Design and Analysis of Urban Land Lease Price Predicting Model Using Batch Gradient Descent Algorithm

Kifle Berhane Niguse[*]

Department of Computer Science and Engineering, Mekelle Institute of Technology, Mekelle University, P.O. Box 231, Mekelle, Ethiopia (*kifle.berhane@mu.edu.et).

ABSTRACT

Standard and econometric models are appropriate for causal relationships and interpretations among facets of the economy. But with prediction, they tend to over-fit samples and simplify poorly to new, undetected data. This paper presents a batch gradient algorithm for predicting the rice of land with large datasets. This paper uses a batch gradient descent algorithm to minimize the cost function, $J(\theta)$ iteratively with possible combinations of θ_0 and θ_1 the number of iterations i=1500 and learning rates, α of 0.01, 0.02, 0.03 for the linear regression case and i = 100, α =0.3, 0.2, and 0.1 for the multiple regression case. The paper uses Octave-4.0.3(GUI) for implementing 129 samples of the lease bid price of Mekelle City as training sets and feature inputs of two and three for linear regression and multiple regressions. Using $\alpha = 0.01$, the best fitting parameters found by training the dataset are $\theta_0 = 6.02$ and $\theta_1 = 2.30$ with a cost of J=67.82. The model predicts with an accuracy of 92.6% using LR and 91.15% using MLR for a 315 m^2 land size. As the learning rate increases, the fitting parameters θ_0 and θ_1 increase and decrease respectively with an equal cost but the model's prediction error increments slowly. With multiple regression, as the learning rate lowers, the model under fits prediction drastically (with an accuracy of 60%) with gradient descent and predicts with an accuracy of 91.5% with ordinary equations. So, prediction with ordinary equations provides the best fit for multiple regressions.

Keywords: Batch Gradient Descent Algorithm, Cost Function, Feature Scaling, Learning Rate, Machine Learning, Regression.

1. INTRODUCTION

Human beings have been passing through different revolutionary ages to make their living standard better than what their ancestors had lived. The industrial revolution age, the information technology age, and now the data science and Artificial Intelligence (AI) era are among the dominant ones, to mention a few. These all brought significant economic changes for every household. Recently, the role of data science and AI, more particularly Machine Learning (ML) along with Artificial Intelligence (AI), has caught the thoughtfulness of economists (Athey and Imbens, 2019). ML is a set of procedures (Algorithms). It is functional for big datasets for prediction, classification, fraud detection, and product recommendation (Athey, 2017). Machine learning applications are ready to lend a hand in attaining improved forecasting that is indispensable for economic policy formulation and targeting (El Naqa and Murphy, 2015). ML is

not much concerned with questions about identification. But it brings better results when the destination is a quasi-parametric approximation or with several covariates in comparison to the size of observations (Giles, 2018). ML is used for development interventions and impact evaluations in measuring outcomes and targeting treatments (McKenzie, 2018). There is a gap in the measurement of fundamental statistics. Machine learning is better for the problem at the macro- and micro-level (Elena Badilo, 2019). Specialists in ML are becoming curious about ML if it helps target interventions, i.e., deciding when and where/for whom to intervene (McKenzie, 2018). ML techniques can be helpful for econometricians. ML approaches are so powerful computationally. Subsequently, the intermingled use of econometric models with ML approaches helps to yield a resounding estimation and hence more multifaceted models (Kleinberg and Ludwig, 2016). The correlation between land lease prices and the macroeconomy is an imperative and inspiring reason for predicting land lease values. Land lease price implications are not only an issue for consumers and suppliers (Mekelle City Administration), but it also signposts the present-day economic condition. Consequently, it is significant to predict these prices without bias to assist both the consumers and suppliers in making their decisions.

Standard economic models are well appropriate to consider causal relationships between different facets of the economy, but when it comes to prediction, they tend to over-fit samples and occasionally simplify poorly to new, undetected data. Econometric models are relatively modest and relaxed to interpret; ML methods address enormous amount of aggregated data, frequently without forfeiting interpretation. Econometrics is to create causation, while ML aims to make accurate and actionable predictions. A significant benefit of ML is that it considers empirical analysis as algorithms that estimate (evaluate) and compare several other models. This approach contrasts with econometrics, where the researcher picks a model based on principles and estimations once (Giles, 2018). However, the growth of ML in social studies is so slow-moving. Social scientists have slowly but steadily begun leveraging ML techniques to gain new insights from data (Kleinberg and Ludwig, 2016).

It is common to see that most of the researches in economics are generally very fascinated causality questions (e.g., how do consumers get into financial distress?) (Varian, 2016). Despite, several policy problems have prediction questions that need to be addressed (e.g., which consumers will become financially distressed?) (Gharehchopogh, 2013). There have been studies on retail sales forecasting, stock market price, gold prices and estimation of GDP with

regression models prediction that check on the goodness of fit, model fit and specification, and statistical significance is explored using linear and multiple linear regressions (Wu et al., 2003). Other methods of ML predicting algorithms are like SVR and flexible ANN (employing a special adaptive regularization term) on travel-time prediction and the daily cash demand for ATM, respectively (Wang and Zhu, 2010). However, none of these studies used batch gradient descent algorithm for optimal parameters estimation and the minimization of the MSE or the cost. That is the reason the author is motivated to apply linear regression, multi-regression and ML algorithms (batch gradient algorithm) with the given training set and feature inputs of the land lease price of Mekelle city administration to contribute on better prediction of macroeconomic questions by blending econometrics and ML approaches. The line of best fit would be, using the model coefficients, such that it would be as close as possible to the actual dataset by minimizing the sum of the squared distances between each dataset and the line $h_\theta(x) = \theta_0 + \theta_1 x$. Once $\theta_0 \; and \; \theta_1$ are known, the model can be used to predict the response (price). The objective of the research is to minimize the cost function iteratively using gradient descent algorithm with optimal learning rate; by taking possible combinations of $\theta_0 \; and \; \theta_1$ and finally, design a price and profit predictive model for better predictability and analysis with new feature inputs of the city administration like that of land grade, land size, land location, bid price and so on.

