Alinteri J. of Agr. Sci. (2021) 36(1): 399-405
e-ISSN: 2587-2249
info@alinteridergisi.com

http://dergipark.gov.tr/alinterizbd
http://www.alinteridergisi.com/
DOI:10.47059/alinteri/V36I1/AJAS21060

RESEARCH ARTICLE

Comparison of Logistic Regression and Generalized Linear Model for Identifying Accurate At - Risk Students

K. Harini[1] • K. Sashi Rekha[2*]

[1]*Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. E-mail: harini.kamjula@gmail.com*

[2*]*Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. E-mail: sashirekhak.sse@saveetha.com*

ARTICLE INFO

Article History:
Received: 22.03.2021
Accepted: 25.04.2021
Available Online: 21.06.2021

Keywords:
Accuracy
Prediction
Significantly
Statistical
Probabilities
At-risk Students
Novel Sigmoid Logistic Regression
GLM Algorithm
Machine Learning

ABSTRACT

Aim: To predict the accuracy percentage of At - risk students based on High withdrawal and Failure rate. **Materials and methods:** Logistic Regression with sample size = 20 and Generalised Linear Model (GLM) with sample size = 20 was iterated different times for predicting accuracy percentage of At - risk students. The Novel sigmoid function used in Logistic Regression maps prediction to probabilities which helps to improve the prediction of accuracy percentage. **Results and Discussion:** Logistic Regression has significantly better accuracy (94.48 %) compared to GLM accuracy (92.76 %). There was a statistical significance between Logistic regression and GLM (p=0.000) (p<0.05). **Conclusion:** Logistic Regression with Novel Sigmoid function helps in predicting with more accuracy percentage of At - risk students.

Please cite this paper as follows:

Harini, K. and Rekha, K.S. (2021). Comparison of Logistic Regression and Generalized Linear Model for Identifying Accurate At - Risk Students. *Alinteri Journal of Agriculture Sciences, 36*(1): 399-405. doi: 10.47059/alinteri/V36I1/AJAS21060

Introduction

The purpose of the study is to predict the accuracy percentage of At-Risk students in Universities. Now-a-days many colleges, Universities and schools have adapted the digital education systems as an alternative to the traditional education system. Everyone prefers the digital education system (Al-Shabandar et al. 2019). Many research studies have proved that High failure rate and withdrawal rate were the major concerns of the digital education systems (J. Chen et al. 2019). In upcoming years, if the problem is not solved it may lead to a low literature rate, because the digital education system is developing more rapidly. In many paid online course platforms still, students withdraw their courses or get less score because of lack of motivation and proper care (Sun et al. 2019).

* Corresponding author: sashirekhak.sse@saveetha.com

If these kinds of students are identified in the early stage it is easy for instructors in Universities to track the students and also students will get proper motivation and guidance (Chaplot, Rhim, and Kim 2016).

Identifying At-risk students process was carried out by many researchers for developing digital education with a high literature rate. Around 20 articles published in IEEE and 16 papers in Google scholar. (Chaplot, Rhim, and Kim 2016) Feed Forward Neural networks was implemented for predicting the accuracy percentage of At-Risk students. The accuracy was predicted on the basis of Clickstream and student sentiments by using Coursera Platform dataset of 2014 and predicted an accuracy of 80.5%. (Xing and Du 2019), Convolution neural network combined with Recurrent neural network to extract features consequently in order to make predictions for predicting the accuracy percentage of At-risk students and the dataset used to predict accuracy in this paper was the online course "XuetangX" dataset. The results demonstrated an accuracy of 92.6%. (Geigle and Zhai 2017), Two-layer hidden Markov Model was used for predicting accuracy percentage of At- risk students on the basis of student's behavioural patterns and results demonstrated that Low motivational status, withdrawal of course at early stage and failure rate are the major concerns and got an accuracy of 80.93%. (Al-Shabandar et al. 2019), At-risk students accuracy was predicted using machine learning algorithms which includes GLM, Gradient Boost machine algorithm, Random Forest. The results of their work demonstrated an accuracy of 91%, 90%, 91% respectively. The most cited article is proposed by the author (Al-Shabandar et al. 2019) focused on GLM with an accuracy of 91%.

