Optimization with Adaptive Learning: A Better Approach for Reducing SSE to Fit Accurate Linear Regression Model for Prediction

Dr Vijay Kumar Verma[1], Dr Umesh Banodha[2], Dr Kamlesh Malpani[3]
Shri Vaishnav Vidyapeeth Vishwavidyalaya Indore, Computer Science and Engineering, Indore, M.P. India[1]
Dr. A.P.J. Abdul Kalam U.I.T. Jhabua, M.P. India[2]
Shri Vaishnav Institute of Management and Science, Indore M.P. India[3]

***Abstract*—The Optimization provides a way through which an optimum can be achieved. It is all about designing accurate and optimal output for a given problems with using minimum available resources. It is a task which refers to minimizing an objective function f(x) parameterized by x or it is the task which refers minimizing the cost function using the model's parameters. In machine learning optimization is slightly different. Usually most of the problems, are very much aware about shape, size and type of data. Such information helps us to know where need improve. In case of machine learning optimization works perfectly when there is no knowledge about new data. The method proposed in this paper is named as Optimization with adaptive learning which is used to minimize the cost in term of number of iterations for linear regression to fit the correct line for given dataset to reduce residual error. In regression analysis a curve or line fit in such a way for the data objects, that the differences of distances between the data points and curve or line is always minimum. Proposed approach Initialize random values for parameters of linear model and calculate Error (SSE). Our objective is minimizing the values of SSE, if SSE is large, need to adjust the selected initial values. The size of the step used in each iteration is direction movement to reach the local minimum for optimal value. After performing certain repetitions of the deviation, minimum value for SSE has found and it has a stable value with no change. Real life data set have been used for expositional analysis.**

***Keywords*—Adaptive learning; regression; optimization; minimum; cost; objective; error; random**

I. Introduction

Optimization is a mathematical technique used to solve quantitative problems in a number of discipline like physics, engineering, computer science etc. Several problems be formulated and solved by combining ideas and methods with the field of optimization. In computer Science mathematical programming contains the study of the mathematical optimization and the study of the mathematical properties of these approaches [21,18]. The advances development in the optimization techniques in computer science are used to solve number of problems in operations research, game theory, and numerical analysis. The problem used for optimization have three basic components shown in Fig. 1, objective function, variables, and set of constraints. The objectives are expected as a return of costs or profits. Variables are the quantities values and can be manipulated so that to objective function can be optimized. Set of constraints for an optimization problem are restricted for values which can choose for the variables or parameters [20]. Research problem that found is how to achieve optimization is linear regression. Adaptive learning provides a way to archives optimization by using random initialization of the parameters [13,15]. Objective of the proposed work is to achieve optimization and reduce SSE for accurate prediction. The proposed work not only helps to students but as help to training institution of JEE examination, to prediction the marks and possibility for selection.

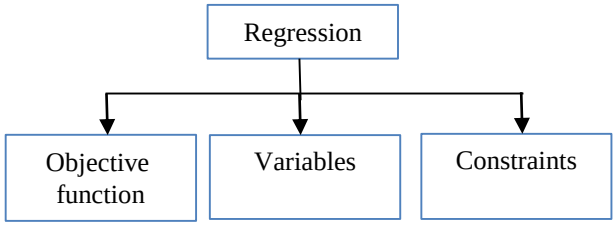

Fig. 1. Components used in regression analysis.

Regression analysis gives us number of benefits like

1) It shows important relations between variables.

2) It gives the effect of multiple explanatory variables.

II. Cost or Objective Function

The objective or cost function for analyzing regression is to find best fitted line for given inputs of x value and correct output for y value. Cost or objective function maps one or more variables on an actual instinctively representation related to predictor. ML models predict according to the new inputs [14,17].

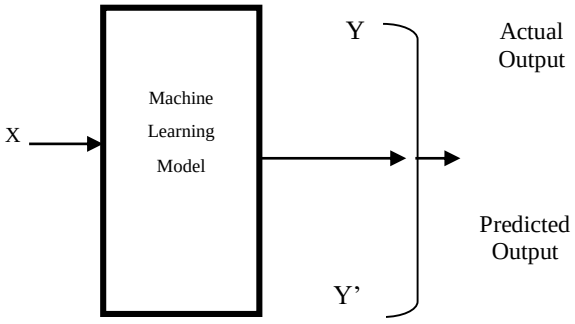

Fig. 2. Machine learning model.

