Original Article

Analyzing the Death Ratio of Covid Patients using Multiple Logistic Regression in Comparison with Lasso Regression for Improving Accuracy

B. Bharath Kumar Raju[1] , N. Deepa [2*]

[1]Research Scholar,Department of Computer Science and Engineering,Saveetha School of Engineering,Saveetha Institute of Medical and Technical Sciences,Saveetha University, Chennai, Tamilnadu, India. Pincode:602105.
[2]Project Guide, Department of Computer Science and Engineering,Saveetha School of Engineering,Saveetha Institute of Medical and Technical Sciences,Saveetha University, Chennai, Tamilnadu, India. Pincode:602105.

Abstract

Aim: The idea of this study is to analyze and improve the death ratio accuracy of covid patients with Novel Multiple Logistic Regression(MLR)and Lasso regression. Both these algorithms fall under supervised learning techniques. **Materials and Method:** Accuracy is analyzed for covid dataset of size 239 places. Analyzingthe death ratio of covid patients is performed by a Novel Multiple Logistic Regression of sample size (N=35) and Lasso regression of sample size (N=35), obtained using the G-power value 80%. These are Supervised learning algorithms. **Result:** Novel Multiple Logistic Regression accuracy is 96% which is comparatively higher than LAS with accuracy of 66%. The significance is determined as $p=0.029$ ($p<0.05$) for obtaining accuracy.
Conclusion: NovelMultiple Logistic Regression performs better in determining accuracy than Lasso Regression.

Keywords: Big Data, Supervised Learning, Death Ratio, Lasso Regression, Novel Multiple Logistic Regression, Machine Learning.

DOI: 10.47750/pnr.2022.13.S04.003

INTRODUCTION

Covid virus has caused drastic changes in the human economy. It was first created in China in Wuhan city which caused a number of deaths and cases in China. The virus spread rapidly from China to different regions which led to several deaths. Atmostly covid spread through every region and caused havoc in every city where it was spreaded. Examining the ratio of pandemic deaths of people suffering through the disease (Nupur 2021). Novel Multiple Logistic Regression is improved in statistical fields and dependent and independent variables are calculated through this model and it was borrowed by machine learning later. It is used to know the ratio of mortality and death that occurred in the pandemic (Noor et al. 2020). It is also used to estimate the widespread of disease throughout all the regions (Ruess, Winton, and Adams 2021).In data analytics it is used in predictive statistics to predict mortality (Jepma et al. 2021).It is used to predict the underlying cause of death in big data (Hassanzadeh, Ying Sha, and Wang 2017).In machine learning it is used to predict future mortality ratio. Various applications of both Lasso and Multiple Logistic Regression are used in performing variable selection to predict the accuracy of patients' health in hospitals.

In the Big data field there are many research articles published on this topic in IEEE and Science Direct. From IEEE Xplore digital library 41 journals are found 23,863 articles were from ScienceDirect 18,700 articles were from GoogleScholar 14,603 articles were from Springer. The most cited article which is published (Shang et al. 2020) has a citation of 62 times. Another article is revised from Science Direct and has a citation of 42 and discusses ML approaches in covid19 for survival analysis (Nemati, Ansary, and Nemati 2020). Another article has a citation of 11 and which is cited in Springer and it discusses how simple-to-use is used to predict deterioration and and survival of patients with covid19 (Zeng et al. 2021). One more article (Zhou et al. 2020) has 8 citations and is cited in Springer and discusses exploiting risks of covid19 patients joining in hospitals. One more article (Laguna-Goya et al. 2020) has a citation of 150. Another article has a citation of 64 (Ying et al. 2022). One more article has a citation of 24 and discusses the mortality ratio of patients (Tsuzuki et al. 2022)Our team has extensive knowledge and research experience that has translate into high quality publications(Bhansali et al. 2021; Jayanth et al. 2021; Sudhakar, Ravel, and Perumal 2021; Sathiyamoorthi et al. 2021; Deepanraj et al. 2021; Raju et al. 2021; Arun Prakash et al. 2020; Kamath et al. 2020; Shanmugam et al. 2021; Rajasekaran et al. 2020; Adhinarayanan et al. 2020; Rajesh et al. 2020; Aurtherson et al. 2021).