Previously our team has a rich experience in working on various research projects across multiple disciplines (Sathish and Karthick 2020; Varghese, Ramesh, and Veeraiyan 2019; S.R. Samuel, Acharya, and Rao 2020; Venu, Raju, and Subramani 2019; M. S. Samuel et al. 2019; Venu, Subramani, and Raju 2019; Mehta et al. 2019; Sharma et al. 2019; Malli Sureshbabu et al. 2019; Krishnaswamy et al. 2020; Muthukrishnan et al. 2020; Gheena and Ezhilarasan 2019; Vignesh et al. 2019; Ke et al. 2019; Vijayakumar Jain et al. 2019; Jose, Ajitha, and Subbaiyan 2020). Now the growing trend in this area motivated us to pursue this project.

In many Universities digital education has been adapted but proper guidance and awareness of students was not followed by the instructors. This leads to a low graduation rate of students. If proper identification of student's performance was done, the overall graduation rate will increase. The main reason for the reduced graduation rate was student withdrawal of course at an early stage and also a high failure rate. To avoid these two drawbacks in digital education, students need to be identified, and proper motivation and effective teaching can be done. The prediction of accuracy percentage of At-risk students can be done with dynamic dataset in order to predict with more accuracy rather than working only with static datasets. Based on the literature review it can be inferred that many machine learning algorithms have been widely used for predicting the accuracy percentage of At-Risk students. Logistic Regression machine Learning algorithm mainly concentrates on classification and also uses Novel sigmoid function for mapping predicted variables to probabilities which improves the prediction of accuracy percentage for At-risk students.

Materials and Methods

The study setting of the proposed work is done in Saveetha School of Engineering. The number of groups identified for the study is two. The group 1 is Logistic Regression and group 2 is GLM. Logistic Regression algorithm and GLM algorithm was iterated at different times with a sample size of 20 (Al-Shabandar et al. 2019), 95% confidence interval and pretest power of 80% ("Sample Size Calculator" n.d.).

The real time dataset used was the Harvard dataset. The input dataset for the proposed work is Harvard dataset.csv collected from kaggle.com (Narang n.d.). The main attributes used to predict the At-risk students accuracy (%) was "grade" (The grade of the student), "Viewed" (No. of Videos Viewed), "certified" (The certification student received). "Course_id" refers to the ID number of the course, "year" refers to the student year of birth, "explored" refers to student view to the home page, "Final_cc_name" refers to student location, "LOE_DI" refers to student degree, "start_time_DI" refers to student start time of activity, "Last_event_DI" refers to students end time of activity, "n_events" refers to number of events a student participated, "nplay_videos" refers to number times a video played by a student, "nchapters" refers to number of chapters, "nforums_post" refers to number of quizzes a student participated, "gender" refers to gender of the student. Above all refers to the attribute description of the Harvard dataset.

After the collection of dataset, it was uploaded and preprocessing was done. All the null values and missing values present in the dataset was removed by cleaning the data. After cleaning, the feature extraction was applied to perform vectorisation, the data which are strings, words and characters was changed to values 0 and 1. Example for female gender attribute the value assigned was 1 and for male gender attribute value assigned was 0. The obtained data set without null values and missing values was well qualified for evaluating the machine learning algorithm. After preprocessing the dataset was splitted into two parts and evaluated as a 25% testing set and 75% training set. The proposed algorithm Logistic Regression was implemented by evaluating the train and test set and the required accuracy percentage was predicted. The learning process of GLM and Logistic Regression is given below.

