JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Supply chain Optimization Technique for biofuels industries using L1 Regularization and eXtreme Gradient Boosting

[1]Dattatray Gandhmal, [2]Nagdatta Gajjam, [3]Farzeen Sheikh

[1]Associate Professor, [2]Research Scholar, [3]Research Scholar
[1,2,3]Computer Science & Engineering,
[1,2,3]N.B.Navale Sinhgad College of Engineering, Solapur, India

Abstract : Recently, a lot of attention has been paid to biofuel energy as an alternative to fossil fuels. To keep this vision alive, we need a strong supply chain that helps get competitive biofuels to end-use markets. This paper first demonstrates what supply chain is. There are many problems are identified in supply chains for biofuels such as Secure Transaction problems, Material Tracking, Farmers profits and Supply chain optimizations. We can optimize supply chain by demand forecasting, Inventory optimization, Predictive maintenance, Route optimization etc. In this paper we will discuss about "supply chain optimization" in demand forecasting, Inventory management and predictive maintenance and its solution. Several techniques are devised in the existing techniques to predict the demands for products according to environment and time of the year. In this paper we will discuss why Reinforcement learning techniques are not suitable for supply chain optimization and along with that we will showcase our new model which is developed using different regularization techniques on Gradient boosting machines. This paper also compares different machine learning algorithms for supply chain optimization and its matrices. The works are analyzed using certain datasets, software tools, performance evaluation measures, prediction techniques utilized, and performance attained by different techniques. The prediction using the proposed Gradient boosting machines using L1 regularization and extreme gradient boosting approach shows the superior performance in terms of mean absolute percentage error and root mean square error.

Keywords—Machine Learning,Gradient Boosting Machines, Regularization,XGBoost,Linear Regression,LASSO

I. INTRODUCTION

Companies with supply chains that include a plant and numerous warehouses, or so-called hub-and-spoke networks (Arnold, 2009), struggle with supply chain optimization. The primary choice is how many goods should be manufactured in the factory and how much stock should be kept.be accumulated in the warehouses. Seasonal demand may make it necessary for businesses to begin accumulating goods early in order to meet future demands (for example, eggnog for Christmas may need to be accumulated in November and December). This complicates the decision-making process further. While small organizations can continue to handle their supply chains manually, automatization is required for large corporations (such as biofuels production). Conventional policies, like the (Q)-policy (Tempelmeier, 2011), are frequently overly straightforward and incapable of adapting to complicated contexts.

Machine learning can be a valuable tool for optimizing supply chain operations by enabling more accurate predictions and better decision-making. Below we can see where we can optimize supply chain for biofuels using machine learning:

1 Demand forecasting: Machine learning algorithms can be used to analyze historical data and identify patterns and trends that can be used to forecast demand for products. This can help supply chain managers better plan inventory levels and optimize production schedules.

2 Inventory optimization: Machine learning can be used to analyze data on product demand, lead times, and supplier performance to optimize inventory levels. This can help reduce excess inventory, improve cash flow, and minimize stockouts.

3 Predictive maintenance: Machine learning can be used to monitor equipment and predict when maintenance is required, reducing downtime and maintenance costs.

4 Route optimization: Machine learning can be used to optimize delivery routes and schedules, reducing transportation costs and improving delivery times.

5 Supplier selection: Machine learning can be used to analyze data on supplier performance and identify the best suppliers based on factors such as quality, lead time, and cost. Overall, machine learning can help supply chain managers make better decisions by providing insights into complex data sets and identifying patterns and trends that might otherwise be difficult to detect.

MDPs are multi-step probabilistic decision problems that rely on the Markov property, which asserts that the only elements that affect the transition probability P(st+1 | st, a) between two states st+1 and st are the current state st and the chosen action an. We characterise the MDP in a manner similar to Moritz (2014) and Powell (2014) by specifying a state-space, a random environment process for the demand, an action space, a set of possible actions, a transition function, a one-step reward function, and a discount factor (2007). Throughout the chapter, we use j = 0,1.2.., K as an identifier for the factory (j = 0), the K warehouses (j = 1,..., K), and t = 1,..., N as an identifier for each period (where N is the terminal period).

