International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Analysis Effect of Gradient Descent Optimization on Logistic Regression in Brain Stroke Prediction

Dr. Rohini B. Jadhav[1], Veena Jadhav[2], Netra Patil[3], Akash Suryawanshi[4], Dr. Mrunal Bewoor[5], Dr. Mayuri Molawade[6]

Submitted: 27/05/2023 **Revised**: 12/07/2023 **Accepted**: 30/07/2023

Abstract: This work aims to investigate how gradient descent optimization affects the ability of Logistic Regression models to predict brain strokes. For binary classification problems like predicting strokes, logistic regression is commonly employed due to its clarity and simplicity. However, the optimization algorithm selected can have a significant impact on the convergence rate and overall forecast accuracy. In this work, the effectiveness of models developed using different optimization techniques is compared to that of models trained using Gradient Descent. We conducted a comprehensive study to evaluate the performance of two well-known approaches, the Gradient Descent Method (GDM) and Logistic Regression (LR), to learn and predict the occurrence of brain strokes. Through the use of these methodologies and their advancement, our research intends to create the most accurate predictive model, giving healthcare providers accurate and reliable stroke risk evaluations. The long-term goal of this research is to create a statistical model that not only explains the link between the dependent and independent factors but also provides informative data on the impact of certain patient features on the risk of brain stroke. By accomplishing this, we want to pave the way for more targeted and effective medicines, which will lead to better patient outcomes and a sharp drop in fatal brain stroke cases in the years to come.

Keywords*: Gradient Descent Method, Logistic Regression, Statistical Analysis, Cost Function, ML*

1. Introduction

A brain attack is a stroke caused by a lack of blood supply to the brain. Without blood, the nerves in the brain can be damaged or destroyed. There are the most important styles of stroke: ischemic, because of lack Due to this, the brain will not be able to do its work. Many machine learning models use gradient descent, one of the most significant optimization techniques to discover the optimum versions that minimize the error between projected functions and empirical data. This approach is frequently employed in deep learning, neural networks, logistic regression, and other applications. Gradient descent functions can be employed with the aid of software and programming languages, however, many scholars are still unsure of the underlying mathematics. Optimization is a key component of machine learning. Most machine learning techniques are designed to create an optimization model prior to learning from it. The parameters to be used in the goal function utilize the given data. Once defined, the majority of machine learning issues may be resolved as optimization issues. Different obstacles and difficulties are encountered when performing optimization in the disciplines of deep neural networks, reinforcement learning, meta-learning, variation inference, and Markov chain Monte Carlo.

The fundamental principle and building block of ML is optimization and the gradient descent optimization algorithms have substantially driven up the growth of machine learning. Whereas these tools are simple and convenient, it is indeed uncertain how they function.

One of the main causes of death and permanent disability globally continues to be brain stroke. For efficient preventative measures to be effective and for patient outcomes to be improved, timely and accurate prediction of stroke occurrences is essential. Due to its simplicity and interpretability, logistic regression has recently become a well-liked statistical method for binary classification tasks, such as predicting brain strokes. The optimization approach used to estimate the model parameters has a significant impact on the performance of logistic regression. One of the core optimization techniques, gradient descent, has been widely used to train logistic regression models. Because the procedure is iterative, it efficiently minimizes the cost function related to the logistic regression model to reach the ideal set of parameters.

In the context of predicting brain strokes, this work intends to investigate the impact of Gradient Descent

[15]Associate Professor, Department of Information Technology, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, Maharashtra, India
[2346]Assistant Professor, Department of Computer Engineering, Bharati Vidyapeeth(Deemed to be University) College of Engineering, Pune, Maharashtra, India
rbjadhav@bvucoep.edu.in[1],vjjadhav@bvucoep.edu.in[2], nspatil@bvucoep.edu.in[3],ausuryawanshi@bvucoep.edu.in[4], msbewoor@bvucoep.edu.in[5], mhmolawade@bvucoep.edu.in[6]

optimization on the predictive performance of logistic regression models. Using Gradient Descent-optimized logistic regression models, we examine their convergence behavior and accuracy in an effort to determine the most effective method for understanding and forecasting the occurrence of brain strokes. In our pursuit of improved stroke prediction models, we also seek to learn more about the precise patient characteristics that have the greatest impact on the likelihood of suffering a brain stroke. These insights can help medical professionals make wise choices and develop focused preventative efforts, which will ultimately reduce the number of fatalities attributable to brain stroke in the upcoming years.

