Detection of natural structures and classification of HCI-HPR data using robust forward search algorithm

ISIAKA, Fatima, MWITONDI, Kassim and IBRAHIM, Adamu M.

International Journal of Intelligent Computing and Cybernetics

DETECTION OF NATURAL STRUCTURES AND CLASSIFICATION OF HCI-HPR DATA USING ROBUST FORWARD SEARCH ALGORITHM

Fatima Isiaka*
Department of Computing, Sheffield Hallam University, Sheffield, S1 1WB,United Kingdom †
fatima.isiaka@student.shu.ac.uk‡

Kassim S. Mwitondi
Department of Computing, Sheffield Hallam University Sheffield, United Kingdom

Adamu M. Ibrahim
School of Computing, University of Leeds, Leeds United Kingdom

Received: 23 August 2015
Revised: 2 September 2015
Accepted: 20 January 2016

Abstract

Purpose - This paper proposes a forward search algorithm for detecting and identifying natural structures arising in Human-Computer Interaction (HCI) and Human Physiological Response (HCI) data.
Design/Methodology/Approach The paper portrays aspects that are essential to modelling and precision in detection. The methods involves developed algorithm for detecting outliers in data to recognise natural patterns in incessant data such as HCI-HPR data. The detected categorical data are simultaneously labelled based on the data reliance on parametric rules to predictive models used in classification algorithms. Data was also simulated based on multivariate normal distribution method and used to compare and validate our original data.
Findings - Results shows that the forward search method provides robust features that are capable of repelling over-fitting in physiological and eye movement data.
Practical implications The authors conducted some of the experiments at individual residence which may affect environmental constraints.

*45 West Point, 58 West Street, S1 4EZ, Sheffield., UK
†Howard Street, Sheffield, S1 1WB.
‡fatima.isiaka@student.shu.ac.uk..

Originality/value - Graphical improvement of the forward search algorithm was done with Deducer package in R to present vivid results that indicates detected stress levels in users.

Keywords: Human physiological response; HCI; Categorical data; Maximum likelihood estimation; Robust forward search.
Paper type: Research paper

1. Introduction

For continuous data, a very fast robust method is often used for the detection of multiple masked effects (outliers) [Atkinson (1994)]. This type of data often requires regression models frequently used on least squares or least median squares (LMS). For the purpose of detecting outliers and predictions, the estimation uses very precise estimates of the parameters and excessive computations. The main goal here is to use the forward search algorithm for the robust estimation to detect natural structures and the possibilities of future Human Computer Interaction-Human Physiological Response (HCI-HPR) modelling. The parameters are chosen based on their capabilities for response detection and p-values obtained from the subset selected based on the generalized additive model. The forward search has the potential for computation of parameter estimates. For large datasets, the multi-tasking approach affords a considerable reduction in computational time.

1.1. *Motivation*

The motivation behind this paper is the need to gain insight into HCI and help to monitor and control stress level, not only directed to interactions with webpage contents, but other visual stimuli. We try to achieve this by investigating how user-generated data be utilised to model and simulate human physiological responses to visual contents of the web. The approach to modeling here is using the forward search algorithm, which is capable of detecting natural structures that might arise in the data.

2. Related work

It is possible to display the pattern of multivariate outliers found during each forward search [Atkinson (1994); Mwitondi and Said (2011)]. Atkinson describes the fast forward algorithm for multivariate outliers and lays emphasis on regression. They also used the stepwise method to study naturally arising patterns in data for labelling of continuous data; the primary focus is on conspicuous and stability between analytical precision and model dependability. To run through this concept, we thought to develop the data collected from the HCI-HPR study, both in optimisation by maximum likelihood estimation. We start by modelling of the criterion data sample acquired from diverse dimension of communication.

A common finding in modelling of cognitive processes [Kretzschmar et al; (2013)], illustrates that a person's subjective perception of their behaviour doesn't always relate to their neural activity. Speculations such as people do not always know what is going on in their minds is asserted, the prospect of directly observing what is going on in the users mind was proposed. An eye tracking study that involves reading [Liversedge et al; (2007)] and provides objective real-time quantitative measures of eye movements, revealed longer fixation times for reading text with transposed letters as compared to reading regular text, but users claimed to spend few seconds on text with transposed letters. Also Electroencephalography (EEG) of language processing [Demiral et al; (2008)], concluded that phrases judged as easy to comprehend and highly acceptable sometimes have larger processing effort on the part of readers. These findings showed that visual attention and engagement are not necessarily linked to user perceptions of experiences, today we struggle to determine the best possible techniques to allow us have a greater understanding of the user's unconscious minds as they interact with elements of a visual stimulus like a webpage.

Physiological measures have evoked enthusiasm regarding this aspect. The primary interest is on mechanisms [Green (1976)][Filipovic et al; (2001)] that tries to answer simple questions on functional phenomena.

Modelling physiological data [Andreassi (2000)] and cognitive state estimate has proved rather complex especially with data scaling and interpreting of results. [Kramer (1990)][Dirican et al; (2011)] describes that there are three types of dimensions that are used in reasoning and anxiety valuation, this includes individual measures, functional measures which are also physiological measures and recital measures [Farmer and Brownson (2003)][Cain (2007); Dirican et al; (2011)]. These are also stated with the mental load [Ganglbauer (2009)][Dirican et al; (2011)] and psychological workload/stress assessment. Though their meaning may show differences, they can also be used in the same context [Kalyuga (2008)][Dirican et al; (2011)]. Cognitive load is a conception that is connected to functioning memory in mental workload concept [Kalyuga (2008)][Dirican et al; (2011)], while on the other hand, mental/stress workload has a more challenging and has multi-dimensional concept attached to it. It is either related to task difficulties or potentials, motivation and the emotional state of users [Kramer (1990); Farmer and Brownson (2003); Dirican et al; (2011)].

The physiological measures used for this study are chosen based on their strength and weakness and also on their analytical assessment founded in literature. For Skin Conductance Response (SCR), the phasic changes were considered since its temporal sensitivity is poor, on tonic changes. Mean Skin Temperature (MST) has an impact on diagnosis; as anxiety intensifies, temperature declines as heat transfers to the forms core (central body). Eye movements (Saccade, Fixations in X and Y coordinate) was chosen based on the fact that the eye gaze is an undeviating quantity of reasoning, and it is a good metric for interface assessment and usability testing. The pupil dilation (pupil changes) of the eye correlates with mental workload (stress) and reacts to the expressive valence. Though hard to relate in everyday context since the eye responds to diverse light settings [Ganglbauer (2009)][Kramer (1990)], but useful when apparatus used combines both the eye movement with changes in pupil size. Physiological measures are continuous and time-varying, and when combined with behavioural measures [Insko (2003)], can lead to boundless stages of making a better design decision for user interface design.

