

# Impact of Machine Learning Applications For Price Prediction Of Onion In India

Sayed Mannan Ahmad  
( UG Student )

Department of Computer Science and Engineering, Kanpur Institute of Technology  
Kanpur, India  
[Mannansarsobiz@gmail.com](mailto:Mannansarsobiz@gmail.com)

Ankush Kumar Gaur  
( Assistant Professor )

Department of Computer Science and Engineering, Kanpur Institute of Technology  
Kanpur, India  
[ankushkumarddm@gmail.com](mailto:ankushkumarddm@gmail.com)

Anil Kumar  
( Assistant Professor )

Department of Computer Science and Engineering, Kanpur Institute of Technology  
Kanpur, India  
[rakumar85@gmail.com](mailto:rakumar85@gmail.com)

## Abstract

Conceptual India is one of the second-biggest maker of onion on the planet. Onion is the most extravagant wellspring of nutrients and cell reinforcements. Furthermore, it contains malignant growth battling mixes, assists with keeping up heart more advantageous and lifts the bone thickness, additionally controls the glucose. Natural just as Climate changes sway on the rural economy of any nation. The creation of onion typically relies upon factors like science, atmosphere, economy, and topography, these elements impact on agriculture. It is been recorded that significant expense chance in the horticulture part constantly bound ranchers to endeavor suicide. The value figures are helpful for homesteads, policymakers, and agribusiness ventures. Utilizing time arrangement information in horticulture, nonstop endeavors are made by the analysts to foresee the costs utilizing different straight and nonlinear anticipating models. These days, Artificial knowledge/Machine learning models are utilized as traditional measurable models in the gauging exercise. Utilizing AI strategies in this way, utilizing authentic cost and information factors. Henceforth this exploration will assume a significant job to anticipate the cost of onion. Just as help to agriculturists to turn out to be financially well and has a more advantageous existence.

**Keywords:** *Indian Agriculture, Precision Agriculture, Machine Learning, Artificial Intelligence, Applications, Demand Forecasting, Micro level, Macro-level.*

## 1. Introduction

The history of **India Agriculture** goes encircling to Indus Ravine Refinement before that in certain spots of Southern India. India has the second position in the world in ranch yields. According to 2018, agriculture utilized moiety of the Indian work power and contributed 17–18% to the nation's GDP [5]. In 2016, agribusiness and unified areas like creature farming, ranger

service, and fisheries represented 15.4% of the GDP (total national output) with about 31% of the workforce in 2014. India positions first on the planet with the most noteworthy net edited territory followed by the US and China. The productive relevance of horticulture to India's Volume is staunchly on the wane close to the nation's wide-based fiscal prepayment. Moreover, farming is geographically the largest financial segment and presumes a critical job in the normal financial architecture of India. India traded \$38 billion worth of agrarian items in 2013, which is making it the 7th biggest rural exporter in the world and the 6th biggest net exporter. The vast majority of its agribusiness trades serve to create and least created countries. Accuracy agribusiness is a proposal for ranch the board that applies data innovation to ensure that crops just as soil get pleasantly what they requirement for standard wellbeing and profitability.

Accuracy agribusiness, as the name infers, implies the utilization of exact and right measures of data sources like water, manure, pesticides, and so on at the right time to the harvest for expanding its profitability and augmenting its yields. Accuracy farming administration practices can fundamentally decrease the measure of supplement and other harvest inputs utilized while boosting yields. Ranchers consequently get an arrival on their venture by saving money on water, pesticide, and compost costs. The second, bigger scope advantage of focusing on inputs concerns ecological effects. Applying perfect measure of synthetic concoctions in the correct spot and at the opportune time benefits yields, soils, and groundwater, and in this manner the whole harvest cycle. Subsequently, exactness horticulture has gotten a foundation of maintainable agribusiness, since it regards harvests, soils, and ranchers. Reasonable horticulture looks to guarantee a proceeded with the supply of nourishment inside the biological, financial, and social limits required to continue creation in the long haul. At last, "Performing right thing on an opportune spot in the right manner on ideal time." To settle on choices for the achievement of above targets, an essential recommendation is to observe the harvest state and their feeling from first to last under various distinctive atmosphere conditions and after examination intentional the gathered information to the AI system, which will utilize the information and in the wake of preparing it will make viable forecasts of yield.[8]

Agribusiness makers are progressively inclined to climate hazard and value dangers than some other area and explicitly little and negligible ranchers who frequently does the trouble offer of their produce. Since a decade ago, high unpredictability in the costs of rural wares has been a worry for the two makers and buyers. Ranchers are in trouble under both the circumstances of yield disappointments and guard creations. During guard creation, because of overabundance in the market, item costs tumble to the absolute minimum hitting the ranchers' pay extremely hard [4][7].