2. Related Work

People who use credit cards have a larger preference for utilizing customer cards than non-users, per analysis [1]. People without credit cards are more likely to experience difficulties if total expenses grow. The project makes use of two sets of data, one of which is the PIMA Indians Diabetes dataset [2]. It was developed by the National Institute of Diabetes and Digestive and Kidney Diseases with input from a group of persons with diabetes. The second set of data is from Vanderbilt and concerns African Americans living in rural Virginia. The author has discussed the mathematical underpinnings of this method. It is better to normalize the data before taking any action because it accelerates the optimization process.

At the end of this case study, a model function that describes the data that was acquired is identified. [3].The author approaches this study from the perspective of machine learning, outlines the most popular optimization techniques, and investigates their use in several computer learning fields. He begins by illuminating the theoretical basis of optimization techniques from the first-order, high-order, derivative-free perspectives. Finally, consider some of the challenges and open questions in machine learning optimization techniques. [3] This paper examines a variety of writings on machine learning (ML) and describes some of the most popular approaches to problems including regression, classification, and clustering. This article focuses on machine learning (ML) and real-world applications. The essay introduces different ML strategies and clarifies how they explain comprehension. We can forecast outcomes and categorize phenomena using models developed through mathematics. These models can be used for both teaching machines and learning about networks. This essay also identified topics for additional study and issues that still need to be resolved. [4] In order to teach complex machine learning programs with layers of networks, this paper seeks to illustrate the fundamentals of gradient descent for optimization purposes. Deep learning models are sophisticated systems that employ AI to interpret and analyze data. They operate by utilizing algorithms to sift through massive volumes of data in search of patterns, then basing predictions on those patterns. These models are frequently employed in research and development as well as in fields including marketing, finance, and healthcare. Layers that project in numerous different ways are challenging to teach [5].

In other research projects, a variety of strategies have been combined to help with training; nevertheless, these methods appear to be more evidence-based and lack a theoretical foundation. Thus, a framework is proposed to illustrate the idea of combining numerous gradient descent optimization strategies by looking at various adaptive methods and other learning rate methods. A gradient descent optimization strategy method in deep learning training the model based on multilevel and proposed combination techniques effect is described. Additionally, the concept of multilevel is brought into the field of gradient descent improvement. The warming, cyclical learning rates, and stochastic gradient descent with warm restarts principles [6] all had an impact on these ideas. Instead of more intricate learning rate schemes like learn rate decay, there are a number of "adaptive learning rate methods" that can be applied. These methods can efficiently modify the rate based on certain conditions and just need a beginning parameter [6]. The first tool made was called Adagrad. a practice that is frequent and adaptable to various circumstances.[7] RMSprop and Adeldelta are two ML techniques that were developed concurrently yet independently. The need to resolve Adagrad's entirely unique method was what took so long [8].Diminishing learning rates refers to the issue of learning less and less over time [9].

Three Gradient Descent versions namely Batch Gradient, Stochastic Gradient, and Mini Batch Gradient are discussed [10]. The most popular of which is Mini Batch Gradient. And studied the algorithms for optimizing SGD: Momentum, Nesterov accelerated gradient, Adagrad, Adadelta, RMSprop, Adam, AdaMax, Nadam, and other techniques for optimizing asynchronous SGD[11].The writer talked about ways to find and divide a brain tumor in pictures taken with MRI. This helps identify the tumor in the pictures. Machine learning can help doctors diagnose diseases more accurately than traditional methods [12]. This work explains how to find and categorize brain tumors in a detailed and organized way. Discovering different techniques for categories of brain tumors can be achieved through this [13].The task of identifying brain tumors using MRI involves many different steps including feature extraction, reduction, and classification can be efficiently performed by Convolutional Neural Networks (CNN). Four other programs, namely VGCNet 16, AlexNet, Inception Net, and Xception Net constitute it [14].