2. METHODOLOGY

2.1. Conceptual Framework

The model comprehends machine learning algorithms, data training sets, and feature inputs to design a predictive model with low cost (MSE), optimal fitting line, and fast learning (Fig 1). Specifically, the cost function, hypothesis function, and batch gradient descent algorithm are presumed for the ML algorithm. The row entities (data training set) are, based on the Mekelle city land lease size of 140 m^2, 175 m^2, 250 m^2, and 300 m^2. The column elements (feature inputs) treated are land size, land grade, and bid price. Pictorial presentation of the whole conceptual framework is presented in figure 1.

2.2. Description of the Study Area

Data used for the study is a secondary data source collected from Mekelle city administration in the form of reports. The study employed correlation and experimental quantitative research. It used correlation methods to show the relation between explained and explanatory variables.
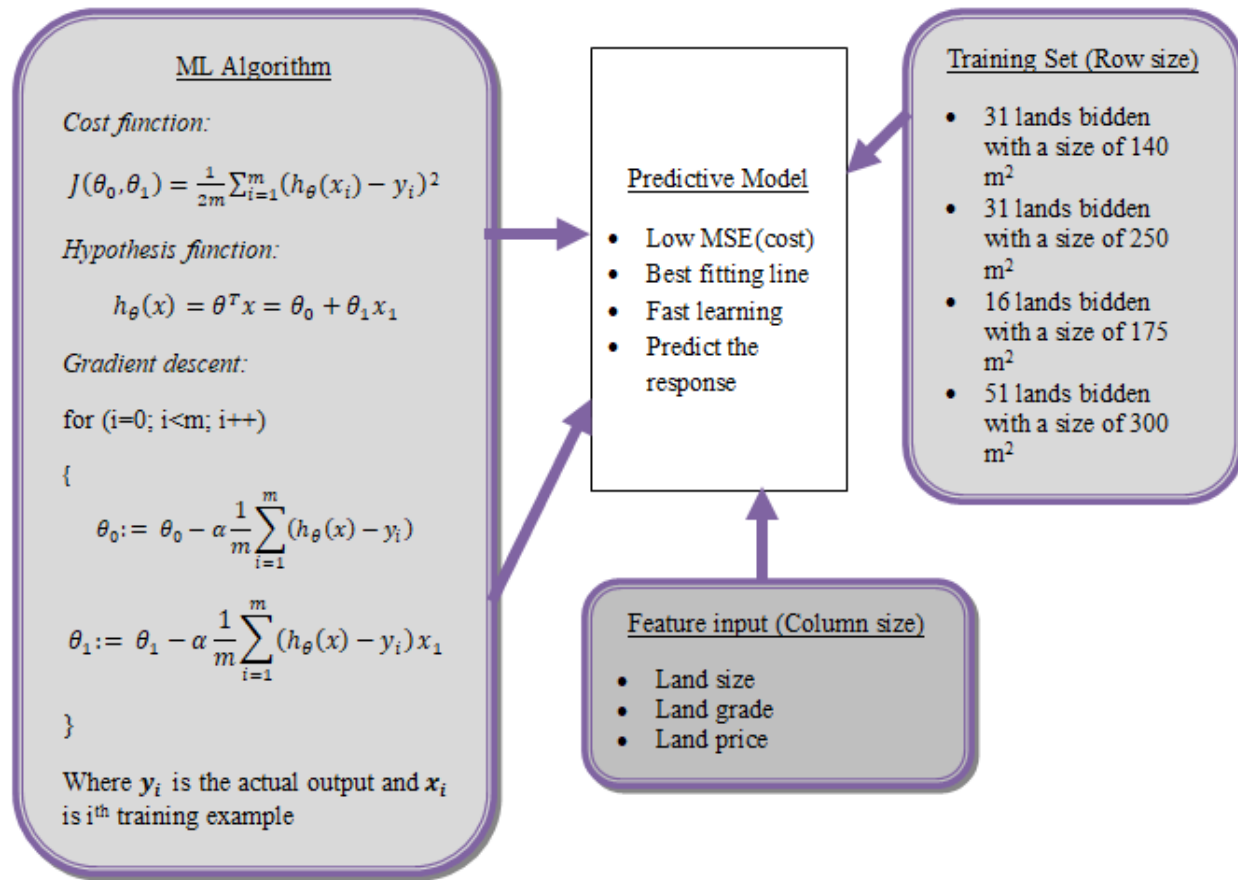
ML Algorithm

Cost function:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2$$

Hypothesis function:

$$h_\theta(x) = \theta^T x = \theta_0 + \theta_1 x_1$$

Gradient descent:

for (i=0; i<m; i++)

{

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x) - y_i)$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x) - y_i) x_1$$

}

Where y_i is the actual output and x_i is i^{th} training example

Predictive Model

- Low MSE (cost)
- Best fitting line
- Fast learning
- Predict the response

Training Set (Row size)

- 31 lands bidden with a size of 140 m²
- 31 lands bidden with a size of 250 m²
- 16 lands bidden with a size of 175 m²
- 51 lands bidden with a size of 300 m²

Feature input (Column size)

- Land size
- Land grade
- Land price

Figure 1. Conceptual Framework.

Quantitative analysis is applied to predict the cause-and-effect relationship between the dependent variable and independent variables. All these procedures use machine learning and econometric tools devised on linear regression and multiple regression models. Because land lease bid in Mekelle city is every three months, that is, discrete and equally spaced time intervals. The study is, thus, conducted using a time series data method.

2.3. Research Approach of the Study

Because predicting prices of the land lease is numeral in their inquiry mode quantitative approach is used to capture the in-depth and enormous data and information for systematic analysis and understanding of the performance. As prices are predicted by linearly relating parameters (features) of the model, the research is correlational by its objective and experimental based on the nature of the investigation. Octave-4.0.3(GUI), NotePad++, and Mathlab are used to implement the experiment.

In this paper, the researcher tested instrument regression (univariate and multivariate) and experienced to see it work on the given dataset. Regression parameters like that of are made to fit the dataset using batch gradient descent algorithm by tuning suitable learning rates and the number of iterations to achieve minimum cost or mean square error. The objective of linear regression is to minimize the cost function in equation (1).

Cost function: $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)^2$ -------------------- (1)

Where, the hypothesis function $h_\theta(x)$ is given by the linear model:

$h_\theta(x) = \theta^T x = \theta_0 + \theta_1 x_1$ -------------------- (2)

Hence, the parameters of the model are the θ_j values. For the regression possible feature input taken are land lease size (m^2), land grade and the bid price. These are the values that are adjusted to minimize cost function $J(\theta)$.

