xIlkogretim Online - Elementary Education Online,2021; Vol 20 (Issue 1): pp.2413-2419
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.01.270

Self-Efficacy And Extensive Reading During covid-19 Using Online Relationship Of Teaching Methodologies

Sivakumar S, Department of Computer Science and Engineering,KoneruLakshmaiah Education Foundation (KLEF), Guntur Dt, Andhra Pradesh, India,sivaiit79@gmail.com
Rajesh Kumar E, Department of Computer Science and Engineering, KoneruLakshmaiah Education Foundation (KLEF), Guntur Dt, Andhra Pradesh, India,rajthalopo@gmail.com
Vidyanandini S, Department of Mathematics, SRM Institute of Science and Technology, Kattankulathur, Chennai
Haritha D, Department of Computer Science and Engineering, KoneruLakshmaiah Education Foundation (KLEF), Guntur Dt, Andhra Pradesh, India
CMAK Zeelan Basha, Department of Computer Science and Engineering, KoneruLakshmaiah Education Foundation (KLEF), Guntur Dt, Andhra Pradesh, India

Abstract-Present-day in reading is believed assignificantability in the EFL education procedure and general understanding to the students reading self-efficacy. It watches members'recognitions under the estimation of wide studying and understanding methods, and furthermore, their observation impactis studied on reading self-efficacy. All things considered relatively are executed by few teachers on a regular basis. The way toward presenting Extensivereading (ER) is observed as moreexpensive, difficult, and time-consuming. The approach would understand the various elements impacting its effective execution to empower its use. This paper has been divided into two factors; initial one is associated with attitude and the second is attitude. The initial one reviewon ERimpact on how the student understandknowledge work. Second one isthatit analyses the student's observation of suitable procedure.The study utilizesquantifiable studyjust as subjective information from students belonging to first year of a systematic reading course and 603 undergraduate students from KL University of Guntur was selected as membersforexisting study. Disclosurespropose that comprehension readingwork was fundamentally the similar including or excludingan ER Program. In any case, the program appeared to positively affectcontributing students. In this paper, we are utilizing the classification techniques likes, decision tree and Mixed Mode Database Miner (MMDBM) and ER Group improved in the post-test compared to pre-test. Additionally, the student'sobservation of ER is very optimistic and MMDBM algorithm has produced higher rate of accuracy in detecting pre-test and post-test.

Keywords: Extensive reading (ER), Mixed Mode Database Miner (MMDBM), Teaching

I Introduction

The motivation of this paper was to see if students' reading self-efficacy can be improved by utilizing Internet-based extensive reading materials. A pre-test and post-test controlclustermodel was utilized. Having shown the language in English for quite a while as a Foreign Language (EFL) setting of Iran, analysts saw learner's issues in English are expected not to present directly in English language; a problem that may prompt major issues in learning the language

This scenario has resulted about reading reliable English texts (for example book, newspapers,English texts in online and so on.) as an open source to get to the real English language. Thus, learners need to build the language contribution for extensive reading in English.

Just as cycle riding is greatlyknowledgeable by riding a bicycle can beanalysed by reading. It is measured as best significant belief basic Extensive Reading's validity [9]. Designed for equally primary and second language learners, ER has been mostly examined and its welfares are well-known [1]. Still, various problems and interruption are reduced byusing it positively and constantly[15]. Among several disadvantage, a lecturer has to overcome, specifically in undeveloped or poor countries like: (a) scarcity of understanding materials; (b) shortage of sufficient training of lecturers; (c) burdens to cover the whole curriculum and text books and no time to perform program such as ER; (d) inflexibility and assessment burdens of necessary assessment deeds; and finally (e) the detail that part of entire guidance, as the means to achieve knowledge and lifetime skills are still held belief. However, reading has been the utmost harassed skill in traditional EFL teaching. Currently, there is vast pressure on EFL reading lecturers to move after teaching texts to teaching [2].

To infer that ER reading without anyone else can support students (who as of now have a level of ability in specific to English) learn and read, there is just enough proof [3]. Thus, ER turns into a treasured, practically predictable tool to attain such an objective. Actually, study on ER system and its execution despite everything misses the mark in conveying adequate data about either its educational angles or its handiness [4].