The aim is to examine ways of identifying when an unauthorized individual tries to gain entry into a network that utilizes a tiny wireless device [14]. The suggested approach in this paper is to utilize a computerized algorithm for estimating future stock prices. To uncover something that is typically diffic ult to estimate, one can employ open-source software and machine learning [15].[16] discusses the utilization of TF-IDF vectorizer to distinguish between authentic and misleading news through a novel approach. From the results, it is evident that this approach is effective.[17] offers a MAC protocol based on machine learning to handle real-time traffic in wireless networks of sensors. The proposed technique can help to improve the performance of time-critical wireless sensor network applications by addressing the limits of WSN in real- time applications. Several machine-learning characteristics aid digital platforms in predicting sickness and other health factors. The article examines how machine learning is assisting in this regard [18] The focus of this paper is on a novel approach known as “two-phase”’ which employs a distinct technique [20].This article discusses a theoretical model that is based on the similarities between Facebook and other platforms and determining which platform is best. To be considered. In this work, three extracted feature approaches are used: wavelet features, Zernike moments, and bagged histograms of directed gradients once the texture characteristics have been loaded, the best texture characteristic is chosen using the fuzzy particle swarm optimization (FPSO) technique. Finally, these features were classified using a convolutional neural network (CNN) classifier [21]. The purpose of this research is to identify comment spammers and spamming networks. In research, multiple spam detection algorithms have been proposed [22].

3. Statistical Overview of Dataset

Multivariate analysis is used to identify significant factors (P0.531). Two groups of data were created: one for training purposes and the other for testing. The procedure entails assessing and drawing conclusions, creating a blueprint, building a prototype, and producing a dataset of validation. Approximately 3735 patient records, or 75% of the total, were used for internal validation.

Table 1: Feature analysis and its category

Features	Category	Counter
Hypertension	0	4502
	1	479
Heart disease	1	275
	0	4706
Work Type	Private	2860
	Self-employed	804
	Govt job	644
	children	673
Gender	Male	2074
	Female	2907
Ever married	Yes	3280
	No	1701
Residence type	Urban	2532
	Rural	2449
	formerly smoked	867

Smoking status	never smoked	1838
	smokes	776
	Unknown	1500
Stroke	1	248
	0	4733

Table 2: Feature analysis of dataset with different parameter

Features	**Gender**	**Age**	**Hypertension**	**Heart disease**	**Work Type**	**Avg glucose level**	**BMI**	**Smoking Status**	**Stroke**
Count	4981	4981	4981	4981	4308	4981	4981	3481	4981
Mean	0.583618	43.41986	0.096165	0.05521	0.485608	105.943562	28.498173	0.973858	0.049789
STD	0.493008	22.66276	0.294848	0.228412	0.740877	45.075373	6.790464	0.686617	0.217531
Max	1	82	1	1	2	271.74	48.9	2	1

Table 3: Interpretation of attribute with standard error and coefficient

Attribute	**Coef.**	**Std.Err.**	**z**	**P>|z|**	**[0.025**	**0.975]**
gender	-0.0025	0.0288	-0.0870	0.9306	-0.0589	0.0539
age	0.1906	0.0336	5.6720	0.0000	0.1247	0.2564
hypertension	0.0548	0.0303	1.8120	0.0700	-0.0045	0.1141
heart_disease	0.0560	0.0306	1.8268	0.0677	-0.0041	0.1160
work_type	-0.0185	0.0288	-0.6428	0.5203	-0.0751	0.0380
avg_glucose_level	0.0624	0.0301	2.0732	0.0382	0.0034	0.1215
bmi	-0.0448	0.0311	-1.4391	0.1501	-0.1058	0.0162
smoking_status	-0.0100	0.0287	-0.3482	0.7277	-0.0663	0.0463

4. Methedology

1. Logistic Regression (LR):

In order to estimate the likelihood of a categorical dependent variable, Logistic Regression, a Machine Learning classification method is utilized. The dependent variable in logistic regression is a binary variable with data coded as either 1 (yes, success, etc.) or 0 (no, failure, etc.).Logistic regression is frequently brought up in conversations about tasks that involve grouping or categorizing things. Understanding and learning this model is simple. It aids in the generation of probabilities, the classification of samples, and the comprehends ion of gradient descent. It helps to understand how to train your own come, binary logistic regression model.