The researcher approached this achievement by

- Trying a number of blends of θ_0 and θ_1, learning rate and number of iterations
- Attaining at the line equation that satisfy most of the dataset
- Use the finalized line equation for predicting the new set of input values

Hence, the cost function rescues in identifying the optimal equation with the right combinations of θ_0 and θ_1 to realize best fitting line and optimized model. For each prospective line equation, the researcher:

- Tried to predict the h_θ value and took the difference from the actual h_θ value which gives the error possible with that particular equation.
- Repeated the same for all the sample datasets available
- Squared all the errors as the interest is in the magnitude of the error, and sometimes the positives and negatives may cancel out leading to no conclusion.
- Took the average by dividing with the number of sample dataset

The value, thus, arrived for each possible line equation is compared to derive the optimal line equation. The one with lowest Mean Square Error (MSE) is the one which can be picked as the Model for future predictions. To minimize the cost function $J(\theta_0, \theta_1)$, one approach to do this is using the batch gradient descent algorithm stated next.

Gradient descent algorithm: Repeat until convergence

{

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1);$$

$$\frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) = \frac{\partial}{\partial \theta_j} \frac{1}{2m} \sum_{i=1}^{m} (\theta_0 + \theta_1 x_1 - y_i)^2$$

For j=0, θ_0

$$\frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 x_1 - y_i)$$

For j=1, θ_1

$$\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 x_1 - y_i) x_1 \}$$

Standard and econometric models are appropriate for causal relationships and interpretations between facets of the economy. But with prediction, they tend to over-fit samples and simplify poorly to new, undetected data. This paper presents a batch gradient algorithm for predicting urban lease prices with large datasets. The paper uses a batch gradient descent algorithm to minimize the cost function, J (θ) iteratively with possible combinations of θ_0 and θ_1, the number of iterations i =1500 and learning rates α = 0.01, 0.02, and 0.03 for the linear regression case and i = 100, α = 0.3, 0.2, and 0.1 for the multiple regression case. The paper uses Octave-4.0.3(GUI) for implementing 129 samples of the lease bid price of Mekelle city as training sets and feature inputs two for linear regression and three multiple regressions. Using α = 0.01, the best fitting parameters found by training the dataset are θ_0=6.02 and θ_1=2.30 with a cost of J=67.82. The model predicts with an accuracy of 92.6% using LR and 91.15% using MLR for 315 m2 land size. It became clear that, as the learning rate increases, the fitting parameters $\theta_0$and θ_1 increase and decrease respectively with equal cost, and the model's prediction error increases slowly. With multiple regression, as the learning rate reduces, the model under fits prediction drastically (i.e., with an accuracy of 60%) with gradient descent and predicts with an accuracy of 91.5% with normal equation. So, prediction with normal equations provides the best fit for multiple regressions.

$$\{ \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 x_1 - y_i)$$

For j=1, θ_1

$$\frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 x_1 - y_i) x_1 \}$$

In batch gradient descent algorithm, each iteration performs the update:

$$\theta_j := \theta_j - \alpha \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)\, x_i^j \quad \text{(Simultaneously update } \theta_j \text{ for all j)} \text{ --------- (3)}$$

For every step of gradient descent, the parameter θ_j comes closer to the optimal values that will achieve the lowest cost value that is $J(\theta_0, \theta_1)$.

For multivariate linear regression, possible feature inputs can be land lease size (m²), land grade, and price per m². The land sizes and the prices in this paper's dataset are much bigger than the land grade, almost 100 times bigger. That deters the gradient descent algorithm does not converge quickly. To avoid this feature scaling is made, so that gradient descent converges more quickly. The working principle of multivariate linear regression is the mean values of each feature must be subtracted from the data set then scale (divide) the feature values by their respective standard deviations. The standard deviation is a way of measuring how much variation there is in the range of values of a particular feature. The gradient descent for multivariate linear regression remains the same as the univariate linear regression except that there will be more features. The hypothesis function and the batch gradient descent update will remain unchanged (See equations, 2 and 3).

In the multivariate case, the cost function can be written in the form indicated by the equation (4).

$$J(\theta_0, \theta_1) = \frac{1}{2m}(X\theta - \vec{y})^T(X\theta - \vec{y}) \qquad \text{--------------------------------- (4)}$$

Where, $\quad X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_m^T \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$

- m is the number of sample datasets (training set).
- $h_\theta(x_i)$ is the predicted output value (price).
- y_i is the actual output value (takes a feature and estimates price).
- α is learning rate.

- x is input variable/features.
- x_i is i[th] training example.
- $\frac{1}{2}$ shows minimizing one half the mean squared error makes the Octave/Mathlab easy.

2.4. Methods of Data Processing and Analysis

The collected dataset is first cleared (edited) at central editing in case of missing values. Next, it is coded and organized in vector-matrix form. Finally, it is processed with LR, MLR, gradient descent algorithm, and normalization procedures. The data processing flow chart of the used algorithm is shown below in figure 2.
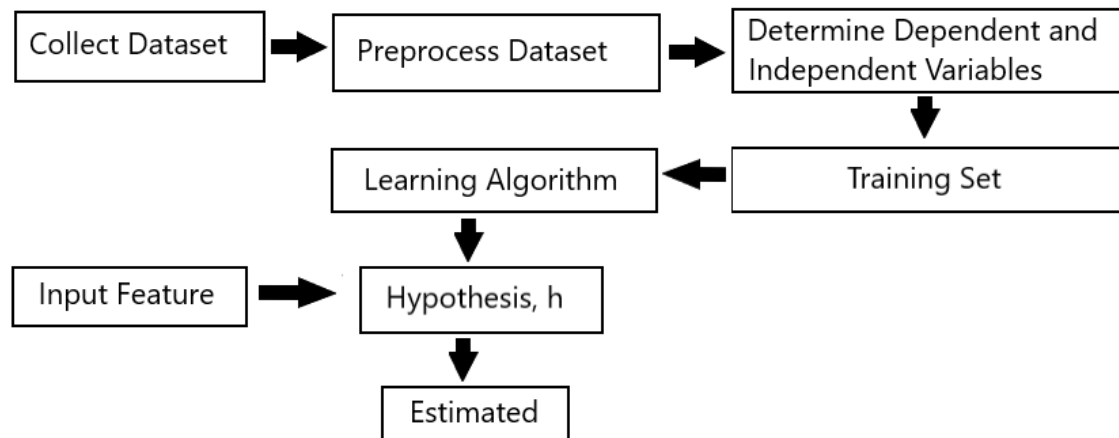

Figure 2. Data processing flow chart.

