Discovering Web Usage Pattern Using Artificial Neural Network Algorithm for Generalized Singular Value Decomposition-based Linea....

Article · December 2018

DOI: 10.7828/jmds.v7i2.1221

3 authors:

Rolysent Paredes
Misamis University

Ruji Medina
Technological Institute of the Philippines

Ariel M. Sison
Emilio Aguinaldo College

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Discovering Web Usage Pattern Using Artificial Neural Network Algorithm for Generalized Singular Value Decomposition-based Linear Discriminant Analysis

Rolysent K. Paredes[1], Ruji P. Medina[2], Ariel M. Sison[3]

[1,2]Graduate Programs, Technological Institute of the Philippines, Quezon City, Philippines
[3]School of Computer Studies, Emilio Aguinaldo College, Manila, Philippines
Corresponding author: Ariel M. Sison, email: ariel.sison@eac.edu.ph

Abstract

With the considerable amount of data collected from the Information Technology systems including computer networks, many institutions and organizations are considering the benefit in discovering user patterns. This paper presents the processes in developing the enhanced algorithm for Linear Discriminant Analysis via Generalized Singular Value Decomposition (LDA/GSVD) which primarily functions to discover web usage patterns. Preprocessing, class labeling using Self-Organizing Map (SOM), computation of the feature subspaces, and the training of the Artificial Neural Network (ANN) architecture are stages in developing the algorithm. After the development, this neural network-based algorithm was used to discover patterns from the university's proxy servers' weblogs. Results showed an appropriate classification of the network users' web usage with the improved algorithm. Hence, results revealed that more users in the campus whether students or employees are utilizing the internet for noneducational rather than educational. Also, simulation results showed that the enhanced algorithm outperformed the current LDA/GSVD algorithm up to 50% regarding computational cost. Through this approach, IT managers will be guided to make better plans to optimize the utilization of the internet. Moreover, this proposed technique will benefit not only the educational institutions but also other organizations which mostly need pattern discovery to their systems' enormous data.

Keywords: computer, development, internet, network, weblogs

Introduction

Pattern discovery becomes an essential management task due to the modern Information Technology (IT) systems which often provide a vast amount of data. One of the approaches to discover patterns is through data mining (Vaarandi & Pihelgas, 2015). Thus, various data mining methods are already developed such as clustering, classification, generalization, visualization, and association (Jothi & Husain, 2015).

Due to the importance of pattern discovery, providing a new mechanism for discovering patterns on the weblogs stored in the network's proxy servers is useful particularly in school campus. These weblogs are composed of necessary data of the network users who requested numerous website resources (Sukumar et al., 2016). Data mining techniques such as classification can be utilized to discover patterns. The classification method includes K-Nearest Neighbor (KNN) classifier, Artificial Neural Network (ANN), ID3, C4.5, Naive Bayes, Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA) (Langhnoja et al., 2013; Nikam, 2015; Zainee & Chellappan, 2016).

The LDA has been used to discover users' web usage patterns from the weblogs since it outperformed other classification models and algorithms and has been utilized widely in the previous years for dimensionality reduction, detection, and supervised learning (Zainee & Chellappan, 2016; Markopoulos, 2017). Further, the LDA has the benefit of looking for projection vector that produces ideal discrimination among different collections of observations (Yu et al., 2017). However, it has a problem in dealing with unlabeled datasets (Liu et al., 2014; Markopoulos, 2017). Also, it fails when a matrix is singular due to small sample size (SSS) problem (Shao et al., 2011; Deng et al., 2012; Wang et al., 2016). Typically weblogs are unlabeled, and due to the dynamic data that composes it, SSS problem can occur.

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

There are several ways to manage the unlabeled datasets on LDA such as PCA-LDA which used Principal Component Analysis (PCA) to separate the classes (Yoo & Pedrycz, 2013). The ULDA which uses target generation process (TGP) for the targets and classes (Lin et al., 2007), LDA-Km (Ding & Li, 2007), TRACK (Wang et al., 2014), Semi-LDC (Liu et al., 2014), and modified-2DLDA (Conka et al., 2015) utilizes k-means clustering to make class labels. Further, LDA-basis sequence calculates the projected class numbers and produces a series of bases that meet practical LDA solutions (Markopoulos, 2017). These approaches are making LDA unsupervised. Hence, introducing a new technique in class labeling by utilizing Self-Organizing Map (SOM) could address the issue. The SOM is essential in creating the clusters or the class labels because it is an unsupervised learning method to determine dataset's patterns and excellently work with high dimensional data (Faigl & Hollinger, 2018). It is fast and robust since it is an example of artificial neural network architecture as well (Sharif et al., 2015; McLean et al., 2017).