Table 1. Components of the one-step reward function

Component	Formular	Variables
Revenue from sold products	$p\sum_{j=1}^{K} d_j$	Price p, demand d_j
Production cost	$\kappa_{pr} a_0$	Unit cost κ_{pr}, production level a_0
Storage cost	$\sum_{j=0}^{K} \kappa_{st,j} \max\{s_j, 0\}$	Storage cost $\kappa_{st,j}$, stock level s_j
Penalty cost	$\kappa_{pe} \sum_{j=1}^{k} \min\{s_j, 0\}$	Penalty cost κ_{pe}, stock level s_j
Transportation cost	$\sum_{j=1}^{k} \kappa_{tr,j} \lceil a_j/\zeta_j \rceil$	Truck cost $\kappa_{tr,j}$,Truck capacity ζ_j Transportation volume a_j

The state space at period t is denoted by st = [s0,..., sK, dt1, dt2], where each sj [0, cj] reflects the stock levels of the factory (s0) and warehouses (s1,..., sK), up to some maximum capacity (cj). We utilise the random stochastic process dj,t to model the individual demand at warehouse j and time t for the demand vector dt = [d1,t,..., dK,t]. We include the most recent demands (dt1, dt2) in the state space to provide the agent some knowledge of the demand history and to help it comprehend how the demand has changed over time. We observe that the real stochastic demand in a period t won't be seen until the following period.

II. PROBLEM

Making important judgements about the production process requires careful consideration of market demands by businesses. Businesses that fail to assess market demands struggle to decide how many things to produce. One of the most frequent supply chain optimization issues that businesses today are dealing with is this one. While making strategic decisions about the manufacture of products, businesses should concentrate on the capacity, quality, and volume of their output. They should create a model for supply chain optimization that works well and make sure to meet client requests.

Fig. 1 Single factory with 5 warehouses

Companies should expect to pay 20% to 40% of their value for inventory management. Because of this, businesses must enhance their inventory control and attain 100% inventory accuracy in order to deliver goods on time to customers. To keep their inventory at the right level, businesses should have a strong warehouse management system or resource planning system. Companies need control procedures that can guarantee the right quantity of supply at order and reorder stages.

III. SOLUTION

The perspective of numerical optimization in function space, as opposed to parameter space, is used to view function estimates and approximation. We establish a link between steepest descent minimization and stagewise additive expansions. Based on any fitting criterion, an all-encompassing gradient descent “boosting” paradigm is created for additive expansions. For least-squares, least absolute deviation, and Huber-M loss functions in regression and multiclass logistic likelihood in classification, specific procedures are described. Tools for analysing such "TreeBoost" models are described, and special improvements are created for the specific scenario when the individual additive components are regression trees. Gradient boosting of regression trees generates competitive, extremely reliable, and comprehensible methods for both regression and classification, which are particularly suitable for mining imperfect data.

- L1-regularization for the Selection of Key Environmental Factors

Due to overfitting and significant correlations between various environmental elements, the prediction task might be difficult when there is a lack of data and a variety of environmental factors. In situations like these, variable selection approaches are frequently employed. In our investigation, LASSO was used to find a suitable variable set of environmental parameters for increasing the predictability of biofuels consumption.

LASSO, which Tibshirani proposed, is frequently used in regression analysis to choose variables and estimate parameters. A specific use of penalized least squares regression that uses the L1-penalty to reduce regression coefficients to zero is the LASSO estimator. Thus, the variables with coefficients other than zero that are the most explicable are chosen. where the first column is the time index and the following columns are predictors of biofuels consumption for a continuous N days, including environmental factors and other channels. The number 1 denotes the daily average of all instances of biofuels consumption over the course of one day. The LASSO estimator is defined as follows given the standardised predictors Xij for i= 1,2,...,N and j= 1,2,...,p and the response values:

$$x_{ij}\hat{\beta}_{LASSO} = arg\,min\{\sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\}$$

Using the LARS package in R, LASSO is implemented with training data to choose the most reliable predictors using the least angle regression algorithm (Version 3.3.2). Based on the related R packages the following approaches are also applied with R such as xgboost, stats, adabag, gbm, etc.

- eXtreme Gradient Boosting:

Extreme Gradient Boosting Package is abbreviated as xgboost. It is a scalable and effective use of Friedman's (2001) gradient boosting paradigm (Friedman et al., 2000). The programme offers a tree learning algorithm and an effective linear model solver. Regression, classification, and ranking are just a few of the objective functions that are supported. The programme is designed to be extensible, enabling users to quickly specify their own goals.

