Opinion mining for user experience evaluation model using kernel-naive bayes classification algorithm

DOI:10.36909/jer.11319

P. Rajkumar*, S.V. Kogilavani**

Email ID: rajkumar.rphd@gmail.com, gogilavani.sv@gmail.com

*Assistant Professor at Mount Zion College of Engineering and Technology, Tamil Nadu, India
**Associate Professor at Kongu Engineering College, Tamil Nadu, India

ABSTRACT

User experience evaluation approach is the major key to adapt the new trends and technology. The product launch is based on the various opinions of users and availability of product. The first impression about the product makes successful sales, which is analysed with UX (User eXperience) design. Before developing / launching the product, have to evaluate the user experience model by online sources. The opinion/sentimental analysis are the way to capture the people’s opinion about the product. Rating, page session, website page views, and number of buyers or users are evaluated as a graph model and predict the requirement of the product. This process makes the product’s benefits. The previous work utilizes the Markov Chain Monte Carlo (MCMC) Method to model the UX design. In this proposed research work, the opinion mining approach is used to get the dataset from Google analytics. This dataset is model using Kernel based Naïve Bayes Classification algorithm and the prior & posterior probability is calculated by MCMC (Markov Chain Monte Carlo) techniques. Classification approach takes the training and testing data. Here the confusion matrix is used to create the UX evaluation model’s accuracy. By this proposed algorithm, it summarized the positive and negative opinion then we can calculate the accuracy of the system and it easily identifies the user opinion. This proposed UX design model improves the result as compared to the previous MCMC method. The data mining based sentimental classification is done with the help of MATLAB 2018a tool.

Keywords—User experience model; opinion mining; Kernel-Naïve Bayes Classification; MCMC; Google Analytics data;

INTRODUCTION

Today’s world, the development of economic society increases the people’s requirement level, which focus comfort level, mental level, demand requirement etc. User experience-evaluation approach is applicable to predict the requirement of product for improving the marketing and satisfy the requirement. For any field, the people’s requirement is satisfied by various sources like collecting feedback, online tracking, Google analytics, and UX report. The UX model is processed with the AI (Artificial Intelligence) algorithms for improving the reliability. New product design is launched by gathering information about the reviewers. Here the data mining technique is used to classify the sentiment analysis of the product, which improves the sales rating. The big data analysis is used in the machine learning for analyzing user experience. The supervised machine learning is preferred to train the classification and regression. The unsupervised approach prefers to clusters and associates the unlabelled dataset. The UI (User Index) model is to interact with the user through machine learning technique.