The Quantitative dataset generated from the sample selected from Mekelle City Administration land lease bid price is analyzed by writing program codes using the Octave software by changing relevant parameters like that of theta (θ), number of iterations and the learning rate alpha (α) using both univariate and multivariate LRs. Predicted price, cost function value and gradient descent values are used in reporting and explaining the results.

3. RESULTS AND DISCUSSIONS

3.1. Results with Linear Regression

3.1.1. Scenario-A: α = 0.01 and i* = 1500

With the learning rate of α = 0.01 and number of iterations, $i^* = 1500$, the fitting parameters and the minimum cost (MSE) value for the sample of land sizes 315 m^2 and 140 m^2 are found to be $\theta_0 = 6.02$, $\theta_1 = 2.30$ and $J(\theta_0, \theta_1) = 67.38$ respectively (Fig 3a, b and c). With these gradient descent fitting parameters, the prices predicted with the experiment for the samples are found to be Birr 13,274.74 and 9245.20 respectively (Table 1) whereas the actual average prices are 14,330.31 and 12,228.97 respectively. The prediction errors are 7.4% and 24.4% for the sample

land sizes of 315 m^2 and 140 m^2. Hence, this model best fits with a prediction accuracy of 92.6% and 75.6% for the first and second samples respectively (Fig 3a). The simulated results for cost and the contour for scenario-A are illustrated in figure 3.
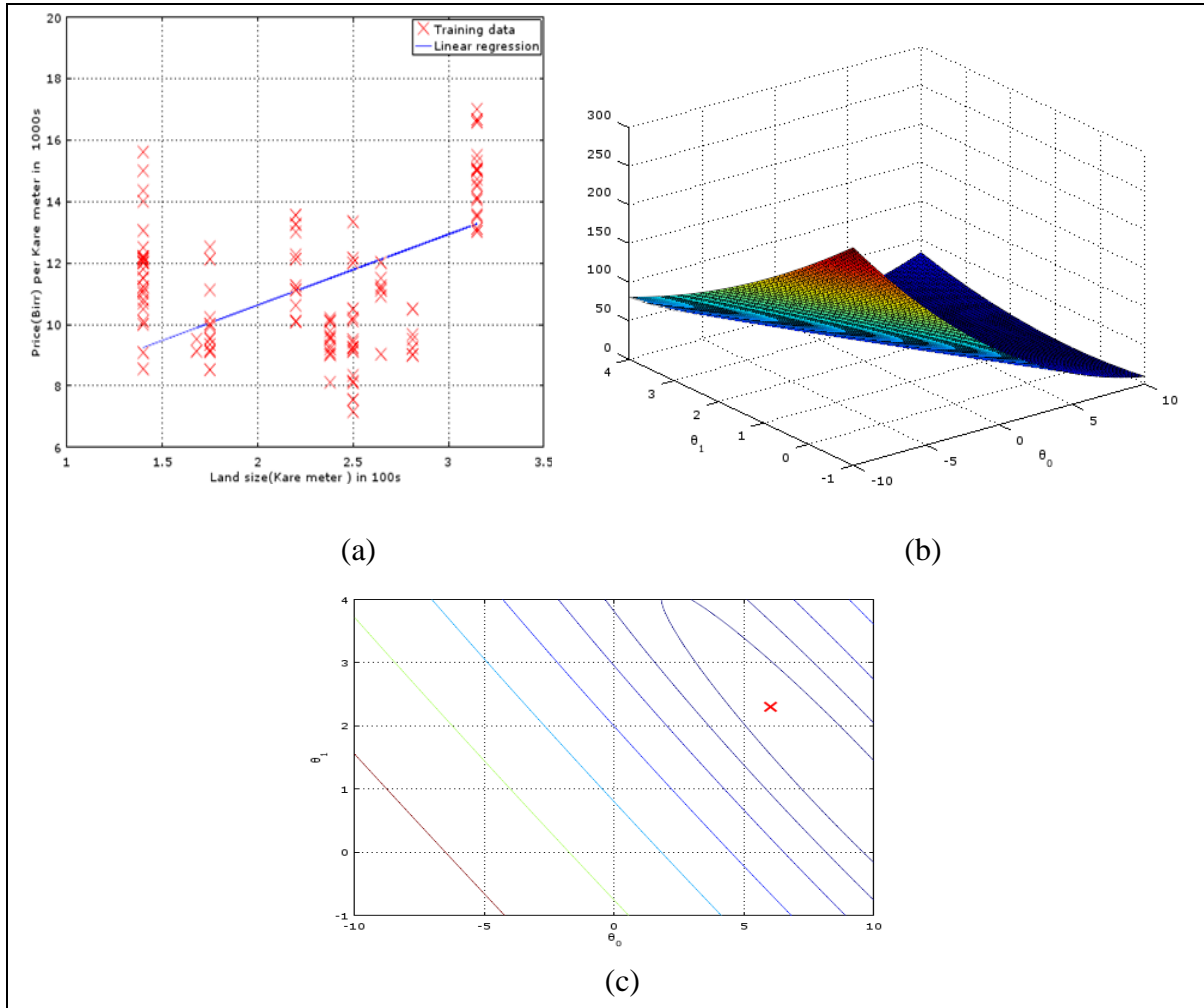

Figure 3. Linear regression fit line, Cost and Contour for α = 0.01 and i^* = 1500 (a) Training data with linear regression fit (b) Surface (c) Contour, showing minimum.