Guelman uses GBDT to create a loss prediction model for insurance and banking companies. When it comes to pricing for insurance and finance, the GBDT model outperforms general linear models built using conventional methods. Chen and Guestrin enhanced the idea of GBDT by proposing the eXtreme gradient boosting tree (XGBoost), which solves real-world classification problems. According to Chen and Guestrin, an improved type of gradient boosting machine is the extreme gradient boosting tree (XGBoost). The primary improvement to GBDT is the normalisation of the loss function to reduce model variability. As a result, there is less likelihood of model overfitness and less modelling complexity.

Table 2. The contrast between the major tree boosting algorithms

System	XGBoost	pGBRT	H2O	scikit-learn
Exact greedy algorithm	Yes	No	No	Yes
Approximate global	Yes	No	Yes	No
Approximate local	Yes	Yes	No	No
Out-of-core computation	Yes	No	No	No
Sparsity aware algorithm	Yes	No	Partially	No
Parallel algorithm	Yes	No	Yes	No

- **eXtreme Gradient Boosting Approach:**

Machine learning models are used to forecast biofuel production for stocks using the predictor subset chosen using LASSO. Empirically, it has been demonstrated that Friedman's gradient boosting model is highly effective . Tree boosting is a very powerful and well-known machine learning technique and one of the most well-liked gradient boosting methods. In this study, biofuels consumption forecasting was carried out using the scalable end-to-end tree boosting method known as eXtreme gradient boosting (XGBoost). Data scientists frequently employ XGBoost to solve challenging machine learning problems at the cutting edge . If a set of data has N samples and p characteristics, then

D $\{(x_i, y_i)\}\ (|\mathcal{D}| = N, x_i \in \mathbb{R}^p, y_i \in \mathbb{R})$,a tree ensemble model uses K additive functions (K trees) to predict the output.

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F}, \qquad (2)$$

Where

$$\mathcal{F} = \{f(x) = \omega_{q(x)}\ (q: \mathbb{R}^p \to T, \omega \in \mathbb{R}^T)$$

stands for the space of regression trees (also known as CART). The structure of each tree is represented by q, which maps an instance to the relevant leaf index. T=no. of leaves in the tree. Each fk represents a distinct tree structure and leaf weights. The decision rules in the trees (offered by) are applied to each instance to classify it into the leaves and generate the final prediction by accumulating the weights in the corresponding leaves (given by q). In equation (3), we lower the following objective to learn the set of functions utilized in the model. .

$$L(\emptyset) = \sum_{i=1}^{N} loss(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)\,, \qquad (3)$$

In equation (3), the first part is the differentiable convex training loss, which measures the variance between the prediction and the target yi. The difficulty of the trees is the second term. The following equation specifies the model complexity (fk).

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \left|\left|\omega_j\right|\right|^2, \qquad (4)$$

where T represents the number of leaves in the tree, represents the regularization weight, is the minimum loss reduction, and ||wj|| signifies the score for matching leaves. The tree (fk(x)) is defined by the formula as follows: (5),

$$f_k(x) = \omega_q(x), \omega \in \mathbb{R}^T, q: \mathbb{R}^d \to 1,2,\ldots,T. \qquad (5)$$

The is trained in an additive manner Functions are parameters in the tree ensemble model in equation (3), which makes it impossible to optimise it using conventional Euclidean-space techniques. Instead, the model is trained in an additive fashion.

Technically, let $\hat{y}_i^{(t)}$ represent the prediction of the i-th training sample at the (i-th iteration, we will need to add ft to minimize the following objective.

$$L^{(t)} = \sum_{i=1}^{N} loss\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \qquad (6)$$

The optimal XGBoost classifier is obtained with the sequential optimization at each scenario.