Vision based interface improves the usability by user experience evaluation model. The satisfaction level of vision-based interfaces like motor-impaired devices for users is improved by UX
model. The automatic face detection scheme selects the feature based on the face expression like sad, happy, and normal. The usability of product design is improved by UX model and it utilized with various algorithms and approaches [1]. The probabilistic parameters are approximated to find the predicted values. This parameter distribution is applicable for group-level estimations from the dataset [2]. The user experience model is used for recommender system of various applications like those that browsing indicates, choice satisfaction, evaluation in feedback, etc. here the human computer interaction strategy is applicable for modelling the user experience. This study proposes to optimize the decision support model. System effectiveness is analyzed by positive and negative feedback of the user [3]. The software usability is evaluated by opinion mining. The subjective usability of software is detected by available data and it is evaluated by opinion about the user. After performing the pre-processing stage of classification, the opinion mining algorithm is used with lexicon base by polarity dictionary, finding maximum entropy level and performing the K-NN (Kernel Nearest Neighbour) to classify and visualize the result [4]. The customer preference is classified using hierarchical Bayesian approach. The qualitative and quantitative modelling is preferred to design the UX model and it depends on the behaviour, feelings, emotions, opinion, and psychological movement [5]. The social network site evaluation is created by Test-model to enhance the UX report. The information of software usability uses Camtasia software. The online social network usability drawback is overcome by evaluating the content, navigations and interactivity [8]. The lexicon based approach uses dictionary based model and corpus based model, which is performed by statistical and semantic analysis [9]. The fragmentation, segmentation, type conversion, normalization, and dimensionality reduction is performed for this stage. Then the feature is identifying for evaluation. Based on the subjective analysis, the feature report is classified for UX model [10]. Heng Luo, et. al has presented the study of online learning using Google analytics. The website data has the three major concerns that are audience data, behaviour data, and acquisition data. The study proposed with the cloud based online course and it collects the geographical information [11]. The positive and negative polarity is based on the statement from the user, which is labelled as weak, strong, and steady. Natural language processing is used for data mining in the sentiment analysis [12]. The quality of user experience is evaluated for mobile communication based on tripping time. Here it uses voice-activity detection algorithm by speech transmission. It supports both online and offline evaluation of user experience [26]. The usability measure is possible for this technique and it designed to satisfy the questionnaires from the users. Archaeometry interactive tool is used in the virtual reality environment. The UX evaluation is modelled for cyber-archaeology [18]. The ML framework is integrated with the opinion-mining task to evaluate the UX model [20]. The naïve bayes classifier is used for slow learner prediction and Swati (2016) discussed it [16]. The statistical analysis is performed by MCMC technique and it a modelled with the graph approach. The twitter user index is analysed by multilayer naïve bayes classifier and it uses the sentiment tendency analysis, which segmented with positive, negative and neutral analysis [21]. The bidding performance adopts the multiphase strategy to find the affective factor of decision-making process. The profitable assessment is created by overcome these risk factors [14]. Sentiment analysis data uses machine-learning classification for various applications. The sequential minimal optimization is performed for synthetic minority over-sampling technique and it is classified using SVM (Support Vector Machine) learning approach and naïve bayes classification approach. Ten fold cross validation is applied on the data mining algorithm. Here the multinomial naïve bayes approach is used classify the data through weka and San Carlos tool [22].

The Google cardboard element is used for virtual reality model and it is cultural game of Indonesian. The playing frequency, gaming questionnaires are survey to form the database. Various features are
evaluated to model the UX design and it utilizes the gender, probability of user, feedback, playing frequencies, personality traits, etc. the gamification results to self-administered development [17]. The game designer provides the permission to access the UX model for improving the quality and quantity of that product. The eye-catching graph model is provided to give the story of game manufacturers. The conventional UX design model is given in the fig1.

![Fig1. Process flow of UX model](image)

The design process of UX model will improve the marketing strategy. The behaviour of the user is understood by gathering information about the usage. The user experience is analyzed by positive, negative and neutral priority. The iterative frame is modelled with the UI features. The virtual reality display is used for analyzing the UX model and it is compared with the other conventional displays [28]. The three dimensional website views is improved by designing the user experience evaluation model. Various attributes are selected to classify the features and it performed with the fuzzy model to evaluate the quality factors. Here the grey correlation analysis is used in the assessment model then it performed with analytic hierarchy process [24]. The user modelling is scheduled by computer interactions and the usability is evaluated to model the UX design, which improves the performance. The persona is described with the user’s wish and it constructed to specify the emotions, feeling, motivations and behaviour [6].

**Literature Survey**

Yaya Heryadi, et al., (2016) has presented the UX design for game flow approach using virtual reality based gamification approach. The dataset has the information about the game players, which contains self-administrated questions. These sample data are analyzed with the user experience evaluation model. Educational games are designed to utilize the frequent players of game, customer traits and associations.

Yan Zhou, et al., (2015) has presented the research on UX design model for smart phone. Here the fuzzy theory is employed to evaluate the affective modeling of user experience evaluation model. The clustering technique is used to evaluate the weight function of the dataset. Degree of membership evaluation model is created based on the fuzzy assessment model. Customer satisfaction and usability evaluation is the major objective to design the UX model.
R B Ravi, et al., (2017) has proposed the classification of opinion mining data using machine-learning techniques. The subjective (positive and negative opinion) and objective (fact) data is used for sentimental analysis. The machine learning is applied to classify the result through training and testing. The testing is validated to model the UX design. Data mining task is employed to identify the opinion about the user for improving their business needs. Offline data is also applicable to train the texture features using machine-learning technique.