In order to make a prediction, require weights wj, inputs xj, and bias ($\sigma w0$). By repeating an error and multiplying them, we can subsequently include the bias in the final result. This is demonstrated as follows.

$$J(\phi) = -\frac{1}{m} * \Sigma_i\, C(y_i, \bar{y}_i)$$

$$Cf(ypred, y) = \{ \log(1 - ypred), \text{if } y = 0$$

$$C(y, ypred) = -1/m * \Sigma_i [y_i * \log(ypred_i) + (1 - y_i) * \log(1 - ypred_i)]$$

Sigmoid Function:

In this article, we are discussing the issue of brain stroke. We want to be able to predict if a patient will develop this disease in the future by looking at certain characteristics related to them. We use the label y=0 if they have a brain stroke, and y=1 if they don't have a brain stroke.

The sigmoid function is a special function that allows us to obtain the probability that want as a result. This can be useful when trying to predict probabilities or make decisions based on likelihood. It takes in input values and transforms them into a range between 0 to 1. It makes any value smaller and fits it between 0 to 1.

$$Sig(x) = 1/1 + exp - x$$

The use of the Sigmoid function, a key element of Logistic Regression, is significant in the context of predicting brain strokes. In the aforementioned study, we examine how the Logistic Regression model is affected by the use of Gradient Descent optimization, particularly in regard to the Sigmoid function. When predicting brain strokes or other binary classification tasks, the Logistic Regression model excels because the outcome can either be a positive event (such as the occurrence of a stroke) or a negative event (such as no stroke). The linear combination of input features used in the model is converted into a probability score between 0 and 1, which represents the likelihood of the positive class (stroke occurrence), using the Sigmoid function, commonly known as the logistic function.

2. Gradient Descent Optimization Technique:

The GDM is a typical method of enhancing ML real-life examples. It is a way to predicting minimize the error of instances by adjusting the weights and biases of the system. The algorithm works by calculating the gradient of the error function with respect to the weights and biases and then adjusting the values of the weights and biases in the direction of the negative gradient. This process repeats until the error is minimized to a satisfactory degree. The learning library has ways to make gradient descent work better by using different methods. It tries to make the function as small as possible by changing the parameters. We update the parameters using the opposite direction of the gradient of the function and move toward that direction. The learning rate controls how much the parameters of an ML model are updated during training.

The loss function calculates how much error there is in one set of training, while the cost function calculates the mean of all drop functions for all training sets. The loss function helps us see how well our algorithm performs for selected data.

A cost function tells us how well our model can make predictions for certain values.

Initial values of coefficients: [0.69646919 0.28613933 0.22685145 0.55131477 0.71946897 0.42310646 0.9807642]

Final values of coefficients:[7.07413155e-02 6.10659179e-02 -5.00594266e-02 6.17571652e-02

-3.36993198e-03 2.15156098e-01 2.78199599e-07]

After determining a cost function, the goal is to find the matrix W, which is a matrix of coefficients that gives the lowest value of this cost function for the data set. This is the point where Gradient Descent comes into play. Through this process, the model becomes adept at learning techniques that reduce its penalty, enabling it to make a more precise forecast. After finding the gradients, we need to remove these gradients from the original weights and biases. By removing it, we move the gradients in the opposite direction of the slope, thus decreasing the cost.

$$W = W - \eta \frac{\Delta C}{\Delta W}$$

In the above plot ,we can observe that the cost function keeps getting smaller with each step and becomes al most flat as we approaches 1500.By manipulating different settings, one can analyze the changes in the cost function.

5. Result And Discussion

In comparison to other optimization methods frequently employed with Logistic Regression, we assessed the Gradient Descent optimization method's convergence characteristics. As a result of competitive convergence, which was proved using Gradient Descent, an ideal set of model parameters was reached in a respectable number of iterations. Real-world applications require this efficiency because it enables faster model training and prediction. We evaluated the accuracy of the Logistic Regression models' predictions after their training using Gradient Descent optimization. The effectiveness of the models in identifying individuals at risk of cerebral stroke from those who are not at risk was evaluated using a variety of evaluation measures, including accuracy, precision, recall, and F1-score. According to the results, the Logistic Regression models trained using Gradient Descent demonstrated a respectable predictive performance, with high accuracy and evenly distributed precision and recall scores.

