PROBLEM STATEMENT

TITLE: ESTIMATION OF CROP YIELD USING DATA ANALYSIS

IDEAL SOLUTION :

It's not a coincidence that the science of teaching computers to learn and create models for predictions is so widely applied. The global economy depends heavily on the agricultural sector. Understanding global agricultural output is essential to addressing issues of food security and minimising the effects of climate change as the human population continues to grow.

Predicting crop yields is a significant agricultural issue. Weather factors (rain, temperature, etc.), as well as pesticides, are the main determinants of agricultural productivity. Making judgments regarding agricultural risk management and forecasting requires accurate knowledge of crop yield history.

Although food is prepared very differently all around the world, the fundamental elements that keep people alive are quite similar. We consume large amounts of rice, corn, wheat, and other basic crops. In this project, the World Data Bank and FAO make their publicly available data used to anticipate the ten most eaten crops.

Regression analysis is a type of predictive modelling technique that looks at how an independent variable (s) and a dependent variable (target) are related (predictor). The project's regression models

- Gradient Boosting Regressor
- Random Forest Regressor
- SVM
- Decision Tree Regressor

It is the proportion of the variance in the dependent variable that is predictable from the independent variable, where it is a statistical measure between 0 and
1 which calculates how similar a regression line is to the data it’s fitted to. If it’s a 1, the model 100% predicts the data variance; if it’s a 0, the model predicts none of the variance.

Factors include humidity, sunlight and factors involving the climate. Environmental factors refers to soil conditions. In this model two climate and one environmental factors are selected, rain and temperature. In addition to pesticides that influence plant growth and development. Rain has a dramatic effect on agriculture, for this project rainfall per year information was gathered from the World Data Bank in addition to average temperature for each country.

REALITY:

In order to address issues with food security and lessen the effects of climate change, it is essential to comprehend global agricultural productivity. One of the top UN Sustainable Development Goals for 2030 is to end hunger, and this goal can assist. In the project, we present a scalable, precise, and lowcost approach to agricultural production prediction utilising openly accessible remote sensing data and machine learning. Several months prior to harvest, our deep learning system can estimate crop yield using just globally accessible covariates with high spatial resolution (county-level). We think our technique may be useful for setting suitable food reserve levels, identifying low-yield locations, and enhancing risk management of derivatives related to crops.

PROBLEM STATEMENT:

1. **Effects of climatic conditions on crops**

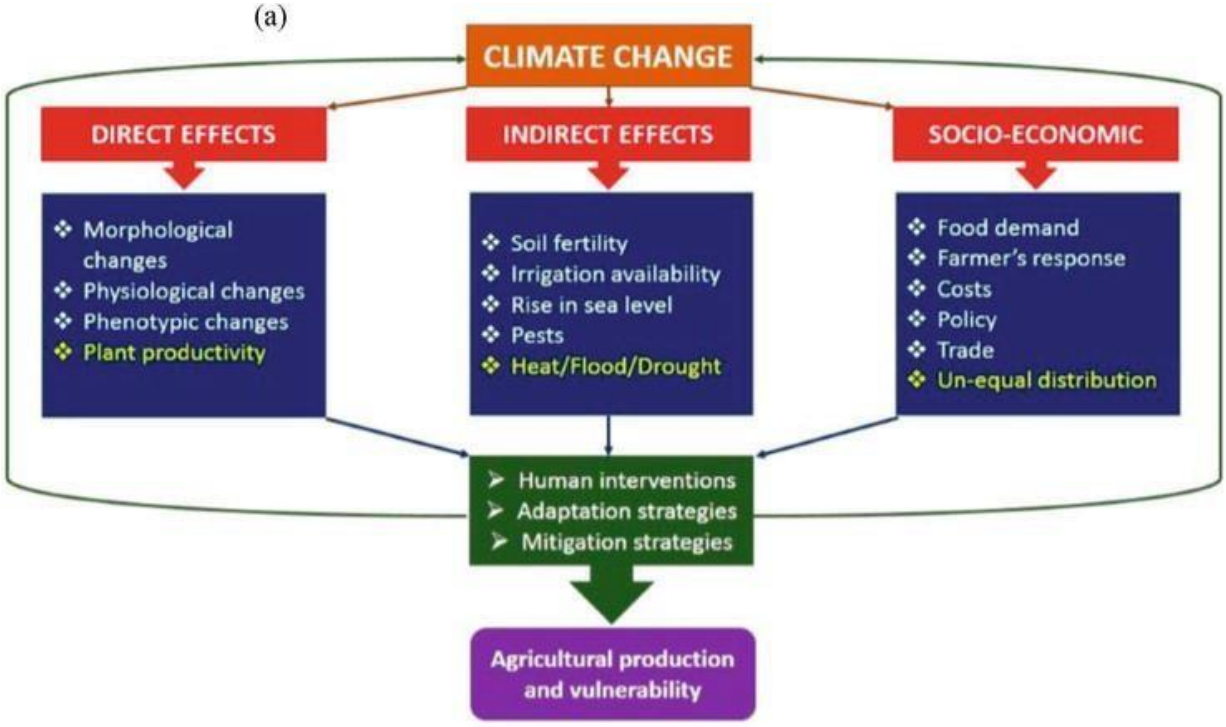
Variations in annual rainfall, average temperature, global increase of atmospheric CO2, and fluctuations in sea levels are some of the major manifestations of climate change, which negatively impact crop yields. Temperature and rainfall changes are expected to significantly have negative impact on wide range of agricultural activities for the next few decades. With the changing of climate, agriculture faces increasing problems with extreme weather events leading to considerable yield losses of crops. Most often, crop plants are sensitive to stresses since they were mostly selected for high yield, and not for stress tolerance. Climate change is the result of global warming. It has devastating effects on plant growth and crop yield which can affects directly, indirectly, and socio-economically reduce crop yields by up to 70% . Weather variations present positive and negative effects in the environment with very high expression of negative effects.