The Error could be calculated by finding the difference between Predicted and Actual value:

Error = Y' (ML model Predicted value) - Y(Actual value)

A. Minimizing Error or Cost

Initial task of every ML model is to reduce cost. Minimizing value of cost functions will result to reduce error between the predicted value by ML model and actual value as shown in Fig. 2. Generally objective or cost function for linear regression is expressed in form of $\mathbf{Y = X^2}$. To minimize Objective or cost function, must find X to produce the correct Y value. The most common function required to minimize the parameters over a dataset is sum of squared error (SSE) and mean squared error (MSE). These errors ensure difference between the estimated value (prediction) and the estimator (the dataset) [16,19].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2$$

$$SSE = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

The above equation can adjust a little to make the calculation down the track a little simpler. Squared differences have used instead using absolute, this helps to derive a regression line [22]. To discover correct line, used first deviation for cost or objective function, and it is difficult to calculate deviation for absolute values as compared to squared values.

III. Problem with Existing System

Objective of regression analysis is to discover types of relation between response and predictor to get correct value for unknown dataset. It is difficult to find most appropriate regression line that attempts to predict and best fit line for given dataset for two points [16,19]. Number of methods have already exist, but the problems is how can optimize the solution Some time it is difficult to select following variables [20,23].

- Response Variable
- Predictor Variable(s)
- Stop
- Coefficients

1) What are the techniques used that the selected model better trained or learn the past relationship?

2) Predicted error should be minimized.

3) How Hyperplane is decided that minimizes the vertical offsets?

In machine learning terminology, optimization is the task of minimizing the cost or values of objective function, function by the model's parameters has following problems:

- Criteria are not fixed for deciding the initial value for parameter.
- Find the lowest possible value of the SSE to minimize the error.

IV. Objectives

For the proposed approach our objectives are:

1) Use random initialization of the pampers, to minimize error, to get optimal output for problem.

2) Continuously apply different updated values for the coefficients, estimate and select new coefficients that have a marginally better value with reduced error.

3) The objective is to deal with real life problems and dataset.

4) Use first order derivation on data for training and check performance and validation new data.

V. Review of Related Works

Several authors have been proposed their work to improve accuracy of prediction for regression analysis and published their work in various journals. In proposed work several related papers have been studied understand the concepts and basic things. Erasmus et al. of proposed method for establishing empirical relationships using linear regression. They describe that assumptions are comparatively poor, the LR model describing a restricted expectation. Specification for model must selected carefully and it is crucial when estimates for the coefficients of interest [1]. Shen Rong et al proposed regression model used in ML. They collected forecast temperature data for one year and the sale of iced products. They created a mathematical regression model which based on theory of data mining. The proposed Linear regression model is based on the practical situation. They implemented proposed model with the latest and most popular Python3.6 [2]. Katarina Valaskova et al proposed Financial Risk Measurement and Prediction Modelling. They used financial data of an enterprises and calculate important financial ratios which is responsible for health of the company. They used important predictors for forthcoming success. Multiple regression analysis where also used and identified the important predictors [3] Darman et al proposed Grade prediction with MLR for Mathematics. They predict the students' score with used of MLR for Final Exam. Score of students used as response variable in the model and the predictor variables are the assessment for tests. They a regression model with the adding SPSS tool [4]. Gaurav Pandeya et al. proposed RM model COVID-19 for. They have collected data for a particular time periods for year 2020. They evaluated performance using 1.75 for the regression model. This was helpful Government and doctors for preparation of their plans for the next two weeks [5]. K. K. Basee et al. proposed analyzing for various Regression Models. They used linear regression, multivariate and nonlinear regression models. They developed exponential, logistic type of regression. They used mat lab software and wrote a code without using pre-defined function [6]. Ira Sharma et al. proposed Linear Regression Model for factors Related with Carbon Stock. The objective is to evaluate the factors related with carbon stock. Data have been taken from department of Forest Research and Survey (DFRS). Linear regression showed a good fit of the model [7]. Samit Ghosal et al proposed Linear Regression to predict the number of deaths due to COVID-19 in India. They trace a trend related to death counts. They interchangeably applied Multiple and