3. Methodology

The method for data collection were based on objective, behavioural and physiological methods. We have a total of 44 participants aged between 18-43. Ten student members was used to carry out the pilot study, with approval from University of Manchesters Committee on Ethics on Research Conducted on Human Beings. Then we have ten each of older students, workers group and average users group. The groups were selected for the purpose of determining the masking effect in their data and variations in modelling data arising from diverse HCI-HPR associations. The stimulus used are webpages with active and deactivated contents. For the purpose of inducing measurable and visible spikes in their physiological measures (Skin Conductance Response (SCR), Skin Temperature (ST) and eye movement). Each participant interacted with six web pages and a total of 264 data was collected and exported to Matlab for data synchronisation and organisation. The software package used for statistical analysis was R, for its capability with graphical and multivariate analysis. The following section discusses the steps for the forward search algorithm and detecting of natural structures for categorising the HCI-HPR data with associated affect states.

3.1. *Forward Search Algorithm*

It is commonly known that objects that are analogous fall in line with each other. Most fundamental theories are built on these concepts. This provides a basis for defining rules that categorize observations as outlying are defined. The method is based on Mahalanobis distances (Algorithm 1) as a measure of the distance between a point and its distribution. The mahalanobis distance is based on a multi-dimensional generalisation of the idea of measuring how many standard deviations away from a point p, given as

$$DT_m(x) = (x-\omega)_T\, \mathcal{S}^{-1}(x-\omega)^{0.5} \tag{1}$$

Where $x = (x_1, x_2, \ldots, x_N)^T$ set of observations, mean $\omega = (\omega_1, \omega_2, \ldots, \omega_{\mathcal{N}})^T$ and covariance matrix S. To detect anomalies, numerous reproductions of data for calibration of residuals are run on datasets with reliant variable simulated from standard normal distribution and the same design matrix. Each simulation calculates the estimated scale parameter given has:

$$\tilde{\sigma}_m^2 = e^2_{[med]}(\beta_m)$$

Where $e^2_{[med]}$ represents the median of the raw residuals σ_m^2. The scale parameter is regular over the initial points for each model taken, modifying for uncommon approximations of the Sigma that standardises the plots. The modifications are used for the building of the regularised residuals. The following equations is used to define cut-off point based on first principles.

$$t_i = \frac{e_i \bar{\sigma}_m}{\tilde{\sigma}_m \sqrt{(1-h_i)}}\ \forall\, i \in\ M_m$$
$$t_i = \frac{e_i \bar{\sigma}_m}{\tilde{\sigma}_m(m) \sqrt{(1+d_i)}}\ \forall\, i \notin\ M_m$$

Where h_i is the hatmatrix diagonal of the same distance $d_i = x_i^T(X^TX)^{-1}x_i$ with X has the explanatorty variable of simulated response Y of a linear model. The forward search moves from fitting m observations to $m+1$ by choosing the $m+1$ observations with the smallest residuals [Atkinson (1994)]. The value of the residual at the starting point serves has the performance for a given model fit.

For our approach, we define the cut-off based on Mahalnobis distance, the following cut-off is used.

$$X_P((n-0.5)/n, p) \tag{2}$$

Algorithm 1 Mahalnobis Squared distances

procedure MAHAL(X, index)
top:
    if $(!is.matrix(X))$ **then return** false
    $Xbar \leftarrow$ *apply(x[index], 2, mean)*
loop:
    if $(is.matrix(X))$ **then**
        $S \leftarrow solve(S)$
        $Xcent \leftarrow t(t(X) - Xbar)$.
        $apply(xcent, 1, function(X)X\% * \%S\% * \%X)$
        goto *loop*.
        close;
    goto *top*.

Given that our input matrix is X with p parameters and size n, the natural structures of the dataset are detected in the steps outlined in Algorithm 2. The goal is to largely condense the masking outcome that may arise as a result of outlying existence on the evaluations of means and covariance calculated from the original data. The main idea here is, given certain distances m observations for the valuation of means and covariance, observations at $m + 1$ positions to be recycled for the valuation in the next phase are probable to $m+1$ lowest space. Hence, an occurrence can be incorporated in the subgroup used for the approximation for say, value m, this can later be exempted as m increases.

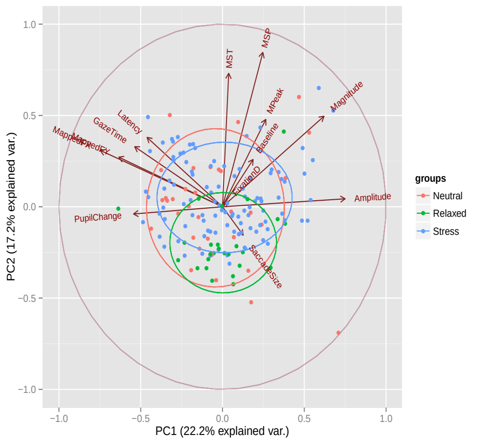
The output plot vividly demonstrates the progress of the set of outliers as the scope of the fixed subgroup m increases. Originally m is usually chosen as $p + 1$ where p is the size of parameters; the smallest number allowing the computation of the requisite mahalnobis distances. The cut-off point commonly used to define an outlier is the extreme predictable value from a set of n size arbitrary parameters with a chi-squared distribution on p degrees of freedom, which is given approximately as the threshold (Equation 2).The proceeding sections illustrates results obtained first for determing natural structures based on Principle Component Analysis PCA) and validating on best fit for different group of data used to run the robust forward search algorithm.