3.1.2. Scenario-B: α = 0.02 and i* = 1500

If the learning rate increases to α = 0.02 with number of iterations, i^* = 1500, the fitting parameter θ_0 increases to 7.76 and θ_1 decreases to 1.57 and the minimum cost (MSE) value remains same for the sample of land sizes 315 m^2 and 140 m^2 (Fig 4 a, b and c). With these gradient descent fitting parameters, the predicted prices are Birr 12,722.51and 9968.26 (Table1) respectively. As a result, this model best fits with a prediction accuracy of 88.78% and 81.5% for the first and second samples respectively (Fig 4a). The simulated results for price, cost and the contour for scenario-B are illustrated in figure 4.
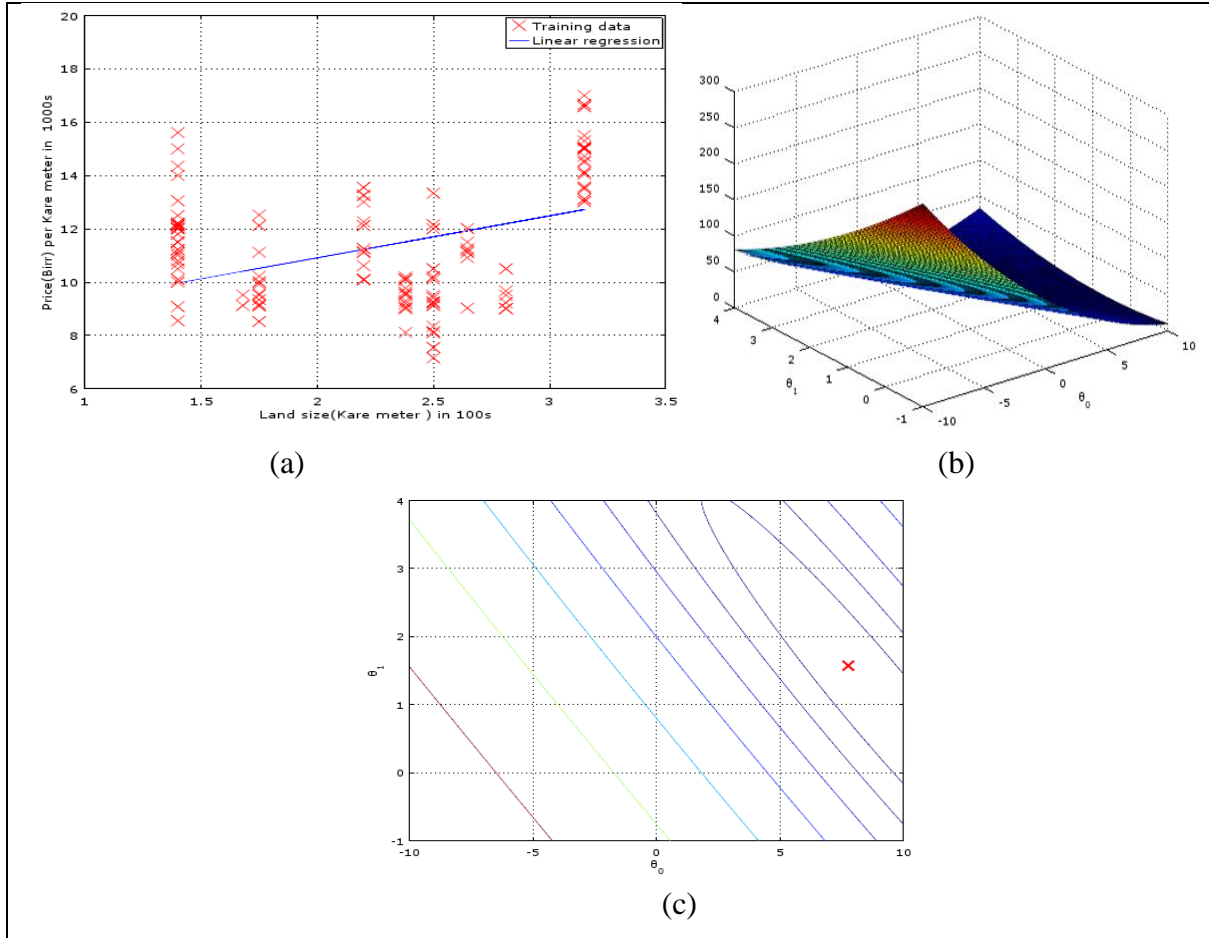
Figure 4. Linear regression fit line, Cost and Contour for α = 0.02 and *i** = 1500 (a) Training
     data with linear regression fit (b) Surface (c) Contour, showing minimum.