linear regression analyses. They employed auto-regression technique to improve the prediction [8]. Yujiang et al. proposed An Adaptive Learning for Regression. They clarified problem and proposed a new methodological framework to forecast targets. The framework contains neural networks, to store newly collected data [9]. M Tanveer et al. proposed LR model for evaluating models to estimate stature. They applied LR and R2 values for preferred model. 100 persona data were used for experimental [10]. Dengyuhan Dai proposed Factors Affecting the Linear Regression Model for Data Analysis. They discussed detail effects of numerous parameters on the linear fit of randomly generated data sets. They include noise, independent variables, and different sample size. They minimized the effectiveness of noise. The optimal choice of each parameter can be obtained by comparing the goodness of fit [11]. Hai-Tao Jin, Fei Wang proposed Linear Regression Analysis of Sleep Quality. Data collected from various demographic area of the participants and psychological scales. They used sample *t*-test one-way ANOVA and showed comparisons [12].

VI. Proposed Approach

Step 1: To achieve rapid optimization, apply adaptive technique, which help to getting faster results with minimum iteration and less efforts. Initialize with random small value to the parameter. Let say parameters are a(c) and b(m) were

$$(Y_{pred} = m\ X+c)$$

a is constant and b slop parameter with random values and calculate error (SSE)

Step 2: For large SSE, apply first derivation and calculate new value for parameter a and b it has very small change from their original initialized value.

$$SSE = \frac{1}{2}(Y - YP)^2$$

This provides direction for movement for what value of a and

b with SSE should minimize.

Step 3: After calculating new value for parameter apply to reach the optimal solution and again calculate new SSE**.**

$$\partial Error/\partial c = -(Y_{actual} - Y_{predicted})$$

$$\partial Error/\partial m = -(Y_{actual} - Y_{predicted}) * X$$

Here value updating in SSE by

$$SSE = \frac{1}{2}(Y_{actual} - Y_{predicted})^2 = \frac{1}{2}(Y_{actual} - (m * X + c))^2$$

Steps 4: During training after specific iterations and make sure error rate must be decreases.

Step 5: Repeat steps 2 and 3 for continuously adjustments of parameters until doesn't significantly reduce the SSE.

Step 6: Sometimes can apply 2nd order derivative which are extremely fast and accurate.

VII. Illustrate with Example using Demo Dataset

TABLE I. Demo Dataset with 10 Records

S. No	No of Hours Study 15 days (120 hours)	Predicted Score (100)
1	10	12
2	23	22
3	25	58
4	34	20
5	36	55
6	45	39
7	44	54
8	58	53
9	94	99
10	101	61

Table I shows 10 sample records. Decimal scaling is used to make calculation easy. Table II shows result after applying decimal scaling.

TABLE II. Result of Decimal Scaling (normalization)

S.No	No of Hours Study per Month (210)	Predicted Score (100)
1	0.10	0.12
2	0.23	0.22
3	0.25	0.58
4	0.34	0.20
5	0.36	0.55
6	0.45	0.39
7	0.44	0.54
8	0.58	0.53
9	0.94	1.00
10	1.01	0.61

Step 1: To better fitted a line

$$Y_{pred} = mb\ X+c,$$

Start with random values for parameter c and m and compute forecast error (SSE). Let first small random values for parameter c and m

$$c=0.46 \text{ and } m=0.74$$

Step 2: Calculate the Error w.r.t the parameters

$$\partial Error/\partial c = -(Y_{actual} - Y_{predicted})$$

$$\partial Error/\partial m = -(Y_{actual} - Y_{predicted}) * X$$

Step 3: Now update value in SSE by

$$SSE = \frac{1}{2}(Y_{actual} - Y_{predicted})^2 = \frac{1}{2}(Y_{actual} - (m * X + c))^2$$

$\partial Error/\partial c$ *and* $\partial Error/\partial m$ give the direction of the movement of c, m w.r.t to Error. Table III shows the result of first derivation and total error.