Algorithm 2 Robust forward search of matrix X

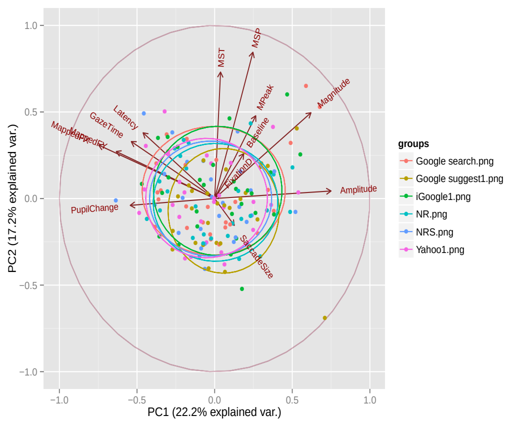
procedure MATRIX(X)
  $X = matrix(X)$
  rn = rownames(X)
top:
  if $(is.null(rn))$ **then return** false
  rn ← *1:nrow(x)*
  n ← *length(X[1])*
  p ← *length(X[1])*
  s ← *1:n*
ind = matrix(0, n - p, n):
  $ind1 \leftarrow 0$
thresh = qisquare((n - 0.5)/n, p): ▷ determine qisquare of parameters
  $index \leftarrow 1 : (p + 1)$
  procedure (i in (p+1):n)
  $ind1 \leftarrow ind1 + 1.$
  if $(i == (p + 1))$ $D \leftarrow Mahal(x, index)$ **then**: ▷ mahalanohis distances
    $index \leftarrow order(D).$
    $i \leftarrow i - 1.$
    $index1 \leftarrow sort(index[1 : i])$
    $D \leftarrow mahal(x, index1)$
    $index2 \leftarrow s[D > thresh]$
    $ind[ind1, index2] \leftarrow ind[ind1, index2] + 1$
    goto *loop*.
    close;
  $y \leftarrow rep(1 : (n - p), rep(n, (n - p))).$
  $x \leftarrow rep(1 : n, (n - p)).$
  $residuals \leftarrow -c("", "*")[as.vector(t(ind[(n - p) : 1,])) + 1]$
  print *x, y*

3.2. *Natural Structures with Principle Component Analysis (PCA)*

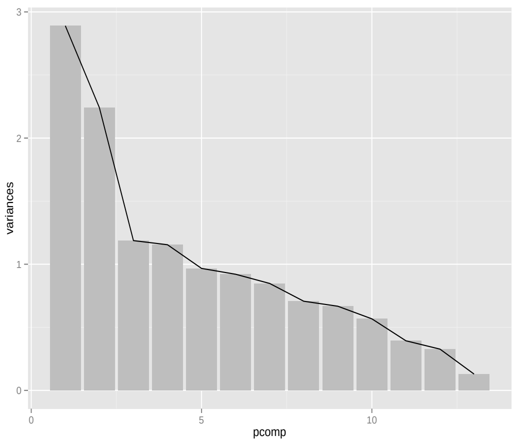
The PCA is another state of the art method used in detecting natural structures in data [Wold et al; (1987)][Martins, et al; (2015)] and it is used here as a comparison and ordination method that reduces the dimensionality of the datasets, this is done by creating different key explanatory variables which are the principle components (PCs). Each of the components accounts for as many variance in the data as possible as far as the principle analysis are unrelated. All the components are independent and quadratical. The components detected are used to determine the relative contribution of each of the original variables to each of the principle analysis. The steps involve standardising the variables to the same relative scale and avoiding some variables to become dominant because of their large measurement units. The forward search algorithm is then used to validate the outcome of the PCA.



(a) PCA of data with stress levels as class labels



(b) PCA of data with websites as class labels



(c) Variance plot showing variations in components

Fig. 1: PCA show variations in components and best features

4. Data Visualisation and Generalised Additive model (GAM)

Before analysis, the data are explored and visualised to see the nature of the data we collected and also predict the nature of possible outcome of the result. Figure 1 above indicates that the data are sparsely distributed. From the result of the PCA, MappedFX and MappedFY are highly correlated. The total percentage of the two sets of principle components (PC1 and PC2) is 39.9% of the datasets variability which might not be a respectable sum for the total variable. To validate this, the forward algorithm is then employed. But first we look at this from the GAM aspect. For physiological parameters, visible spikes in response typically indicate a phasic change and reactions to stimulus especially in SCR; for the purpose of determining HCI-HPR relations, the average peaks on response is used as the response variable. Considering negatively correlated parameters (Figure 2) such as average peaks (MPeak) with Fixation Duration (FixationD) and Latency. This sort of variables obeys the reversal paradox [Good et al; (1987)][Simpson (1951)] where a trend that appears in different groups of data and disappears or reverses when combined. It can also be related to situations such as the more the eye fixates on a particular object of interest the less the spikes on physiological data or the reverse may be the case.

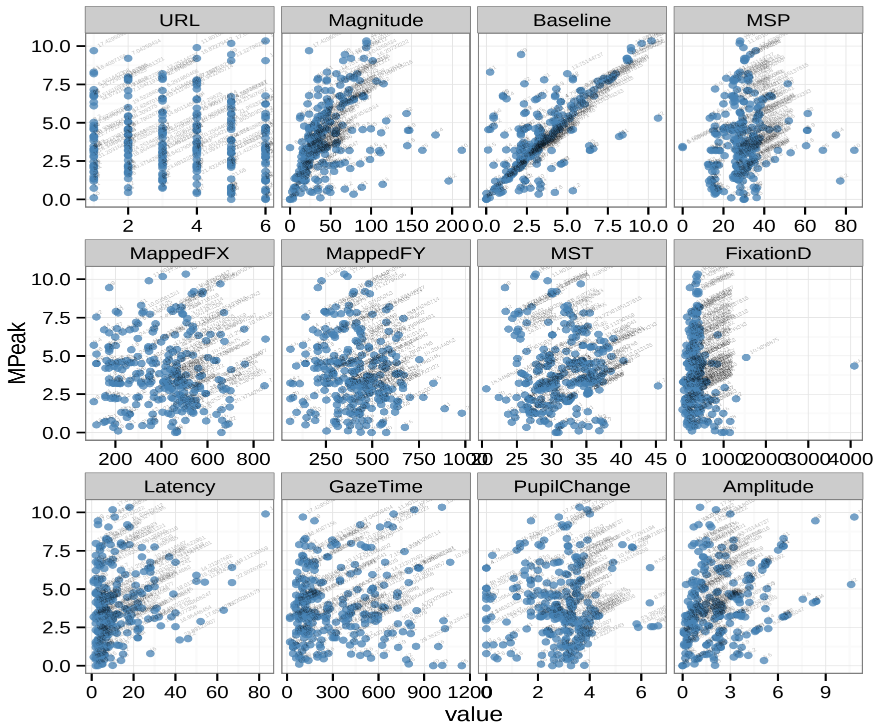


Fig. 2: Scatter plot showing correlation of response variables and predictors

The additive model considers all parameters as smooth. Not all parameters obey the linear regression fit as shown in Figure 3. This is based on the GAM that treats all parameters as stable, especially amplitude, mean skin potential (MSP) the eye movement data (FixationD and Pupil change). The GAM replaces the linear form of each parameter by a sum of smooth functions, which are unspecified and estimated using scatter plot by local scoring. It applies to any likelihood-based regression model and all predictors are replaced by additive predictors [Hastie and Tibshirani (1990)][Beck and Jackman (1998)]. To determine the nature and masking effects of our datasets, the steps in Algorithm 2 was adopted, and the following section discusses on results obtained.