The drawback in the existing system is that it consists of less accuracy and it is only considered in a few regions but there is more than one region. In the proposed work accuracy is more than the existing work and the proposed system has more places than compared to the existing system. Also studied online courses about data analytics and read the base paper thoroughly with my guide and came to an end to improve accuracy than the proposed model. Hence the proposed method aims at comparing algorithms to know which algorithm was giving more accuracy than the Lasso Regression. The aim of this work is to show more accuracy than lasso regression while comparing the death ratio.

MATERIALS AND METHODS

The research work was carried out in a data analytics lab, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences where the laboratory facilitates higher configurations to perform better experimental results.Two groups are used for this study with an sample size of (N=35) (Smithson and Shou 2019; Liang et al. 2020). The computation is done with G-power with 80% with alpha value 0. 05 and beta value is 0. 95 with a confidence interval at 95%.

In sample preparation group 1 Multiple logistic regression which is a supervised learning technique and this algorithm is used to train the statistical model at the back end. It gives us good accuracy for large and small data sets. It has a resistance to overfitting. This model will help us to get accuracy. Death ratio of individuals is provided with this model. In the field of statistics Multiple Logistic Regression is a specification method that gives Multiple logistic Regression to multi-class problems. It is used to predict probability on independent variables. The independent variable may be binary or continuous. Multiple Logistic Regression's goal is to find equations that give output. Table 1 represents pseudocode Multiple Logistic Regression and it is also a machine learning technique.

In sample preparation group 2, Lasso regression is used for training this model with statistical analysis at the back end and it comes under supervised learning technique. Lasso regression is a regularization technique.It is used over regression methods for accurate precision.This model uses shrinkage. Lasso regression uses L1 regularization technique. It is used to perform various feature selections.It is also used to find accuracy. Linear Regression comes under Lasso Regression which uses shrinkage and it is a machine learning technique. LASSO stands for Least Absolute shrinkage and Selection Operator. The Lasso regression has a simple and sparse model. It is used to get predictor subsets which are used to reduce prediction error for quantitative response.

For this proposed model the Jupyter notebook is used as an implementation tool. I have implemented codes in that itself. Hardware configurations of the system which I worked on consists of 8GB RAM and ROM of 1TB HDD+256 SSD with a processor of 11th gen intel(R) core i5-1135G7 @2. 40 GHZ. There are two groups: group 1 consists of Multiple Logistic Regression and group 2 consists of Lasso Regression.

For the proposed system dataset was taken from kaggle. The dataset contains several places. It consists of different regions and has a data of total number of cases registered and total number of deaths occured and total number of vaccines available due to covid.

STATISTICAL ANALYSIS

IBM SPSS of version 26.0 is used as Statistical software to find standard deviation, mean, standard error mean, mean difference, sig and F value. Unnamed and State/UTs are the Dependent variables and Active Ratio, Death Ratio and Discharge Ratio are the independent variables. Independent T-Test analysis is carried out in this research work. It is also used in checking the widespread of disease in different regions of people suffering with the disease (Sookaromdee and Wiwanitkit 2020).

RESULTS

Table 1 represents the pseudocode of Multiple Logistic Regression. At first declare libraries which are necessary for the program.Then perform some basic operations. Then the data is splitted into two types: training set and testing set.then assign some values to the model and by performing different mathematical operations and functions required accuracy is met.

Table 2 represents the pseudocode of Lasso regression. At first libraries are declared for the model. Libraries play a key role for every model. The data is splitted into training and testing sets. And perform some basic operations and assign some values to the model and find the accuracy for the total number of deaths caused due to covid.

Table 3 represents mean, Standard Deviation and Standard error mean of Multiple logistic regression and Lasso regression. Accuracy of Multiple logistic regression is greater when compared with Lasso regression.

Table 4 represents the raw data which is used to calculate accuracy for both Multiple Logistic Regression and Lasso Regression.

Table 5 represents Standard error difference and significance of data.The accuracy of significance value for both Multiple logistic regression and Lasso regression is 0.029.