TABLE III. CALCULATING SSE WITH RANDOM VALUE

Error	$\partial Error/\partial c$	$\partial Error/\partial m$
0.102	0.45	0.00
0.076	0.39	0.09
0.001	0.05	0.01
0.126	0.50	0.17
0.017	0.18	0.07
0.077	0.39	0.18
0.031	0.24	0.11
0.063	0.35	0.20
0.09	0.14	0.13
0.175	0.59	0.59
Total (Error)=0.668	3.210	1.455

TABLE IV. CALCULATING ERROR GRADIENT WITH RESPECT TO WEIGHT

$Y_{Predicted}$ = m X+c	Error	$\partial Error/\partial c$	$\partial Error/\partial m$
0.42	0.087	0.42	0.00
`0.58	0.064	0.36	0.08
0.59	0.000	0.01	0.00
0.66	0.107	0.46	0.15
0.69	0.010	0.14	0.05
0.74	0.063	0.36	0.16
0.74	0.021	0.20	0.09
0.84	0.048	0.31	0.18
1.10	0.005	0.10	0.09
1.15	0.148	0.54	0.54
	Total (Error)=0.543	2.830	1.410

Table IV shows the reduced error with respect to weight. It is found with random values of c and m, total Error=0.668. In this situation, need to update these values.

Step 4: Adjust value of parameter with the initial random value to reach the optimal values and minimized the error.

Now, need to update parameter of c and m so that move in the direction for optimization. Now need to calculate

$$c - \partial Error/\partial c$$

$$m - \partial Error/\partial m$$

Step 5: The adaptive rules are:

$$c = c - R * \partial Error/\partial c$$

$$m = m - R * \partial Error/\partial m$$

Here, is the Adaptive rate (R) have small value which can be adjustable. In this case adaptive rate value is 0.01.

Step 6: Now use updated value of c and m for prediction are c=0.43 and m=0.72. Now again calculate Error

It is clear that new prediction, the total Error has reduced form (0.668 to 0.543). Accuracy of the prediction has been improved.

Step 7: Repeat step 3 to 5 further adjustments in the value of c and m until significantly reduce the error. At a situation when no change in the error value stop the repetition at last, got the optimal and highest prediction accuracy.

Table V shows the reduced error with updated value.

TABLE V. CALCULATING ERROR WITH RESPECT TO UPDATED VALUE

S.No	**Predicted value**	**Error**
1	0.45	0.087
2	0.62	0.064
3	0.63	0.000
4	0.70	0.107
5	0.73	0.010
6	0.78	0.063
7	0.78	0.021
8	0.88	0.048
9	1.14	0.005
10	1.20	0.148
		Total= 0.553

VIII. EXPERIMENTAL ANALYSIS AND RESULT

A. Description Bout Dataset

Performance of proposed approach were evaluated with real life dataset. Dataset has been collected from 4 JEE entrance examination preparation institute Indore M.P. (ALLEN, KALPVIKSHA, FIITJEE, NARAYANA COACHING). More 1000 students' data have been collected. Datasets contain number of hours study and score during preparation. Based on the score in preparation for JEE they got selected in final JEE examination. Dataset used number of hours students' study in 15 days and appeared in JEE test series of 100 marks. Decimal scaling is used for number of hours. Python language is used to implementation and dataset stored in CSV data.

B. Description about Implementation

Python language to implement the roped approach, pandas and matplotlib library is used for implantation. Plot regression function have created and passed three parameters x, m and c, for unknown values to predict the correct value of y with high accuracy. In a new array predicted value are appended. Another function is created for updating error and update the value of m and c by using first derivation and calculate new value of error. Scatter plot is used to show number of hours study by the student and predicted score. Data points are very close so that they are overlapped. Fig. 3 shows the positive relationship between the variables.