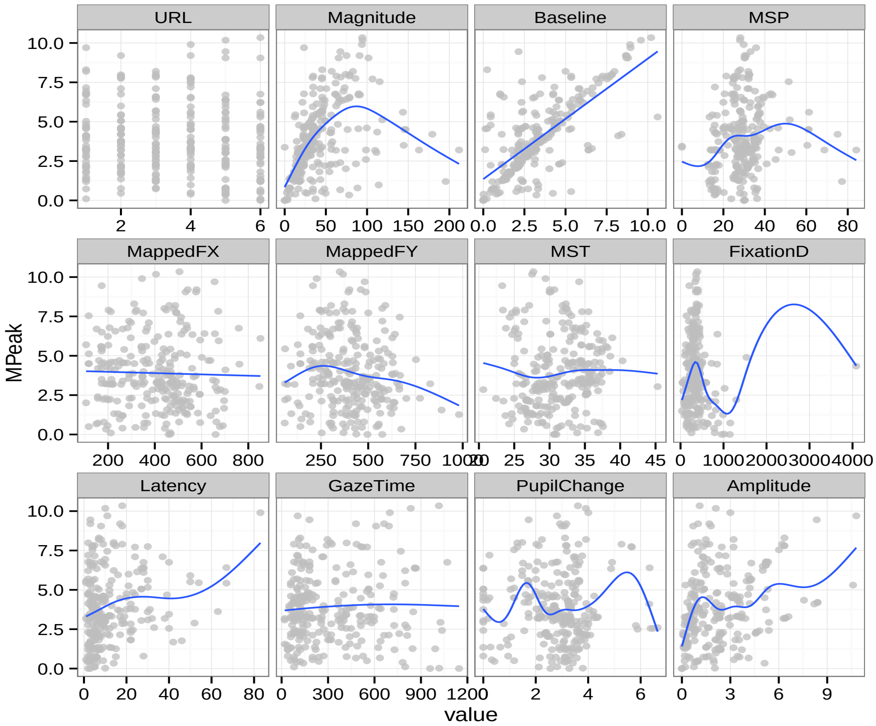


Fig. 3: GAM shows parameters that adopts linear fit and un-linear fit

5. Results

For the output of the plots, the cut-off point as mentioned earlier determines the starting point for the subsets based on the least absolute residual value for each given parameter and the variations in the data. The following sections discusses results obtained from a group of dataset.

5.1. *Pilot data*

The pilot data obtained contains sixty observations with fifteen physiological parameters. Comparing the linear model fit of the data with R-squared ($r^2 = 0.85$ *and* $p < 0.01$), and the plot in Figure 4 showing the result of the forward search algorithm, the best possible fit for the affect state with starting point at observation 15 gives 85%. This can be termed an even set for this particular dataset.The behaviour of the search is illustrated by the blurry dots, and the first detection of the darker dots appear as the largest residuals for an observation, indicating an outlier. For instances $m > 41$, observation 22, 39, 40, 47, 52 and 59 are outliers. Figure 5 shows the detected stress level based on the forward search algorithm. The red cross shows observation with residuals greater than the cut-off while the blue cross represents residuals less than the cut-off. For each case the presence of each point shows the categorical data for observations that are capable of been labelled. It also shows the variation of highly stress participants found in the NRS (National Rail enquiry page); Google suggest page (have active content), Google search page (contents static) and iGoogle/Yahoo page. For these participants, the amplitude and changes in pupil size have the most impact with p-values ($p > 0.05$).