In dealing with SSS problem on LDA, there are many techniques proposed such as Regularized LDA (RLDA) which offers a computation concerning the relationship among the dilemmas of multi-class discriminant analysis and multivariate regression (Zhang et al., 2010). An exponential discriminant analysis (EDA) technique was suggested to solve the undersampled problem (Zhang et al., 2010) while Spectral Regression Discriminant Analysis (SRDA) casts discriminant analysis into a regression framework. Direct LDA (D-LDA) (Yu & Yang, 2001), LDA/QR (Ye & Li, 2005), and multiple between-class linear discriminant analysis (MBLDA) (Wang et al., 2016) are for classifying multidimensional data with fast learning capability. A two-stage method was utilizing bidirectional LDA and RLDA, and designed for two-dimensional data only (Zhao et al., 2015). A split and combined approaches for LDA (SC-LDA) was developed to replace the full eigenvector decomposition (Seng & Ang, 2017). But among these methods, the widely used approach is the application of Generalized Singular Value Decomposition (GSVD) (Howland & Park, 2004) which is generally applied by various discriminant analysis approaches (Park & Park, 2005; Jing et al., 2012). It is also a typical method for computing the matrix singular problems in different mathematical solutions (Jing et al., 2012; Chen & Chan, 2017). Further, GSVD on LDA (LDA/GSVD) provides extraordinary recognition accuracy (Wu & Ahmad, 2009) that is why many researchers used and developed variance of it. However, GSVD suffers from computational cost (Zhao & Yuen, 2008; Wu & Ahmad, 2009; Bahrami, 2017) which can cause a longer time in classifying datasets when applied to LDA. Moreover, if there is a new instance for classification in the existing LDA/GSVD, the whole process of the algorithm will be repeated from the very start.

Developing the ANN algorithm to enhance the existing LDA/GSVD algorithm could overcome the issue of computational cost. With the use of the ANN algorithm, it eliminates the classical mathematical computations and numerous iterations that are involved in the current LDA/GSVD algorithm which compromises time complexity making it less efficient (Bahrami, 2017; Wu & Ahmad, 2009; Zhao & Yuen, 2008). Further, with this method, learning can be done from the weblog and classification to each instance. Whether new or previously part of the training or testing, learning will be faster because it will not go back to the start of the whole procedure. The use of ANN in developing the algorithm has the benefit of accuracy. Also, ANN has the uniqueness of concurrent processing. It can learn and recall data relationships, and map non-linear instances (Dash & Dubey, 2012; Ranhotra et al., 2017).

The widely employed method for ANN is the Back Propagation Neural Network (BPNN). It is composed of input layer, hidden layers, and output layer (Dalipi & Yayilgan, 2015). An example of BPNN is Bayesian Regularization Back Propagation (BRBP). The BRBP offers robust approximation for difficult and noisy inputs. It works excellently by removing network weights which have no impact on the problem solving and presents improvements in evading the issues of local minima (Jazayeri et al., 2016). Furthermore, it delivers weights into a

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

training function while advancing the simplification performance of the old BPNN automatically (Dalipi & Yayilgan, 2015).

This study is an attempt to provide a new mechanism for discovering patterns on the weblogs stored in the school network's proxy servers that can address the issues of LDA in dealing with unlabeled datasets and SSS problem. This study introduces a new technique in class labeling by utilizing Self-Organizing Map (SOM). It also highlights the development of the ANN algorithm to enhance the existing LDA/GSVD algorithm to overcome its issues most especially its computational cost. The main objective of this study is to discover web usage pattern using the ANN algorithm for generalized singular value decomposition-based linear discriminant analysis. Specifically, this study identified active users for the First Semester of the school year 2017-2018, developed the enhanced LDA/GSVD algorithm or ANN for LDA/GSVD, and compared the output of the existing and enhanced algorithms in terms of pattern discovery and performance.

Hence, the study could provide a technique to the network administrator or those who particularly manage the network in school in identifying users who are inclined in accessing educational contents and those who are not. This pattern discovery approach can also support the educational institutions in policy-making regarding the utilization of the internet. Moreover, organizations which handle enormous data can also benefit from this technique.

Materials and Methods

The study simulated the existing and the enhanced LDA/GSVD algorithms in discovering the groups of users who are inclined to accessing educational sites and those who are not. The time complexity of both algorithms was recorded and compared. But before simulating the enhanced algorithm or the ANN for LDA/GSVD, the algorithm was developed first. Figure 1 presents the process of developing the algorithm which has four essential stages: preprocessing, class labeling, computation of feature subspaces, and ANN training.