As of now, **Artificial Intelligent / Machine learning models** have built up themselves as genuine contenders to old-style measurable models in the estimating network. For value forecasts, explicitly **AI** gives a novel method for consolidating specialized investigation and principal examination strategies. While specialized investigation exclusively takes a gander at chronicled costs, crucial examination comprises of understanding the outer and inside components that impact the cost of a specific stock. Independently they have qualities; specialized examination gives exact transient expectations while the crucial investigation is increasingly fit toward a long haul viewpoint. By joining the two, more precision in forecasts can be accomplished. Thus this exploration depends on anticipating the costs of Onion Applying AI procedures in this way utilizing recorded cost and information factors [9].

Underneath measurement speak to that the absolute territory under onion good and bad times across India from monetary years 2010 to 2019. In the budgetary year 2012, the complete cost was extremely low, however, in the monetary year, all-out cost becomes high. It is seen that the cost of onion is incrementing and diminishes capriciously that is the reason we need computerized reasoning and AI with the goal that we can foresee the cost across the whole nation from the following year's.[1][2]

Market Year	Maximum price In Rupees (Per quintal)	Minimum price In Rupees(Per quintal)	Model price In Rupees (Per quintal)
2010	1500	1200	1390
2011	3600	3000	3360
2012	750	500	665
2013	1200	850	1080
2014	1300	1030	1180
2015	1700	980	1420
2016	1500	830	1370
2017	1300	640	825
2018	3100	2730	2900
2019	1160	925	1050

Fig (1): Price of the onion from 10 years.

By utilizing the above table we can anticipate the up and coming cost of the onion by utilizing AI calculations.

As the above table contain the most recent 10 years' information about onion cost. The above table additionally is helpful for the graphical portrayal of the given information with the goal that one can without much of a stretch comprehend the distinctions in cost.

We draw a graphical portrayal of the costs of the onion since the most recent 10 years (figure [2]). with the assistance of this chart, we can without much of a stretch comprehends the good and bad times of the cost. AI depends on the past information on the off chance that we have earlier year's information so we can foresee the future costs of onion all the more precisely.

As we probably are aware the forecast depends on recently utilized information and expectation might be valid. Be that as it may, on the off chance that we use AI, at that point, it will expand the odds to anticipate the more precise worth.

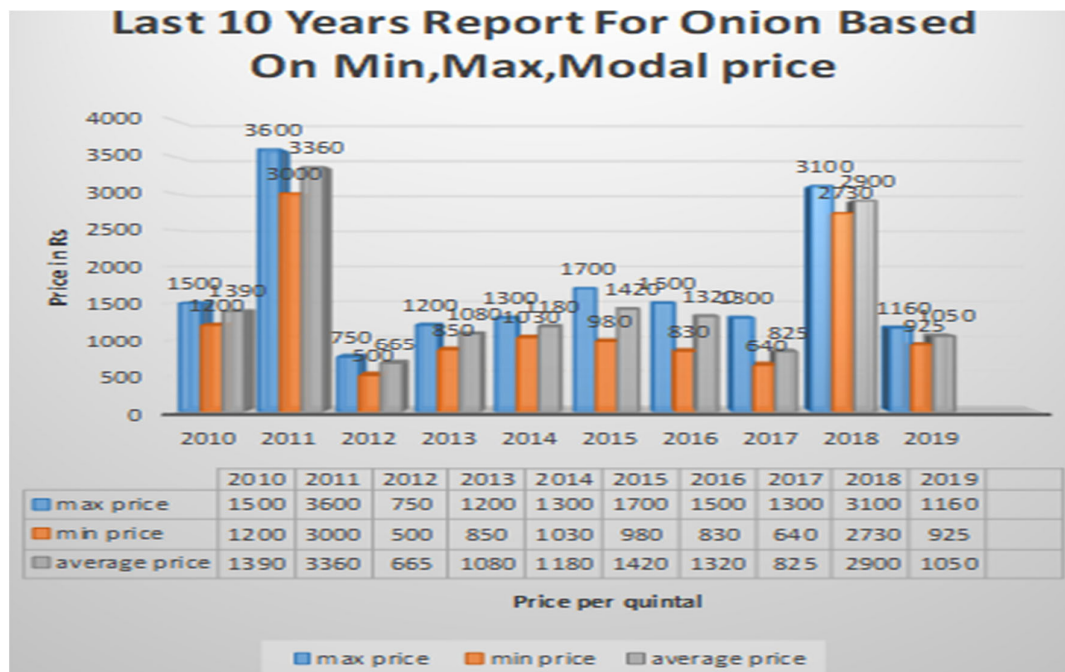


Fig (2): Last 10 years Report for onion based on min, max, Modal price [2]