Revising the literature elucidates the part of inspiration in reading both English and EFL as a Second Language (ESL) settings and the effect of self-efficacy have a solid relationship with one another [5]. It is demonstrated that understanding inspiration and self-efficacy reading have a sturdy association among each other [6]. The casual comments of the existing study in their classes as EFL instructors on dissimilar degrees and their reading student's activities of inspiration are in interpretation exercises which prompted to consider the effect of self-efficacy on the student's inspiration in comprehensive reading. [7]

RwzeRaissi estimated self-efficacy as a scheme merging behavioural variations that students preserve in their psyche about their ability in acting the errands well and carrying on reasonably. The self-efficacy level has reflected as anissue that has impacts on specific behaviours. [8]

II Methodology

In the existing study, researchers utilized a measurable way to deal with above questions and answer. it was a semi test concentrate in which treatment and control groups were not changed. [10] Data were achieved from two significant scores as bases: a perusing comprehensive test and a perusing self-efficiency survey. Boththe tests have offered students who were reading general English courses in K L University over the span of the study in Vijayawada, Guntur, in single semester.[11] Applicants are considered for study from 60 different college understudies partitioned into two significant groups, to be specific, the experimental groups and control group. 30 understudies in the testing group got the solution of express guidance of understanding procedures and broad understanding guidance while the next 30 understudies from the control group got simply broad perusing guidance without unequivocal guidance of understanding systems. [12] The entire data assortment technique occurred inside one academic semester for 4 months.

The gathered data has been evaluated descriptively by using Weka 3.6 tools for our proposedexperimentation to carry out. In this study, researchers used classification methods. Classification is the most important techniques in data mining concept likes decision tree, Bayesian classification, Regression models, Best first (BF) Tree, Functional Trees (FT) tree, RandomForest (RF) tree, J48,RandomTree (RT), Alternative Decision (AD) tree. Additional decision tree contains several algorithms like SLIQ, SPRINT and MMDBM have been extensively used for researchin recent days. Decision tree algorithm is compared to MMDBM algorithm and calculated precision, recall, F-measure and accuracy.

A Decision Tree

Decision Tree (DT) constructs regression or classification models in the procedure of a tree structure. They can be utilized to deal with both regression and classification problems. The DT utilizes the tree demonstration to explain the problem in which each leaf node coincides to a class label and attributes are denoted on the internal node of the tree Refer the Figure1. We can show any Boolean function on discrete attributes utilizing the decision tree [13]

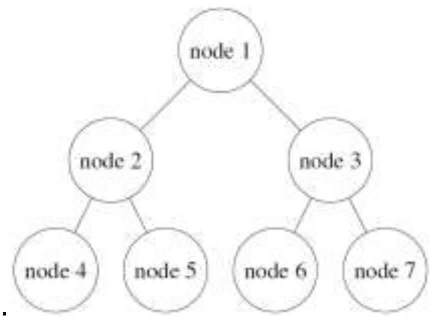

Figure 1. Decision tree algorithm

To check whether extensive reading approaches is useful for each group, results of comprehensive reading at initial and semester end. All participant in the proposed work attended comprehensive reading and self-efficacy test in semester beginning and at semester end to check their knowledge level about their reading skills and whether there is any increase in self-efficacy reading or not as discussed in the following section [14]

III Decision Tree Splitting Point Methods

A Variance

Variance is a techniquefor mid-point of the node utilized when the object variablesare continuous, i.e., regression algorithm. It is referred becausetheusageof variance is as a quantity for determining the feature on node is divided into child node or terminalnode[16].

$$Varience = \frac{\sum (\mathrm{X} - \mu)^2}{N} \qquad (1)$$

Similarity of node calculation for exclusively equal node, variance is usedwith value zero.

B. Information Gain

In case, what if objective variable of categorial attribute appears? Variation reduction will never cut it. The only solution is to use information gain for splitting the node points when final node is categorial, this is done by using entropy given by:

$$Information\ Gain = 1 - Entropy \qquad (2)$$

Node transparency can be computed using entropy. “Lower the value of entropy, higher is the purity of the node”. Equal node entropy value is zero. So, subtract 1 from entropy; for all purer node the information gain is greater with value 1 as maximum[13,14,12]. To calculate the entropy the following formula is used:

$$Entropy = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad (3)$$

C. Gini Index

For node splitting Gini index is used if the variable is categorical. It is simple and best one for splitting node points of a DT. The Gini index can be calculated by:

$$Gini\ Index = 1 - Gain \qquad (4)$$

It is the probability to appropriately classify elements that are selected randomly if it was considered for label distribution in the nodes. Gini index can be known by:

$$Gini\ Index = 1 - \sum_{i=1}^{n} p_i^2 \qquad (5)$$

“Lesser the Gini Index, greater is the homogeneity of the node.” Pure node Gini index value is zero.