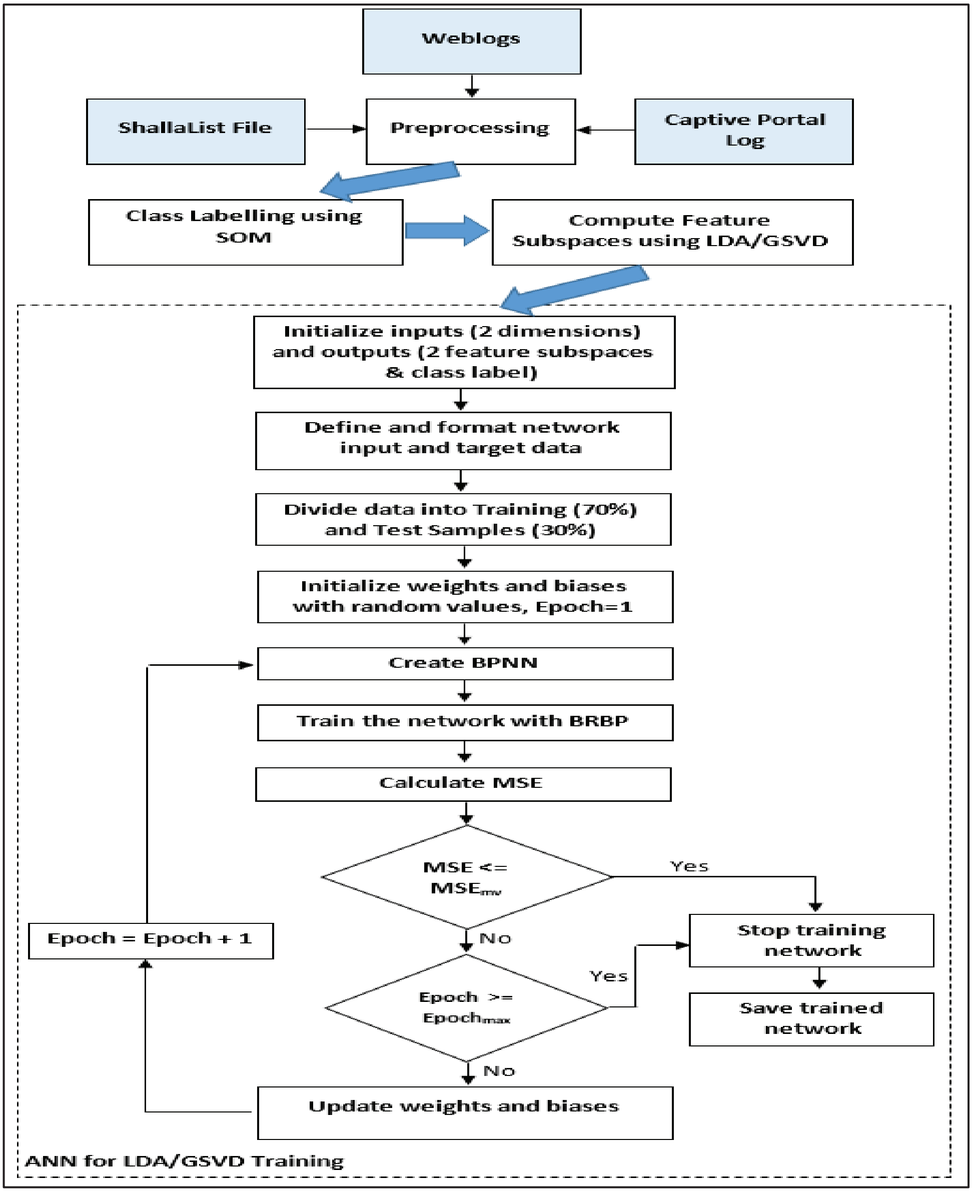

Figure 1. The training process for the development of the ANN for LDA/GSVD algorithm.

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Steps in developing the ANN for LDA/GSVD algorithm

1. Actual dataset identification through preprocessing

The primary function of this step is to remove the irrelevant instances of the weblogs. This step also includes the identification of the website category and the user who requests for it. There are three objects involved in preprocessing. These are the weblog, captive portal log, and shallalist file. The weblog and captive portal log were from the proxy servers in Misamis University, Ozamiz City, Philippines. It is composed of the records in the First Semester of the School Year 2017-2018.

Weblogs contain the websites accessed by the network users (students and employees) while the captive portal log includes the date and time when the users log in or use the internet (Aryeh et al., 2016). This captive portal log was used to identify the specific user who requests the particular website. The shallalist file contains the category of websites (Squicciarini et al., 2014; Shalla Secure Services KG, 2018) which utilized in determining where the requested site belongs. The actual dataset is generated through preprocessing. Table 1 presents the algorithm for the preprocessing.

2. Class labeling

After the preprocessing, class labeling is necessary for the LDA/GSVD to work because the generalized data are unlabeled. The SOM was employed to create these clusters which became the class labels.

Table 1. Preprocessing algorithm.

Algorithm: Preprocessing Algorithm

```
1.   Read log entry from weblog file
2.    if(weblog.mime="text/html" AND weblog.status=200 AND weblog.method="GET")
      {
3.        site=GetDomain(weblog.URL);
4.        user=GetUserAccount();
5.         if(IsNewLogin(user)=true) {
6.            if(GetSiteCategory(site)="Education") {
7.                num_educ=GetNumEduc(user)+1;
8.            } else {
9.                num_not_educ= GetNumNotEduc(user)+1;
10.           }
11.        UpdateSitesAccessed(user);
12.       }
13.   }
14.  Repeat 1 and 2 until EOF encountered
15.  Generalize all the users with respect to their number of
     educational and noneducational sites.
End the process
```

3. Computation of feature subspaces

In this stage, the dataset has been labeled and ready for actual calculation of the feature subspaces. For the ANN architecture to learn, predict, and classify the users based on their web usage patterns, the feature subspaces must derive from that current LDA/GSVD algorithm. These feature subspaces are actual points where the users belong and used for further visualization of the classification of the users. Thus, it is the existing algorithm used to get the feature subspaces so that the accuracy of the developed ANN algorithm is the same. Table 2 presents the existing LDA/GSVD algorithm.

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Table 2. Existing LDA/GSVD Algorithm.

Algorithm: Existing LDA/GSVD Algorithm

For the matrix $A \in R^{m\times n}$ with k groups, it calculates the matrix's columns $G \in R^{m\times(k-1)}$, which maintains the configured cluster dimensionally narrowed space, and determines $(k - 1)$-dimensional depiction Y of A.

Step 1: Calculate $H_w \in R^{m\times n}$ and $H_b \in R^{m\times k}$ from A
Step 2: Solve the $K = (H_b, H_w)^T \in \mathrm{R}^{(k+n)\times m}$ for its orthogonal decomposition.

$$P^T K Q = \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}$$

Step 3: Let $t = \mathrm{rank}(K)$.
Step 4: Calculate W from the SVD of $P(1:k, 1:t)$, which is $U^T P(1:k, 1:t)\, W = \Sigma_A$.
Step 5: Solve the first $k - 1$ columns of

$$X = Q \begin{pmatrix} R^{-1}W & 0 \\ 0 & I \end{pmatrix}$$

and allocate those to G.
Step 6: $Y = G^T A$.