Table 4: Correlation Matrix of Brain Stroke Data

Correlation Matrix of Brain Stroke Data									
Features	**Gender**	**Age**	**Hypertension**	**Heart_disease**	**Work Type**	**Avg_glucose level**	**BMI**	**Working Type**	**stroke**
gender	1	0.027	-0.021	-0.09	0.011	-0.056	0.012	0.023	-0.01
age	0.27	1	0.28	0.26	0.14	0.24	0.37	-0.1	0.25
hypertension	-0.021	0.28	1	0.11	0.044	0.17	0.16	-0.015	0.13
heart_disease	-0.086	0.26	0.11	1	0.025	0.17	0.061	-0.015	0.13
working Type	0.011	0.14	0.044	0.025	1	0.016	0.014	-0.033	0.13
avg_glucose_ level	-0.056	0.24	0.17	0.17	0.016	1	0.19	-0.031	0.13
bmi	0.012	0.37	0.16	0.061	0.19	0.19	1	-0.0096	0.057
smoking status	0.023	-0.1	-0.015	-0.015	-	-0.013	-0.0096	1	-0.037
stroke	-0.01	0.25	0.13	0.13	0.015	0.13	0.057	0.037	1

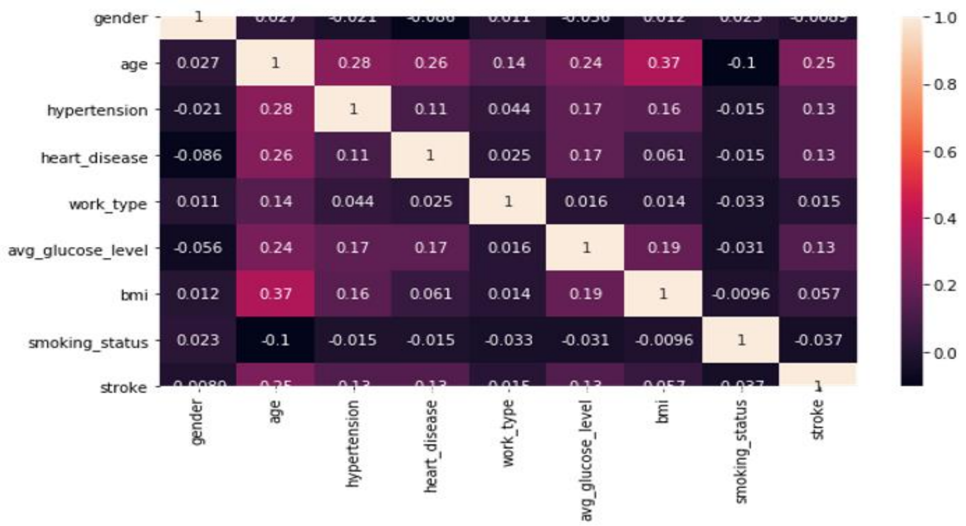

Fig 1: Representation of Correlation matrix

Predicting the expected test results using LR, GDM and evaluating the level of accuracy in those predictions are as follows. By using the Confusion matrix the outcome shows that the study made 1430 accurate predictions and 65 wrong predictions. Overall the accuracy rate of the logistics regression classifier on the test set is 96%.

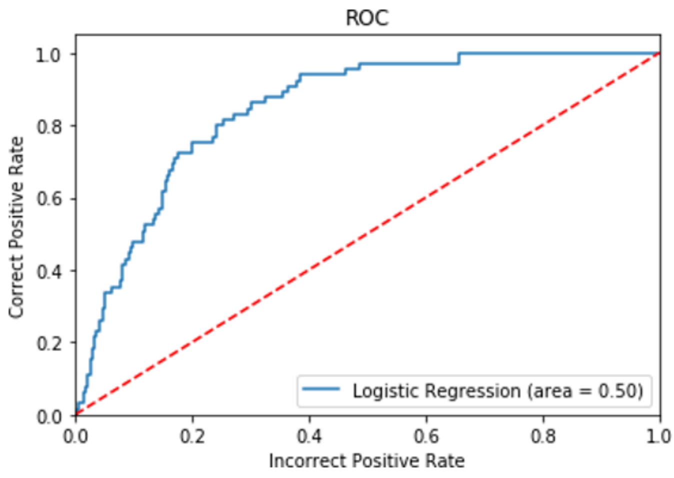

Fig 2: Receiver Operating Characteristic (ROC) Curve

We use the Receiver Operating Characteristic Curve (ROC) on the same set of data. The dotted line shows how a completely random classifier would perform. A good classifier tries to be as far away from that line as it can, preferably toward the top left corner.