3.1.3. Scenario-C: α = 0.03 and i* = 1500

If the learning rate is further increased to α = 0.03 with number of iterations, i* = 1500, the
fitting parameter θ_0 keeps increasing to 8.47 and θ_1 decreases to 1.28 and the minimum cost
(MSE) value remains same for the sample of land sizes 315 m2 and 140 m2 (Fig 5 a, b and c).
With these gradient descent fitting parameters, the predicted prices are Birr 12,499.09 and
10,260.81 (Table 1) respectively. Therefore, the prediction accuracies are 87.22% and 83.9% for
the first and second samples respectively (Fig 5a). The simulated results for price, cost and the
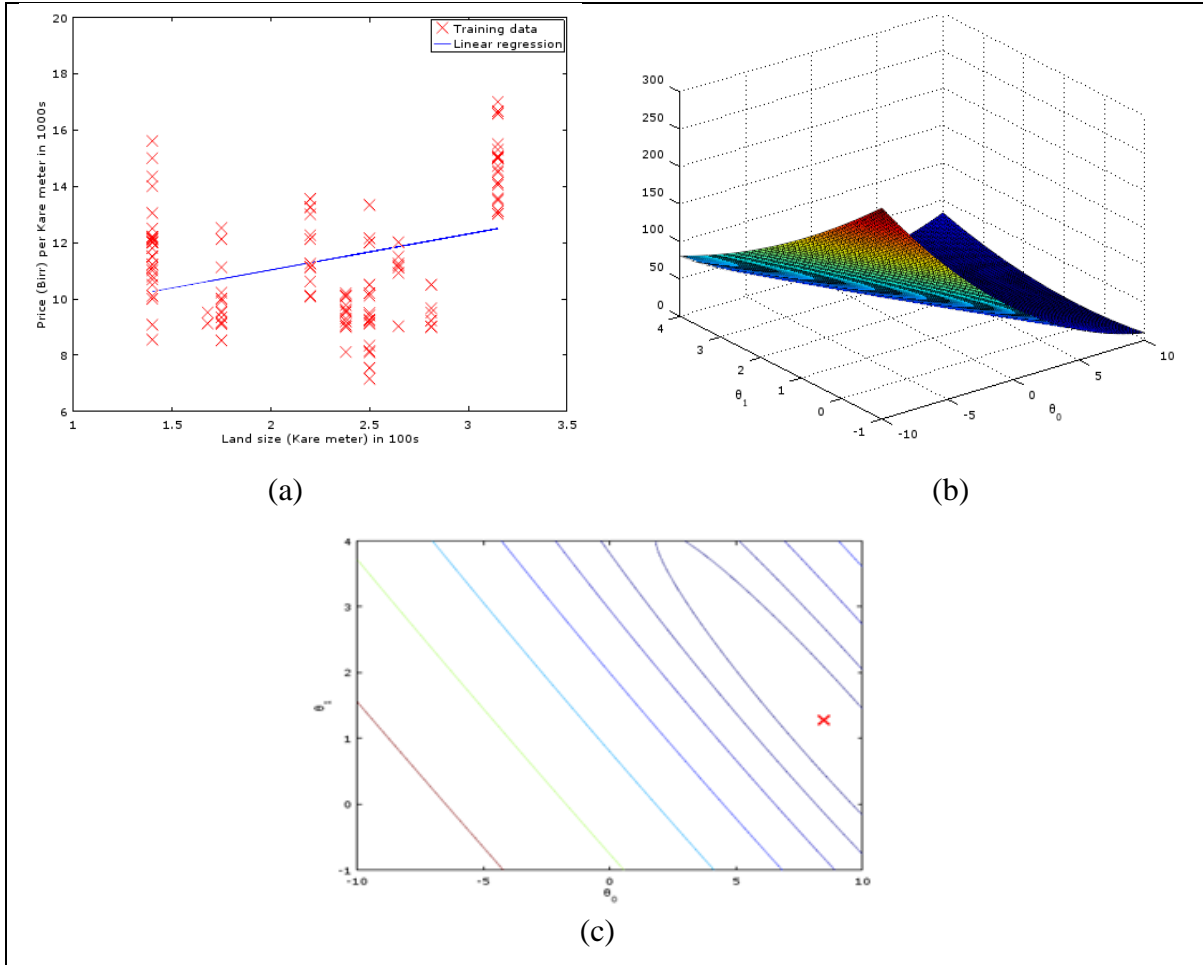contour for scenario-C are illustrated in figure 5.

Figure 5. Linear regression fit line, Cost and Contour for α = 0.03 and i^* = 1500 (a) Training
data with linear regression fit (b) Surface (c) Contour, showing minimum.

3.1.4. Scenario-D: α = 0.01 and i* = 750

If the learning rate is further increased to α = 0.01 with number of iterations lowered to, i^* = 750, the fitting parameter θ_0 decreases to 4.35 and θ_1 increases to 3.00 and the minimum cost (MSE) value remains same for the sample of land sizes 315 m² and 140 m² (Fig 6 a, b and c). With these gradient descent fitting parameters, the predicted prices are Birr 13,805.13 and 8550.72 (Table 1) respectively. Therefore, the prediction accuracies are 96.34% and 70% for the first and second samples respectively (Fig 6 a). The simulated results for price, cost and the contour for scenario-D are illustrated in figure 6.
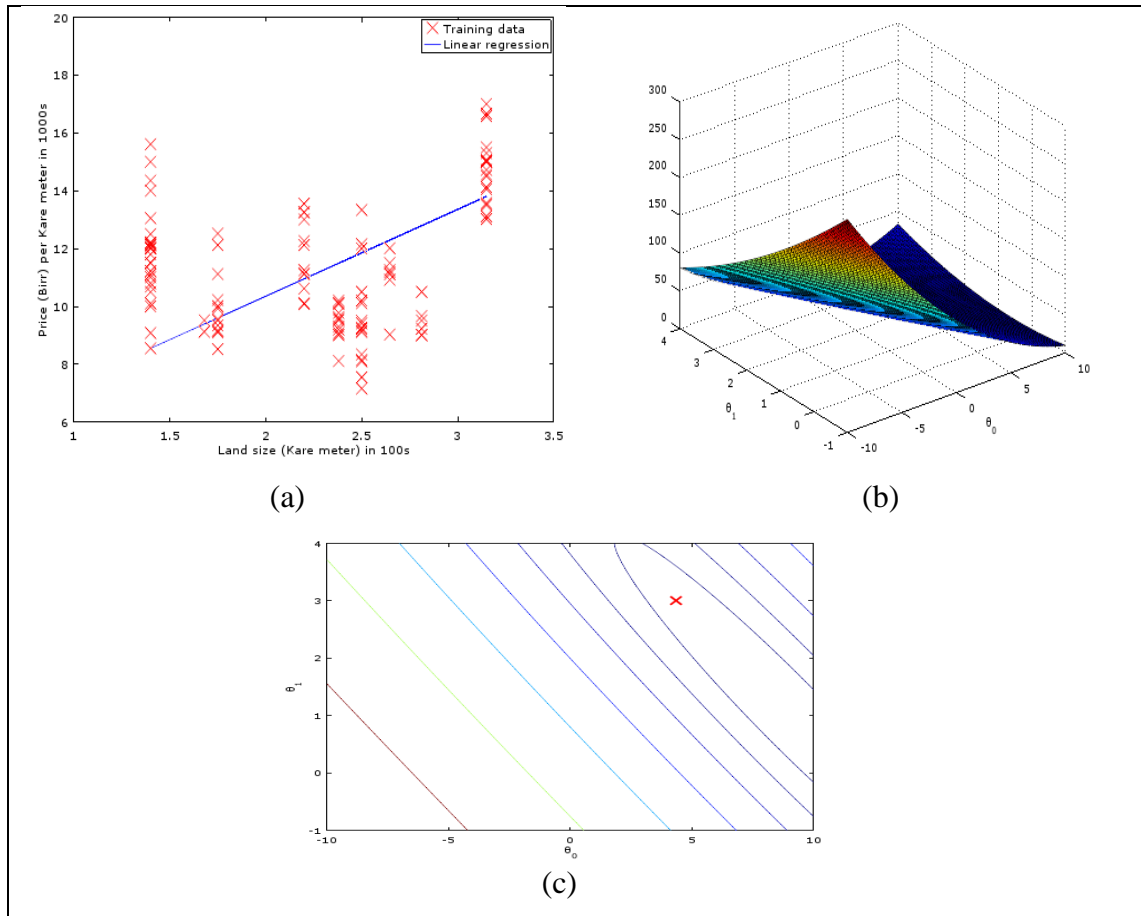
(a)   (b)

(c)

Figure 6. Linear regression fit line, Cost and Contour for α = 0.01 and i* = 750 (a) Training data with linear regression fit (b) Surface (c) Contour, showing minimum (*i** = number of iterations).