4. ANN training

The ANN architecture has two input variables which are the dimensions or the frequency of the educational and noneducational websites per user, while the corresponding three output variables are the expected feature subspaces and class label. These dimensions, feature subspaces, and class labels are for training and testing purposes. For the sampling, 70% of the instances of the dataset allocated were for training, and 30% for the testing. The number of hidden neurons is seven which utilized tansigmoid transmission function as the activation function. Moreover, in training of the network, Bayesian Regularization Back Propagation (BRBP) was employed.

After saving the trained network, it becomes a module or subroutine that is used to solve the expected new feature subspaces of the inputs. Thus, the algorithm shown in Table 3 is for discovering patterns which separate the users according to the group they belong. Further, the architecture of the trained network was employed to separate the groups of users according to their web usage patterns.

Table 3. Pattern Discovery Algorithm.

Algorithm: Pattern Discovery Algorithm
1. Open Dataset 2. Get each user's total number of educational and noneducational websites. 3. Compute the corresponding feature subspaces and determine the class label using the ANN for LDA/GSVD. 4. Classify and visualize the user's web usage pattern. 5. Repeat steps one and two until EOF encountered 6. End Process

As shown in Figure 2, the inputs to the trained ANN for GSVD are the dimensions or the total number of educational and noneducational websites accessed by each user. For the outputs, the trained network computes the feature subspaces and identifies the class label. In viewing the actual separation or classification of the users, visualization is presented as well to find out the density of the users who managed to browse educational websites and those who did not.

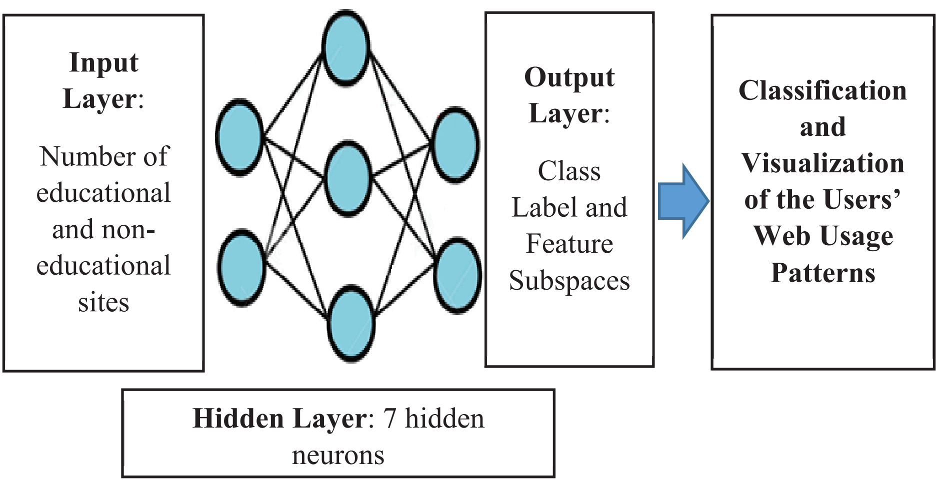

Figure 2. The architecture of the trained ANN for LDA/GSVD used to discover Users' Web Usage Pattern.

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Comparing the existing and enhanced algorithms

The output of LDA/GSVD algorithm and ANN for LDA/GSVD algorithm was compared in terms of pattern discovery and performance. Comparison of the two algorithms was carried out with regard to the accuracy in identifying the user's web usage patterns, the time to finish the entire process, and the computational costs.

Results and Discussion

A. Enhanced LDA/GSVD algorithm or ANN for LDA/GSVD

Actual Dataset

After preprocessing, a total of 104 active users was identified in specific network nodes for the First Semester of the School Year 2017-2018. Therefore, this study used a dataset that has 104 instances and two dimensions. Figure 3 shows the graph that corresponds to the total number of educational websites and noneducational websites that were accessed by the users. Blue bars are for the educational sites while the orange bars are for noneducational.

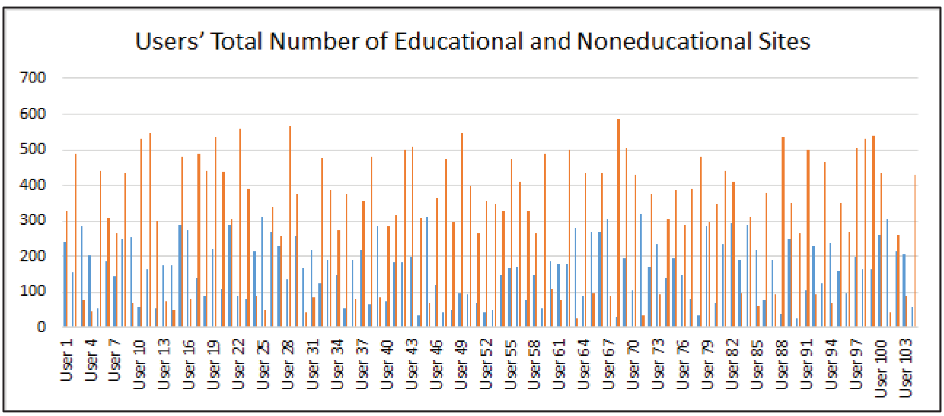

Figure 3. The total number of educational and noneducational websites accessed by the users.

Class Labeling

Table 4 presents the number of users belonging to each class label. The number of users in Class label 2 dominates.