D. Chi-Square

It is one of the methods to calculate two splitting point for a node in a DT for object values that are categorical in nature. The splitting point can be calculated by using a statistical difference between child and parent node using the formula[13,14,12,10].

Chi-Square value is:

$$Chi - Square = \sqrt{\frac{(\mathrm{Actual} - \mathrm{Expected})^2}{Expected}} \qquad (6)$$

Expected is a predictable class value from the child node among the classes established and distributed in parent node. And finally, with actual value that is a real class value in terminal nodes.

IV MMDBM ALGORITHM

MMDBM algorithm is based on decision tree and splitting point of each node of the datasets. In this method is based on the two models, specifically SPRINT and SLIQ. Both algorithms are based on the Decision Tree algorithm, it deals with numeric and categorical attributes of the datasets. MMDBM model constructs a tree that is a perfect DT. It develops an improved arrangement system to reduce numeric attributes evaluation cost in tree constructing stage.In this approach, it is synchronised with a breadth-first tree evolvingtechnique to constructdisk-resident classification datasets. For every attribute SLIQ develops a fastalgorithm to classify split point which is categorical in features. Alternatively,DTmethod is used to control numerical and categorical data in massive datasets known- (MMDBM).

A. Evaluation Results

The dataset was generated manually for reading course from students belonging to first year and 603 undergraduate students from K L University of Guntur for proposed study. For the experimental purpose, the results are compared with decision tree algorithm and MMDBM algorithm to calculate precision, recall, F-Measure and accuracy

Evaluation metrics

Data mining approaches requires evaluation technique. "This technique is utilized to authenticate the model produced by MMDBM algorithm. In supervised learning data labels can be used for evaluation of classification models using different metrics like f-measure, recall and precision, confusion matrices and performance.

Table 1. confusion matrix

P- Pre-test control Q- Post-test control

R- Pre-test exp S- Post-test exp

	P	Q	R	S
P	tpP	ePQ	ePR	ePS
Q	eQP	tpQ	eQR	eQS
R	eRP	eRQ	tpR	eRS
S	eSP	eSQ	eSR	tpS

The confusion matrix is shown in Table1. The diagonal elements represent the features are classified appropriately and the remaining features represents incorrect classification of data. Precision is represented as the ratio between both false positive and true positive, true positive.

Precision: This is data retrieval;precisionrepresents retrieved data from connected dataset applicable for search. It indicates how many instances are classified correctly in confusion matrix as shown in Table.1. (true positive shows correct data and error positive indicates incorrect data classification.

$$\Pr ecision = \frac{tpP}{tpP + eQP + eRP + eSP} \qquad (7)$$

Where*tpP*is indicates true positive for the class *P* and*eQP, eRP and ePS*implies as false negative.

Recall: This is a data retrieval; recall indicates the portion of retrieved data from connecteddatasets that are applicableforqueries that are successfully. Recallrepresents the total number of instancesare classified correctly and the ratio between both true positive, error negative and true positive.

$$\operatorname{Re} call = \frac{tpP}{tpP + ePQ + ePR + ePS} \qquad (8)$$

Where*tpP*is represents true positive for the class P and*ePQ, ePR and ePS* are representing error positive.

F-measure: This is evaluated by the harmonic mean among precision and recall.

$$F - Measure = 2 * \frac{\Pr ecision * \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \qquad (9)$$

Accuracy: From all data; it is calculated by the amount of true negative, true positive and true outcomes.

$$Accuracy = \frac{(tpP + tpQ + tpR + tpS)}{\begin{pmatrix} tpP + eQP + eRP + eSP + ePQ + tpQ + eRQ + eSQ \\ + ePR + eQR + tpR + eSR + ePS + eQS + eRS + tpS \end{pmatrix}} \tag{10}$$

The decision tree classification algorithms run on the test datasets and process each record. It classifies the records into pre-test control and experiment and post-test control and experiment. The results are verified for the above proposed classifier. These results are shown in table-2 and table-3. The total records employed in proposed model is shown in Table 3.

The evaluated results by the precision, recall, and f-measure are shown in Figure4. The proposed method (MMDBM) is able to classify the data into classified nodes.This is shown in Table2 and 3.
Table2 and Figure2 is describing confusion matrix for decision tree algorithm and comparing class A datasets using Pre-test, Post-test control and experiment".