GLM is a statistical method that comes under supervised machine learning algorithms which assumes every number of observations has a distribution like polynomial, binomial, gamma, average (Al-Shabandar et al. 2019). GLM gives a continuous output of dependent variables by evaluating independent variables and performs linearly.

$$\eta_i = \beta 0+\beta 1X1 + \beta 2X2+\ldots\ldots\beta nXn \qquad (1)$$

Where $X1, X2\ldots..Xn$ are the observations and predictive variables, η_i is the dependent variable as shown in equation (1). $\beta 0$ is the intercept and βi are the coefficients

The main attributes of the dataset grade, viewed and certified, was used for predicting the accuracy percentage of At-risk students using GLM Machine Learning Algorithm as shown in Fig. 1.

```
Import the dataset and the required packages.
Define X as grade, certified and Y as viewed
Define train and test sets
Log←glm (df, data=data_train, family=binomial)
Table←table (data$certified, data$grad, data$viewed, pred>0.5)
Accuracy←sum(diag(table)/sum(table))
Predicted accuracy of At-Risk students
```

Fig. 1. Algorithm for GLM (X,Y)

Logistic Regression is one of the most important supervised machine learning algorithms. Logistic Regression mainly concentrates on classification example 0 or 1 and pass or fail. It helps in predicting categorical dependent variables using independent variables. It uses the Novel sigmoid function for mapping the predicted values to probabilities and decides on which values to pass as output and not to pass. The general equation for Logistic regression was shown in equation (2).

$$log[y/y-1] = b0 + b1x1 + b2x2 + \ldots\ldots + bnxn \qquad (2)$$

Where $x1, x2, \ldots\ldots xn$ are the observations and predictive variables

$log[y/y-1]$ is the sigmoid function

$b0$ is the intercept and bi are the coefficients

The main attributes of the dataset grade, viewed and certified, was used for predicting the accuracy percentage of At-risk students using Logistic Regression Machine Learning Algorithm as shown in Fig. 2.

```
Import the dataset and the required packages.
Define X as grade, certified and Y as viewed
Define train and test sets
x_train, x_test, y_train, y_test←train_test_split (x, y, test_size=0.25)
Log←LogisticRegression ()
Log. Fit (x_train, y_test)
y_pred ← Log. Predict(x_test)
Score ←accuracy_score (y_test, y_pred)
Predicted accuracy of At-Risk students
```

Fig. 2. Algorithm for Logistic Regression (X,Y)

Accuracy was calculated for Logistic Regression and GLM based on equation (3).

$$Accuracy = TP + TN/TP + TN + FP + FN \qquad (3)$$

Where TP, is the number of true positives classified by the algorithm

TN, is the **number** of true negatives classified by the algorithm

FP, is the number of false positives classified by the algorithm

FN, is the number of false negatives classified by the algorithm

The software tool used to evaluate the Logistic Regression and GLM algorithms was Google colab© with python programming language. The hardware configuration was intel core i5 processor with a RAM size of 8GB. The system type used was 64-bit, OS, X64 based processor with HDD of 917 GB. The software configuration includes windows 10 operating system.

From the total sample size 75% of the data with features extracted is trained in the Logistic Regression and GLM. For training the model involves a number of iterations to get better performance. After training the algorithm, random test data is given to the algorithm.

The work was statistically analysed using the Statistical Package for the Social Sciences (SPSS) ("SPSS Software" n.d.) and Goggle collab©. Descriptive statistics for mean, standard deviation and standard error was carried out for Logistic Regression and GLM algorithm. The independent variables are course id, user id, year, semester, viewed, certified, grade, explored, final_cc_name, start_time_Dl, end_time_Dl, nevents, ndays_act, nplay_videos, nchpaters, nforums_post, incomplete, age. The dependent variable was Accuracy. Independent Sample t-test was performed to compare the performance of algorithms.