Table 4. The frequency of each class after class labeling.

Class label	Number of users
1	27
2	77

Computation and visualization of the feature subspaces using the existing algorithm

It took four seconds for the existing LDA/GSVD to finish calculating the feature subspaces after all the instances of the dataset were labeled. Figure 4 shows the graph for the feature subspaces separating the classes in which each point represents a user. Thus, Class label 1 composed of fewer users compared to Class label 2. It is noticeable that all data points appropriately separated for better visualization.

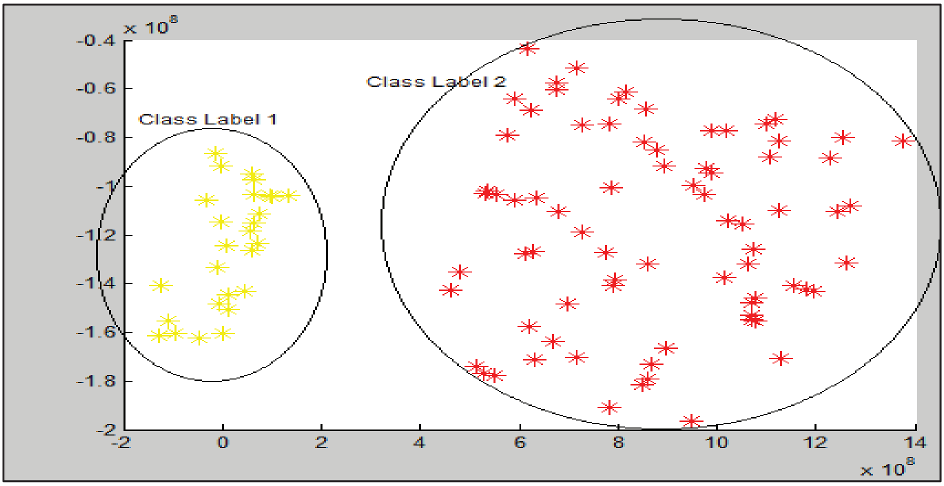

Figure 4. Graph of the subspaces after applying the existing LDA/GSVD.

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Training for the development of ANN for LDA/GSVD Algorithm

After getting all the necessary values such as class labels and feature subspaces, training for the development of the enhanced algorithm is the next step. This study used the performance functions which includes the Mean Squared Error (MSE) and Regression (R) to evaluate the performance of the ANN for LDA/GSVD algorithm. The MSE is the average squared difference between experimental output values and the fed targets in training.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(t_i - a_i)^2$$

Where n is the sample set's size, a_i is the ANN experimental or observed output and t_i is the matching targets. A small MSE value means that the developed algorithm can obtain less error. Regression (R) computes the outputs and targets' correlation. When R is 1, it signifies a good or close relationship, otherwise a random or no association at all (Dalipi & Yayilgan, 2015).

Figure 5 depicts the performance of training and test samples using BRBP algorithm. The graph shows that the test and training samples almost overlap with each other. The training and test curves stabilized at epoch 328,612 which obtained an MSE error value of approximately 0.6945, and that implies that the trained network attained insignificant error.

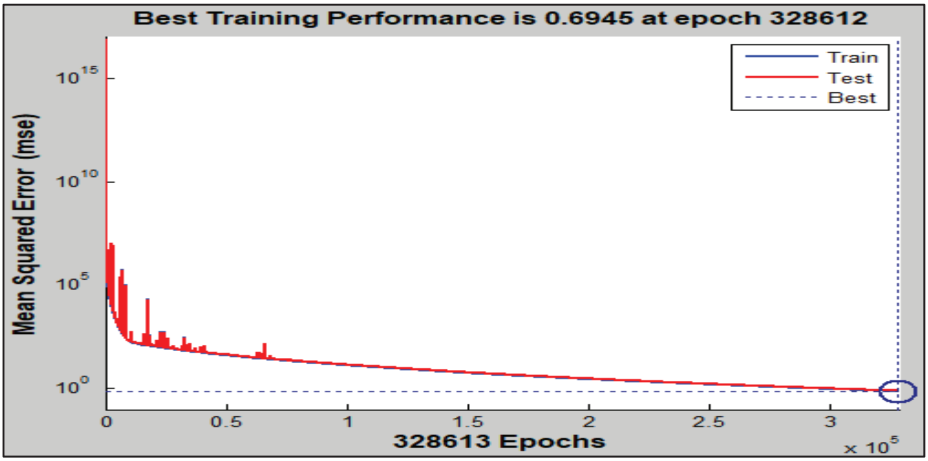

Figure 5. BRBP's prediction result.

The histogram in Figure 6 presents the frequency of the instances per error. The measurement of the error is by subtracting the targets and the resultant outputs. The most significant error in the training was at around 3.789 but still very small.

