



HAL
open science

Measuring Successful Digital Services by Identifying Active Users

Sarwar Jahan Morshed, Juwel Rana, Marcelo Milrad

► **To cite this version:**

Sarwar Jahan Morshed, Juwel Rana, Marcelo Milrad. Measuring Successful Digital Services by Identifying Active Users. 6th IFIP International Workshop on Artificial Intelligence for Knowledge Management (AI4KM), 2018, Stockholm, Sweden. pp.48-64, 10.1007/978-3-030-52903-1_5 . hal-03478006

HAL Id: hal-03478006

<https://inria.hal.science/hal-03478006>

Submitted on 13 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Measuring Successful Digital Services by identifying Active Users

Sarwar Jahan Morshed¹, Jewel Rana^{1,2} and Marcelo Milrad¹

¹ Department of Computer Science and Media Technology, Linnaeus University, Vaxjo, Sweden

² Telenor Group, Oslo, Norway

¹sarwar.morshed@lnu.se, jewel.rana@lnu.se, marcelo.milrad@lnu.se

²jewel.rana@telenor.com

Abstract. This paper is an extended version of the work published in [1]. We extended our previous work by the proposed model representation and more experimental result. Almost every service in the current era are either managed or operated with the help of digital services to manage their customer's activities. So digital service providers are engaging in different ways with the users to know their feedback about the service. Traditional methods for understanding user satisfaction from their direct feedback or survey. Both of these methods might produce biased result due the limited user participations. Therefore, knowledge management of user activities with the digital service could be an alternative approach for identifying satisfied users. In the previous work [1], we have presented a data science model. In this paper, more result using the proposed data science model have been presented to prove that the proposed model is workable to measure successful digital services.

Keywords: User feedback, digital service; user satisfaction; data science, prediction model; knowledge management, service evaluation, machine-learning algorithm;

1. Introduction

There is an increasing demand for evolving digital mobile services around the world to facilitate human life and activities. Technology-based startup companies are pioneers in forming digital services in all over of the world. One of the key issues of current businesses is not only to recognize who the customers are – rather to understand users' contextual information such as their location, real-time activi-

ties and ways of communication and interaction with the different services. If the organizations do have relevant customer (of their services/products) information and behavior patterns, it will help them in making better decisions regarding the product or service development. The success of the digital services is represented by an increasing number of loyal customers. On the other hand, the loyalty of users relies on how they utilize and like the features provided by a digital service. However, understanding the powerful features as well as the digital services that are liked by a user is a challenging issue. Besides, users' behaviors are changing continuously along with the rapid moving of the digital age. Therefore, it is required to analysis user activities related to the digital service in order to understand the motivation or level of interests towards it (i.e. how often the user access this service, how the user the service per day, week or month) [2].

It has been a common trend to identify user's gratification with the digital services by getting the users' direct feedback on the determinants of the user satisfaction. In general, user's feedback is collected using any of the traditional forms like filling feedback forms, conducting surveys, oral questionnaires carried out by the survey team, or by getting a rating from the user after using a specific service or the product. In the age of a nomadic lifestyle, all users do not have such patience or time to give these kinds of feedback. Besides, a small number of users usually participate in these forms of feedback. Therefore, a decision based on this feedback may lead to developing the service to a wrong direction. In contrast, users of the different digital mobile services generate a large volume of user activity data that can be used to measure user's satisfaction instead of relying only on user's direct feedback data.

Analyzing such large amount of data using statistical models is difficult. Besides, current businesses require instant analysis of user feedback - which is not possible using the conventional manner of statistical analysis. In this case, appropriate Machine Learning algorithms can be used to analyze data in real time. To the best of our knowledge, there are very few proposals that use activity data to find user satisfaction. Given these circumstances, we are exploring how to develop a data science model that could be able to measure the success of a digital mobile service. The proposed model exploits user's activity data rather than user's direct feedback data for identifying his/her level of satisfaction. Thus, the main contributions of this paper can be enumerated as follows:

- We develop an innovative data science model for discovering satisfied users from digital mobile service unlabeled access login data.
- We demonstrate that the volume of the dataset is not the main prerequisite for building such a predictive model.
- We observe that satisfied users are the key indicator for measuring the success of a digital mobile service.

The remaining of the paper is structure as follows: in section 2, related research efforts are described. Section 3 illustrates the methodology and method of the proposed user satisfaction model. Section 4 depicts the experimental results based on a sample dataset used in the proposed user satisfaction model and shows that the

proposed user satisfaction model could resolve the research problem mentioned in Section 3.3. Finally, future efforts and conclusions are presented and discussed in section 5.