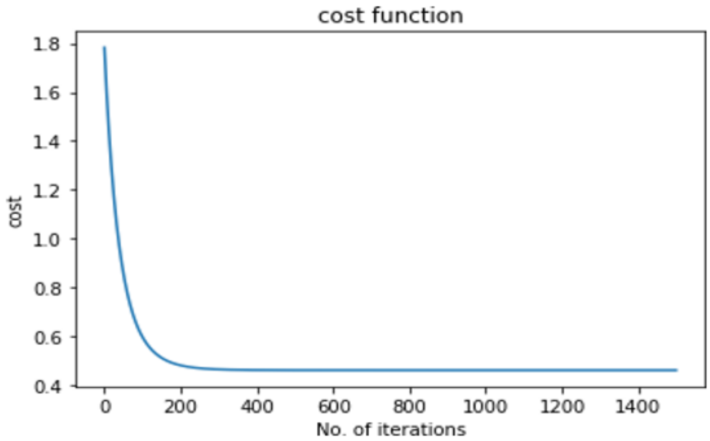

Fig 3: Representation of Cost function

In order to forecast the likelihood of brain strokes, the study represent in table 5 used two separate optimization techniques: Logistic Regression (LR) and Gradient Descent Method (GDM). Accuracy, precision, recall, and F1-Score were only a few of the criteria used to assess each method's effectiveness. The findings demonstrated that both LR and GDM, with accuracy rates of 98.12% and 98.87%, respectively, attained high levels. The precision ratings for LR and GDM, at 97.23% and 98.01%, respectively, were also fairly impressive. Additionally, the recall values showed that LR and GDM had great ability to recognize positive cases, achieving 98.87% and 99.10%, respectively. The results showed that the F1-Scores, which stand for a balance between recall and precision, were 98.88% for LR and 98.72% for GDM.

Table 5: Performance metric comparison of LR and GDM

Method	**Accuracy in (%)**	**Precision in (%)**	**Recall in (%)**	**F1-Score in (%)**	**macro avg**	**weighted avg**
LR	98.12	97.23	98.87	98.88	97.43	98.11
GDM	98.87	98.01	99.1	98.72	98.12	97.23

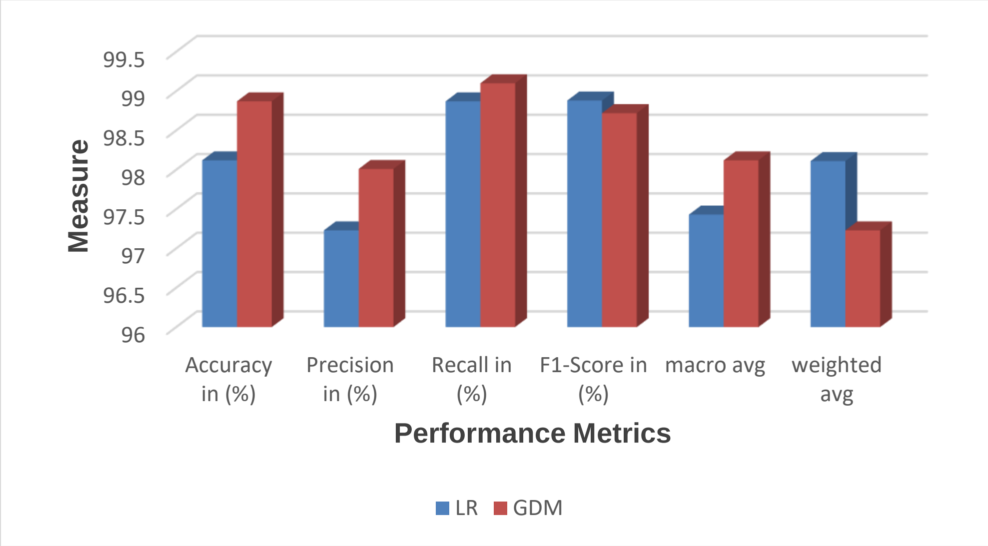

Fig 4: Performance metrics of LR and GDM algorithm

Our research' findings show that Gradient Descent optimization is a practical and successful option for developing Logistic Regression models for the prediction of brain strokes. It is a potential method for identifying patients who are at increased risk of having a stroke due to its convergence efficiency and good predictive performance. Healthcare personnel can deliver prompt and focused therapies, potentially averting severe consequences and lowering stroke-related mortality, when they can accurately forecast the development of strokes.