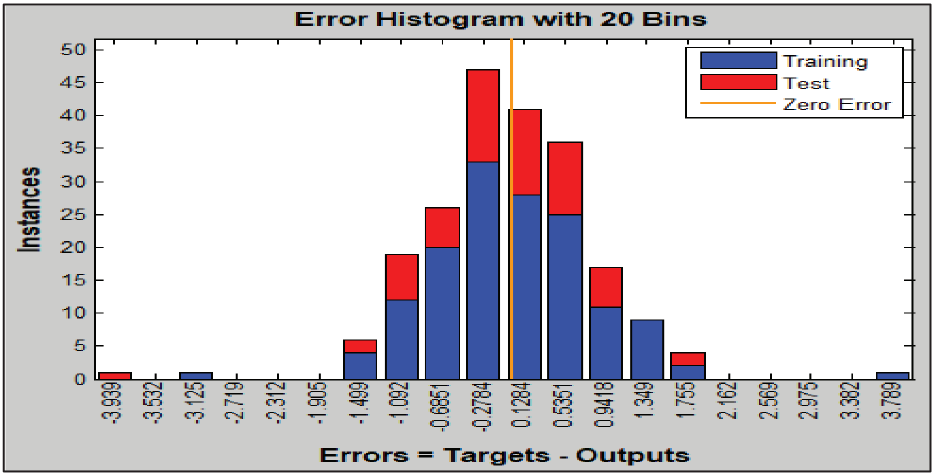

Figure 6. BRBP's Histogram of error sequences.

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Table 5 shows the performance of the ANN for LDA/GSVD using BRBP. Figure 7 shows BRBP algorithm's correlation which indicates that there is a relationship between the outputs and targets. Thus, the algorithm is accurate and better because the MSEs are in smaller values, and the *R* for the training, test, and overall analysis is 1.

Table 5. Performance of the ANN for LDA/GSVD using BRBP.

Dataset Sample	Mean Squared Error (MSE)	Regression (*R*)
Training	0.6945	1
Testing	0.7336	1

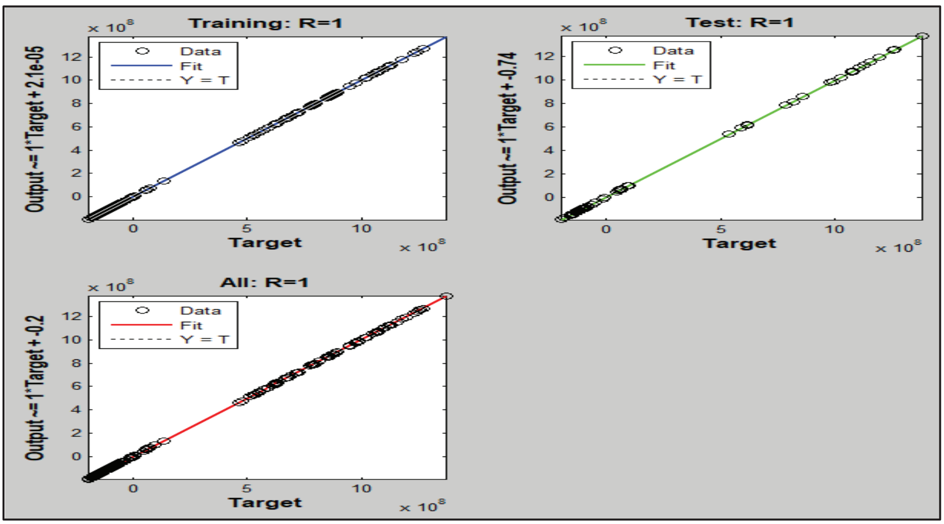

Figure 7. BRBP's Regression Analysis.

B. Comparison of the existing and enhanced algorithms

Computation and visualization of the feature subspaces using the enhanced algorithm

Figure 8 presents the classification of the data using the enhanced algorithm which is very much similar to Figure 4 that utilized the existing LDA/GSVD. It is a manifestation that the improved LDA/GSVD maintains the accuracy of current LDA/GSVD algorithm. Thus, if a new user joins the network, this developed algorithm can be used directly in identifying the user's web usage patterns whereas the existing LDA/GSVD executes the whole algorithm and entire process depends on all the instances again. With that, the current algorithm can be exhaustive especially when there is a new instance added to the dataset. The ANN for LDA/GSVD algorithm finished the whole process in two seconds.

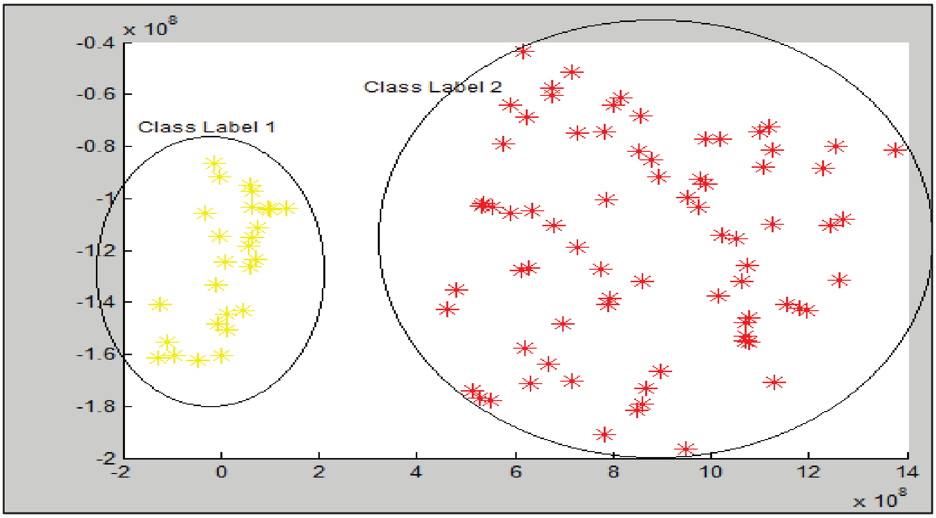

Figure 8. Graph of the subspaces after applying the ANN for LDA/GSVD Algorithm.