Figure 1 represents the comparison of accuracy with a significance level of Multiple logistic regression and Lasso regression.The accuracy of Multiple logistic regression is 96.10 whereas Lasso regression accuracy is 66.37. Therefore can conclude that Multiple Logistic regression has better accuracy and standard deviation when compared to Lasso regression.

DISCUSSION

Based on the results obtained in independent sample T-test analysis the significance value is determined as 0.029 ($p<0.05$) between the two groups for the selected dataset. LASSO Regression algorithm which has a mean accuracy of 66.37%. Multiple Logistic Regression which means accuracy is 96.10%. When compared with the algorithms Multiple logistic regression has a better accuracy than Lasso regression.

The proposed work has mentioned the analysis of health records of covid patients who all are suffering with the pandemic disease (Riley et al. 2019). Proposed work has discussed the comparison of clinical traits portrayed by the different patients who suffered with that disease (Shang et al. 2020). This Proposed work has discussed the validation of patients suffering with covid and predicting 30 days mortality and CCI (Elmoheen et al. 2021). The Proposed work discusses the score of risk to predict the mortality of patients who are suffering with covid-19 (Her et al. 2021). LASSO demonstrated high sensitivities (> 90%) and almost perfect NPV (99.7%) in predicting mortality, which is clinically important because identifying and detecting at-risk patients is more significant than reducing false positive prediction. When it comes to death ratio analysis, the accuracy of Multiple Logistic Regression is better when compared with other supervised learning algorithms. Multiple Logistic Regression accuracy depends on the inputs taken from training and testing data sets. In our study, accuracy appears to be better than Lasso Regression. However, the average error appears to be more in our proposed work which should be reduced.

The deficiency of medical instruments and deficiency of food and treatment and lack of cleanliness are the factors that affect the death ratio. The limitations that are faced while dealing with the death ratio is that the data may not be appropriate to perform the task for finding the death raio. Due to lots of data sometimes miscalculation can be done so that may face difficulties to find the death ratio. And these are the limitations which are found when dealing with the death ratio. The death ratio in the future is used to know the number of people who lived and the number of people who passed away in that year so that they can calculate the difference in the population ratio of each year. In this way the death ratio is used in the future.

CONCLUSION

Novel Multiple Logistic Regression and Linear Regression are both machine learning techniques which use averaging to improve the accuracy. The work shows the death ratio accuracy of people suffering in the pandemic through covid disease using Multiple Logistic Regression, Linear Regression, Lasso Regression,Logistic Regression and Bayesian Linear Regression. It is found that Multiple Logistic Regression gained more accurate results than Linear Regression, Logistic Regression, Lasso Regression, Bayesian Linear Regression. Hence, it is concluded that Multiple Logistic Regression provides more accuracy when compared with other algorithms.

DECLARATION

Conflict of interest
No conflict of interest in this manuscript.

Authors Contributions
Author BKN was involved in methodology, text analysis and writing the manuscript. Author ND was involved in review and editing, supervision and validation.

Acknowledgements
The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) 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. Cyclotron Technologies, Chennai
2. Saveetha University
3. Saveetha Institute of Medical and Technical Sciences
4. Saveetha School of Engineering