2. Related Efforts

Several interesting proposals for discovering the level of user satisfaction have been presented in the last couple of years [3-5]. Almost all of them worked with data science models based on the user direct feedback data. Proynova and Paech [6] identified the factors that influence user satisfaction. Nourikhah and Kazem Akbari [7] present the impact of service quality on user satisfaction. They present a model estimating the distribution of quality of experience using Bayesian data analysis, *Electronic Commerce Research and Applications*. These authors present a method that correlates Quality of Experience (QoE) and Quality of Services (QoS). In their approach, they use opinion score distribution instead of the mean opinion score. Besides, they use Bayesian data analysis instead of linear regression as they find two shortcomings of linear regression: (i) linear regression assumes that dependent variable complies with the Gaussian distribution and that the predictor variables are independent. (ii) The dataset used in linear regression should be metric, while other forms of data cannot be used. Their approach depends on the user feedback on QoE using which they develop their model to find the user satisfaction on QoS. Kiseleva et al., [8] proposed a model that could predict user satisfaction by using the interactive dialogue with the intelligent systems like Microsoft's Cortana, Google Now, and Apple's Siri. Kim et al., [9] present in their paper which and how variables should be considered in the data science model for user satisfaction. The authors present a forecast model for the user satisfaction on searches using the clicking data (dwell time) on the link. Hsieh and Tang [10] presents an approach to use Neural Network in meteorology and oceanography. The authors show how to use different statistical methods by improving and modifying various parameters to adjust to the problem Noh et al., [11] present the influential factors for digital home services. O'Leary [17] analyses the prediction market where he indicates that there are limited works, based on user centric data. He outlines that user centric data could provide more effective and transparent pattern of market prediction. Although most of these techniques use the users' direct feedback data, they could impact other methods on identifying the factors for measuring the level of satisfaction with the different digital services.

3. Methodology and Proposed Model

The deductive reasoning method [12] has been selected to guide and carry out our research efforts. It is a common understanding that if a user interacts with the service frequently than the average interaction (by all users) - we can claim that the

user is satisfied with it. On the other hand, if the number of satisfied users is higher, it can be said that the service has a success factor. Therefore, the proposed model first estimates the satisfied users of the service by descriptive statistical analysis. Then, the user with this statistical estimation will be measured using the machine learning data science model. In this section, we present a data science model which facilitates an interpretable, multi-granular analysis of the decision-making process. We begin by discussing the problem setting and relevant details, then dive into the details of modelling and inference.

3.1. Definition of a Satisfied User

Three issues affect the level of satisfaction of a digital service user according to [13]:

- Impact of the service to the user (this impact may be social, personal, professional, or financial)
- The opportunity provided by the service. For example, social ties through the service (e.g., like, sharing contents, status update, etc.) may lead the increasing customer satisfaction.
- Usability of the service (e.g. technical difficulties are faced by the user)

All of these issues can only be measured by empirical studies (survey, questionnaire, rating, etc.). Besides, there is no benchmark value for satisfaction. Therefore, we rely on Trait theory to identify and define a satisfied user. Trait theory posits that a person's behavior will be generated consistently with his or her personality traits. There were empirical studies [15-16] reported that personality traits have a significant relationship with customer-oriented behavior. Consumer's attitude, behavior, and thoughts are reflected through the service/product that is being used by them. They also consider these services as the brand. When a user considers and recognizes a product or service as a brand, he/she becomes active to the service/product that means that active user is an indicator of the happy user. So, from different researchers' point of view, a satisfied user means a loyal user who repeatedly uses the service or the product.

3.2. Success of the Digital Service

Customer loyalty is vital to the success of business organizations. The success of a digital service relies on the user satisfaction, while higher satisfaction makes a user loyal to a digital service. Chen et al., [18] present and analyze the factors that affect the success of a digital service. These factors are revenue efficiency, distribution model, the competition of product, service/implement model, strategic alliance, and market segment. These authors illustrated domain specific business factors rather than common factors for all IT based business or services in this paper. Therefore, these factors cannot be used to develop a generic predicting model that could measure the success of a digital service. But how can these success factors

be measured? The determinants that lead most to measure the success of the digital service are (i) number of users and (ii) revenue earning [19]. These two determinants of success are influenced by the user level of satisfaction, which relies on the quality of the service. Popular services and products become brand while increasing its number of users. Considering the definition of the quality of any digital service, user happiness and benchmark of the success of the digital service, we assume that there is a hierarchical relationship among the quality of the service, the user happiness and success of that digital service – which can be presented as the Figure 1.

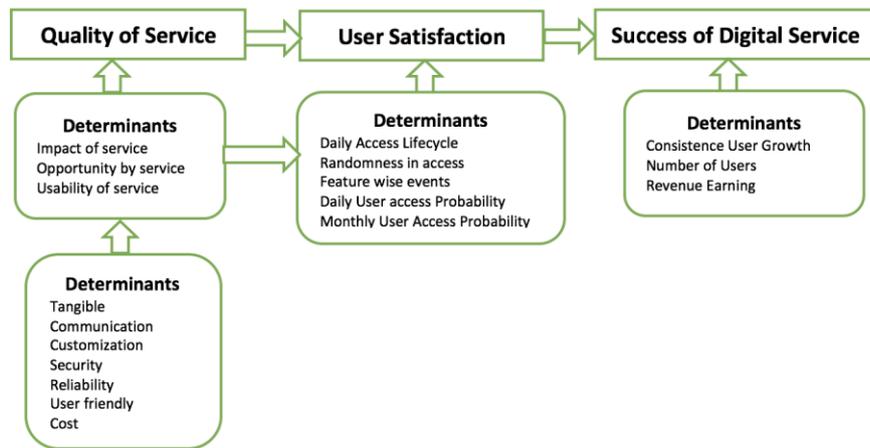


Fig. 1. Relation between user satisfaction and success of digital services

3.3. Formalization of the Problem & Research Question

Consider a decision space S holding m number of users of a digital service. There are p number of features provided by the service (i.e. the decision space is p dimensional). S contains N number of multivariate observations $\{x_1, x_2, \dots, x_n\}$ interacting by the users with different features of the given service. We consider that these observations are sampled from a normal distribution. A hyperplane (decision boundary) l is required to be identified to separate the solution S' space for classifying the users into two solution space: S'_h (*satisfied*) and S'_u (*unsatisfied*). Based on the above problem and the discussions provided in the previous sections we formulate our main research question as follows:

How to measure user's satisfaction of a digital mobile service by analyzing access service log data?