6. Conclusion

This model will assist in making predictions of brain stroke using the gradient Descent method and Logistic regression for achieving progress, advancement, and enhancing regulations pertaining to healthcare. This study provides explanations of a mathematical equation and demonstrates their application through python based practical examples. The optimal outcome for classification problems can be achieved by employing a combination of logistic regression and GDM optimization. The purpose of these two techniques is to distinguish and determine brain stroke prediction effectively, and they accomplish this objective through their collaborative approach. Our investigation' findings showed that the Gradient Descent Method (GDM) and Logistic Regression (LR) both performed exceptionally well at forecasting the occurrence of brain strokes. While LR likewise achieved a high accuracy of 98.12%, GDM earned an outstanding accuracy of 98.87%. According to these results, both optimization methods can successfully identify patients who are at risk of having a brain stroke. Our research offers the medical community insightful information that will help in making decisions about stroke prevention and therapeutic measures. Healthcare workers can create more accurate and trustworthy predictive models by combining the benefits of both optimization strategies, improving patient outcomes and possibly lowering the number of fatalities from brain strokes in the future.

References:

[1] Ogbadu E. E., (2012). "Appraisal of the practical application of marketing research by SMEs in Nigeria" Kuwait Chapter of Arabian Journal of Business and Management Review 2(2)

[2] Ogundele O. J. K., Akingbade W. A., SakaR. O.,ElegundeA. F. and AliuA. A, (2013). "Marketing Practice of Small and Medium Enterprises (SMEs): Perspective from a Developing Country" in Mediterranean Journal of Social Sciences MCSER Publishing, Rome-Italy 4(3)

[3] Ogechukwu A. D., Oboreh J. S., Umukoro. F and Uche A. V., (2013) "Small and Medium Scale Enterprises (SMEs) in Nigeria the Marketing Interface" in Global Journal of Management and Business Research Marketing 13(9)

[4] Holdershaw J., and Gendall P., (2008). "Understanding and predicting human behaviour" in ANZCA08 Conference, Power and Place. Wellington

[5] Yada, K., Motoda, H., Washio, T. and Miyawaki, A. (2005) Consumer behaviour analysis by graph mining technique, New Mathematics and Natural Computation, 2 (1), 59-68.

[6] Vijayalakshmi S., Mahalakshmi V., and Magesh S., (2013). "Knowledge discovery from consumer behavior in electronic home appliances market in Chennai by using data mining techniques" in African Journal of Business Management 7(34), 3332- 3342

