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

Suggestive Crop Modeling System Incorporating Statistical Inference Analysis

Dr Nagaraj G Cholli
Department of Information Science & Engineering
RV College of Engineering,
Mysore Road
Bengaluru-560059, India

Dr. Amarjeet Singh
Department of Computer Science & Engineering
New Horizon College of Engineering
Ring Road, Bellandur Post,
Bengaluru-560103, India.

INTRODUCTION

It is more beneficial and fruitful to represent the underlying data in the form of graphical plots such as Bar Graphs, Line Graphs, Pie Charts, etc. than to represent it the form of a simple table of numbers and raw data which is bound to confuse the end user.

Over the past recent years, there has been a growing trend of Data aggregation and analyis has proven its potential for effective communication of information. This trend is fuelled by increased availability of statistical inference tools and a spike in its use. News outlets experimented with new and attractive ways of presenting information to the public. This has raised questions on what benefit the data aggregation can offer in order to communicate information more effectively. In order to begin any Data modeling, the agricultural statistics is the main foundation to begin with.

The world population is increasing day by day and it is estimated that in the near future, if proper planning is not done, we may run into a food shortage. This is especially true for developing countries whose population is very high compared to the rest of the world and which do not have access to advanced infrastructure compared to the rest of the developed countries in the world.

This problem can be resolved by utilizing the available data in order to analyse the food shortage problem and make proper preparation well in advance in order to meet the agricultural production requirements. Combining the existing data available to the public and the data on weather patterns, we can forecast the future requirements for agricultural production and plan accordingly.

An agricultural census is needed to improve estimates between censuses and provide detailed data on small areas not provided in the survey data. To do this, the data must be comparable, that is, standardized definitions, registration units and data collection methods.

PROBLEM STATEMENT

Crops require a certain temperature, soil, rainfall amount and water supply in order to grow properly and in order to produce the maximum yield output. We have come up with an idea to design an application which analyses the previous year's datasets of temperature, soil and rainfall precipitation in order to come up with a modeling system which suggests the most feasible set of crops to grow. It is up to the cultivator to decide based on the suggested statistics in order to maximize his yield and profit. Complete documentation is useful in finding out what may cause estimation errors and help identify uncertain ways to improve the estimation

LITERATURE SURVEY

Over the past recent years, there has been a growing trend of Data aggregation and analysis has proved its potential for effective communication of information. This trend is fuelled by increased availability of data visualization tools and an increase in its use.

While in the past, statistical data was regarded as an essential analytical device for researchers, it is quickly being identified as an important factor of effective research communication. Although statistics is fairly new for development researchers, it affords possibilities to both remodel and display statistics (Crop yield statistics)[3]. Aggregation proponents additionally highlight that those capabilities are extremely beneficial within complicated and changing environments, which are comparable to the contexts surrounding IDRC-supported projects. Data visualization offers precise statistical information for self-organizing maps [2].

The significance of statistics visualization is further heightened through the growing digitization of the world, which has created statistical-overloads in a time-deprived coverage and improvement sector. For an instance of diverse handling strategies, see e.g. [5]. Our technique of the use of SOMs is influenced through the need to better understand the available yield data and extract knowledge from those records. However, whilst the collection of this records is frequently ground-breaking and innovative, the expounded findings nevertheless should be heard within saturated statistics markets.

Data aggregation and analysis is both an art and a technological know-how. [3] It is considered as a branch of descriptive statistics by some, but also as a grounded concept development tool by others. Increased quantities of records created through Internet activity and an increasing number of sensors in the environment are referred to as "big data" or Internet of things. Processing, studying and communicating this statistics present moral and analytical challenges for statistical inference. The area of data science and practitioners called data scientists help deal with this challenge. [4]

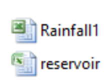
Today, the world is experiencing another surge in statistics visualization popularity. This interest can be partially connected to the increased availability of new technology and software products which allow each user to dabble in the world of visualization. However, these resources have not come about on their own, but had been the by-products of years of studies and development from an worldwide community of scholars and practitioners. In a recent publication, that said, exploring these 3 streams provides a stronger understanding of the rich and diverse scholarship which has contributed to the sector of statistical inference.

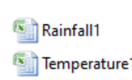
Farms are consolidating at an increasing rate as technology supports automation and economies of scale. Input applications are primarily based on factual statistics and investments into farming tech are funded by profit saved by data-driven efficiency. While the advantages of digital agriculture are compelling, it has been met by significant challenges, for example, difficulty using software, data utilization concerns, disparate and propriety data formats and an unclear return on investment. Agribusiness has struggled to offer immediate, tangible outcomes from digital agriculture equipment and software.

DESIGN

The Statistical analysis application assumes the following data sets mentioned below:

- Temperature datasets for the present and next year
- Rainfall data sets for the present and next year
- Reservoir water levels and planned outflow cusecs

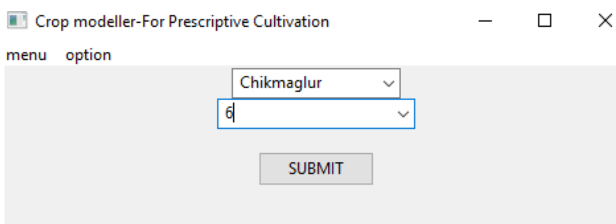




These datasets contain the temperature readings and rainfall in enimeters and reservoir outflow amount information in each month for the present and following year. The duration would be selected based on the month of information publication and the statistical inference is carried out for the feasibility.