REFERENCES

1. Adhinarayanan, Rajesh, Aravindh Ramakrishnan, Gopal Kaliyaperumal, Melvinvíctor De Poures, Rajesh Kumar Babu, and Damodharan Dillikannan. 2020. "Comparative Analysis on the Effect of 1-Decanol and Di-N-Butyl Ether as Additive with diesel/LDPE Blends in Compression Ignition Engine." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, June, 1–18.
2. Arun Prakash, V. R., J. Francis Xavier, G. Ramesh, T. Maridurai, K. Siva Kumar, and R. Blessing Sam Raj. 2020. "Mechanical, Thermal and Fatigue Behaviour of Surface-Treated Novel Caryota Urens Fibre–reinforced Epoxy Composite." *Biomass Conversion and Biorefinery*, August. https://doi.org/10.1007/s13399-020-00938-0.
3. Aurtherson, P. Babu, Bhanu Teja Nalla, Karthikeyan Srinivasan, Kulmani Mehar, and Yuvarajan Devarajan. 2021. "Biofuel Production from Novel Prunus Domestica Kernel Oil: Process Optimization Technique." *Biomass Conversion and Biorefinery*, May. https://doi.org/10.1007/s13399-021-01551-5.
4. Bhansali, Karan J., Kamlesh R. Balinge, Subodh U. Raut, Shubham A. Deshmukh, M. Senthil Kumar, C. Ramesh Kumar, and Pundlik R. Bhagat. 2021. "Visible Light Assisted Sulfonic Acid-Functionalized Porphyrin Comprising Benzimidazolium Moiety for Photocatalytic Transesterification of Castor Oil." *Fuel* 304 (November): 121490.
5. Deepanraj, B., N. Senthilkumar, D. Mala, and A. Sathiamourthy. 2021. "Cashew Nut Shell Liquid as Alternate Fuel for CI Engine—optimization Approach for Performance Improvement." *Biomass Conversion and Biorefinery*, February. https://doi.org/10.1007/s13399-021-01312-4.
6. Elmoheen, Amr, Ibrahim Abdelhafez, Waleed Salem, Mohamed Bahgat, Ali Elkandow, Amina Tarig, Nauman Arshad, et al. 2021. "External Validation and Recalibration of the CURB-65 and PSI for Predicting 30-Day Mortality and Critical Care Intervention in Multiethnic Patients with COVID-19." *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases* 111 (August): 108–16.
7. Hassanzadeh, Hamid Reza, Ying Sha, and May D. Wang. 2017. "DeepDeath: Learning to Predict the Underlying Cause of Death with Big Data." *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2017 (July): 3373–76.
8. Her, Ae Young, Youngjune Bhak, Eun Jung Jun, Song Lin Yuan, Scot Garg, Semin Lee, Jong Bhak, and Eun Seok Shin. 2021. "A Clinical Risk Score to Predict In-Hospital Mortality from COVID-19 in South Korea." *Journal of Korean Medical Science* 36 (15): e108.
9. Jayanth, Bellappu Venkat, Melvin Victor Depoures, Gopal Kaliyaperumal, Damodharan Dillikannan, Dilipsingh Jawahar, Kumaran Palani, and Ganesha Prasad Meravanigee Shivappa. 2021. "A Comprehensive Study on the Effects of Multiple Injection Strategies and Exhaust Gas Recirculation on Diesel Engine Characteristics That Utilize Waste High Density Polyethylene Oil." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, June, 1–18.
10. Jepma, Patricia, Lotte Verweij, Arno Tijssen, Martijn W. Heymans, Isabelle Flierman, Corine H. M. Latour, Ron J. G. Peters, Wilma J. M. Scholte Op Reimer, Bianca M. Buurman, and Gerben Ter Riet. 2021. "The Performance of the Dutch Safety Management System Frailty Tool to Predict the Risk of Readmission or Mortality in Older Hospitalised Cardiac Patients." *BMC Geriatrics* 21 (1): 299.
11. Kamath, Manjunath, Subha Krishna Rao, Jaison, Sridhar, Kasthuri, Gopinath, Sivaperumal, and Shantanu Patil. 2020. "Melatonin Delivery from PCL Scaffold Enhances Glycosaminoglycans Deposition in Human Chondrocytes – Bioactive Scaffold Model for Cartilage Regeneration." *Process Biochemistry* 99 (December): 36–47.
12. Laguna-Goya, Rocio, Alberto Utrero-Rico, Paloma Talayero, Maria Lasa-Lazaro, Angel Ramirez-Fernandez, Laura Naranjo, Alejandro Segura-Tudela, et al. 2020. "IL-6-Based Mortality Risk Model for Hospitalized Patients with COVID-19." *The Journal of Allergy and Clinical Immunology* 146 (4): 799–807.e9.
13. Liang, Li-Lin, Ching-Hung Tseng, Hsiu J. Ho, and Chun-Ying Wu. 2020. "Covid-19 Mortality Is Negatively Associated with Test Number and Government Effectiveness." *Scientific Reports*. https://doi.org/10.1038/s41598-020-68862-x.
14. Nemati, Mohammadreza, Jamal Ansary, and Nazafarin Nemati. 2020. "Machine-Learning Approaches in COVID-19 Survival Analysis and Discharge-Time Likelihood Prediction Using Clinical Data." *Patterns (New York, N.Y.)* 1 (5): 100074.
15. Noor, Aziz Ullah, Farhana Maqbool, Zulfiqar Ahmed Bhatti, and Asmat Ullah Khan. 2020. "Epidemiology of CoViD-19 Pandemic: Recovery and Mortality Ratio around the Globe." *Pakistan Journal of Medical Sciences*. https://doi.org/10.12669/pjms.36.covid19-s4.2660.
16. Nupur, Priyanka. 2021. "Cities in Crisis: Examining the Pandemic through Urban Planning and State Capacity." *India Review*. https://doi.org/10.1080/14736489.2021.1895567.
17. Rajasekaran, S., D. Damodharan, K. Gopal, B. Rajesh Kumar, and Melvin Victor De Poures. 2020. "Collective Influence of 1-Decanol Addition, Injection Pressure and EGR on Diesel Engine Characteristics Fueled with diesel/LDPE Oil Blends." *Fuel* 277 (October): 118166.
18. Rajesh, A., K. Gopal, De Poures Melvin Victor, B. Rajesh Kumar, A. P. Sathiyagnanam, and D. Damodharan. 2020. "Effect of Anisole Addition to Waste Cooking Oil Methyl Ester on Combustion, Emission and Performance Characteristics of a DI Diesel Engine without Any Modifications." *Fuel* 278 (October): 118315.
19. Raju, P., K. Raja, K. Lingadurai, T. Maridurai, and S. C. Prasanna. 2021. "Glass/Caryota Urens Hybridized Fibre-Reinforced nanoclay/SiC Toughened Epoxy Hybrid Composite: Mechanical, Drop Load Impact, Hydrophobicity and Fatigue Behaviour." *Biomass Conversion and Biorefinery*, March. https://doi.org/10.1007/s13399-021-01427-8.
20. Riley, Richard D., Danielle van der Windt, Peter Croft, and Karel G. M. Moons. 2019. *Prognosis Research in Healthcare: Concepts, Methods, and Impact*. Oxford University Press.
21. Ruess, Roger W., Loretta M. Winton, and Gerard C. Adams. 2021. "Correction: Widespread Mortality of Trembling Aspen (Populus Tremuloides) throughout Interior Alaskan Boreal Forests Resulting from a Novel Canker Disease." *PLOS ONE*. https://doi.org/10.1371/journal.pone.0253996.
22. Sathiyamoorthi, Ramalingam, Gomathinayakam Sankaranarayanan, Dinesh Babu Munuswamy, and Yuvarajan Devarajan. 2021. "Experimental Study of Spray Analysis for Palmarosa Biodiesel-diesel Blends in a Constant Volume Chamber." *Environmental Progress & Sustainable Energy* 40 (6). https://doi.org/10.1002/ep.13696.Shang, Yufeng, Tao Liu, Yongchang Wei, Jingfeng Li, Liang Shao, Minghui Liu, Yongxi Zhang, et al. 2020. "Scoring Systems for Predicting Mortality for Severe Patients with COVID-19." *EClinicalMedicine* 24 (July): 100426.Shanmugam, Rajasekaran, Damodharan Dillikannan, Gopal Kaliyaperumal, Melvin Victor De Poures, and Rajesh Kumar Babu. 2021. "A Comprehensive Study on the Effects of 1-Decanol, Compression Ratio and Exhaust Gas Recirculation on Diesel Engine Characteristics Powered with Low Density Polyethylene Oil." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 43