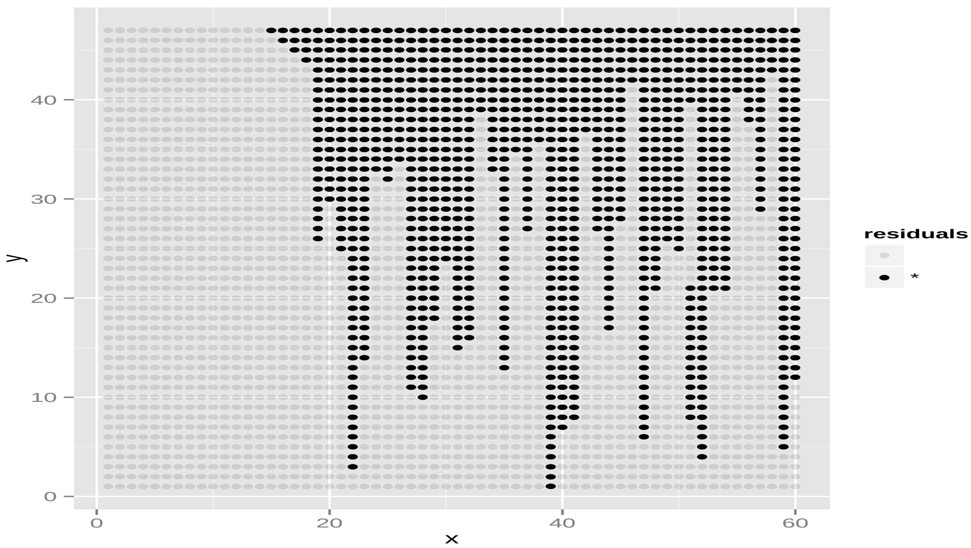


Fig. 4: Best fit for pilot data

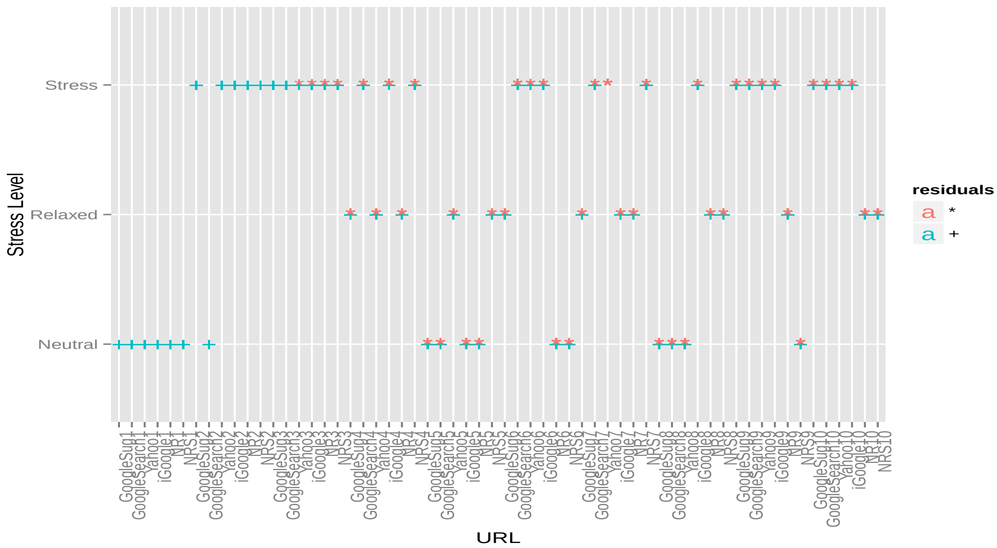


Fig. 5: Detected categorical class for pilot data

5.2. *Older Student data*

The dataset for older student contains 60 observations and 15 explanatory variables just like the pilot study. This part of reflecting of society more accurately by including older participants. It is known that being young, technically conversant and extremely cultivated is not signifying demographic certainties. In research, it is more feasible to include older participants to compare the differences in interaction and model evaluation from other groups. The result shows best fit for running the algorithm on the dataset.

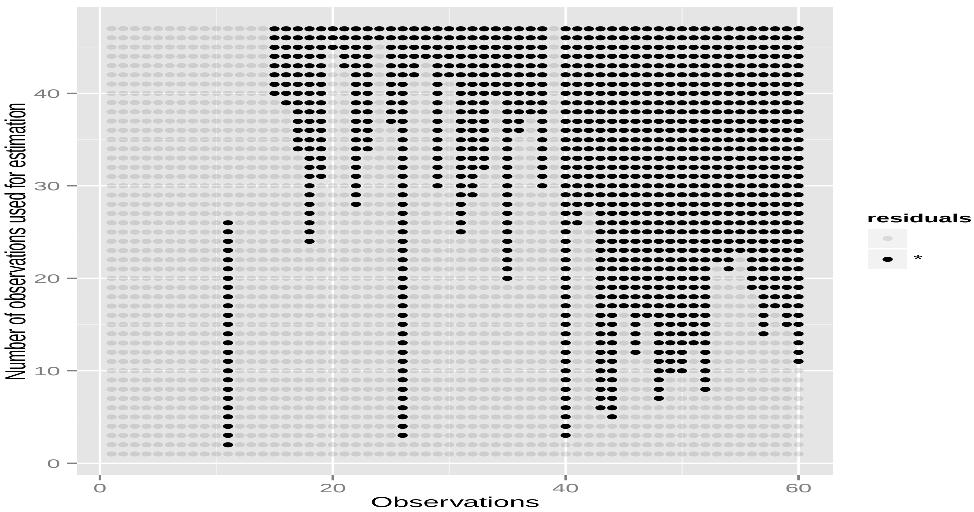


Fig. 6: Best fit for older student data

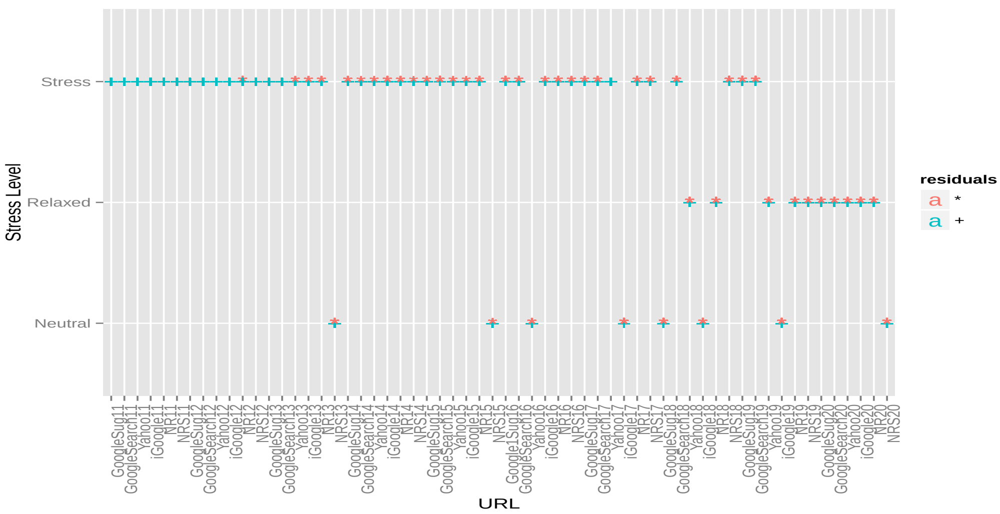


Fig. 7: Detected categorical class for older student data

For this plot, we have two starting points that reflected in the Figure 7. The blue cross appear only on all blank area of space of the fit in both models and before starting points of Figure 6. The starting points are noted to be instances 11 and 15 with performance (-0.9 and 1.5). Comparing this with R-square value ($r^2 = 0.83\ and\ p = 1.663e^{-14}$) it seem to be more feasible for this set of data. For this model, the magnitude of response and intercept seem to be of high significance. The class value for the high-stress level is also dominating, and maximizing contents that appear to contribute to relaxed should be of high priority. Associating this to pilot data, it gives an impression to be a good model proposition for prediction.

5.3. *Workers Group*

Specifically we considered highly engaged workers, to determine if behaviour, tools and practices that drive their engagement can influence and improve model performance. We can use this to define prediction of users interaction in both areas of interests and expression attributes towards web pages. To determine the level of model performance, based on workplace, we run the algorithm on the data collected for this group. Below is the best fit for the model. Highly stressed users are still the motivating class factor for this set of interaction, which could also relate to the amount of work added to their cognitive process during office hours as compared to pilot data.