Table 2: Decision tree for confusion matrices, precision, recall, and accuracy

Data set		Confusion Matrices				Results			Accuracy %
		Pre-test control	Post-test control	Pre-test exp	Post-test exp	Precision %	Recall %	F-Measure %	
Class A	Pre-test control	114	9	5	7	84.4	84.5	84.7	83.4
	Post-test control	7	120	6	6	84.5	86.3	85.4	
	Pre-test exp	6	7	113	6	85.0	85.6	85.3	
	Post-test exp	8	6	9	174	90.2	88.3	89.3	

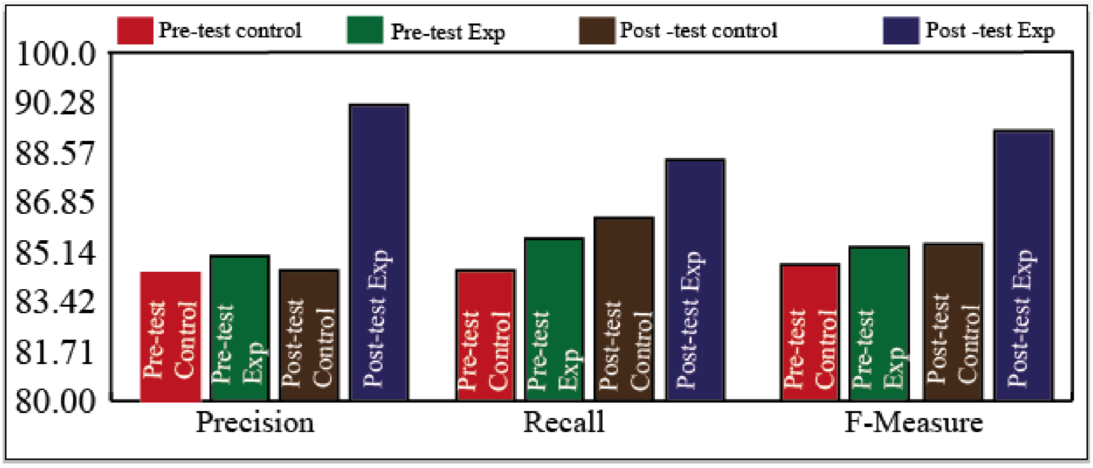

Figure 2: Pre-test and Post-test using decision tree and compared with precision, recall, F-Measure and accuracy

Table 3 and Figure 3 is describing about the confusion matrix for decision tree algorithm and comparing class A datasets using Pre-test, Post-test control and experiment.

Table 3: MMDBM Algorithm for confusion matrices, precision, recall, and accuracy

Data set		Confusion Matrices				Results			Accuracy %
		Pre-test control	Post-test control	Pre-test exp	Post-test exp	Precision %	Recall %	F-Measure %	
Class A	Pre-test control	112	8	8	7	87.5	85.5	86.5	88.8
	Post-test control	8	118	8	5	86.8	86.8	86.8	
	Pre-test exp	7	8	111	6	88.1	88.1	88.1	
	Post-test exp	9	7	9	172	91.5	96.6	94.0	

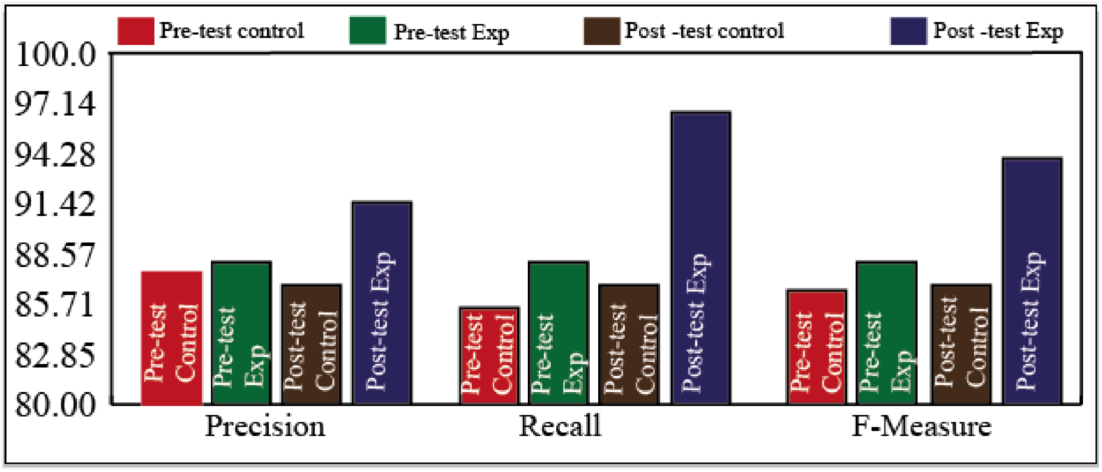

Figure 3: Pre-test and Post-test using MMDBM algorithm and compared with precision, recall,F-Measure and accuracy