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Computational costs of existing and enhanced algorithms

Table 6 presents the computational costs of two algorithms. It is evident that the enhanced LDA/GSVD improved the computational cost by 50%. The values for the computational costs may be too small because there are only 104 instances that composed the dataset.

Table 6. Computational costs of the existing and enhanced algorithms.

Algorithm	Computational Cost or Total Execution Time
LDA/GSVD	4 seconds
ANN for LDA/GSVD	2 seconds
Improvement of the Enhanced LDA/GSVD	**50%**

Conclusion and Recommendations

Simulation results showed that discovering users' web usage patterns can be done effectively using the ANN algorithm for LDA/GSVD. Also, the enhanced LDA/GSVD algorithm outperformed the existing LDA/GSVD algorithm on computational cost during classification or discovering patterns. Therefore,the new approach is an efficient way of doing LDA/GSVD. It is also evident in the simulation that the new technique using BRBP can obtain the best performance of accuracy. With that, this new approach in discovering network users' web usage patterns is highly recommended to determine the utilization of the internet in the campus. This method is applicable as well to the datasets that are composed of many instances and dimensions due to lower computational cost. Hence, it is also recommended to test the method together with the weblogs of the Second Semester of the School Year 2017-2018 to further improve the performance of the developed algorithm since it is best to have many instances during the training. Moreover, implementation of the improved LDA/GSVD algorithm to bandwidth allocation and big data will be the next research to be done.

Acknowledgment

The authors are thankful for the full support given by the Technological Institute of the Philippines, Quezon City, Philippines. Misamis University in Ozamiz City is also instrumental for the completion of this study.

Literature Cited

Aryeh, F. L., Asante, M., & Danso, A. E. Y. (2016). Securing wireless network using pfsense captive portal with radius authentication– A case study at UMaT. *Ghana Journal of Technology*, *1*(1), 40-45.

Bahrami, S. (2017). Three-dimensional inverse scattering approach using analytical singular value decomposition method. *2017 18th International Radar Symposium (IRS)*. doi:10.23919/irs.2017.8008247

Chen, Z., & Chan, T. H. (2017). A truncated generalized singular value decomposition algorithm for moving force identification with ill-posed problems. *Journal of Sound and Vibration,401*, 297-310. doi:10.1016/j.jsv.2017.05.004

Conka, D., Viszlay, P., & Juhar, J. (2015). Improved class definition in two dimensional linear discriminant analysis of speech. *2015 25th International Conference Radioelektronika (RADIOELEKTRONIKA)*. doi:10.1109/radioelek.2015.7129025

Dalipi, F., & Yayilgan, S. Y. (2015). The impact of environmental factors to skiing injuries: Bayesian regularization neural network model for predicting skiing injuries. *2015 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. doi:10.1109/icccnt.2015.7395218

Journal of Multidisciplinary Studies
Vol. 7, Issue No. 2, pp. 222-245, December 2018
ISSN 2350-7020 (Print)
ISSN 2362-9436 (Online)
doi: http://dx.doi.org/10.7828/jmds.v7i2.1221

Dash, Y., & Dubey, S. K. (2012). Quality prediction in object oriented system by using ANN: A brief survey. *International Journal of Advanced Research in Computer Science and Software Engineering*, *2*(2).

Deng, W., Hu, J., & Guo, J. (2012). Extended SRC: Undersampled face recognition via intraclass variant dictionary. *IEEE Transactions on Pattern Analysis and Machine Intelligence,34*(9), 1864-1870. doi:10.1109/tpami.2012.30

Ding, C., & Li, T. (2007). Adaptive dimension reduction using discriminant analysis andK-means clustering. *Proceedings of the 24th International Conference on Machine Learning - ICML 07*. doi:10.1145/1273496.1273562

Faigl, J., & Vana, P. (2016). Self-organizing map for data collection planning in persistent monitoring with spatial correlations. *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. doi:10.1109/smc.2016.7844737

Howland, P., & Park, H. (2004). Generalizing discriminant analysis using the generalized singular value decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 26*(8), 995-1006. doi:10.1109/tpami.2004.46

Jazayeri, K., Jazayeri, M., & Uysal, S. (2016). Comparative analysis of Levenberg-Marquardt and bayesian regularization backpropagation algorithms in photovoltaic power estimation using artificial neural network. *Advances in Data Mining. Applications and Theoretical Aspects Lecture Notes in Computer Science,* 80-95. doi:10.1007/978-3-319-41561-1_7

Jing, X., Dong, Y., & Yao, Y. (2012). Uncorrelated optimal discriminant vectors based on generalized singular value decomposition. *International Conference on Automatic Control and Artificial Intelligence (ACAI 2012)*. doi:10.1049/cp.2012.1444

Jothi, N., Rashid, N. A., & Husain, W. (2015). Data mining in healthcare – A review. *Procedia Computer Science,72*, 306-313. doi:10.1016/j.procs.2015.12.145

Langhnoja, S. G., Barot, M. P., & Mehta, D. B. (2013). Web usage mining using association rule mining on clustered data for pattern discovery. *International Journal of Data Mining Techniques and Applications*, *2*(01).