## 2. Research Methods

### 2.1.1 Machine Learning

Machine Learning (ML) is the coherent examination of counts and quantifiable models that the PC uses to play out a specific task without using unequivocal rules, contingent upon past information. AI calculations fabricate a numerical model dependent on test information, known as "training data", to settle on forecasts or choices without being expressly modified to play out the assignment. AI calculations are utilized in a wide assortment of utilizations, for example, email separating and PC vision, where it is troublesome or infeasible to build up an ordinary calculation for adequately playing out the assignment.

Machine learning is identified with computational statistics, which focuses on making predictions using computers. The investigation of numerical enhancement conveys techniques, hypotheses, and application areas to the field of AI. Information mining is a field of study inside AI and spotlights on exploratory information investigation through unaided learning. In business problems, machine learning is also used as predictive analytics. Applying AI strategies with suitable parameters government will have the option to heighten their onion value creation. AI approaches which are existing today can be utilized in the value forecast of onion. The decision for the classes of calculations depended on a writing audit for value estimating techniques. For time arrangement gauging, we chose ARIMA (Auto-Regressive Integrated Moving Average), for choice trees, we applied Random forest and LASSO (least absolute shrinkage and selection operator), SVM (support vector machine) and GLM (generalized linear model) for Regression.

### 2.1.2 Types Of Machine Learning

We can arrange Machine learning calculations dependent on the supported result of the calculation. General calculation types comprise of:

#### a. Supervised Machine Learning

Supervised machine learning algorithms are based on the past value to gain new information utilizing marked guides to anticipate future occasions. Beginning from the examination of a known preparing dataset, the learning calculation delivers a gathered capacity to make forecasts about the yield esteems. The system can train based on the input data. The learning calculation can likewise contrast its yield and the right proposed yield and discover mistakes to alter the model as needs.

#### b. Unsupervised Machine Learning

Unsupervised machine learning algorithms are used at the point when the data used to prepare is neither classified nor labeled. Unsupervised learning concentrates on how frameworks can gather a capacity to depict a concealed structure from unlabeled information. The system doesn't find the correct output, but it finds the data and can predict from datasets to describe hidden structures from unlabeled data.

#### c. Semi-Supervised Machine Learning

Semi-directed AI figuring's fall somewhere at the hour of overseeing and solo learning except if they utilized both stamped and unlabeled data for planning – consistently a humble amount of named data and a ton of unlabeled data. The systems that use this method can altogether improve learning precision. For the most part, semi-supervised learning is picked when the obtained marked information requires gifted and pertinent assets to prepare it/gain from it.

#### d. Reinforcement Machine Learning

Reinforcement machine learning algorithms is a learning strategy that speaks with its condition by making exercises and discovers bumbles or remunerates. Experimentation search and delayed prize are the most appropriate properties of support learning. This technique licenses

machines and programming administrators to, therefore, choose the ideal direct inside a specific setting to expand its introduction. Direct prize information is required for the administrator to acknowledge which movement is perfect; this is known as the support signal. Basically in support learning, five sorts of components exist let me show you with the assistance of figure (3).



Fig (3): Elements of reinforcement learning

When an agent interacts with the environment, the environment gives him a state then the agent performs a particular action towards the environment based on the action environment rewards him. That could be positive or negative [9].

### 3 Result and Discussion

#### 3.1.1 Algorithm Of Machine Learning In Price Prediction Of Onion

The performance of different ML algorithms strongly depends on the size and structure of the input data. Thus, we have to test many algorithms so that based on the trails we can choose the correct algorithm which gives a more accurate result. For Onion, the total dataset consisted of data points from 2010 to 2019. To shortlist the right algorithm, we trained them on the dataset from 2010 to 2018 and tested the results on 2019 data. For time series forecasting, we use ARIMA (Auto-Regressive Integrated Moving Average), for decision trees, we select Random forest and LASSO (least absolute shrinkage and selection operator), SVM (support vector machine) and GLM (generalized linear model) for Regression. The Root Mean Square Error (RMSE) was calculated during the training and testing of each algorithm. RMSE is based on the standard deviation of the residuals (prediction errors). Residuals are calculated with the help of data points that data points are how far from the regression line. In other words, it tells us how concentrated the data is around the line of best fit [6].