Raja H Ali, et al., (2017) has described the MCMC algorithm for graphical and statistical analysis of data. Here the Visual MCMC technique is proposed for statistical analysis. Posterior distribution parameters are estimated to determine the large scale parameters. Graphical tree mining technique estimates the convergence curve to model with the post processing technique. GUI (Graphical User Interface) based applications are supporting to the graphical analysis of data mining techniques. The continuous parameter of the tree mining technique is estimated on MCMC analysis.

Andrew C Flores, et al., (2018) has presented the sentimental analysis model using SVM and Naïve Bayes classifier. The user experience-evaluation model is performed with the SMOTE (Synthetic Minority Over-Sampling Technique). Ten fold cross validation is applied for texture feature evaluation on the machine learning technique. The pre-processing of dataset is performed for reducing the noisy data. Sequential mining optimization technique is used for continuous sample in the database, which performs the feature analysis.

Jing Liu, et al., (2018) has proposed the user experience design of 3D websites, which is performed by using gray correlation analysis. Statistical assessment model is created to analyze the feedback of 3D website user. Here the gray correlation analysis model prefers the subjective and objective factors form the dataset. The minimum and maximum sequence sequences are correlated with the interconnect degree finding method.

Kudakwashe Zvarevashe, et al., (2018) have proposed the UX design model using sentimental analysis and opinion mining techniques. Hotel reviews are collected to form the dataset and it is analyzed by UX design. Hotel customer feedback is collected and it analyzed with the opinion-mining task. The labeled datasets are performing training and testing by extracting the unbiased opinions. Here the naïve bayes multinomial classification with minimal sequential optimization is performed to analyze the service of hotel.

Lei Chen, et al., (2017) have proposed the UX model for multimedia communications. Here the lightweight end-side user experience dataset is used for improving the evaluation model. Communication network in the multimedia application is performed to evaluate for quality measurement. Multimedia mobile communication data is used for UX evaluations. QoE is estimated based on the functional application of UX model to improve the multimedia services. Mobile services and user behavior is analyzed to provide the result of QoS (Quality of Service) and QoE (Quality of Experience).

Qi Yu, et al., (2018) has designed the hybrid UX model for predicting the mobile game usability. The user experience-evaluation model helps improve the product. Here the improved ripper algorithm is proposed to obtain the analysis result of game event. The test dataset contains 300 tester data of least five minutes user who played the parkour mobile game. Here the ripper algorithm is performed with missing attributes, which specified their subjective values.

Quinate Ihemedu-Steinke, et al., (2018) has proposed the virtual reality display model for UX design. The experiment is done with the statistical model of virtual reality affects. The display evaluation compares the flat screen, television monitoring, and quality factor. Here the 55 inches smart
HD TV is used to compare UX model with conventional work. HMD product design is used for VR technology to model the UX design, which improves the quantity requirement from the user.

Yingying Xie, et al., (2019) have described about the UX design for personal cloud storage services. Here the user experience-evaluation scheme is presented to analyze the quantification model of personal cloud storage service. Linear and prospect theory evaluation is designed for compensatory approach. The non-compensatory approach includes the conjunctive and disjunctive approaches. The cloud storage service is described for china case study dataset. It uses prospect theory to evaluate the UX model.

**EXISTING WORK**

The user experience-evaluation model is designed to perform the Google analytic data using advanced Markov chain Monte Carlo algorithm. The website page view, session, bounce rate are collected to form the dataset. The posterior distribution function is done by MCMC algorithm, which provides the probability factor of distribution parameters. The existing work evaluates the user experience model for Google merchandise store, which is collected from the Google analytic websites. Prior and posterior distribution function of MCMC is designed to estimate the UX parameters from the dataset. Page view and page session for particular time of user evaluates the probability of usability of the particular page. The product design specification is clearly viewed by the user who searches the item. The statistical data estimates the number of user for particular website makes the usability of that page. The affective factor for UX model shows the particular website result. The conventional UX design model is improved by adding machine-learning classification using kernel based naïve bayes classification. This process utilizes the Google analytic data to model the statistical data and it utilizes the advanced Markov Chain Monte Carlo method to perform the posterior distribution function. Here the texture feature classification is performed to train the data of Google merchandise store.