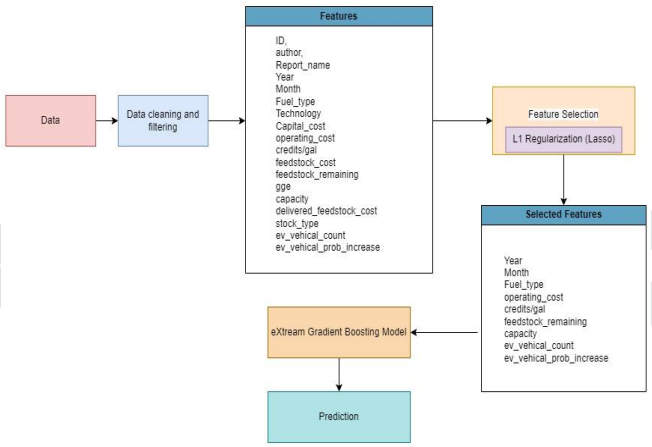


Fig 2. Block diagram of proposed method for biofuels consumption prediction

IV. EVALUATION

The 2 indicators root mean square error (RMSE) and mean absolute percentage error (MAPE) are used to evaluate the errors. The RMSE, which is the square root of the median of the squared errors and is scale-dependent, is a popular metric of the variances between values predicted by a model and values observed. In statistics, the MAPE is common metric for determining how accurate a prediction method is. Every predicted point in time is added together, and the calculation's absolute value is divided by the number of fitting points. The MAPE differs from the RMSE in that it is independent of scale. The calculating formulas are defined by the following terms:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \qquad (7)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i}, \qquad (8)$$

Where N is the number of data samples, yi (i = 1,2, … ,N) is the observed number of biofuels production sales and ŷi (i= 1,2, … ,N)is the predicted daily number of total sales of biofuels products.

V. RESULTS AND DISCUSSION

- **Benchmark Algorithms for Comparative Analysis**

We put benchmark algorithms into practice using their typical implementations and contrasted them with the XGBoost.

- **Linear Regression (LR):** The categorization is done by calculating the error, and the dependent variable is thought to be categorical in nature [16].
- **Support Vector Regression with Linear Kernel (SVR):** SVR seeks to achieve generalized performance by minimizing the generalization error constraint. SVR is based on computing a linear regression function in a high-dimensional feature space with a nonlinear kernel to transfer the input data to the space of features.

Table 5.1 Performance of Multiple Models on Test Dataset

Model	Period(In days)	RMSE	MAPE%

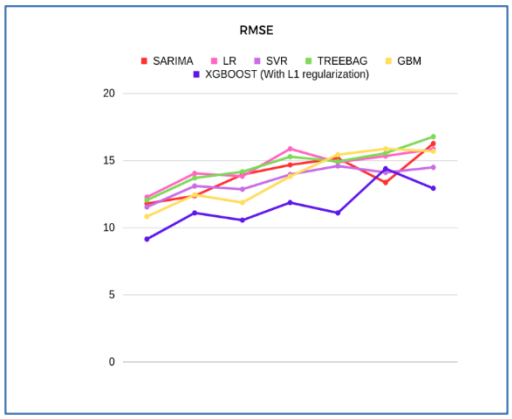
SHRIMA	5	11.82	20.14
	10	12.39	20.84
	15	13.97	21.56
	20	14.69	22.92
	25	15.19	24.24
	30	13.38	21.92
LR	5	12.29	25.89
	10	14.06	29.65
	15	13.85	28.10
	20	15.89	31.56
	25	14.89	30.23
	30	15.35	31.62
SHRIMA	5	11.56	22.87
	10	13.12	26.85
	15	12.88	25.98
	20	13.98	27.89
	25	14.61	28.17
	30	14.13	29.71
TREEBAG	5	12.07	21.32
	10	13.72	23.02
	15	14.18	25.44
	20	15.30	26.40
	25	14.95	25.03
	30	15.57	27.21
GBM	5	10.85	21.09
	10	12.47	24.23
	15	11.89	24.89
	20	13.84	26.04
	25	15.45	27.44
	30	15.89	28.67
XGBOOST (with L1 Regularization)	5	9.17	19.15
	10	11.12	21.14
	15	10.58	20.30
	20	11.89	18.04
	25	11.12	19.06
	30	14.41	21.50