Results

In Table 1, it was observed that Logistic Regression algorithm and GLM algorithm were run at different times in Google colab© with a sample size of 20 and accuracy was calculated. Logistic Regression appears to have better accuracy than the GLM algorithm. In Table 2, an Independent Sample T-test was performed to compare the accuracy of Logistic Regression and GLM algorithms and a statistically significant difference was noticed $P < 0.001$ with 95% confidence interval showed that our hypothesis holds good. The mean difference of Accuracy was identified as 2.13200. In Table 3, the statistical analysis of 10 samples was performed, Logistic Regression obtained 0.68733 standard deviation with 0.21735 standard error while GLM algorithm obtained 1.41033 standard deviation with 0.44599 standard error. With respect to changes in the input values (independent variables) the corresponding output values (dependent variables) also changes (Table 2). Accuracy percentage of Logistic Regression (94.48) and GLM (92.76) infers that Logistic Regression appears to have better accuracy than GLM (Fig. 3) and simple mean bar graph shows the standard deviation of Logistic Regression is slightly better than GLM (Fig. 4).

Table 1. Predicted accuracy of At-risk students (Logistic Regression accuracy of 94.48% and GLM accuracy of 92.46%)

Sl. No	Sample Size	Logistic Regression Accuracy (%)	GLM Accuracy (%)
1	21	93.79	93.45
2	31	93.82	90.03
3	41	93.50	93.14
4	51	94.18	93.19
5	61	92.11	92.38
6	71	94.12	89.76
7	81	94.48	92.76
8	91	93.82	93.85
9	100	93.89	92.03
10	120	93.97	92.24

Table 2. Independent Sample T - test Result is applied for dataset fixing confidence interval as 95% and level of significance as 0.05 (Logistic regression appears to perform significantly better than GLM with the value of p=0.000)

		Levene's test for equality of variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std.Error Difference	95% confidence interval of the difference: Lower	95% confidence interval of the difference: Upper
Accuracy	Equal variances Assumed	6.602	.019	4.29	18	.000	2.13200	.49613	1.089	3.17433
	Equal variance not Assumed			4.29	13.04	.001	2.13200	.49613	1.060	3.20343
Loss	Equal variances Assumed	6.569	.020	-4.29	18	.000	-2.13200	.49638	-3.175	-1.0901
	Equal variance not Assumed			-4.29	13.06	.001	-2.13200	.49638	-3.204	-1.0611

Table 3. Group Statistical analysis of Logistic Regression and GLM. Mean accuracy value, Standard deviation and Standard Error Mean for Logistic Regression and GLM algorithms are obtained for 10 iterations. It is observed that the Logistic Regression algorithm performed better than the GLM algorithm.

Algorithm		N	Mean	Std. Deviation	Std. Error Mean
Accuracy	Logistic Regression	10	94.3730	.68733	.21735
	GLM	10	92.2410	1.41033	.44599
Loss	Logistic Regression	10	5.6260	.68733	.21735
	GLM	10	7.7590	1.41033	.44599

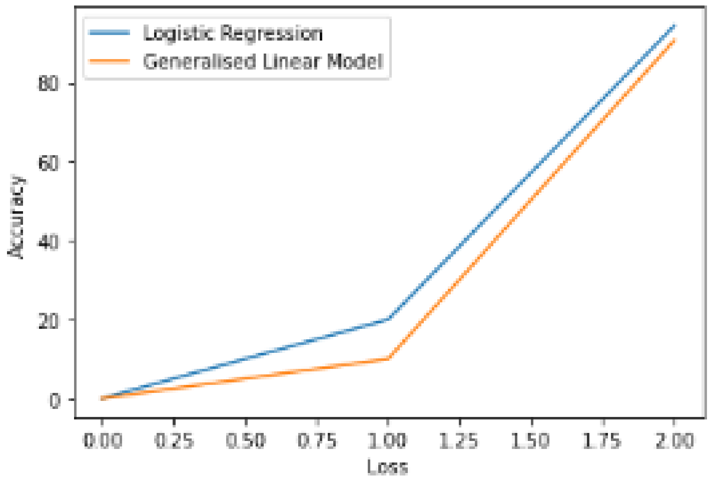

Fig. 3. Comparison of accuracy percentage (Logistic Regression got more accuracy of 94.48% than GLM accuracy of 92.46%)