(23): 3064–81.
23. Smithson, Michael, and Yiyun Shou. 2019. *Generalized Linear Models for Bounded and Limited Quantitative Variables*. SAGE Publications.
24. Sookaromdee, Pathum, and Viroj Wiwanitkit. 2020. "COVID-19: Ratio of Occurrence in General People and Medical Personnel, an Observation from the Second Country That Disease Emerged." *International Journal of Preventive Medicine*. https://doi.org/10.4103/ijpvm.ijpvm_66_20.
25. Sudhakar, M. P., Merlyn Ravel, and K. Perumal. 2021. "Pretreatment and Process Optimization of Bioethanol Production from Spent Biomass of Ganoderma Lucidum Using Saccharomyces Cerevisiae." *Fuel* 306 (December): 121680.
26. Tsuzuki, Shinya, Kayoko Hayakawa, Yukari Uemura, Tomohiro Shinozaki, Nobuaki Matsunaga, Mari Terada, Setsuko Suzuki, et al. 2022. "Efficacy of Remdesivir in Hospitalized Nonsevere COVID-19 Patients in Japan: A Large Observational Study Using the COVID-19 Registry Japan." *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases*, February. https://doi.org/10.1016/j.ijid.2022.02.039.
27. Ying, Piaopiao, Jiajing Chen, Yinchai Ye, Jianzhong Ye, and Weiyang Cai. 2022. "Adipose Tissue Is a Predictor of 30-Days Mortality in Patients with Bloodstream Infection Caused by Carbapenem-Resistant Klebsiella Pneumoniae." *BMC Infectious Diseases* 22 (1): 173.
28. Zeng, Zhiyong, Chaohui Wu, Zhenlv Lin, Yong Ye, Shaodan Feng, Yingying Fang, Yanmei Huang, et al. 2021. "Development and Validation of a Simple-to-Use Nomogram to Predict the Deterioration and Survival of Patients with COVID-19." *BMC Infectious Diseases* 21 (1): 356.
29. Zhou, Yiwu, Yanqi He, Huan Yang, He Yu, Ting Wang, Zhu Chen, Rong Yao, and Zongan Liang. 2020. "Exploiting an Early Warning Nomogram for Predicting the Risk of ICU Admission in Patients with COVID-19: A Multi-Center Study in China." *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine* 28 (1): 106.