Fig. 8: Best fit for Workers group data

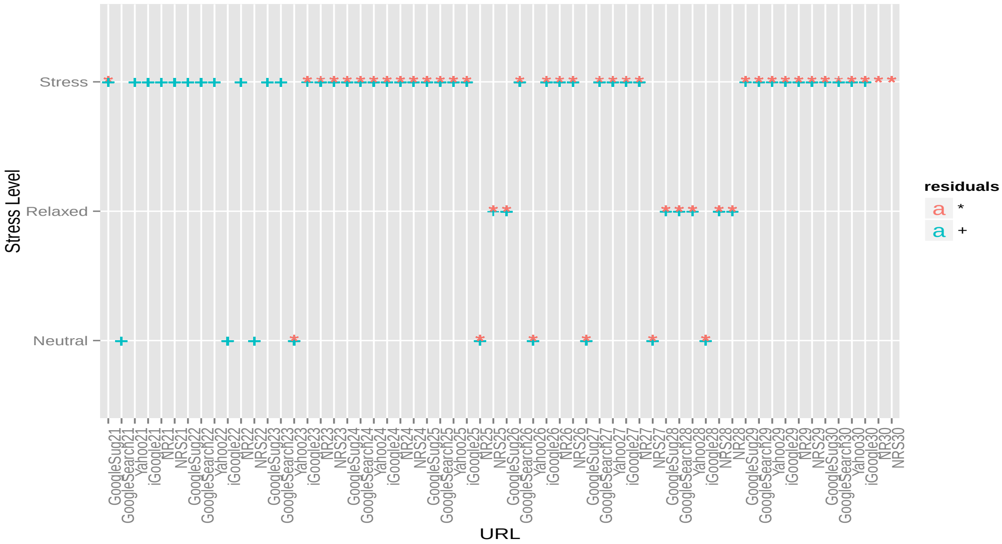


Fig. 9: Detected categorical class for Workers group data

The starting point for this is the first instance for this group and observation 15 on the iGoogle page. A residual value and performance of 0.06 and 0.10 compared to R-squared of ($r^2 = 0.9826 \; and \; p = 2.2e^{-16}$), is enhanced from previous datasets and also for advancing subsequent models. The model is practicable in terms of prediction of future occurrences such as events and behaviour relating to interactions with the web contents. Mostly where users are likely to look at and what will optimize marking strategy and boost sales based on the heuristic analysis.

5.4. *Regular Users Group*

The regular users group is a combination of users from groups categorized as average, a form of simplification and influencing of model performance based on generalized population. This group contains 64 instances with 15 explanatory variables. The starting point for this model after running the forward search algorithm are three (8, 12, and 15) with residuals 0.01, 0.24 and 0.1. Comparing this to the R-squared value ($r^2 = 0.97 and p = 2.2e^{-16}$), the best solution found for this model is highly significant but the workers group superseded with 0.01.

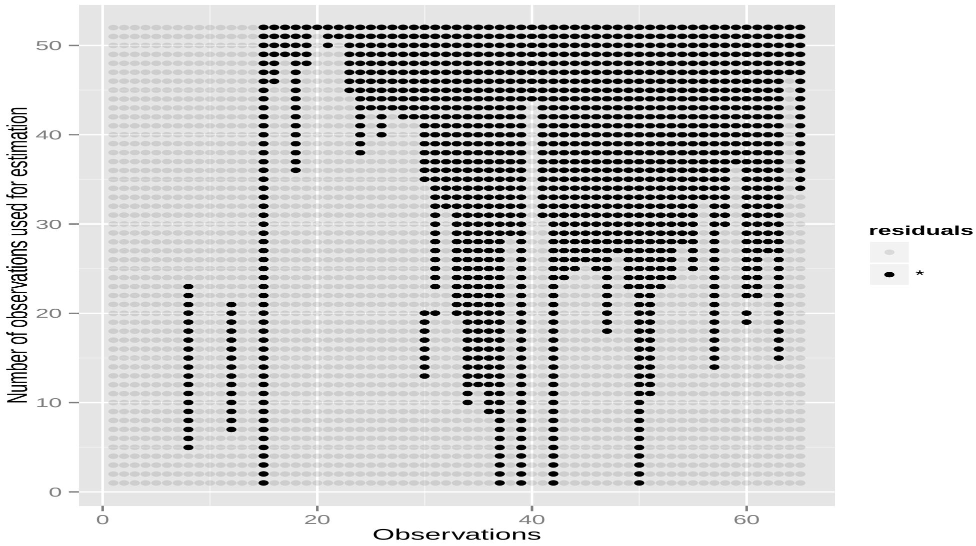


Fig. 10: Best fit for pilot regular users data

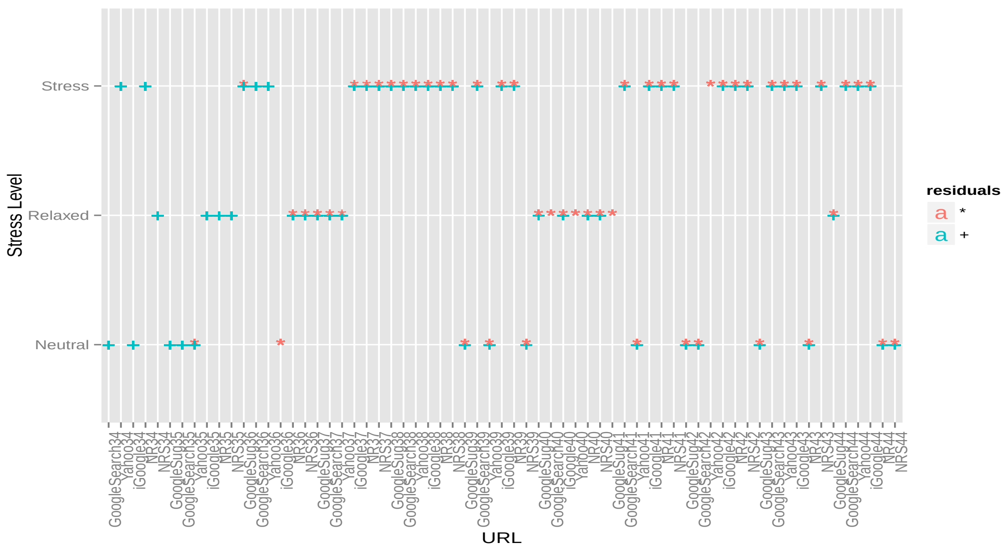


Fig. 11: Detected categorical class for regular users

For observation $m > 8$, instance 15, 37, 39, 42, and 50 are outlying. This corresponds to webpages where the contents are mostly deactivated and those with interactive contents like video and picture contents that can be an incentive to induce reaction in users. For the overall group of data, the workers and regular group seem to perform significantly high enough to indicate potentials of model reliability.

Comparing most of this findings from robust forward search algorithm to the PCA, the remote cases are usually found to be the outliers in the PCA analysis, indicating stability of all models. To also determine if it is conceivable to model HCI-HPR data instinctively, we simulated dataset from a multivariate normal distribution for optimization based on maximum likelihood estimations to compare with real data.