Lin, G., Wang, C., & Wang, W. (2007). An unsupervised linear discriminant analysis approach to multispectral MRI images classification. *2007 International Conference on Machine Learning and Cybernetics*. doi:10.1109/icmlc.2007.4370478

Liu, C., Hsaio, W., Lee, C., & Gou, F. (2014). Semi-supervised linear discriminant clustering. *IEEE Transactions on Cybernetics, 44*(7), 989-1000. doi:10.1109/tcyb.2013.2278466

Markopoulos, P. P. (2017). Linear discriminant analysis with few training data. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi:10.1109/icassp.2017.7953033

Mclean, R., Walker, A. J., & Bright, G. (2017). An artificial neural network driven decision-making system for manufacturing disturbance mitigation in reconfigurable systems. *2017 13th IEEE International Conference on Control & Automation (ICCA)*. doi:10.1109/icca.2017.8003144

Nikam, S. S. (2015). A comparative study of classification techniques in data mining algorithms. *Oriental Journal of Computer Science and Technology*, *8*(1), 13-19.

Oh, S., Yoo, S., & Pedrycz, W. (2013). Design of face recognition algorithm using PCA -LDA combined for hybrid data pre-processing and polynomial-based RBF neural networks : Design and its application. *Expert Systems with Applications,40*(5), 1451-1466. doi:10.1016/j.eswa.2012.08.046

Park, C. H., & Park, H. (2005). A relationship between linear discriminant analysis and the generalized minimum squared error solution. *SIAM Journal on Matrix Analysis and Applications,27*(2), 474-492. doi:10.1137/040607599

Ranhotra, S. S., Kumar, A., Magarini, M., & Mishra, A. (2017). Performance comparison of blind and non-blind channel equalizers using artificial neural networks. *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*. doi:10.1109/icufn.2017.7993785

Seng, J. K., & Ang, K. L. (2017). Big feature data analytics: Split and combine linear discriminant analysis (SC-LDA) for integration towards decision making analytics. *IEEE Access,5*, 14056-14065. doi:10.1109/access.2017.2726543

Shalla Secure Services KG. (2018). Shalla list website blacklist database. Retrieved from http://www.shallalist.de/

Shao, J., Wang, Y., Deng, X., & Wang, S. (2011). Sparse linear discriminant analysis by thresholding for high dimensional data. *The Annals of Statistics,39*(2), 1241-1265. doi:10.1214/10-aos870

Sharif, S. M., Kusin, F. M., Asha'Ari, Z. H., & Aris, A. Z. (2015). Characterization of water quality conditions in the Klang River Basin, Malaysia using self organizing map and k-means algorithm. *Procedia Environmental Sciences,30*, 73-78. doi:10.1016/j.proenv.2015.10.013

Squicciarini, A. C., Petracca, G., Horne, W. G., & Nath, A. (2014). Situational awareness through reasoning on network incidents. *Proceedings of the 4th ACM Conference on Data and Application Security and Privacy - CODASPY 14*. doi:10.1145/2557547.2557562

Sukumar, P., Robert, L., & Yuvaraj, S. (2016). Review on modern data preprocessing techniques in web usage mining (WUM). *2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*. doi:10.1109/csitss.2016.7779441

Vaarandi, R., & Pihelgas, M. (2015). LogCluster - A data clustering and pattern mining algorithm for event logs. *2015 11th International Conference on Network and Service Management (CNSM)*. doi:10.1109/cnsm.2015.7367331

Wang, D., Nie, F., & Huang, H. (2014). Unsupervised feature selection via unified trace ratio formulation and k-means clustering (TRACK). *Machine Learning and Knowledge Discovery in Databases Lecture Notes in Computer Science,*306-321. doi:10.1007/978-3-662-44845-8_20

Wang, Z., Shao, Y., Bai, L., Li, C., Liu, L., & Deng, N. (2016). MBLDA: A novel multiple between-class linear discriminant analysis. *Information Sciences,369*, 199-220. doi:10.1016/j.ins. 2016.05.031

Wu, W., & Ahmad, M. O. (2009). Orthogonalized linear discriminant analysis based on modified generalized singular value decomposition. *2009 IEEE International Symposium on Circuits and Systems*. doi:10.1109/iscas.2009.5118084

Ye, J., & Li, Q. (2005). A two-stage linear discriminant analysis via QR-decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence,27*(6), 929-941. doi:10.1109/tpami.2005.110

Yu, H., & Yang, J. (2001). A direct LDA algorithm for high-dimensional data — With application to face recognition. *Pattern Recognition,34*(10), 2067-2070. doi:10.1016/s0031-3203(00)00162-x

Yu, S., Cao, Z., & Jiang, X. (2017). Robust linear discriminant analysis with a Laplacian assumption on projection distribution. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi:10.1109/icassp.2017.7952620

Zainee, N. B., & Chellappan, K. (2016). A preliminary dengue fever prediction model based on vital signs and blood profile. *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*. doi:10.1109/iecbes.2016.7843530

Zhang, T., Fang, B., Tang, Y. Y., Shang, Z., & Xu, B. (2010). Generalized discriminant analysis: A matrix exponential approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),40*(1), 186-197. doi:10.1109/tsmcb.2009.2024759

Zhang, Z., Dai, G., Xu, C., & Jordan, M. I. (2010). Regularized discriminant analysis, ridge regression and beyond. *Journal of Machine Learning Research*, *11*(Aug), 2199-2228.

Zhao, H., & Yuen, P. C. (2008). Incremental linear discriminant analysis for face recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),38*(1), 210-221. doi:10.1109/tsmcb.2007.908870

Zhao, J., Shi, L., & Zhu, J. (2015). Two-stage regularized linear discriminant analysis for 2-D sata. *IEEE Transactions on Neural Networks and Learning Systems,26*(8), 1669-1681. doi:10.1109/tnnls.2014.2350993