#### 3.1.2 Models For Price Prediction

- a. ARIMA : It is an abbreviation that represents the Auto-Regressive Integrated Moving Average. It is an off kind of model that catches a suite of various standard worldly structures in time arrangement data.[10]
- b. RANDOM FOREST : These are a gathering learning technique for arrangement, relapse, and different undertakings, that work by developing a huge number of choice trees at preparing time and yielding the class that is the method of the classes (characterization) or mean expectation (relapse) of the individual trees.
- c. LASSO : It is an abbreviation for the least outright shrinkage and choice administrator. It is a relapse investigation strategy that performs both variable determination and regularization to improve the forecast precision and interpretability of the factual model it produces.

- d. SVM : Support Vector Machine is a directed AI calculation that can be useful for both characterization and reduce difficulties. In any case, it is generally utilized in grouping issues. Right now, data thing is plotted as a point in n-dimensional space (where n is the number of features) with the estimation of every component being the estimation of a facility. At this, the order is performed by finding the hyper-plane that separates the two classes well overall.
- e. GLM : the summed up straight model (GLM) is adaptable speculation of common direct relapse that takes into account reaction factors that have blunder appropriation models other than an ordinary conveyance. The GLM sums up direct relapse by permitting the straight model to be identified with the reaction variable employing a connection work and by permitting the greatness of the fluctuation of every estimation to be an element of its anticipated worth.

#### 4. Conclusion

This examination exhibited that AI systems can be effectively applied to figure costs of onion if the solid memorable information at the different cost determinants is accessible. In my exploration, I am utilizing 10 years of authentic information (for all impacting factors) from 2010 to 2019[3]. The examination incorporated the determination of the correct sort of calculation from various classes. Tether (a class of relapse techniques) has demonstrated to be the most appropriate calculation for the 15-and 30-day interims. Going ahead we expect to research how to additionally improve the calculation by investigating gathering techniques. The extent of this exploration was the advancement and testing of models for onion. Going ahead the accompanying advances has been talked about:

1. Make prescient models for other significant harvests, for example, Mustard, Arahara, Urad, and Chana.
3. Improve the present onion value expectation model by incorporating extra affecting components and research on how to improve oneself learning capacity of the calculation.
4. The job of MSP in affecting the costs ought to be explored further since it is the main deciding element that the policymakers can legitimately impact. This should be possible by giving recreations to evaluate the adjustment in costs and this manner could be utilized as a dynamic device for computing the perfect MSP.
5. A free investigation ought to be attempted to gauge the impact of the value expectations at the smaller scale level on ranchers. An examination will help see how the conduct of a rancher could change and give an approval of the forecast's utility.

#### References

- [1] <https://agmarknet.gov.in/>
- [2] <https://www.indexmundi.com/agriculture>
- [3] [https://en.wikipedia.org/wiki/Precision\\_agriculture](https://en.wikipedia.org/wiki/Precision_agriculture)
- [4] <https://agropedia.iitk.ac.in/content/precision-agriculture-glance>
- [5] [https://en.wikipedia.org/wiki/Economy\\_of\\_India](https://en.wikipedia.org/wiki/Economy_of_India)
- [6] [https://en.wikipedia.org/wiki/Machine\\_learning](https://en.wikipedia.org/wiki/Machine_learning)
- [7] <https://www.omicsonline.org/open-access/agriculture-role-on-indian-economy-2151-6219-1000176.php?aid=62176>
- [8] Subhadra Mishra, Debahuti Mishra and Gour Hari Santra "Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper" Indian Journal of Science and Technology, Vol 9(38), October 2016

[9] Ankush Kumar Gaur, Dr. J. Arul Valan, "Impact of Machine Learning applications for intensification of rice production in India", IEEE sponsored 3<sup>rd</sup> International Conference on Engineering & Technology 2018 organized by the Department of CSE & IT, Rathinam Technical Campus, Coimbatore 30<sup>th</sup>& 31<sup>st</sup>, August 2018.

[10]<http://agropedia.iitk.ac.in/content/precision-agriculture-glance>