**PROPOSED METHODOLOGY**

The proposed research work employed with the utilizing naïve bayes classification and MCMC technique to estimate the user experience. Here the Google analytics database is used to evaluate the UX model, which estimates the prior and posterior results. The web page usability is evaluated to improve the product through online. Here the ML algorithm is proposed with the kernel function to estimate the parameters. The dataset contains information about the website/page visitors; number of visitors, timing, review, etc. in this, the web page is improved for user satisfaction. The feedback of user is evaluated to enhance the product through UX design model. The Google analytics provides the data of flow report, review report, user’s activity, and in-page analysis report. This makes the online learning behavior. The UX evaluation model appears when it requires additional page improvement. This statistical information is processed with the MCMC technique to estimate the UX model. Here the classification technique is used to classify the statistical data from Google analytics. The behavioral detection approach utilized with data of user’s feelings, emotions, attentions, etc., which evaluated by UX model. This proposed UX model performed by kernel based naïve bayes classifier with MCMC technique. the block diagram of proposed UX model is designed using machine learning and MCMC methods, which is given in the figure 2.
In this proposed research work, the decision making process is performed on the UX model, which optimizes the classification technique. The gain and loss value is evaluated to determine the affective state of UX model. The mapping of design parameter is to identify the user experience’s affective states. The product design is improved based on the user’s review from UX design. UX process evaluates the statistical data from the Google analytics page. The user’s transfers the opinion through online and it collected as a Google analytics data for respective product/pages. These data is estimated to make the interaction between the customer and the employees, which results to provide the better product to satisfy the customer. This proposed system takes the dataset of Google merchandise store from the Google analytics, which contains details such as page session, visited user, new users, page viewers, uniqueness in page viewers, average-time of staying of the page, entrances, bounce rate, page exits and page values.

A. Dataset Description

The user experience model is evaluated for Google merchandise store and it is downloaded from the Google analytic page, which provides the information about the usability of web page. The dataset contains, page view, page session, number of users, bounce rate, timing of page sessions, page value, etc. The active user’s of the Google analytics page is shown in the Fig3.
Here the analytics reports are measured to increase the usability rate through UX model. Here the kernel-based approach is used to track the data on peak point. The feedback of the customer is evaluated by cumulative prospect theory and it is functioned with the Markov Chain Monte Carlo method, which performs the prior and posterior distribution function.

**Fig4. Behavior flow model**

The behavioral flow model is shown in the fig4. Here the behavior flow model of Google analytics merchandise store is evaluated to rare the review and it helps to model the UX index. The interaction model uses decision-making approach in the UX design and it evaluated using MCMC method. The kernel function based machine learning classification provides the better result and it is utilized with the MCMC technique for performing the prior and posterior distributions. The usability of the product is estimated to improve the review rate. The web page usage rate is modeled from Google analytics data, which is shown in the Fig5.

**Fig5. Page usage estimation**

The usability of product/webpage is estimated to model the UX design, which efficiently optimized with the MCMC technique. The cumulative prospect theory is applied to estimate the affective model and it utilizes the classified data from naïve bayes approach. This improved method is called advanced MCMC technique.
B. Kernel based Naïve Bayes Classification

Affective model uses naïve bayes predictor, which provides the result based on the current state of user experience. The possibility if estimated to model the graph. The classification is performed with opinion mining techniques, which results are send to the MCMC for evaluate the UX model. Initially, the data pre-processing technique is performed to eliminate the unwanted data from the dataset. The data preparation task gives the knowledge about the decision-making. The training and testing is performed after initializing the input parameters of naïve bayes classifier. The post processing stage analyzes the data attributes and performs the advanced MCMC technique.