3.4. Data Description

In our experiment, event-based access log datasets are used. The dataset contains 10 million of events for 21 attributes. These datasets contain timestamp which indicates when an event has been occurred. An event is generated if a user successful login, successful authorization and so on has taken place. That means the dataset is limited within the service access authorization, however it neither contain any demographic data nor service featuring data such as feature viewed or clicked. If a particular access event is generated, '1' is assigned to the relevant attribute of that event, if not then '0' is assigned to the feature for that event. The experimental dataset contains 18 such events-oriented attributes among the 21 attributes. It is notable that these attributes cannot be used directly for recognizing the user satisfaction pattern. Therefore, we derive four features from these 18 attributes through a feature engineering mechanism. Later these derived features have been used as predictors for estimating user satisfaction.

3.5. Data Science Model Description

Fig 2 presents the block diagram of the proposed data science model. Having our research problem defined, our proposed model uses Binary classification, since the solution space is to divide into two classes (satisfied and unsatisfied) by a decision boundary. The supervised learning model is used to train the classifier. In this case, a learning algorithm is provided a set of N labeled training examples $s\{u_i, c_i : i = 1, \dots, N\}$ from which it must produce a classification function $f: U \rightarrow C$ that maps target variables (here users) to classes. Here u_i denotes the i^{th} training user and c_i is the corresponding classes [20]. In our given research problem, the data is not a labeled data; instead, it used unlabeled data. Therefore, the proposed model is required to develop a labeled training dataset from the given unlabeled data. Training the proposed data science model using the training dataset and discover on the test dataset are carried out in the following three stages:

- Latent variables (predictors) derivation
- Label Estimation on training dataset
- Prediction Model

Latent Variables Derivation

This stage is applied on both the training dataset and test (new) dataset. It is hard to find a pattern of the user behavior from the access log data if the dataset is unlabeled. Especially, if the user generates the observations in the sample dataset as events and the value of those events are presented in binary form. Therefore, a data science model is required to identify the unobservable variables that are used as the predictor in the data science model. The steps below are followed in order to identify the predictor variables.

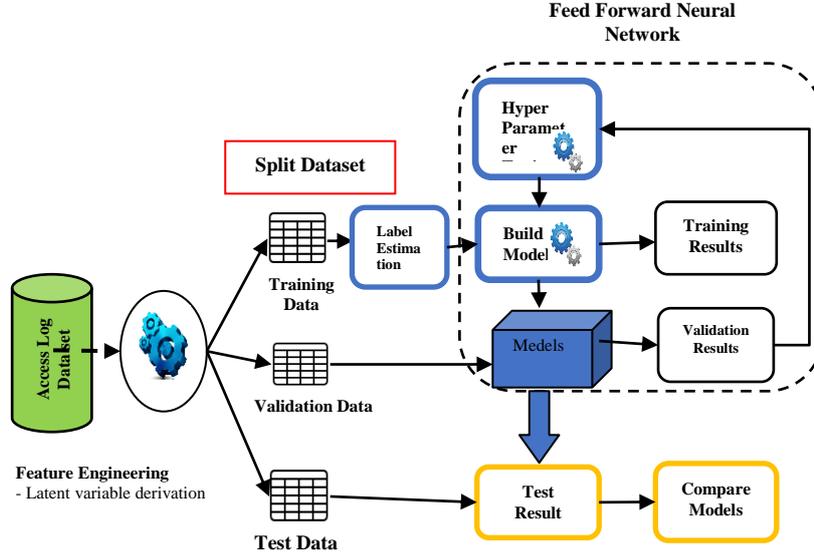


Fig. 2. Proposed Data Science Model

Input: Experimental dataset is an unlabeled dataset F that consists of both binary data and categorical data. It is converted into a data frame (we use matrix notation to represent the data frame, as it is common to represent a spreadsheet or a tabular data frame of q rows and r columns as a matrix [21]), where q is the number of observations, and r is the number of features provided by the digital service. $f_{i,j}$ is any value (either 1 or 0, since we are using access log data) representing an event related to any feature of an individual user. Below we describe the steps used to derive the latent variables.

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & \dots & f_{1,r} \\ f_{2,1} & f_{2,2} & \dots & f_{2,r} \\ \dots & \dots & \dots & \dots \\ f_{q,1} & f_{q,2} & \dots & f_{q,r} \end{bmatrix} \text{ and } X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,q} \\ x_{2,1} & x_{2,2} & \dots & x_{2,q} \\ \dots & \dots & \dots & \dots \\ x_{p,1} & x_{p,2} & \dots & x_{p,q} \end{bmatrix}$$

Derivation of Predictors: From the dataset F , find the following latent/unobservable variables for each user and store into any data frame X that contains the derived predictors of p users and q (in this case, $q = 4$ as we assumed four

predictors that are highly correlated derived features from our experimental dataset) is the number of derived feature. In this approach of happy user identification, we derive four different latent variables (predictors) from the raw users' access log dataset such as i) Slot wise Average Spent Time ii) Daily Interaction iii) Day wise Life cycle and iv) Feature Wise Interaction Ratio. Methods to derive these four predictors have been explained in the following. These methods are influenced by the work of Rana et al., [15]:

Slot wise Average Spent Time, $IR_{slotwise}$: We divided the 24 hours into four timeslots 00:00 to 06:00, 06:01 to 12:00, 12:01 to 18:00 and 18:01 to 23:59 to identify individual user's access pattern on a daily basis. Slot wise Average Spent Time f_s means the summation of the individual user's total events in each slot and total number of events of all user of the similar slot.

$$f_s = \frac{\sum_{i=1}^N \sum_{j=1}^{s_u} e_{i,j}}{\sum_{p=1}^N \sum_{q=1}^m \sum_{r=1}^s e_{p,q,r}} \dots\dots\dots(1)$$

Where N is the number of days in the dataset, s_u is the number of slots of user u , m is the total number of users in the dataset.

Daily Average Interaction IR_{Daily} : IR_{Daily} is the individual user's daily average events (i.e. daily interaction ratio) generation and can be calculated as dividing the daily events (generated by individual user in each day) by the total events per day generated by all users.

$$f_{di} = \frac{\sum_{i=1}^{F_{ind}} e_i}{d_l - d_f} \dots\dots\dots(2)$$

- $e \rightarrow$ events or observation in the dataset
- $F_{ind} \rightarrow$ Total number of individual events in the dataset
- $d_f \rightarrow$ Individual user's first day
- $d_l \rightarrow$ Individual user's last day

Day wise Lifecycle, $Lifecycle_{Daywise}$: It means the average difference between the individual user's first access and the last access in the day.

$$f_{lifecycle} = \frac{\sum_{i=1}^{N_u} (t_{l_i} - t_{f_i})_i}{N_u} \dots\dots\dots(3)$$

- Where,
- $t_l \rightarrow$ last access of a user in the service

$t_f \rightarrow$ first access of a user in the service

$N_u \rightarrow$ total number of days accessing the service of a user

Feature wise ratio, $IR_{Featurewise}$: It means the summation of the average ratio of individual user's number of events divided by the total number of events generated for a specific feature.

$$f_f = \frac{\sum_{i=1}^{F_{ind}} e_i}{\sum_{j=1}^m \sum_{k=1}^F e_{j,k}} \dots \dots \dots (4)$$

Where $F_{ind} \rightarrow$ individual user's feature wise event

Normalization of Predictors: To scale the derived features, this model normalizes the derived predictors and transform the normalized values into another data frame $\hat{\mathcal{X}}$. Feature scaling is used to normalize the values of each derived predictor variable between 0 and 1. Following the feature scaling formula [21] is used in this regard:

$$\hat{x} = \frac{x_{i,j} - x_{\min}}{x_{\max} - x_{\min}} \dots \dots \dots (5)$$

The resultant data frame $\hat{\mathcal{X}}$ contains the normalized values of the predictors (as shown in the following):

$$\hat{\mathcal{X}} = \begin{bmatrix} \hat{x}_{1,1} & \hat{x}_{1,2} & \dots & \hat{x}_{1,N} \\ \hat{x}_{2,1} & \hat{x}_{2,2} & \dots & \hat{x}_{2,N} \\ \dots & \dots & \dots & \dots \\ \hat{x}_{m,1} & \hat{x}_{m,2} & \dots & \hat{x}_{m,N} \end{bmatrix}$$

If this derived data frame $\hat{\mathcal{X}}$ is generated from the train dataset, it is used as the input for label estimation to train the prediction model. In contrast, if $\hat{\mathcal{X}}$ derived from the test dataset, it is used as the input in the prediction model to classify the target variables.

Label Estimation on Training Dataset

This step is applied only on the training dataset since the classification algorithms require a labelled dataset as the input with the target (class) variable column to train the prediction model. We used the following label estimation function:

$$h(x) = \sum_{i=1}^F W_i x + b \dots \dots \dots (6)$$

Where x is the feature matrix, $h(x)$ is the label vector, F is the number of features, W is the weight matrix, and b is the noise. It is notable that weight and noise are randomly generated values by pre-defined function. The estimated label for training the dataset is calculated using the following conditions:

$$y = 1 \text{ if } h(x).mean() - h(x).var() \in h^3 h(x).mean()$$

And $y = 0$ otherwise.

Finally, we get the training dataset as follows and it consists of the latent variables. This dataset is used for training the model.

$$d = \begin{bmatrix} \hat{x}_{1,1} & \hat{x}_{1,2} & \dots & \hat{x}_{1,N} & y_1 \\ \hat{x}_{2,1} & \hat{x}_{2,2} & \dots & \hat{x}_{2,N} & y_2 \\ \dots & \dots & \dots & \dots & \dots \\ \hat{x}_{m,1} & \hat{x}_{m,2} & \dots & \hat{x}_{m,N} & y_m \end{bmatrix}$$

Prediction Model of a Satisfied User

To predict a satisfied user, we have method a Feed Forward Deep Neural Network (FFDNN) algorithm. This proposed method is influenced by a paper of Andrew Ng [22]. The model contains *four* hidden layers and each of these layers contains a weight matrix W of the input feature and a bias vector b . We compute vector h_1 which is the input layer and first activation layer of the proposed model.