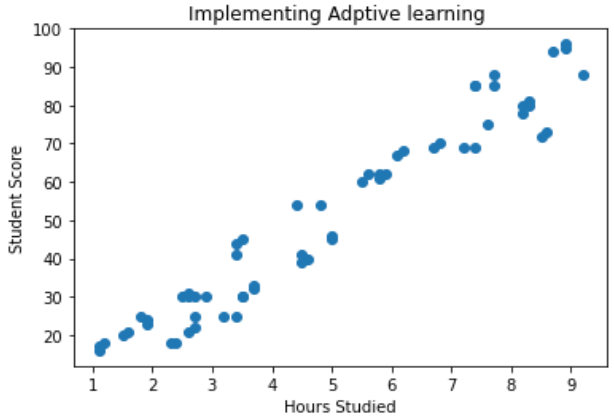

Fig. 3. Scatter plot for given dataset.

Initially the value selected for the m and c are 0.31 and 0.23, respectively. The selected valued used adaptive learning. From Fig. 4 staring with orange line and finally fitted the correct regression line for the given dataset. Fig. 4 shows the working process of adaptive learning. By the minimum number of iterations, got the best fitted line for the given dataset.

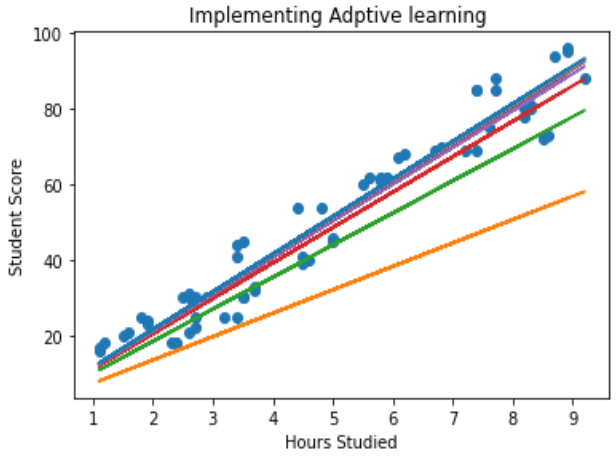

Fig. 4. Working process of adaptive learning.

In every iteration the reduced value of error has calculated. By the experimental analysis it is found that after 10 iteration the value of the error got minimized and has stable value (no minimization) at this point stop the process. Fig. 5 shows the change in value of parameters and the reduction in the error. Fig. 5 shows how the value of error gets reduced. In seconds iteration error is minimized to -5.33346109 just half of the first iteration. In the third iteration, it is minimized to -3.48410 and for fourth iteration it is minimized -2.762388. Similarly, in fifth, sixth, seventh, up to the tenth iteration the values of the error get minimized with very minor changes. So, after tenth iteration stop the process and got the best fitted line for given dataset.

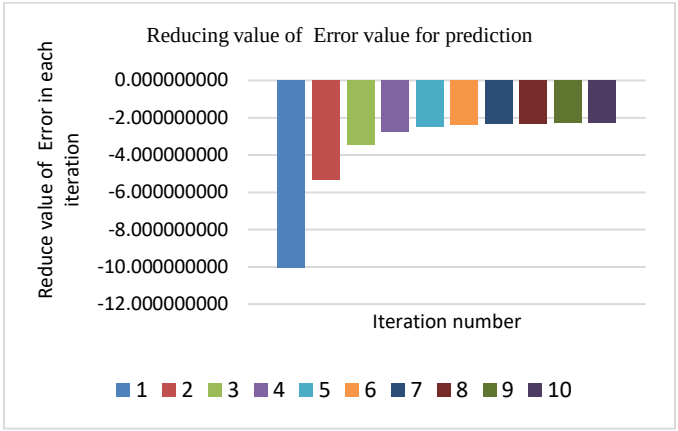

Fig. 5. Reducing error value in each iteration.

The value of the parameter m and c has been updated during each iteration. Fig. 6 shows the updated value of parameter m and c, from the Fig. 6 it is clear that after 5th iteration there are very minor changes are made in the value of m and c.