![Flow diagram of K-naïve bayes classifier](image)

The kernel tricks are applied on the naïve bayes classifier to predict the affective states. It is a supervised classification algorithm, which is used to predict the affective states effectively on the UX design. The probabilistic evaluation is determined on the prior and posterior distribution function. The positive and negative attributes are selected on the training set and it performed with kernel function to cluster the given vectors. In the data alignment process, the input parameters are initialized based on the given attributes set. Confusion matrix provides the predicted class values of classifier. Assume the input vector of naïve bayes classifiers as follows,

\[ x_1, x_2, x_3, \ldots, x_n \]  \hspace{1cm} (1)

The probability of predicted class is \( Ca \) and it is processed with confusion matrix that is given by,

\[ p \left( \frac{Ca}{x} \right) > p \left( \frac{Cb}{x} \right) \quad \forall \ 1 \leq b \leq c \text{ and } b! = a \]  \hspace{1cm} (2)

By maximizing this predicted class, the posterior function is given as,

\[ p \left( \frac{Ca}{x} \right) = \frac{p(x/Ca) p(Ca)}{p(x)} \]  \hspace{1cm} (3)

The constant class value obtains the result of different probability index and it assumes the same class for each predicted levels, which is given by,

\[ p(c_1) = p(c_2) = p(c_3) = \ldots = p(cn) \]  \hspace{1cm} (4)

The maximized level of training process evaluates the probability of predicted class, which is expressed as,

\[ p(x|Ca)p(Ca) \]  \hspace{1cm} (5)

The opinion mining is applied for naïve bayes classifier, which is expressed as follows. Here the feature vectors are trained based on the class value from the confusion matrix. The negative words similarity is found to make the test set from the classifier. Then probability of data and respective class is expressed as,
The feature probability is expressed by trained data from the classifier.

\[ P(x|Ca) = p(x_1, x_2, \ldots, x_n|Ca) \] (6)

\[ C_{conf} = \arg\max_{c \in Ca} p(c|x) \]

\[ = \arg\max_{c \in Ca} \frac{p(x|c)p(c)}{p(x)} \] (8)

\[ = \arg\max_{c \in Ca} p(x_1, x_2, x_3, \ldots, x_n|c) x p(c) \] (9)

The probability estimation is done with labeled data of supervised machine learning. The positive and negative polarity is varied based on the attributes of database. Opinion mining is performed on this classification approach to model the affective graph.

C. Advanced Markov Chain Monte Carlo Method

The prior and posterior distribution function is performed on the MCMC method, which is mainly used for statistical data analysis. The sample data is processed to find the probability of classified results, which is send to the A-MCMC to evaluate the UX design. The Monte Carlo technique solves the complex problem in mathematical analysis. The input data is assigned as a vector and it is used to analyze the prior and posterior function. After performing normalization on classifier, the MCMC will perform the distribution function. The probability event of statistical data is the posterior probability distribution function. The prior probability distribution function is previous states data to be processed on the prior function. By adding prior probability function and the current data probability function is known as posterior probability distribution function, which is expressed as,

\[ \frac{d\pi}{d\pi_0}(x) = p(b/k) \] (11)

Where, \( k \) denotes the error removed data and \( b \) is the data with likelihood of prior function. The approximated value is estimated to apply the proposition statement to the MCMC, which utilizes continuous and finite-difference approximation method. Here the user feedback is evaluated to improve the usability of the product through web page. This supervised machine learning approach is applied through the advanced Markov chain Monte Carlo method to estimate the affective model of customer’s data. This technique is result to the prior and posterior distribution function.

RESULTS AND DISCUSSION

Thus, the design of UX model is used to estimate the affective graph by using novel techniques. The dataset of Google analytic page is collected to form the statistical data to perform the user experience model. The statistical data is shown in fig7. Here it performed with the kernel based naïve bayes classifier and advanced MCMC technique. The MI algorithm utilizes the opinion mining for classification, which obtains the result user’s feedback either it is positive or negative. In the kernel based naïve bayes classifier, predicted classes form the confusion matrix, which is shown in the fig8.
Fig7. Statistical data from Google analytics

The training data is selected randomly to process with the statistical features. The labeled class value is predicted by input vector, which assigns from the classifier. The confusion matrix formed based on the true positive, true negative, false positive and false negative values, which determines the accuracy of the classifier. In A-MCMC technique, the statistical data is analyzed to provide the results as below. The page return level versus density is estimated for 50 pages of data, which is given in the fig9.