TABLES AND FIGURES

Table 1. Pseudo code for Multiple Logistic Regression algorithm. Multiple Logistic Regression is a classification method that generalizes logistic regression to multi-class problems. Multiple Logistic Regression is used to predict the probability of category membership on a dependent variable based on independent values.

Input: Covid Dataset
Output: Accuracy
Step 1: Import and read the dataset.
Step 2: Select some features from the dataset.
Step 3: Generate the parameter.
Step 4: Analyze the dataset by changing the dependent and independent variables.
Step 5: Predict the output in numerical variables.
Step 6: Predict the output by using functions.

Table 2. Pseudo code for LASSO Regression algorithm. Lasso regression uses L1 regularization technique. It is used to perform various feature selections.It is also used to find accuracy. Lasso regression is a type of linear regression that uses shrinkage. LASSO stands for Least Absolute shrinkage and Selection Operator. The Lasso regression has a simple and sparse model.

Input: Covid Dataset

Output: Accuracy

```
function FastLasso (ipy, ipx, λ, N)
stop_thr
p-length(ipy)
beta-0 with length p
gc 0 with length p
do
dif Betamax <-- 0
for j = 1 p do
z (ipyl)-gc)/N + beta]
beta_tmp-max(0, z-A)-max (0,-z-λ)
dif Beta beta_tmp - beta[j]
 difabs abs(dif Beta)
 if difabs > 0 then
betalj]beta_tmp
gcgc + ipx[j] x dif Beta
dif Betamax = max(dif Betamax, difabs)
end if
end for
while dif Betamax 2 stop_thr
end do-while
return beta
end function
```

Table 3. Independent sample test calculation is done among Multiple logistic regression and Lasso regression. Accuracy of mean for Multiple logistic regression is 96.10 whereas Lasso regression mean accuracy is 66.37 whereas standard deviation for both Multiple logistic regression and Lasso regression are 1.834 and 2.032.

	Algorithm	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**

Accuracy	MLR	35	96. 10	1.834	.310
	LAS	35	66.37	2.032	.343

Table 4. MultipleLogistic Regression andLasso Regression raw data is taken from SPSS tools.