The quantitative output (for the model parameters, MSE and price) obtained after considering scenarios A - D are summarized in table1 for the case of LR.

Table1. Model parameters and results for LR.

Scenario		Theta found by gradient descent		Cost (MSE)	Predicted price (Birr) for sample land sizes			
		θ_0	θ_1	$J(\theta_0, \theta_1)$	315 m^2	250 m^2	175 m^2	140 m^2
A	α = 0.01 i* = 1500	6.02	2.30	67.38	13274.74	11778.05	10051.10	9245.20
B	α = 0.02 i* = 1500	7.76	1.57	67.38	12722.51	11699.51	10519.11	9968.26
C	α = 0.03 i* = 1500	8.47	1.28	67.38	12499.09	11667.73	10708.46	10260.81
D	α = 0.01 i* = 750	4.35	3.00	67.38	13805.13	11853.49	9601.60	8550.72

Note: *i** = number of iterations   and α = learning rate.

3.2. Results with Multiple Regression

For $\alpha = 0.3$ and $i* = 100$, the fitting parameters computed from the experiment are $\theta_0 = 11110.98$, $\theta_1 = 640.15$, and $\theta_2 = 674.27$ (Table 2) for the minimum cost value. The predicted prices for the two sample land sizes (using batch gradient descent algorithm) are Birr 13062.20 and 9745.72. The prediction accuracy is, therefore, 91.15% for the 315 m^2 sample land size and 80% for 140 m^2. This prediction accuracy and values of best fitting parameters remain almost the same for learning rates of 0.2, and 0.1 (Table 2).

Table 2. Model parameters and results for MLR.

Scenario	Theta found by gradient descent					Predicted price (Birr) for sample land sizes	
			θ_0	θ_1	θ_2	315 m^2	140 m^2
A $\alpha = 0.3$ $i* = 100$	Using gradient descent		11110.98	640.15	674.27	13062.20	9745.72
	Using normal equation		7701.69	10.25	304.58	13062.20	9745.72
B $\alpha = 0.2$ $i* = 100$	Using gradient descent		11110.98	640.15	674.27	13062.20	9745.71
	Using normal equation		7701.69	10.25	304.58	13062.20	9745.72
C $\alpha = 0.1$ $i* = 100$	Using gradient descent		11110.69	640.36	674.06	13061.89	9745.29
	Using normal equation		7701.69	10.25	304.58	13062.20	9745.72
D $\alpha = 0.01$ $i* = 100$	Using gradient descent		7044.00	513.94	525.91	8587.19	5959.41
	Using normal equation		7701.69	10.25	304.58	13062.20	9745.72

Note: $i*$ = number of iterations; α = learning rate.

However, when the learning rate is small, i.e., when $\alpha = 0.01$, the prediction accuracy is found to be so low, i.e., 60% and 49% for land size of 315 m^2 and 140 m^2 respectively with gradient descent algorithm. Conversely, for $\alpha = 0.01$, the model predicts with an accuracy of 91.15% for the 315 m^2 sample land size and 80% for 140 m^2, if normal equation is used. Hence, the model predicts best with gradient descent algorithm when the learning rate, $\alpha = 0.3$ and number of iterations, $i* = 100$ and it predicts best with normal equation with same accuracy as batch gradient descent algorithm when $\alpha = 0.01$ and $i* = 100$. It can be inferred that at $\alpha = 0.3$, 0.2, 0.1, and $i* = 100$, the batch gradient descent algorithm learns fast with an appropriate

learning rate as shown in figure 7(a to d) in order of preference. When α = 0.01 and i^* = 100, the algorithm learns so slowly even though the performance of price prediction is the same as with what the batch gradient algorithm does.
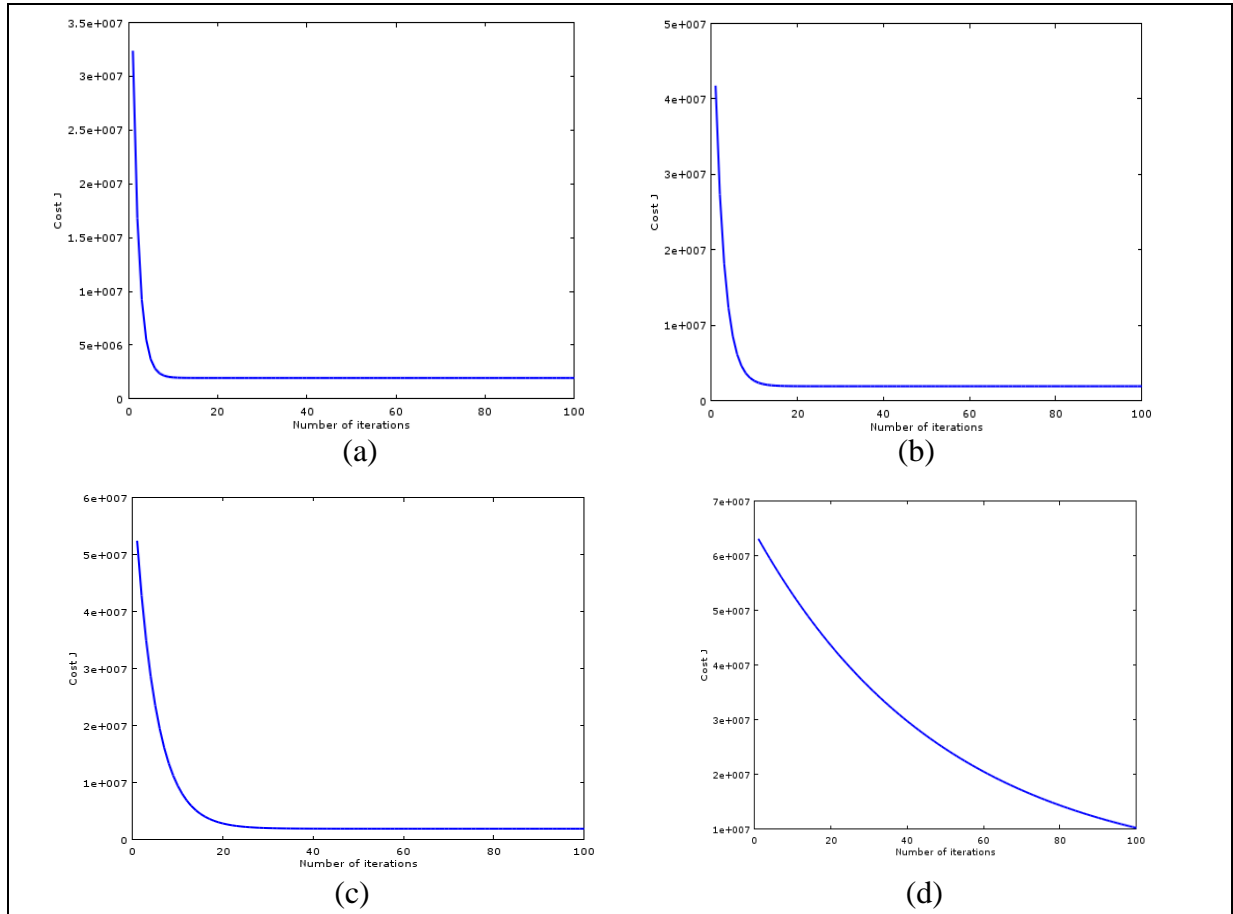