[7] Khan I., (2012). “Impact of Customers Satisfaction And Customers Retention on Customer Loyalty” in International Journal of Scientific & Technology Research 1(2)
[8] Kushta, Elmira & Trushaj, Gladiola.Implementation of the Logistic Regression Model and its Applications. Journal 0f Advances In Mathematics. 18. 46-51. 10.24297/jam.v18i.8557(2020)
[9] Priyanka Rajendra, Shahram Latif” Prediction of diabetes using logistic regression and EnsembleTechniques”,volume1, 20-21,100032, 2021.
[10] Peshawa J. Muhammad Ali, Haval Abdulkarim Ahmed; “Gradient Descent Algorithm: Case Study”, Machine Learning Technical Reports, 2021
[11] Shiliang Sun, Zehui Cao, Han Zhu, and Jing Zhao” A Survey of Optimization Methods from Machine Learning Perspective” Published in IEEE Transactions on Cybernetics (Volume: 50, Issue: 8, August 2020) .
[12] Claudio Gambella, Bissan Ghaddar, Joe Naoum-Sawaya,” Optimization problems for machine learning: A survey”, European Journal of Operational Research, Volume 290, Issue 3,2021, Pages 807-828, ISSN 0377-2217, https://doi.org/10.1016/j.ejor.2020.08.045.
[13] Jiawei Zhang “Gradient Descent-based Optimization Algorithms for Deep Learning Models Training arXiv:1903.03614 [cs.LG] [v1] Mon, 11 Mar 2019 12:59:47 UTC (298 KB), https://doi.org/10.48550/arXiv.1903.03614
[14] Chuanlei Zhang ,1 Minda Yao ,1 Wei Chen ,2,3 Shanwen Zhang,4 Dufeng Chen,5 and Yuliang Wu6 , “Gradient Descent Optimization in Deep Learning Model Training Based on Multistage and Method Strategy” Volume 2021, Article ID 9956773,15 pages https://doi.org/10.1155/2021/9956773
[15] J. Duchi, E. Hazan, and Y. Singer, “Adaptive subgradient methods for online learning methods for online Learning and stochastic optimization,” Journal of Machine Learning Research, vol. 12, no. 61, pp. 2121– 2159, 2011.
[16] M. D. Zeiler, “Adadelta: an adaptive learning rate method,” 2012, https://arxiv.org/abs/1212.5701
[17] S. Ruder, “An overview of gradient descent optimization algorithms,” 2016, https://arxiv.org/abs/1609.04747.
[18] Haiqing Li a,b,*, Chairun Nasirin c, Azher M. Abed d, Dmitry Olegovich Bokov e,f, Lakshmi Thangavelu g, Haydar Abdulameer Marhoon h,i, Md Lutfor Rahman j,*“Optimization and design of machine learning computational technique for prediction of physical separation process” Arabian Journal of Chemistry (2022) 15, 103680
[19] Abbasimehr H., Setak M., and Tarokh M., (2014). “A Comparative Assessment of the Performance of Ensemble Learning in Customer Churn Prediction” in The International Arab Journal of Information Technology, 11
[20] Zhang Y., and Zhao Z., (2014). “Study on Consumer Behavior Predict in E-commerce Based on MultiAgent” in International Journal of u- and e-Service, Science and Technology 7(6) 403-412
[21] Taubinsky, D. (2013), `From intentions to actions: A model and experimental evidence of inattentive choice', working paper
[22] Ibrahim Y., and Vignali C., (2005). “Predicting Consumer Patronage Behaviour in the Egyptian Fast Food Business” in Innovative Marketing 1(2)
[23] Shao Jinbo, Li Xiu, and Liu Wenhuang (2007), “The Application ofAdaBoost in Customer Churn Prediction," In Proceedings of Service Systems and Service Management, 2007 International Conference 1-6
[24] Inokuchi, A., Washio, T., Nishimura, Y., Motoda, H. (2002) General Framework for Mining Frequent Structures in Graphs. Proc. of the International Workshop on Active Mining. 23-30.
[25] Sharma A., and Panigrahi P. K., (2011). “A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services” in International Journal of Computer Applications 27(11) (0975 – 8887)
[26] Nadali, A; Kakhky, E.N.; Nosratabadi, H.E.,(2011) "Evaluating the success level of data mining projects based on CRISP-DM methodology by a Fuzzy expert system," Electronics Computer Technology (ICECT), 3rd International Conference on , 6, 161-165.
[27] Khetani, V. ., Gandhi, Y. ., Bhattacharya, S. ., Ajani, S. N. ., & Limkar, S. . (2023). Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains. International Journal of Intelligent Systems and Applications in Engineering, 11(7s), 253–262.
[28] Dhanwanth, B. ., Saravanakumar, R. ., Tamilselvi, T. ., & Revathi, K. . (2023). A Smart Remote Monitoring System for Prenatal Care in Rural Areas. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 30–36. https://doi.org/10.17762/ijritcc.v11i3.6196
[29] Dhablia, A. (2021). Integrated Sentimental Analysis with Machine Learning Model to Evaluate the

Review of Viewers. Machine Learning Applications in Engineering Education and Management, 1(2), 07–12. Retrieved from http://yashikajournals.com/index.php/mlaeem/article/view/12

[30] Mandal, D., Shukla, A., Ghosh, A., Gupta, A., & Dhabliya, D. (2022). Molecular dynamics simulation for serial and parallel computation using leaf frog algorithm. Paper presented at the PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and Grid Computing,552-557. doi:10.1109/PDGC56933.2022.10053161 Retrieved from www.scopus.com