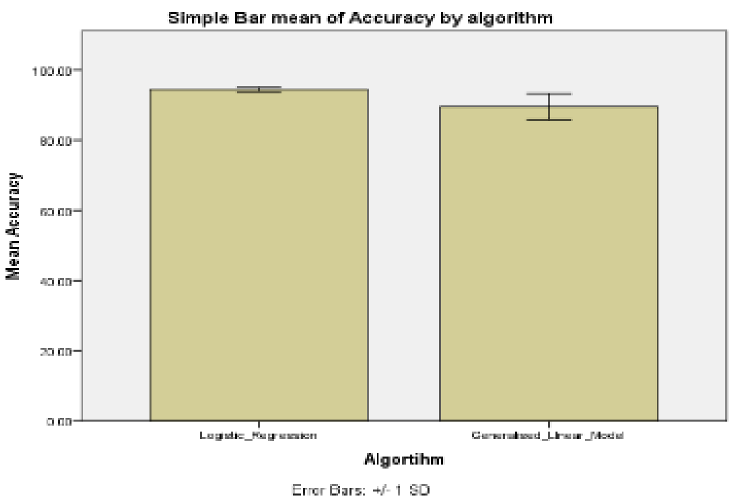

Fig. 4. Comparison of Logistic Regression and GLM in terms of mean accuracy. The mean accuracy of Logistic Regression is better than GLM and the standard deviation of Logistic Regression is slightly better than GLM. X Axis: Logistic Regression vs GLM Algorithm, Y Axis: Mean accuracy of detection ± 1 SD

Discussion

In this research work Logistic Regression and GLM algorithm was analysed for predicting the accuracy percentage of At-risk students. It is observed that Logistic Regression appears to have better accuracy (94.48 %) compared to GLM (92.76 %). The Novel sigmoid function maps the predictions to the probabilities of At-Risk students based on Certified, grade attributes of the dataset which helped in improving the accuracy percentage. The results show the evidence there is a statistically significant difference between the Logistic Regression and GLM algorithms.

In this paper (Al-Shabandar et al. 2019; Chaplot, Rhim, and Kim 2016; Xing and Du 2019), feedforward neural networks was implemented with an accuracy percentage of 80.5%. (Sun et al. 2019; Al-Shabandar et al. 2018), CNN combined with RNN and predicted an accuracy percentage of 92.6%. (Al-Shabandar et al. 2018), 75% of accuracy was predicted using the Data Driven approach. (Minaei-Bidgoli et al., n.d.), KNN was implemented with an accuracy of 82.3%. (Y. Chen et al. 2020), explained the prediction model and the results demonstrated an accuracy of 90.4%. (Geigle and Zhai 2017), 90.3% of accuracy was predicted using two-layer hidden Markovs. (Gardner and Brooks 2018) SVM and GLM were implemented with an accuracy of 89.2% and 91.4% respectively. (J. Chen et al. 2019) 85.96% of accuracy was predicted using a decision tree and extreme learning machine.

Our institution is passionate about high quality evidence based research and has excelled in various fields ((Vijayashree Priyadharsini 2019; Ezhilarasan, Apoorva, and Ashok Vardhan 2019; Ramesh et al. 2018; Mathew et al. 2020; Sridharan et al. 2019; Pc, Marimuthu, and Devadoss 2018; Ramadurai et al. 2019). We hope this study adds to this rich legacy.

The attributes that affect the accuracy percentage of At-risk students in this research work are course_ID, final_cc_cname_DI, viewed, year, nforums_post, explored, gender, Last_event_DI, LOE. The attributes that mainly concentrated to increase the accuracy percentage were Certified and grade.Logistic Regression appears to have better Accuracy compared with previous research articles discussed. This proposed work can be implemented in universities, schools and in online course platforms for giving intensive support to students. The limitation of the proposed work is that one of the attributes in the Harvard dataset used for predicting the accuracy of At-risk students is "Viewed" which doesn't predict the exact accuracy of At-risk students. In future work, if the dataset has attributes like attendance, student_feedback, student_interest there might be a chance to predict more accuracy percentage of At-risk students.