Figure 7. Gradient descent convergence for (a): α = 0.3, i^* = 100, (b): α = 0.2, i^* = 100, (c): α = 0.1, i^* = 100, and (d): α = 0.01, i^* = 100.

Therefore, the best price predictor mode would be when α = 0.3 and i^* = 100 (Fig 7a) in θ_0 = 11110.98, θ_1 = 640.15, and θ_2 = 674.27 with a minimum cost value as shown in figure 7(a) with MLR. Figure 7 shows how and when the gradient descent converges or diverges. The quantitative output (for the model parameters and price) obtained after considering scenarios A - D is summarized in table2 for the case of MLR.

4. CONCLUSION

Each iteration in this paper is performed simultaneously using the gradient descent update (algorithm). For every step of gradient descent, each fitting parameter gets closer to the optimal

values that achieve the lowest cost value. The cost decreases with each step and iteration; the cost function has never increased and proved that the gradient descent algorithm functioned appropriately and converged to a steady state value that the algorithm terminates on. The final parameters' values helped to make price predictions on new input features. When the sample training set becomes large, the cost function increases. Learning rates are on a log scale such as 0.01. 0.02, 0.03, 0.1, 0.2, and 0.3 were in a better range. A Too small learning rate led the batch gradient descent algorithm to converge slowly. A learning rate that is too small blows up and increases the value of the cost function. When the learning rate is too large, gradient descent overshoots the minimum failed to converge or diverge to some steady state cost value; the cost function was with a value of infinity. In the case of LR, using a learning rate of 0.01, the best fitting parameters found by training the dataset are 6.02 for theta1 and 2.30 for theta2 in order with a cost of 67.82; the model predicts with an accuracy of 92.6%. However, for MLR, the batch gradient descent algorithm converges fast with a learning rate of 0.3 and some iterations amounting to 100. The model achieves a prediction accuracy of 91.15% for the 315 m^2 sample land size and 80% for 140 m^2 with the best fitting parameters of 11110.98, 640.15, and 674.27 for theta0, theta1, and theta2 in order.

This model predicts the price of urban land leases for Mekelle city administration based on time series data trained with machine learning and econometric models unified for better optimality. Designing such economic models enhances the predictability of important macroeconomic terms. As a result, consumers may be interested in how much to bid per square meter of the land lease in the city. The output of this research may support the government's macroeconomic decisions, interventions, policy formulation, and targeting.

6. REFERENCE

Athey, S. 2017. The Impact of Machine Learning on Economics. In: Ajay Agrawal, Joshua Gans, and Avi Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda,* University of Chicago Press, ISBN: 978-0-226-61333-8, http://www.nber.org/books/agra-1, Conference volume, pp.507-547.

Athey, S & Imbens, G. W. 2019. Machine Learning Methods That Economists Should Know About. *Annual Review of Economics,* **11**: 685-725.

El Naqa, I & Murphy, M. J. 2015. What Is Machine Learning? In: El Naqa, I., Li, R., Murphy, M. (eds), *Machine Learning in Radiation Oncology*, **1**: 3-11.

Elena Badilo. 2019. The Impact of Machine Learning on Economics: What Machine Learning Can (and Cannot) Do for Economic Research. *Chicago Policy Review*. Available at: https://chicagopolicyreview.org/2019/01/21/the-impact-of-machine-learning-on-economics- what-machine-learning-can-and-cannot-do-for-economic-research/

Gharehchopogh, F. S. 2013. A Linear Regression Approach to Prediction of Stock Market Trading Volume: A Cases Study. *International Journal of Managing Value and Supply Chains*, **4(3)**: 1-7.

Giles, A. 2018. Machince learning and Economics. Financial Conduct Authority. Availae at: https://spe.org.uk/site/assets/files/5204/adam_giles_slides.pdf

Kleinberg, J & Ludwig, J. 2016. Prediction Policy Problems. *American Economic Review,* **105(5)**: 491-95.

McKenzie, D. 2018. How can machine learning and artificial intelligence be used in development interventions and impact evaluations? World Bank Blogs. Available at:https://blogs.worldbank.org/impactevaluations/how-can-machine-learning-and-artificial-intelligence-be-used-development-interventions-and-impact.

Varian, H. R. 2016. Causal inference in economics and marketing. *PNAS.org*, **113 (27)**: 7310-15.

Wang, L & Zhu, J. 2010. Financial market forecasting using a two-step kernel learning method for the support vector regression. *Annals of Operation Research,* **174**: 103-120

Wu, Chun-Hsin., Chia-Chen, Wei., Da-Chun, Su., Ming-Hua, Chang & Jan-Ming, Ho. 2003. Travel time prediction with support vector regression. Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems : **2(2)**: 1438-1442.