$$h_1(x) = S(W_1^T x + b_1) \dots \dots \dots (7)$$

where x is input vector that represents correlated $k \times k$ features, W_i is the weight matrix for input layer and b_i is the bias vectors for the first hidden layer. W can be pre-initialized or randomly generated by a customized function. In order to train our model, we use a pre-defined function that initializes the value of W . The number of neurons in each layer is considered as the bias b . For all other intermediate hidden layers, vector h_i can be defined as;

$$h_i(x) = S(W_i^T x + b_i) \dots \dots \dots (8)$$

In this equation each column of the matrix W_i corresponds to the weights of i^{th} hidden layer for a specific timestamp T . In all of these hidden layers, we use a non-linearity function $f(z)$ for computing the hidden layers. In our experimental implementation, the FFDNN which consists of input layer, 3 hidden layers, output layers with 350 neurons, 26401 parameters have been used. As shown in the Fig 3, five derived features from our dataset are used in the input layer as the input neurons and outputs 200 neurons. These 200 neurons are used in the first hidden layer to produce 100 neurons which feeds into second hidden layer as input. Third hidden layer contains 50 neurons which produce the output neuron.

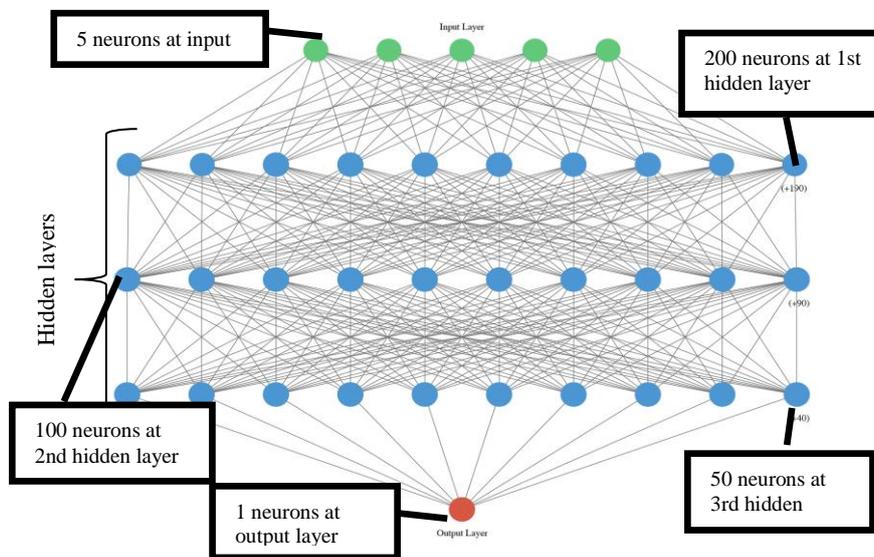


Fig. 3. Neurons in the proposed Feed Forward Neural Network

Parameters in DNN are generated by the formula below:

$$\text{Parameters} = \text{input values} * \text{neurons in the first layer} + \text{bias values}$$

In the first layer 1200 parameters ($5*200+200$). Similarly, 20,100 in the second layer, 5050 parameters in the third layer and the output layers contains 51 parameters. In our model, we utilize the Rectifier Linear Unit (ReLU) as the activation function for all of the hidden layers except the output layers. The reason behind using ReLU in our model is to get better performance for active user prediction. Rectifier is an activation function which is defined with the positive argument of this function and can be shown as in eq. (9).

$$S(z) = \max(0, z) \dots \dots \dots (9)$$

Where $z \rightarrow$ is the input to a neuron.

On the other hand, we have used sigmoid as the activation function in the output hidden layer as sigmoid function takes real values from the previous hidden layer and output any value of either 0 or 1. A Sigmoid function can be defined as:

$$S(z) = \frac{1}{1 + e^{-z}} \dots\dots\dots(10)$$

We use the sigmoid function for the following reasons:

- i) Non-linear relationship between the input
- ii) convert the input into a more useful output (in our case, between 0 and 1)

Again, we use the optimizer function Adam (Adaptive moment of estimation) [23]. It is a stochastic gradient method. Since Adam does not need any stationary objective $f(x)$ might change with respect to time and still the algorithm will maintain converge. The output hidden layer takes input fifty neurons which produced from the third layer. It provides the predicted value of the label for each user. We used different epochs with 10 batch-size for training the model. We have used ‘binary cross entropy as loss function’ (in the proposed DNN model) as it measures the performance of a classification algorithm whose number of output is a probability value between 0 and 1. The last layer will provide an output vector y which is the probability distribution over the N output classes.

4. Experimental Results

In order to validate our concept, we have implemented the proposed model. Our experimental user access log dataset size was of 10704949 observations with 21 features. Features in this dataset are binary and timestamp. We divided the raw dataset into two before any data pre-processing: 1st 5000000 observations were used for the trained dataset and the remaining 5704949 observations for the test dataset. The first 5 000 000 observations were divided into 70%:30% ratio for training and evaluation respectively during training the model. Then, the trained model is stored to evaluate the performance of the model with the test dataset. In our experimental settings we simulate with four popular optimization algorithm such as Adam, SGD, RMSprop and Adagrad. Since any artificial neural network can defined as a stochastic process, we need to tune the hyper parameters of the optimization algorithms.