FRONT END

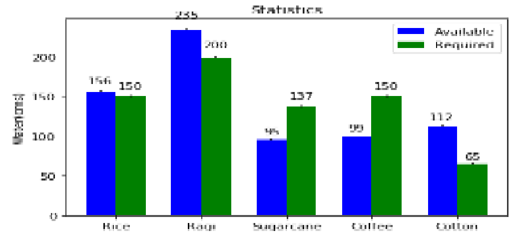
The view provides a facility to choose the district, an option to browse and choose the data file and the option to choose the month from which the cultivatable statistics have to be taken into account and visualized. This is being designed using *wxPythonPhoenix.*



The above figure shows the selection being made on the application GUI. The interface consists of two windows which give general information of Crop modeling statistics and also a selection tab containing drop down menus. The district is selected from the list and the results of statistical inference are displayed in console.

STATISTICAL INFERENCE

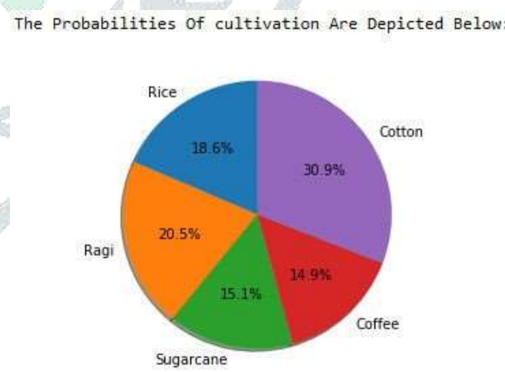
The Modeling system filters the feasibility of growth of crops based on the input datasets. In the first stage we perform removal of non-cultivable crops based on soil availability in the region assumed. In the second stage, we eliminate more crops based on temperature requirements and availability obtained from the datasets. The third stage involves analysing water availability (aggregating both rainfall and reservoir availability) and obtaining the final list of cultivable crops. The statistical data would be subjected to analysis which calculates the standard deviation of availability vs requirement and this can be used to obtain a feasibility score. A ranking score is assigned, which work as follows- If the deviation is more from the standard value, less is the probability of successful cultivations and vice versa.



The above figure shows the analysed results of water statistics as a bar graph plot. If the deviation of availability is more than the required range, the respective crops are eliminated from the list. The others are considered feasible and displayed.

RESULTS

The data thus analysed would be projected and visualized using bar graphs and pie charts that indicate the required and available values of water availability and temperature readings. The ratios of the available and required values would indicate the probability of growth. If the value is nearer to 1, that particular crop would have the highest chances of cultivation.



The above figure shows the inference analysis result carried out based on the temperature water requirements, and is displayed as a probability ratios for growth. The input for the modeling system is the district name selected from the drop down menu in the interface of the application which here, is selected as Chikmaglur, Bengaluru, India. The value with a higher probability ration is inferred as the crop with the higher chances of sustainable cultivation. The standard deviation of the requirement from availability is inferred as a measure of success for the crop. Lesser the difference, then more is the feasibility of growth and vice versa.

CONCLUSION

This survey aims at promoting smart farming mechanisms by prescribing possible crops by incorporating statistical data visualization techniques. This will benefit cultivators in a viable manner based on the regional parameters. This promotes prescriptive crop cultivation based on smart farming and digital analysis methods. The key role of this project is to promote more success rates of cultivation and prevent losses arising from unplanned cultivation while also benefiting the farmer with commercial profits. Thus it helps in establishing a balance of growth and profit rates of crop cultivation.

The charts and graphs developed and presented in this paper were produced using a series of frameworks and python libraries. These programs and frameworks are all useful technologies, but they are most useful when information designers think critically about how best to display the data of interest. When looking at historical data, researching and applying best principles and appropriate technologies to create well-designed graphics can help data-users find the information they want, and learn more about the big picture, and the story behind the data.

REFERENCES

([1] NASS Research, Science and Technology website. (http://www.nass.usda.gov/Research_and_Science/)

[2] Timo Honkela, SamuelKaski, KristaLagus, and TeuvoKohonen. WEBSOM—self- organizing maps of document collections. In Proceedings of WSOM'97, Workshop on Self- Organizing Maps, Espoo, Finland, June 4-6, pages310–315. Helsinki University of Technology, Neural Networks Research Centre, Espoo, Finland, 1997.2.

[3] T.Jensen, A.Apan, F.Young, and L.Zeller. Detecting the attributes of a wheat crop using digital imagery acquired from a low-altitude platform. Compute. Electron. Agric., 59(1-2):66– 77,2007.

[4] Timo Honkela, SamuelKaski, KristaLagus, and TeuvoKohonen. WEBSOM—selforganizing maps of document collections. In Proceedings of WSOM'97, Workshop on Self Organizing Maps, Espoo, Finland, June 4-6, pages310–315. Helsinki University of Technology, Neural Networks Research Centre, Espoo, Finland, 1997.

[5] M.SchneiderandP. Wagner. Prerequisites for the adoption of new technologies - the example of precision agriculture. In Agricultural Engineering for a Better World, Dusseldorf, 2006.¨VDIVerlagGmbH.