Fig8. Simulated result of kernel based naïve bayes classifier

Fig9. Page return level estimation in Advanced MCMC
The web page evaluation is performed by proposed UX design and it optimized with the advanced MCMC technique. Number of customer to visit the website and its page return level is shown in the fig 10. Here the prospect theory is applied to perform the posterior distribution function. Prior and posterior distribution is performed to evaluate the user experience data. Page value versus bounce rate is plotted on the fig 11. The statistical data analyzed with the report of behavior, web page details, user’s feedback, page visitors, etc. in this the A-MCMC calculates the probability to perform posterior distribution function. A page with maximum users is measured to find the most likely web page.

The statistical data from the Google analytic report is evaluated to improve the product through UX design. The minimum page value is taken to improve the product/page through feedback from the user. The UX evaluated report of statistical data is shown in fig 12.
In this proposed research work, the UX design model is constructed with the kernel based naïve bayes classifier and advanced MCMC technique. The web page evaluation is done along with the use of cumulative prospect theory. Here the naïve bayes classification provides the 65% of accuracy for this proposed method. The previous work only analyzed with the MCMC since it not providing accurate result since the proposed method achieves the result classification accuracy. The comparison result of proposed and existing work is given in the table 1. Behavior flow rate, scaling factor, website visitor and most likely product’s rates are evaluated to improve the web page/product.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Existing work</th>
<th>Proposed work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate of customer flow</td>
<td>0.351</td>
<td>0.904</td>
</tr>
<tr>
<td>Scaling factor</td>
<td>0.4493</td>
<td>0.0754</td>
</tr>
<tr>
<td>% of same website used</td>
<td>2.5718</td>
<td>1.579</td>
</tr>
<tr>
<td>% of most preferable brand</td>
<td>0.66</td>
<td>0.73</td>
</tr>
<tr>
<td>Elapsed time (sec)</td>
<td>5.221</td>
<td>4.566</td>
</tr>
</tbody>
</table>

Thus, the proposed result of UX design estimation is done with the help of MATLAB 2018a. Here the statistical data is used to analyze the performance of the webpage, which improves the page through feedback of user and it is evaluated by UX design. The computation time of proposed method is 4.566seconds, which reduces than the existing work. The accuracy of classified data is measured through kernel based naïve bayes classifier. The ML algorithm and MCMC method improves the performance of the proposed UX design. Here the affective graph models provided the gain and loss of the product, which is obtained through MCMC approach. Therefore, the machine learning and MCMC based UX design model gives the best result than the existing work.

**CONCLUSION AND FUTURESCOPE**

Thus, it concluded with the proposed UX design model for user experience evaluation of web page/product by statistical data from the Google analytics. The quantitative information of user experience is evaluated to improve the product through online page. The UX design approach is applicable for various industries, organizations, institutions, etc. The kernel trick is applied on the naïve bayes algorithm to getting efficient classified result. Specifically, training process utilizes the kernel function to select the training data and then it performs testing to provide the classification result.
Opinion mining task is performed for feedback analysis and it sets the review rate on Google analytics. The affective graph model shows the user experience report for improving the product through respective web page. Therefore the proposed opinion mining based UX model is designed to evaluate the user experience using kernel based naïve bayes classifier and advanced MCMC technique, which results to accuracy of the classifier and different parameters about the statistical data. This method achieves the result of accuracy and computation time than the existing method. In future work, the proposed method is enhanced by utilizing support vector machine learning technique and other optimization algorithm to efficiently perform the statistical data than this method. The MCMC technique is improved by adding some other mathematical approach.

REFERENCES