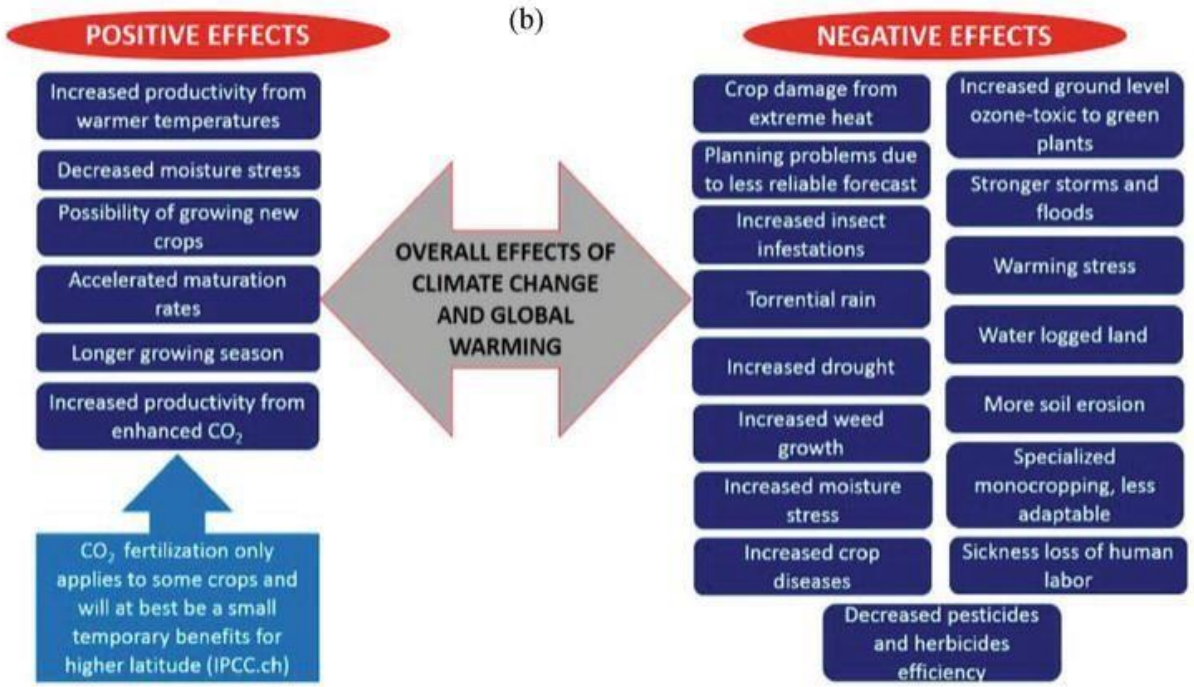
SOLUTION:

The regression analysis model between historical climatic data and yield data for food crops over the last 30 years in Nepal showed an increase in temperature of approximately 0.02–0.07°C per year in different seasons and a mixed trend in precipitation. Additionally, no significant impact of climate variables on yields of all crops was observed and the regression

analysis revealed negative relationships between maize yield and summer precipitation, between wheat yield and winter minimum temperature, and finally positive relationship was observed between millet yield and

summer maximum temperature.