6. Maximum Likelihood Estimation

To determine the expected maximum likelihood for unknown set of explanatory variables, we simulated data based on multivariate normal distribute sample of the following order.

$$x \sim \mathcal{N}(\mu \sum)$$

Where $x = [X_1, X_2, X_3, \ldots, X_k]$ is k-dimensional random vector, $\mu = [\mathcal{E}(X_1), \mathcal{E}(X_2), \ldots, \mathcal{E}(X_k)$ is the k-dimentional mean vecter and a $k \times k$ covariance matrix $\sum = [Cov(X_i, X_j)],\ i = 1, 2, \ldots, k; j = 1, 2, \ldots, k$. The maximum likelihood (MLE) estimates the parameters of the model when applied to the simulated datasets. We are interested in determining the probability of obtaining categorical values from the dataset that will best define larger size of participants. The MLE values for the mean and standard deviation are close to the corresponding sample for the original data, they closely conforms with the values used to obtain the data. We achieved an R-squared equivalent to 1 which is as perfect as the values we used in generating the data. The results were also as well as that of the original data. The MLE for the slope (beta0) and (beta1) (Table 1 and Figure 12) are reasonable enough to validate this.

Table 1: Coefficient estimates for simulated data problem.

	Estimate (Rad/s)	Std. Error (Rad/s)	z value	Pr(z)
beta0	2.5540	0.0838	30.4423	$< 2.2e^{-16}$
beta1	4.9516	0.2931	16.8886	$< 2.2e^{-16}$
mu	0.5540	0.0838	6.6036	$4.014e^{-11}$
sigma	0.8782	0.0621	14.1421	$< 2.2e^{-16}$

The estimations can be applied to models of illogical density. If the residuals are expected to be normally distributed then the log-likelihood function can be used. On the other, if it conforms to different distribution then the appropriate density function is applied.

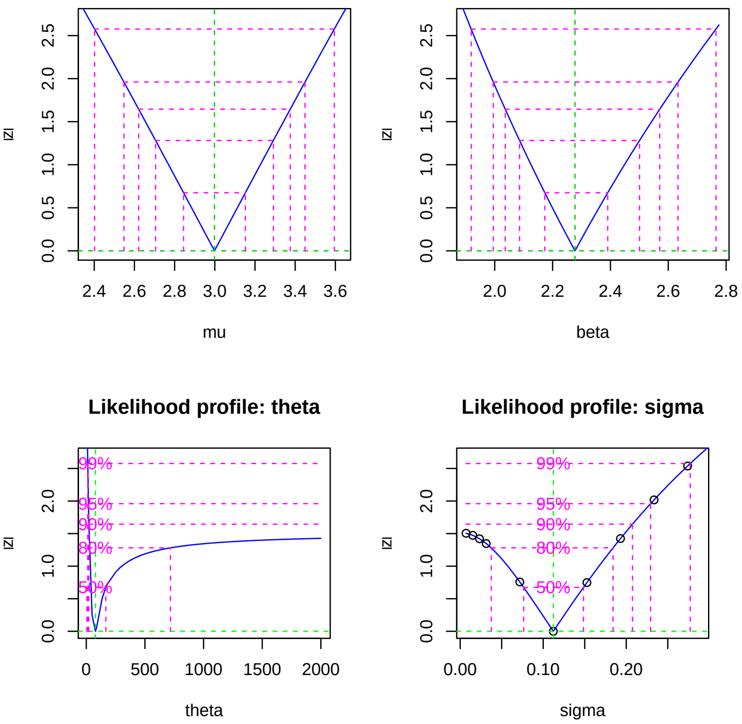


Fig. 12: Maximum likehood for parameters

The MLE assumes that the data for each class in each group are normally distributed and computes the likelihood that a given point belongs to a particular group. All points are classified unless a probability threshold is selected. Each point are assigned to a group that possess the maximum probability which is the maximum likelihood. A point remains unclassified if the highest probability is less than the specified threshold.

The density distribution function below (Figure 13) shows that the overall probability of relaxed participants supersedes that of the stressed participants. So in order to balance the ratio and set the threshold high the contents on webpages that cause stress reaction are replaced by the number of contents on pages that participants are found to be relaxed, or having less complex content.

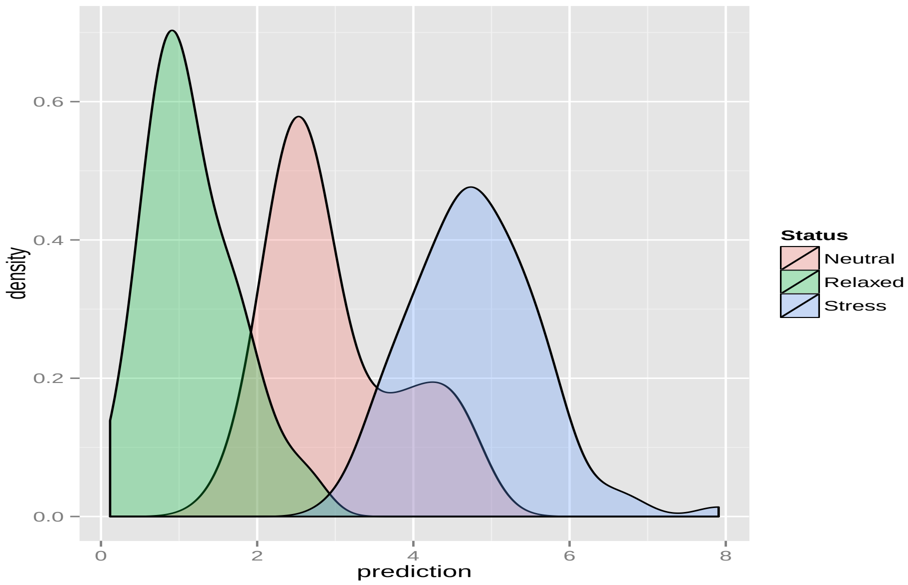


Fig. 13: Density distribution of class labels

7. Conclusion

The algorithm developed in this paper has the following modules, a robust forward search, and standardisations of the residuals outputs by simulation and visualisation of the result in a plot. If outliers are detected, they form a single distinct set. In the case of where there are more than one distinct sets of outliers that results to similar values of least mean squares, this will be defined in the plot. An important feature of the forward search method is that attention is not focused on a particular fit to the data, but rather to any conceivable effects which are ordered and compared.