V Conclusion

The final outcomes must be discussed and comparison of the Control Group and experimental group outcomes from the post-test and pre-test are calculated using decision tree algorithm and *MMDBM* algorithm. For both algorithm confusion matrix is generated and*Recall, precision, F-measure and Accuracy* are calculated. The proposed algorithm called *MMDBM* for classification method based on the *DT* algorithm. The outcomes of the *MMDBM* model achieved accuracy 88.8% greater than *DT* algorithm based on the pre-test, post-test control with comparative analysis.

References

1. Elley, W. B., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly, 19,* 53-67.
2. Haas, C., & Flower, L. (1988). Rhetorical reading strategies and the construction of meaning. *College Composition and Communication, 39,* 167-83.
3. Rajesh Kumar. E, K.V.S.N. Rama Rao, Soumya Ranjan Nayak & Ramesh Chandra. (2020). Suicidal ideation prediction in twitter data using machine learning techniques, Journal of Interdisciplinary Mathematics, 23(1), 117-125.
4. Rajesh Kumar .E & K.V.S.N. Rama Rao. (2020). Suicide Prediction in Twitter Data using Mining Techniques: A Survey, International Conference on Intelligent Sustainable Systems (ICISS 2019), 122-131.
5. Prasanth Yalla, Venkata Naresh Mandhala, ValavalaAbhishiktha, ChitturiSaisree, KandepiManogna (2019), Machine Learning Techniques to Predict Defects by using Testing Parameters, International Journal of Recent Technology and Engineering (IJRTE), 8(4), 7829- 7834.
6. Prasanth, Y., Sreedevi, E., Gayathri, N., & Rahul, A. S. (2017). Analysis and implementation of ensemble feature selection to improve accuracy of software defect detection model. Journal of Advanced Research in Dynamical and Control Systems, 9(18), 601-613.

7. Pradeepini, G., Pradeepa, G., Tejanagasri, B., & Gorrepati, S. H. (2018). Data classification and personal care management system by machine learning approach. International Journal of Engineering and Technology (UAE), 7(32), 219-223.
8. Reza Raissi, Mostafa Roustaei, On the relationship of reading strategies, extensive reading and self-efficacy, 6th International Conference on University Learning and Teaching (InCULT 2012), Procedia - Social and Behavioral Sciences 90 (2013) 634 – 640.
9. Smith, F. (1985). *Reading without nonsense* (2nd ed.). New York, NY: Teachers College Press.
10. Siva Kumar Pathuri, ViswaGanesh.A, Ravi Teja.P and V.Rishi Chowdary. (2019). Feature-Based Opinion Mining for Amazon Product's using MLT, International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(11), 4105- 4109.
11. Siva Kumar. P, N.Anbazhaghan, Shaik Razia, M Sivani, S Pravalika, A S Harshini, Prediction of cardiovascular disease using classification techniques with high accuracy, JARDC, 12(2), 1134-1139.
12. S.Sivakumar,Soumya Ranjan Nayak, Ashok Kumar, S. Vidyanandini, Gopinath Palai, "An empirical study of supervised learning methods for Breast Cancer Diseases", " International Journal for Light and Electron Optics ", 175, pp.105-114, 2018.
13. Sreedevi, E., & Prasanth, Y. (2017). A novel class balance ensemble classification model for application and object oriented defect database, Journal of Advanced Research in Dynamical and Control Systems, 9, 702-726.
14. Sreedevi. E, PremaLatha. V, Sivakumar. S. (2019). A Comparative Study on New Classification Algorithm using NASA MDP Datasets for Software Defect Detection, *2nd International Conference on Intelligent Sustainable Systems (ICISS 2019),* 312-317.
15. Susser, B., & Robb, T. (1990). EFL extensive reading instruction: Research and procedure. *JALT Journal*, *12* (2).
16. Tremblay, P. F., & Gardner, R. C. (1995). Expanding the motivation construct in language learning. *Modern Language Journal, 79*, 505-518.