2. impacts due to season changes

In India crop yield is season dependent and majorly influenced by the biological and economic cause of an individual crop. Reporting of progressive agriculture yield in all the seasons is an ample task and an advantageous task for every nation with the respect to assesses the overall crop yield and prediction and estimation at present a common issue world wide is farmers are stressed in producing higher crop yield due to influence of unpredictable climatic changes and significant reduction of water resources world wide. A study was carried out to collect the data onworld climatic changes and the available water resources which can be used to encourage advanced and novel approaches such as big data analytics to retrieve the information of the previous results to the crop yield prediction and estimation. Study imported that the selection and usage of the most desirable crop according to the existing conditions, supports to achieve the higher and enhanced crop yield,

SOLUTION

The accurate prediction of crop yield certainly benefits the farmers in choosing the right method to reduce the crop damage and gets best prices for their crops. A research group conducted a work with an objective of accurate prediction of crop yield through big data analytics to assess various crop yield influencing factors such as Area under Cultivation (AUC) interimsof hectors, Annual Rainfall (AR) rates and Food Price Index (FPI) and to develop relationship among these parameters. Regression Analysis (RA)

methodology was applied to examine the selected factors and their impact on crop prediction and final yield. RA methodology is a multivariable investigation practice which can categorize the factors in to groups such as explanatory and response variables and helps to assess their interaction to obtain a resolution. All the selected factors of the present study design known as AR, AUC and FPI were measured for a period of 10 years between the years 1990-2000. A novel method called Linear Regression (LR) is applied to analyze the relationship between explanatory variables (AR, AUC, FPI) and the crop yield considered as response variable. Study reported that the R2 value for the studied factors clearly indicate that crop yield is principally depends on AR. Study also reported that the other two factors (AUC and FPI)screened were also found to have significant impact after the AR. Study shall be continued to analyze the impact of for other substantial factors like Minimum Support Price (MSP), Cost Price Index (CPI), Wholesale Price Index (WPI) etc. and their relationship on the yields of different crops [12].

Crop yield gaps, measured as difference between expected yields based on the potency and actual farm yield received. In order to achieve the higher crop yield, farmers must need to tackle the influencingfactors such as influence of change in climate conditions on the prospects of crop yields, andchange in the usage of agricultural land to assess and ultimately reduce the crop yield gaps. Several researchers reported the applications of bio simulation models to estimate the crop yield gaps in the last decade. The impact of the crop yield gaps assessment studies conducted through bio simulation based methodologies were negatively influenced by quality and resolution of climate and soil data, as well as unscientifically expectations about crop yield prediction systems and crop yield assessment modeling

designs calibration method. An explicit rationale model which can effectively applied at various levels of the availability of quality information for identifying data sources to analyze crop yield and measuring yield gaps at definite geographical locations and works based on the rise in titer approach. The model is highly helpful in retrieving the useful data from the available, poor quality, less rigorous data sources or if the data is not available. A case study was discussed on the application of selected model design to quantify the yield gaps of maize crop in the state of Nebraska (USA), and also at the different geographical locations representing the nations Argentina and Kenya at national scale level. Different geographical locations such as Nebraska (USA), Argentina and Kenya were identified to symbolize the distinct scenarios of Agri based data availability and the quality for the selected variables assessed to predict and estimate the crop yield gaps. The definitive aspiration of the planned method is to afford transparent, easily accessible, reproducible and technically sound and strong guidelines for predicting the yield gaps. The proposed guidelines were also relevant for understanding and to simulate the influence of change in climate conditions and usage of cultivable land changes from national to global scales. As indicated, the better understanding of data.

CONSEQUENCES:

The crop maize, often known as corn, is farmed all over the world, with regions closer to the equator producing the bulk of the crop. As average temperatures rise in these breadbasket countries, greater stress will be placed on the plants, which might lead to a drop in maize yields across North and

Central America, West Africa, Central Asia, Brazil, and China in theupcoming years and beyond.

As temperatures rise, a wider region, including the Northern United States and Canada, the North China Plains, Central Asia, Southern Australia, and East Africa, may be able to grow wheat, which thrives in moderate climes. However, these gains may level off by the middle of the century. The models don't just focus on temperature while simulating

CONCLUSION:

In conclusion, the introduction of technology into the agricultural sector has led to a minor rise in productivity. Technology improvements have made new concepts like precision agriculture, smart farming, and digital agriculture viable. In the literature, it has been highlighted that assessments of agricultural productivity, the discovery of concealed patterns using data sets connected to seasons and crop yields data, have been carried out. We have observed and analysed the several crops that are grown, as well as their area and production rates in various states and districts, using IBM Cognos. A few of these are:

Production-rate averaging seasons. This analytics teaches us which seasons experience a rise in productivity on average and which experience a drop.
Yield-average seasons.
Our understanding of which seasons see an increase in average productivity and which experience decreases comes from this data. Production, second, by crop year. Our goal in conducting this research is to boost crop productivity.