Table 1: Accuracy & RMSE for different optimization algorithms

	Accuracy on validation Dataset	Accuracy on Test Dataset	RMSE on Validation Dataset	RMSE on Test Dataset	Last Epoch
Adam	0.8837209302	0.7837837837	0.273	0.465	194
SGD	0.5639534883	0.1756756756	0.508	0.908	1
RMSprop	0.7674418604	0.6756756756	0.404	0.569	181
Adagrad	0.7965116279	0.6351351351	0.400	0.604	193

We know that value of hyper parameters varies from model to model. Therefore, we need to tune these hyper parameters during training the model to find the optimal trained model. Although, each of this optimization algorithms have different set of hyper parameters, learning rate and learning rate decay are two common parameters. We simulate the proposed model by tuning these two parameters with different values to see the impact of the accuracy of the model. We find the highest accuracy by considering the learning rate 0.005, 0.01, 0.001, 0.01 for adam, SGD, RMSprop and Adagrad respectively. We did not find significant variation in the accuracy for learning rate decay. Therefore, we kept the value of decay 0.0 for all of these optimization algorithms. The simulated results of our experiments are shown in Table 1. The result of accuracy and RMSE in Table 1 shows that the Adam optimization algorithms provides better optimized result. In the following sub-sections, we discuss the salient result we found after analyzing this experiment.

4.1. Result Analysis and Discussion

As we know that performance of any data science model depends on three issues: i) setting the optimal model parameters ii) accuracy and error of the model and iii) how do the model produce result on dataset for particular decision making. The goal to train a model is to set the model parameters that have low loss. If the value of loss for both training data set and test dataset are consistent after certain iterations, model parameters are considered of that iteration. On the other hand, value of accuracy of a model indicates how the model performs. Besides. Value of AUC curve indicates how the model predicts correctly.

We have iterated the model 200 times to train it. From the simulation, we find that the model learns until 194 epochs when optimization algorithm is ‘adam’. This indicates that the proposed model learns consistently. Based on the ‘adam’ optimization algorithm, the experimental result shows that the RMSE is 0.273 with evaluation data set while 0.465 with test dataset. Besides, we get 88% accuracy with validation data while 78% with test dataset. Figure 2, presents the accuracy for the given train dataset and test dataset. In the case of neural networks, the

loss is negative log-likelihood. The model becomes better when loss becomes lower. Besides, the value of loss refers to how good or poor the model performs after each or a several number of epochs (iterations). Figure 3 shows that loss is decreasing as the Figure 2 shows that the accuracy is increasing, since accuracy is inversely proportional to the loss of any model with respect to the iteration. From both of these figures, we find that the accuracy and loss does not vary significantly after 100 iterations. This indicates the generalization of the proposed model.

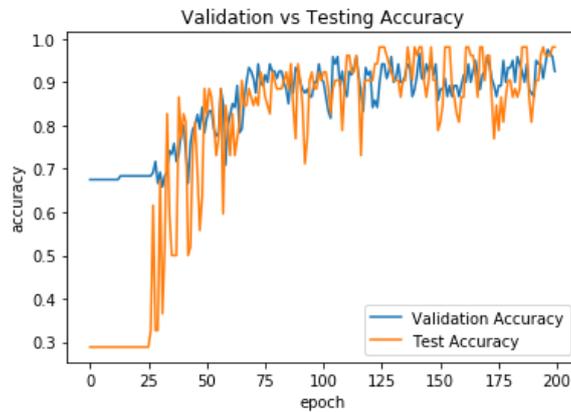


Figure 2: Summarize history for accuracy

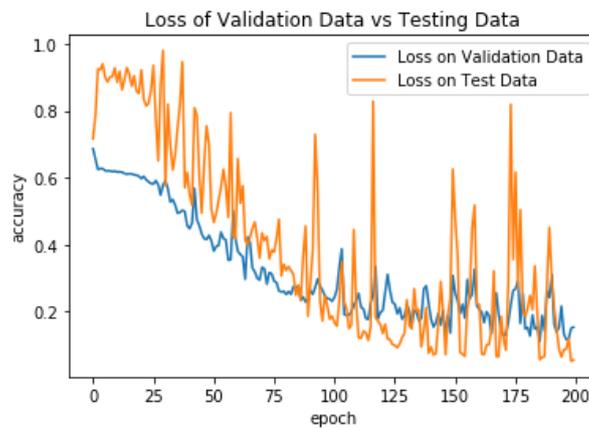


Figure 3: Summarize history for loss

We also evaluate the model using the ROC curve. We find that the AUC score is 0.96 in the evaluation data as shown in the Figure 4. High AUC ROC score indicates the classifier currently can perform the classification correctly. However, it requires to search the threshold for which it can classify more better. On the other

hand, low AUC ROC score indicates that indicates the classifier currently cannot perform classification correctly, and even fitting a threshold will not improve its performance. In this case, our proposed model could predict an active and satisfied user correctly, as we found the AUC score in test data set is 0.93:

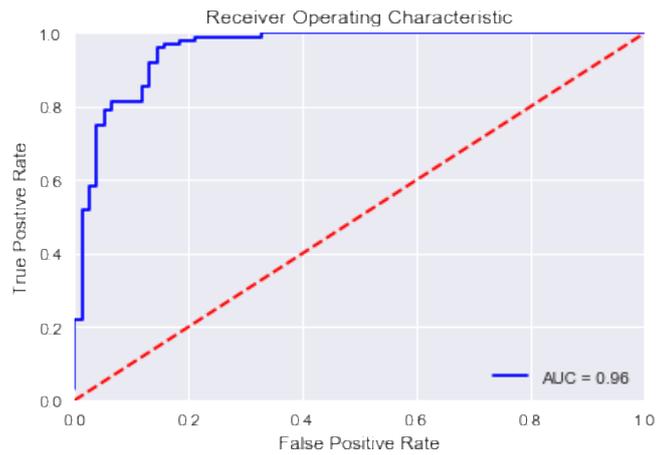


Figure 4: ROC curve on Validation Dataset

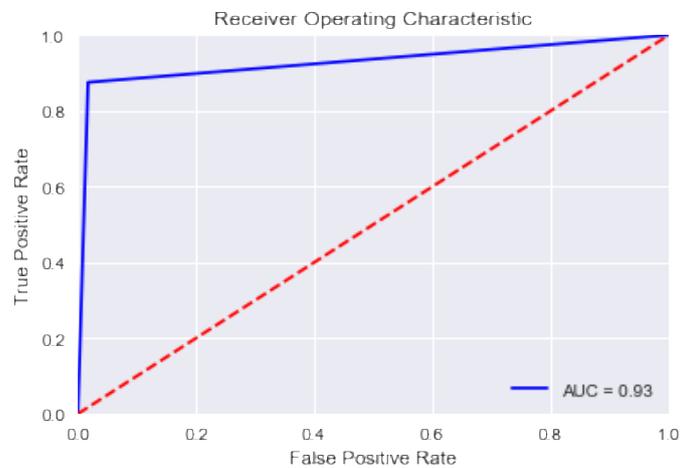


Figure 5: ROC curve on Test Dataset

We also evaluated the model by varying the number of predictors. We found that a significant number of users have a very low interaction with the service. As a result, it significantly deviates the overall accuracy, precision and AUC. Therefore, we filtered out the users those who have interactions less than the mean

of total users' interactions. This improves the accuracy as well as the AUC of the model.

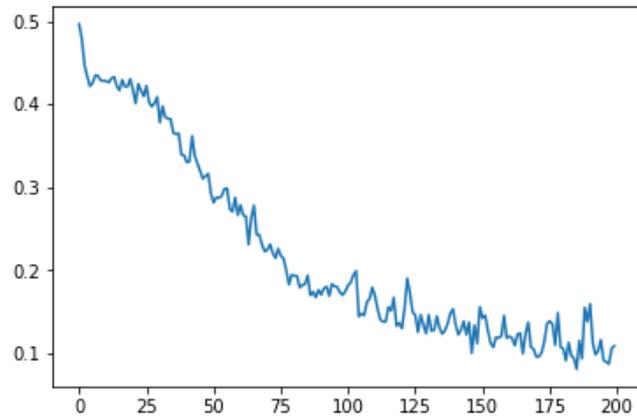


Figure 6: Summarize history of Mean Absolute error

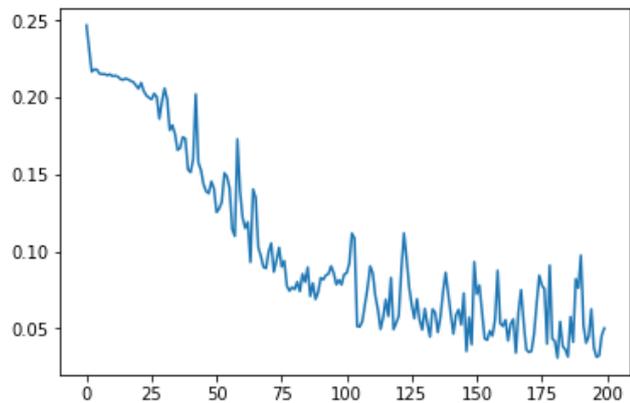


Figure 7: Summarize history of mean squared error

Figure 6 shows the decreasing mean absolute error and finally the value becomes around 0.1 which indicates the proposed model would predict accurately. As we know that the MSE indicates the difference between the predicted value and targeted value. Figure 7 shows that the model very small mean squared error that shows the correctness of the model.

Besides, all predictors that we have derived from the raw dataset are slot wise interaction in each day, daily interaction, feature wise interaction, and user's Lifecycle in each day in the system. Our experiment shows that all of these features are highly correlated to measure user's activeness in the digital system. Therefore,

we can argue that if someone spend a countable time regularly with the digital mobile service, one can be a satisfied user.

5. Conclusion

In this extended version of the paper of [1], we present a data science model to discover satisfied users of a digital mobile service. The novelty of this approach is that it uses the user's access log data of the given service rather than user's direct feedback data. Knowledge Management paradigm is used during designing the proposed data science model. We show that the model is able to classify users as satisfied and un satisfied. We also evaluate the performance of the model using performance matrices such as accuracy, and ROC-AUC curve. The accuracy scores derived using these matrices above 90% for both validation and test dataset – which indicates a good and acceptable value for a data science model. We use a dataset of 10 million of user observations (events) in our investigation for the proof of concept of the proposed model. This proposed model shows that any size of dataset can be used to develop a promising prediction model using a Deep Neural Network. In this way, we also show that active and satisfied users can measure the success of digital service, if we could classify them. Besides, a large scale of user access log dataset can be used to see the performance of the proposed model, and we left open this task of scalability of the model as the future work. We have been also working with other users access log dataset to see how this proposed neural network model works

References.