Group_Id	MLR	LAS
1	92	63
2	98	68
3	94	66
4	95	65
5	93	63
6	90	64
7	98	67
8	94	66
9	97	65
10	96	68
11	97	66
12	95	69
13	98	63
14	96	64
15	97	63
16	96	65
17	97	63
18	98	67
19	95	65

20	97	64
21	95	69
22	95	64
23	96	68
24	95	63
25	97	64
26	94	65
27	95	66
28	94	63
29	98	65
30	98	67
31	96	69
32	95	66
33	94	64
34	95	63
35	96	69

Table 5. The independent sample T-test performed between both Multiple logistic regression and Lasso regression has a confidence interval of 95%. Value of significance is determined as 0.029 ($p<0.05$) for obtaining accuracy.

		F	Sig.	T	df	Sig((2-trailed)	Mean diff	Std. Error diff	Lower	Upper
Accuracy	Equal variances assumed	1.128		65.26	68	<. 001	30.200	.463	29.277	31.123
	Equal variances not assumed		.029	65.26	67.29	<. 001	30.200	.463	29.277	31.124

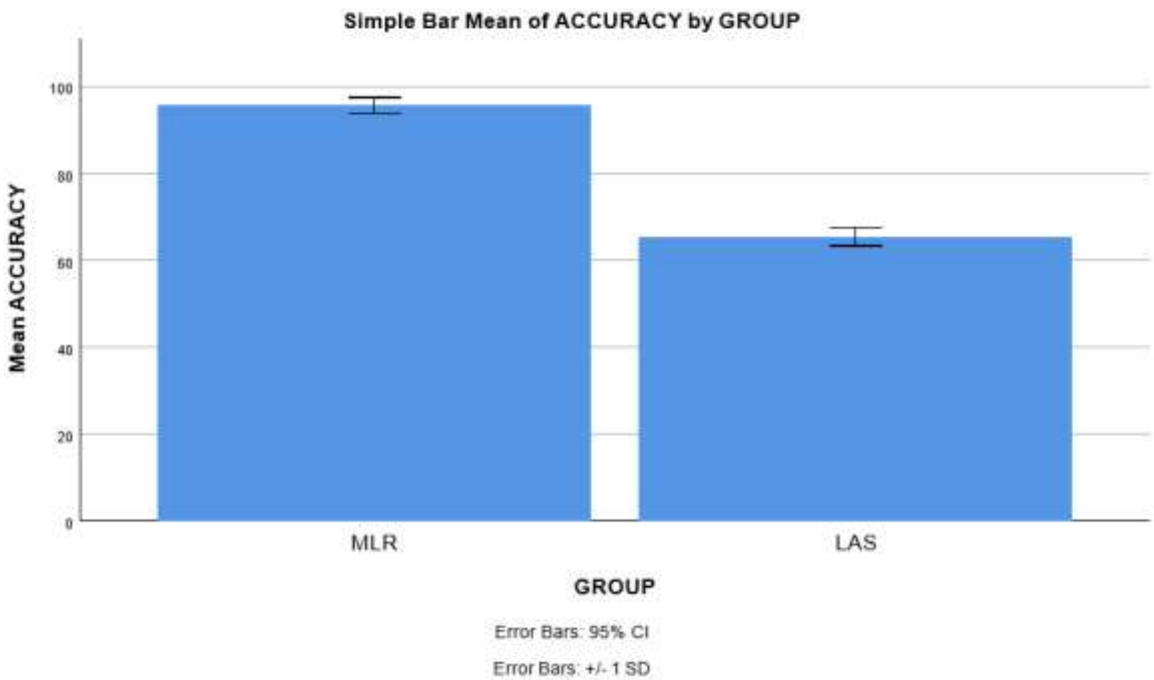

Fig. 1. Comparison of Algorithm and Accuracy for Simple Bar Graph is calculated . Mean accuracy of MLR is better than LAS and standard deviation of MLR is slightly better than LAS. X axis: MLR vs LAS Y axis: Mean accuracy of detection ± 1 SD.