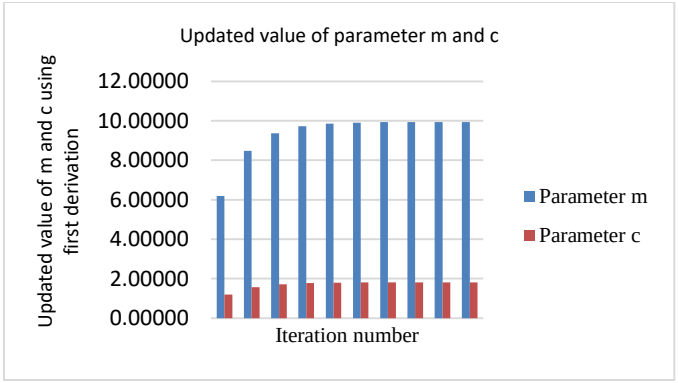

Fig. 6. Updating value of m and c during each iteration using first derivation.

So, prediction for unknown value of x the parameter has the value m=9.943 and c=1.816

C. Average prediction Error

Different number of records and calculated average prediction error have been used and found that Adaptive learning-based approach comparatively gave better prediction as other traditional approach. Table VI shows the prediction accuracy and average prediction error with number of students.

TABLE VI. PREDICTION ACCURACY AND AVERAGE PREDICTION ERROR

S No	No of Students	Prediction Accuracy	Average Prediction Error
1	890	96.31	3.69
2	2014	95.67	4.33
3	3022	97.23	2.77

Fig. 7 shows the graphical representation of number of students with prediction accuracy and average prediction error. It is also found that proposed approach is work well and proposed approach can scale up to any size of the dataset without effecting the performance.

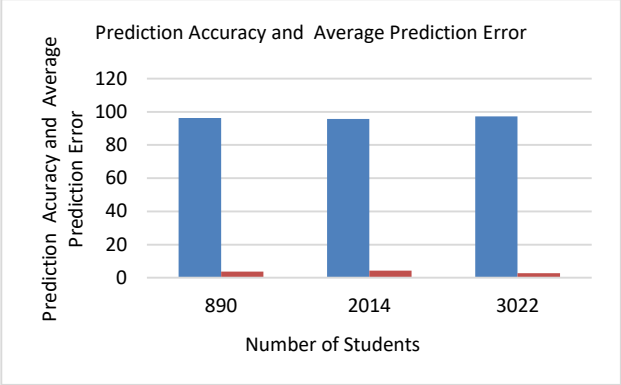

Fig. 7. Comparing prediction accuracy and average prediction error.

IX. Advantages and Limitations

The proposed approach has two advantages over the previous approach, it uses a minimum number of iterations and gives better accuracy with minimum error. The limitation of the proposed approach is that it uses only a single independent variable, there are several other parameters which can also be considered. The proposed approach is not suitable for nonlinear repression.

X. Conclusion and Future Work

Prediction is an important data analysis technique. Prediction helps us to predict what happens in future. In this paper, optimization with adaptive learning approach is proposed. Proposed approach improves accuracy of prediction by using adaptive learning and give optimal solution with minimum number of iterations. Real life data set is used for experimental analysis. Correct score can be predicted of the students based on number of studies for JEE exam. Dataset is collected from different JEE preparing centers of Indore city of Madhya Pradesh. By experimental analysis it is found that the accuracy of proposed approach for prediction is better and scalable. The proposed work helps students as well as training institution of JEE examination, to prediction the marks and possibility for selection. In future this approach can be applied for other type of regression and used for different datasets.

References

[1] Marno Verbeek Using linear regression to establish empirical relationships IZA World of Labor 2017: 336 doi: 10.15185/izawol.336 | Marno Verbeek © | February 2017 | wol.iza.org IZA World of Labor | February 2017 | wol.iza.org.

[2] Shen Rong Zhang Bao-wen The research of regression model in machine learning field MATEC Web of Conferences 176, 01033 (2018) https://doi.org/10.1051 2018 IFID 2018.