Conclusion

In this proposed work, prediction of accuracy percentage of At-risk students based on failure and withdrawal rate is performed using Logistic Regression appears to have improved accuracy of 94.48% when compared to GLM algorithm.

Declarations

Conflict of Interests

No conflict of interest in this manuscript.

Authors Contributions

Author K Harini was involved in data collection, data analysis, manuscript writing. Author Sashi Rekha K was involved in conceptualization, data validation, and critical review of manuscript.

Acknowledgments

The authors would like to express their gratitude towards Saveetha School of Engineering for providing the necessary infrastructure to carry out this work successfully.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1. Wells Fargo Pvt. Ltd., Chennai.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

References

Al-Shabandar, Raghad, Abir Jaafar Hussain, Panos Liatsis, and Robert Keight. 2018. Analyzing Learners Behavior in MOOCs: An Examination of Performance and Motivation Using a Data-Driven Approach. *IEEE Access*. https://doi.org/10.1109/access.2018.2876755

Al-Shabandar, R., Hussain, A.J., Liatsis, P., & Keight, R. (2019). Detecting at-risk students with early interventions using machine learning techniques. *IEEE Access, 7,* 149464-149478 https://doi.org/10.1109/access.2019.2943351

Chaplot, Devendra Singh, Eunhee Rhim, and Jihie Kim. 2016. Personalized Adaptive Learning Using Neural Networks. *Proceedings of the Third (2016) ACM Conference on Learning @ Scale* https://doi.org/10.1145/2876034.2893397

Chen, Jing, Jun Feng, Xia Sun, Nannan Wu, Zhengzheng Yang, and Sushing Chen. 2019. MOOC Dropout Prediction Using a Hybrid Algorithm Based on Decision Tree and Extreme Learning Machine. *Mathematical Problems in Engineering*. https://doi.org/10.1155/2019/8404653

Chen, Yan, Qinghua Zheng, Shuguang Ji, Feng Tian, Haiping Zhu, and Min Liu. 2020. Identifying at-Risk Students Based on the Phased Prediction Model. *Knowledge and Information Systems* https://doi.org/10.1007/s10115-019-01374-x

Ezhilarasan, Devaraj, Velluru S. Apoorva, and Nandhigam Ashok Vardhan. 2019. Syzygium Cumini Extract Induced Reactive Oxygen Species-Mediated Apoptosis in Human Oral Squamous Carcinoma Cells. *Journal of Oral Pathology & Medicine: Official Publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology* 48(2): 115–21.

Gardner, Josh, and Christopher Brooks. 2018. Student Success Prediction in MOOCs. *User Modeling and User-Adapted Interaction*.

https://doi.org/10.1007/s11257-018-9203-z

Geigle, Chase, and Chengxiang Zhai. 2017. Modeling MOOC Student Behavior with Two-Layer Hidden Markov Models. *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale*.

https://doi.org/10.1145/3051457.3053986

Gheena, S., and D. Ezhilarasan. 2019. Syringic Acid Triggers Reactive Oxygen Species-Mediated Cytotoxicity in HepG2 Cells. *Human & Experimental Toxicology* 38(6): 694–702.

Jose, Jerry, Ajitha, and Haripriya Subbaiyan. 2020. Different Treatment Modalities Followed by Dental Practitioners for Ellis Class 2 Fracture – A Questionnaire-Based Survey. *The Open Dentistry Journal* 14(1): 59–65.

Ke, Yang, Mohammed Saleh Al Aboody, Wael Alturaiki, Suliman A. Alsagaby, Faiz Abdulaziz Alfaiz, Vishnu Priya Veeraraghavan, and Suresh Mickymaray. 2019. Photosynthesized Gold Nanoparticles from Catharanthus Roseus Induces Caspase-Mediated Apoptosis in Cervical Cancer Cells (HeLa). *Artificial Cells, Nanomedicine, and Biotechnology* 47(1): 1938–1946.