The main advantage of the algorithm compared to other method of outlier detection is that it investigates values of possible instances of the first parameter to the last based on the Mahalanobis square distances, and returns subsets of residuals starting with the list value. The datasets used to run the algorithm was chosen to associate and compare the different dimension in HCI-HPR associations, rather than limiting ourselves to the young, theoretically familiar and particularly cultivated set of participants, we refined to make a more generality approach to the more older and worker groups model which turn out to be a more reasonable group for testing that implements enhanced instance for optimization in the forward search approach as compared to the R-squared result from a linear model fit. To define the general predictive ability of the HCI-HPR model instinctively, datasets where generated based on multivariate normal distribution with properties of the original data. The MLE was used to estimate the parameters of the simulated model using a negative log function of the data.

The likelihood is the probability of obtaining the original data given the selected probability distribution of the data. The results of the betas closely relates to the estimated coefficient of the original data. The MLE is applied to stress and estimate magnitude of response as well as human behaviour scaling. One improvement of the forward search algorithm compared to others is that no alteration of parameters are involved and can be applicable to larger size of dataset and reveals natural structures and outliers has fast as the limit of intervals adopted for the search. The limitations is that when the number of digits for residuals value is more than the expected size for stack flow, it normally yields an error caution; to counter this, the datasets are normally standardised by taking the logarithmic function of the model before running the algorithm. The improvement made to the algorithm is application of more graphical display and rendering of the residual plot.

Acknowledgments

The authors would like to thank Web Ergonomics Lab in University of Manchester where the study was conducted and data collected. And also the Petroleum Technology Develop Fund Nigeria (PTDF) for sponsoring this research.

References

Andreassi, John L.. (1977).Psychophysiology: Human behavior and physiological response. *Psychology Press,*.

Atkinson, AC. (1994).Fast very robust methods for the detection of multiple outliers. *Journal of the American Statistical Association.*.

Beck, Nathaniel and Jackman, Simon. (1998).Beyond linearity by default: Generalized additive models. *American Journal of Political Science, JSTOR,* **19**:596–627.

Cain, Brad. (2007). A review of the mental workload literature. *DTIC Document*.

Chhikara, R. S. and Folks, J. L. (1977). The inverse Gaussian distribution as a lifetime model. *Technometrics,* **19**:461–468.

Demiral, Şükrü Barış and Schlesewsky, Matthias and Bornkessel-Schlesewsky, Ina. (2008). The inverse Gaussian distribution as a lifetime model. *Cognition,* **106**:484–500.

Dirican, Ahmet Cengizhan and Göktürk, Mehmet. (2011).Psychophysiological measures of human cognitive states applied in human computer interaction. *Procedia Computer Science,* **3**:1361–1367.

Farmer, Erick and Brownson, Adam. (2003). Review of workload measurement, analysis and interpretation methods. *European Organisation for the Safety of Air Navigation,* **33**.

Filipovic, Sasa R and Andreassi, John L. (2001). Psychophysiology: Human Behavior and Physiological Response. *Journal of Psychophysiology,* **15**:210–212.

Ganglbauer, Eva and Schrammel, Johann and Deutsch, Stephanie and Tscheligi, Manfred. (2009).Applying psychophysiological methods for measuring user experience: possibilities, challenges and feasibility. *Workshop on user experience evaluation methods in product development,* Citeseer.

Good, IJ and Mittal, Y and others. (1987).The amalgamation and geometry of two-by-two contingency tables. *Institute of Mathematical Statistics, The Annals of Statistics,* **15**: 694–711.

Green, J.H. (1976).An introduction to human physiology.

Hastie, Trevor J and Tibshirani, Robert J. (1990). Generalized additive models. *CRC Press,* **43**

Insko, Brent E. (2003). Measuring presence: Subjective, behavioral and physiological methods. *Emerging Communication,* **5**:109–120.

Kalyuga, Slava. (2008). Managing cognitive load in adaptive multimedia learning. *IGI Global*.

Kramer, Arthur F. (1990). Physiological metrics of mental workload: A review of recent progress. *DTIC Documents,*.

Kretzschmar, Franziska and Pleimling, Dominique and Hosemann, Jana and Füssel, Stephan and Bornkessel-Schlesewsky, Ina and Schlesewsky, Matthias. (2013). Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. *PloS one,* **8**:e56178.

Liversedge, Simon P and Blythe, Hazel I. (2007). Lexical and sublexical influences on eye movements during reading. *Language and Linguistics Compass,* **1**:17–31.

Martins, M., and Santos, C., Costa, L., and Frizera, A. (2015). Feature Reduction with PCA/KPCA for gait classification with different assistive devices. *International Journal of Intelligent Computing and Cybernetics* 8(4).

Mwitondi, K. S and Said, R. A. (2011). A step-wise method for labelling continuous data with a focus on striking a balance between predictive accuracy and model reliability. *international Conference on the Challenges in Statistics and Operations Research (CSOR)*

Simpson, Edward H. (1951).The interpretation of interaction in contingency tables. *JSTOR,*:238–241.

Wold, Svante, and Kim Esbensen, and Paul Geladi. (1987). "Principal component analysis." *Chemometrics and intelligent laboratory systems* 2.1: 37-52.

About the authors

Fatima Isiaka obtained her B.Sc Mathematics from Federal University of Technology Bauchi, Nigeria in 2003, Masters in Computer Science at Federal University of Agriculture, Makurdi, Nigeria in 2009. She is currently studying for a PhD in Computer Science in Sheffield Hallam University, United Kingdom. Her research interest includes: Human Computer Interaction, Data mining, Programming and Affective Computing.

Kassim S. Mwidondi acquired his Msc (Statistics/Informatics) in Sofia, Bulgaria 1991, MSc (Financial Modelling) in Strathclyde Graduate Business School, Glasgow, United Kingdom, 1997. He obtained his PhD Statistics/Data Mining, School of Mathematics, University of Leeds, United Kingdom, 2003. He currently works and lectures in Sheffield Hallam University, United Kingdom. His research interest includes, Data mining (bioinformatics and business intelligence), Data migration and integration, Database Management systems and Applied Statistics (biostatistics and business intelligence).

Adamu M. Ibrahim obtained his B.Sc Statistics from University of Abuja Nigeria in 1995, MBA (Finance and Management) in Bayero University Kano, Nigeria, 2000. Masters in Computer Science at Federal University of Agriculture, Makurdi, Nigeria in 2009. He completed his PhD in Computer Science in University of Leeds, United Kingdom. His research interest is based on Artificial Intelligence, Machine Learning, Bashian and Geospatial modeling, GIS, Statistics and Mineral exploration.