[3] Katarina Valaskova , Tomas Kliestik Financial Risk Measurement and Prediction Modelling for Sustainable Development of Business Entities Using Regression Analysis Sustainability 2018, 10, 2144; www.mdpi.com/journal/sustainability doi:10.3390/su10072144.

[4] Hazlina Darman, Sarah Musa Predicting Students' Final Grade in Mathematics Module using Multiple Linear Regression International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878 (Online), Volume-7 Issue-5s, January 2019.

[5] Gaurav Pandeya, Poonam Chaudhary SEIR and Regression Model based COVID-19 outbreak predictions in India https://doi.org/10.48550/arXiv.2004.00958 arXiv:2004.00958

[6] Baseer, K. K. and Neerugatti, Vikram and Tatekalva, Analysing Various Regression Models for Data Processing (June 30, 2019). International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-8 June, 2019.

[7] Ira Sharma and Sampurna Kakchapati Linear Regression Model to Identify the Factors Associated with Carbon Stock in Chure Forest of Nepal Hindawi Scientific Volume 2018, Article ID 1383482, 8 pages https://doi.org/10.1155/2018/1383482

[8] Samit Ghosal a Sumit Sengupta Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14th 2020) Clinical Research & Reviews 14 (2020) 311e315 journal homepage: www.elsevier.com/locate/dsx

[9] Yujiang He and Bernhard Sick CLeaR: An Adaptive Continual Learning Framework for Regression Tasks arXiv:2101.00926v4 [cs.LG] 16 Jul 2021 Intelligent Embedded Systems (IES) Group, University of Kassel, Wilhelmshoher Allee 71 - 73, Kassel, Germany.

[10] M Tanveer Hossain Parash and Mohammad Mostafizur Simple linear regression approach for evaluating models to estimate stature based on upper limb dimensions of adult Bangladeshi Males Hossain Parash et al. Egyptian Journal of Forensic Sciences (2022) 12:20 https://doi.org/10.1186/s41935-022-00277-3.

[11] Dengyuhan Dai Several Factors Affecting the Linear Regression Model in Data Analysis ICMML 2023, November 24–26, 2023, Nanjing, China © 2023 ACM ISBN 979-8-4007-1697-3/23/11 https://doi.org/10.1145/3653724.3653734.

[12] Sharyn O'Halloran Linear Regression: A Model for the Mean Spring 2005

[13] D.S.G. POLLOCK: ECONOMETRICS The Classical Linear Regression Model basic theory of the classical statistical method of regression analysis.

[14] Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining Introduction to Linear Regression Analysis Fifth Edition.

[15] John O. Rawlings Sastry G. Pantula David A. Dickey Applied Regression Analysis: A Research Tool, Second Edition

[16] Robert Nau Fuqua Notes on linear regression analysis School of Business, Duke University (c) 2014 by Robert Nau, all rights reserved. Last updated on 11/26/2014.

[17] Walter A. Shewhart and SAMUEL S. WILKS Editors: Ruey S. Tsay, Introduction to Linear Regression Analys Wiley Series In Probability And Statistics Established By Sanford Weisberg Editors Emeriti: Vic Barnett, J. Stuart Hunter, Joseph B. Kadane, Jozef L. Teugel.

[18] Rebecca Bevans. Revised on July 17, 2020.An introduction to simple linear regression Published on February 19, 2020 .

[19] Julien I.E. Hoffman Variations Based on Linear Regression, in Biostatistics for Medical and Biomedical Practitioners, 2015.

[20] Jean-François Dupuy A Brief Overview of Linear Models, in Statistical Methods for Over dispersed Count Data, 2018.

[21] Claudia and Angelini Regression Analysis, in Encyclopedia of Bioinformatics and Computational Biology, 2019

[22] Astrid Schneider, Gerhard Hommel, and Maria Blettner "Linear Regression Analysis" Department of Medical Biometrics, Epidemiology, and Computer Sciences, Johannes Gutenberg University, Mainz, Germany.

[23] B. Van Schaeybroeck and S. Vannitsem "Post-processing through linear regression" Nonlin. Processes Geophys., 18, 147–160, 2011 doi:10.5194/npg-18-147-2011.