Krishnaswamy, Haribabu, Sivaprakash Muthukrishnan, Sathish Thanikodi, Godwin Arockiaraj Antony, and Vijayan Venkatraman. 2020. Investigation of Air Conditioning Temperature Variation by Modifying the Structure of Passenger Car Using Computational Fluid Dynamics. *Thermal Science* 24(1 Part B): 495–498.

Malli Sureshbabu, Nivedhitha, Kathiravan Selvarasu, Jayanth Kumar V, Mahalakshmi Nandakumar, and Deepak Selvam. 2019. Concentrated Growth Factors as an Ingenious Biomaterial in Regeneration of Bony Defects after Periapical Surgery: A Report of Two Cases. *Case Reports in Dentistry* 2019 (January): 7046203.

Mathew, M.G., S.R. Samuel, A.J. Soni, and K.B. Roopa. 2020. Evaluation of Adhesion of Streptococcus Mutans, Plaque Accumulation on Zirconia and Stainless Steel Crowns, and Surrounding Gingival Inflammation in Primary …. *Clinical Oral Investigations*. https://link.springer.com/article/10.1007/s00784-020-03204-9

Mehta, Meenu, Deeksha, Devesh Tewari, Gaurav Gupta, Rajendra Awasthi, Harjeet Singh, Parijat Pandey, et al. 2019. Oligonucleotide Therapy: An Emerging Focus Area for Drug Delivery in Chronic Inflammatory Respiratory Diseases. *Chemico-Biological Interactions* 308 (August): 206–215.

Minaei-Bidgoli, B., D.A. Kashy, G. Kortemeyer, and W.F. Punch. n.d. Predicting Student Performance: An Application of Data Mining Methods with an Educational Web-Based System. *33rd Annual Frontiers in Education, 2003. FIE 2003*. https://doi.org/10.1109/fie.2003.1263284

Muthukrishnan, Sivaprakash, Haribabu Krishnaswamy, Sathish Thanikodi, Dinesh Sundaresan, and Vijayan Venkatraman. 2020. Support Vector Machine for Modelling and Simulation of Heat Exchangers. *Thermal Science* 24(1 Part B): 499–503.

Narang, Kanika. n.d. MOOC Dataset. Accessed March 24, 2021. https://www.kaggle.com/kanikanarang94/mooc-dataset.

Pc, J., T. Marimuthu, and P. Devadoss. 2018. Prevalence and Measurement of Anterior Loop of the Mandibular Canal Using CBCT: A Cross Sectional Study. *Clinical Implant Dentistry and Related Research*. https://europepmc.org/article/med/29624863.

Ramadurai, Neeraja, Deepa Gurunathan, A. Victor Samuel, Emg Subramanian, and Steven J.L. Rodrigues. 2019. Effectiveness of 2% Articaine as an Anesthetic Agent in Children: Randomized Controlled Trial. *Clinical Oral Investigations* 23 (9): 3543–50.

Ramesh, Asha, Sheeja Varghese, Nadathur D. Jayakumar, and Sankari Malaiappan. 2018. Comparative Estimation of Sulfiredoxin Levels between Chronic Periodontitis and Healthy Patients - A Case-Control Study. *Journal of Periodontology* 89(10): 1241–1248.

Sample Size Calculator. n.d. Accessed March 23, 2021. https://www.calculator.net/sample-size-calculator.html.

Samuel, Melvin S., Jayanta Bhattacharya, Sankalp Raj, Needhidasan Santhanam, Hemant Singh, and N. D. Pradeep Singh. 2019. Efficient Removal of Chromium (VI) from Aqueous Solution Using Chitosan Grafted Graphene Oxide (CS-GO) Nanocomposite. *International Journal of Biological Macromolecules* 121(January): 285–292.