1. Sarwar, J. M. Juwel, R, and Marcelo, M., Active and Satisfied Users as a Key to Measure the Success of a Digital Mobile Service, AI-based Innovation in Digital Services, Proceedings of 6th AI4KM, July 15, 2018, Stockholm,
2. Top 12 Key Performance Indicators for Maximizing Mobile App Revenue, June 11, 2011.
3. Engelbrech, K. P., Gödde, F., Hartard, F., Ketabdar, H., and Möller, S., 2009. Modeling user satisfaction with Hidden Markov Model. In Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL '09), Matthew Purver (Ed.). Association for Computational Linguistics, Stroudsburg, PA, USA, 170-177.
4. Sarukkai, R. R., 2013. Real-time user modeling and prediction: examples from youtube. In Proceedings of the 22nd International Conference on World Wide Web (WWW '13 Companion). ACM, New York, NY, USA, 775-776.
5. Salehan, M., Dan J. Kim, 2016. Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics, Decision Support Systems, Volume 81, January 2016, Pages 30-40, ISSN 0167-9236.

6. Proynova, R., Paech, B., 2013. Factors Influencing User Feedback on Predicted Satisfaction with Software Systems. In: Doerr J., Opdahl A.L. (eds) Requirements Engineering: Foundation for Software Quality. REFSQ 2013. Lecture Notes in Computer Science, vol 7830. Springer, Berlin, Heidelberg.
7. Nourikhah, H., Akbari, M., K., 2016. Impact of service quality on user satisfaction: Modeling and estimating distribution of quality of experience using Bayesian data analysis, *Electronic Commerce Research and Applications*, Volume 17, May–June 2016, Pages 112-122.
8. Kiseleva, J., Williams, K., Awadallah, A. H., Crook, A. C., Zitouni, I. and Anastasakos, T., 2016 Predicting User Satisfaction with Intelligent Assistants. In Proceedings of the 39th International ACM SIGIR conference (SIGIR '16). ACM, New York, NY, USA, 45-54.
9. Kim, Y., Hassan, A., Ryen W., and Zitouni, I., 2014. Modeling dwell time to predict click-level satisfaction. In Proceedings of the 7th ACM international conference on Web search and data mining (WSDM '14). ACM, NY, USA, 193-202.
10. Hsieh, W.W., Tang, B., Applying neural network models to prediction and data analysis in meteorology and oceanography *Bull. Am. Meteorol. Soc.*, 79, pp. 1855–1870.
11. Noh, M. J., Ju Seong Kim, Factors influencing the user acceptance of digital home services, *Telecommunications Policy*, Volume 34, Issue 11, December 2010, Pages 672-682.
12. Johnson-Laird, Philip, N.: "Deductive reasoning." *Annual review of psychology* 50.1 (1999): 109-135.
13. Zepernick, H. J., Iqbal, M. I., and Khatibi, S.: Quality of experience of digital multimedia broadcasting services: An experimental study, *IEEE Sixth International Conference on Communications and Electronics (ICCE)*, Ha Long, 2016, pp. 437-442.
14. Touré, Carine; Michel, Christine; Marty, Jean-Charles, Re-designing knowledge management systems: Towards user-centred design methods integrating information architecture, *Knowledge Management and Information Sharing*, Oct 2014, Rome, Italy.
15. Rana, J., Kristiansson, J. and Synnes, K., 2010. "Enriching and Simplifying Communication by Social Prioritization," *2010 International Conference on Advances in Social Networks Analysis and Mining*, Odense, 2010, pp. 336-340.
16. Brown, T. F., Mowen, J. C., Donovan, D. T., and Licata, J. W.: The customer orientation of service workers: Personality trait effects on self-and supervisor performance ratings, *Journal of Marketing Research*, vol. 39, no. 1, pp. 110-119, 2002.
17. Daniel E. O'Leary, User participation in a corporate prediction market, *Decision Support System*, 78 Science Direct, 2015, page 28-38.
18. Chen, M.K., Shih-Ching Wang.: The critical factors of success for information service industry in developing international market: Using analytic hierarchy process approach, *Expert Systems with Applications*, Volume 37, Issue1, January 2010, Pages 694-704,.
19. Jong-Hyun Park, Y. B. Kim, Moon-Koo Kim. 2017. Investigating factors influencing the market success or failure of IT services in Korea, *International Journal of Information Management*, Volume 37, Issue 1, Part A, February 2017, Pages 1418-1427, ISSN 0268-4012.
20. Peng, F., Schuurmans, D.: 2003. Combining Naive Bayes and n-Gram Language Models for Text Classification. In: Sebastiani F. (eds) *Advances in Information Re-*

- trieval. ECIR 2003. Lecture Notes in Computer Science, vol 2633. Springer, Berlin, Heidelberg
21. Gareth, J., Daniela, W., Trevor, H., Robert, T.: A Introduction to Statistical Learning with Applications in R, Springer, 2013, DOI 10.1007/978-1-4614-7138-7.
 22. Quoc V.Le, , Andrew Y.Ng and et cl., Building High-level Features Using Large Scale Unsupervised Learning, Proceedings of the 29 th International Conference on Machine Learning, Edinburgh, Scotland, UK, 2012.
 23. Diederik P. Kingma & Jimmy L. Ba, ADAM: A Method for Stochastic Optimization, ICLR, 2015.
 24. J. D. Applen, "Tacit knowledge, knowledge management, and active user participation in Web site navigation," in *IEEE Transactions on Professional Communication*, vol. 45, no. 4, pp. 302-306, Dec. 2002.