Electric Vehicles are now a trending vehicle type in Sri Lanka because of its environmental friendly approach, cost effective fuel consumption, comfortable interior design and cutting edge technology. According to the PUSL report related to electric vehicles (Public Utilities Commission of Sri Lanka, n.d.), from 2012 to 2017 there have happened more than 4000 EV registrations. Considering the size of Sri Lanka 4000 is a great number. But most of the people interested in EVs are not willing to buy an EV and the second hand market of EVs is rapidly decreasing. Industry research has uncovered that this is caused due to "anxiety felt by many drivers about the remaining driving range their vehicle can run before the next charge". This anxiety is mainly because of the current range and power consumption estimation algorithms do not accurately estimate the remaining driving range and required power to complete a journey. Range estimation technologies use limited data to calculate the estimated range such as battery health, state of charge, acceleration information, fixed auxiliary device power consumption and aggregated trip data. And current technologies do not use external environmental data such as traffic condition, speed limits, altitude and weather information to predict range. To overcome this anxiety, it is necessary to have an accurate power consumption prediction mechanism integrated into EVs.
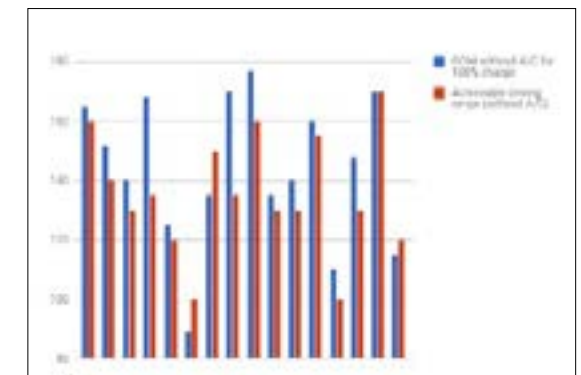

Figure 1. Difference between Estimated Range vs Achievable Range (km)

According to the survey based on Nissan Leaf 24kWh owners in Sri Lanka, Figure 1 shows that there's a considerable deviation on predicted range by inbuilt default prediction model in the vehicle with the achievable range for a full charge.

In order to predict accurate required power consumption to achieve a journey in an electric vehicle, it's necessary to have a prediction model which analyses battery consumption against achieved driving distance including various dynamic facts.

Newspapers and Matrimonial Websites". [online] Bing.com.

shaadi.com. (n.d.). [online] Available at: http://www.shaadi.com [Accessed 2 March 2018].

Pal, J. (2011). [online] Available at: https://www.researchgate.net/publication/256484632 [Accessed 2 March 2018].

Smith, A. (2005). Exploring online dating and customer relationship management. *Online Information Review*, 29(1), pp.18-33.

En.wikipedia.org. (n.d.). Cluster analysis. [online] Available at: https://en.wikipedia.org/wiki/Cluster_analysis [Accessed 15 April 2018].

Gal, M. (n.d.). *Online Dating and Marketing Using Machine Learning*. [ebook]

Tausch, B., d'Amato, C., Staab, S. and Fanizzi, N. (n.d.). Efficient *service matchmaking using tree-structured clustering*. [online] Academia.edu. Available at: http://www.academia.edu/2884228/Efficient_service_matchmaking_using_tree-structured_clustering [Accessed 18 April 2018].

Pei, S. (2008). *A Comparative Study of Record Matching Algorithms*. Edinburgh.

Segaran, T. (2007) *Programming Collective Intelligence*, O'Reilly, pp. 197-224 [10]

Wassermann,B and Zimmermann,G. (2011) "*User Profile Matching: A Statistical Approach*", in The Fourth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services, p. 4.