Samuel, Srinivasan Raj, Shashidhar Acharya, and Jeevika Chandrasekar Rao. 2020. School Interventions-Based Prevention of Early-Childhood Caries among 3-5-Year-Old Children from Very Low Socioeconomic Status: Two-Year Randomized Trial. *Journal of Public Health Dentistry* 80(1): 51–60.

Sathish, T., and S. Karthick. 2020. Wear Behaviour Analysis on Aluminium Alloy 7050 with Reinforced SiC through Taguchi Approach. *Journal of Japan Research Institute for Advanced Copper-Base Materials and Technologies* 9(3): 3481–3487.

Sharma, Parvarish, Meenu Mehta, Daljeet Singh Dhanjal, Simran Kaur, Gaurav Gupta, Harjeet Singh, Lakshmi Thangavelu, et al. 2019. Emerging Trends in the Novel Drug Delivery Approaches for the Treatment of Lung Cancer. *Chemico-Biological Interactions* 309 (August): 108720.

SPSS Software. n.d. Accessed March 28, 2021. https://www.ibm.com/analytics/spss-statistics-software.

Sridharan, Gokul, Pratibha Ramani, Sangeeta Patankar, and Rajagopalan Vijayaraghavan. 2019. Evaluation of Salivary Metabolomics in Oral Leukoplakia and Oral Squamous Cell Carcinoma. *Journal of Oral Pathology & Medicine: Official Publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology* 48(4): 299–306.

Sun, Di, Yueheng Mao, Junlei Du, Pengfei Xu, Qinhua Zheng, and Hongtao Sun. 2019. Deep Learning for Dropout Prediction in MOOCs. *2019 Eighth International Conference on Educational Innovation through Technology (EITT)*. https://doi.org/10.1109/eitt.2019.00025.

Varghese, Sheeja Saji, Asha Ramesh, and Deepak Nallaswamy Veeraiyan. 2019. Blended Module-Based Teaching in

Biostatistics and Research Methodology: A Retrospective Study with Postgraduate Dental Students. *Journal of Dental Education* 83(4): 445–450.

Venu, Harish, V. Dhana Raju, and Lingesan Subramani. 2019. Combined Effect of Influence of Nano Additives, Combustion Chamber Geometry and Injection Timing in a DI Diesel Engine Fuelled with Ternary (diesel-Biodiesel-Ethanol) Blends. *Energy* 174 (May): 386–406.

Venu, Harish, Lingesan Subramani, and V. Dhana Raju. 2019. Emission Reduction in a DI Diesel Engine Using Exhaust Gas Recirculation (EGR) of Palm Biodiesel Blended with TiO2 Nano Additives. *Renewable Energy* 140 (September): 245–263.

Vignesh, R., Ditto Sharmin, C. Vishnu Rekha, Sankar Annamalai, and Parisa Norouzi Baghkomeh. 2019. Management of Complicated Crown-Root Fracture by Extra-Oral Fragment Reattachment and Intentional Reimplantation with 2 Years Review. *Contemporary Clinical Dentistry* 10(2): 397–401.

Vijayakumar Jain, S., M.R. Muthusekhar, M.F. Baig, P. Senthilnathan, S. Loganathan, P.U. Abdul Wahab, M. Madhulakshmi, and Yogaen Vohra. 2019. Evaluation of Three-Dimensional Changes in Pharyngeal Airway Following Isolated Lefort One Osteotomy for the Correction of Vertical Maxillary Excess: A Prospective Study. *Journal of Maxillofacial and Oral Surgery* 18(1): 139–146.

Vijayashree Priyadharsini, Jayaseelan. 2019. In Silico Validation of the Non-Antibiotic Drugs Acetaminophen and Ibuprofen as Antibacterial Agents against Red Complex Pathogens. *Journal of Periodontology* 90(12): 1441–1448.

Xing, Wanli, and Dongping Du. 2019. Dropout Prediction in MOOCs: Using Deep Learning for Personalized Intervention. *Journal of Educational Computing Research*. https://doi.org/10.1177/0735633118757015