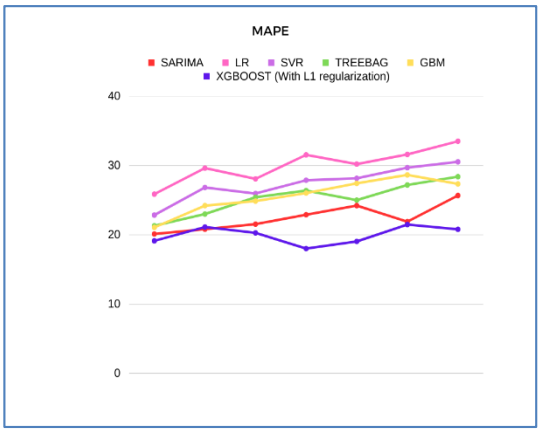
- **Bagging (Treebag):** A technique for producing numerous iterations of a predictor in order to obtain an aggregated predictor. When forecasting a numerical outcome, the aggregate takes an average across the variations.
- **Gradient Boosting Model (GBM):** This model iteratively learns from losses that occur at each stage by calculating a negative gradient. It then combines the outcomes of many prediction models.

We further compared the 5 days span consumption of biofuels forecasting performances obtained from other machine learning models including linear regression (LR), support vector regression with linear kernel (SVR), bagging (treebag), gradient boosting model (gbm), and XGBoost model. Table 5.1 shows the forecasting performances obtained on test dataset.

Both SARIMA and XGBoost were able to produce a steady and reliable prediction precision in terms of RMSE and MAPE, depending on the forecasting horizon. The XGBoost prediction errors stayed constant as the forecasting horizon expanded, however they rose in tandem with the SARIMA model.

Fig 5.1 Biofuels future needs for our sale prediction with Different Machine Learning Techniques





VI. Conclusion

For the purpose of predicting the trend of sales of biofuels products, this research provides a regularization-based model based on the eXtreme gradient boosting technique. The LASSO method was used to shrink the predictors and find the crucial variables that would affect the production of biofuels. Environmental elements, such as weather markets of vehicles, time of the year, were included in the prediction model. With a lowest MAPE of 19.06% for 1 month ahead forecasting, the real example displayed a satisfactory and solid performance in terms of 1 month ahead prediction. The approach does not, however, significantly enhance the ability to forecast environmental factors

References

[1] Regression Shrinkage and Selection Via the Lasso Robert Tibshirani First published: 1996
[2] K. Farid, Applied Linear Regression Models, Technometrics,32(2004), 352-353.
[3] A.J. Smola and B. Schölkopf, A tutorial on support vector regression, Stat comput 14 (2004), 199-222
[4] L. Breiman, Bagging predictors, Mach Learn 24 (1996), 123-140.
[5] Artificial intelligence for supply chain resilience: learning from Covid-19 Sachin Modgil, Rohit Kumar Singh, Claire Hannibal The International Journal of Logistics Management ISSN: 0957-4093
[6] Systematic analysis and review of stock market prediction techniques Author links open overlay panelDattatray P. Gandhmal, K. Kumar
[7] Predicting supply chain risks using machine learning: The trade-off between performance and interpretability
[8] J.H. Friedman, Greedy function approximation: A gradient boosting machine, Ann Stat 29 (2001),1189-1232.
[9] A Regularization-Based eXtreme Gradient Boosting Approach in Foodborne Disease Trend Forecasting Authors :Shanen Chen, Jian Xu, Lili Chen, Xi Zhang, Li Zhang, Jinfeng LiPages930 - 934DOI10.3233/SHTI190360
[10] Chronological penguin Adam-based deep long short-term memory classifier for stock market prediction Dattatray P. Gandhmal and K. KumarPublished Online:May 3, 2021 pp 215-238
[11] Development and implementation of an optimisation model for biofuels supply chain Christiana Papapostolou a , Emilia Kondili a , John K. Kaldellis.
[12] Richard, T. L. (2010). Challenges in Scaling Up Biofuels Infrastructure. Science, 329(5993), 793–796. doi:10.1126/science.1189139
[13] Stadtler, H. (2014). Supply Chain Management: An Overview. Supply Chain Management and Advanced Planning, 3–28. doi:10.1007/978-3-642-55309-7_1
[14] Lambert, D. M., & Cooper, M. C. (2000). Issues in Supply Chain Management. Industrial Marketing
[15] Management, 29(1), 65–83. doi:10.1016/s0019-8501(99)00113-3.
[16] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, Proceeding of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco (2016) 785-794
[17] L. Guelman, Gradient boosting trees for auto insurance loss cost modeling and prediction, Expert Systems with Application 39 (3) (2012) 3659-3667