Analysis of Geolocation Dataset and Fertiliser Availability to Farmers at Minimum Cost

Chintan Rajvir, Rajasekaran Rajkumar, Jolly Masih, Paviter Singh Matharu

Abstract: ***According to World Bank's survey, India has the largest area of arable land of approximately 156.4 million hectares which is about 57% of total land in the country. The demand for agricultural products is high in India which corresponds to high use of fertilisers. Both natural and synthetic fertilisers are equally predominant. The primary issue with fertilisers is how to buy the fertilisers with lesser cost where cost of travelling and transportation plays a major role. The poor market research and unawareness of the farmers makes them vulnerable while buying fertilisers. They levy unnecessary costs on fertilisers and increase their expenditure for the yield. We analysed this growing concern for a vast agricultural country like India and came up with proper insights and solution. Our solution involves analysis of distances between geolocations or places where fertilisers are available (e.g. fertiliser shop or retail store for fertiliser) and allowing the farmers to go select amongst those available storesthe nearest and best place to purchase the fertilisers.***

Keywords: Geolocation Dataset, agricultural products

I. INTRODUCTION

Farming is one of the the major occupation in India. Farming is one of the oldest economic pillar of our country. Different regions have different farming practices. However, the agricultural and farming practices have considerably evolved over the years with changes in weather conditions, technological innovations and socio-cultural practices.
1

Farming strategies prevailing in India may be classified as Primitive Subsistence Farming, Intensive Subsistence Farming, and Commercial Farming. Another technique of economic farming is 'plantation'. Plantation farming may be a mix of agriculture and trade, practiced across a massive space of land. It's a labour-intensive farming technique that additionally uses latest technological support for sustaining, cultivating and yielding. The crop yield obtained from plantationare used as raw materials in their several industries e.g. cotton plantation, tea plantation, coffee plantation etc..

Various Farming techniques demands need of different types of fertilizers like:

1) Straight Nitrogenous – Ammonium Sulphate and Chloride, Urea, Calcium Ammonium Nitrate

2) Straight Phosphatic – Single and Triple Super Phosphate

3) NP/NPK Complex fertilizers – Urea Ammonium Phosphate, Ammonium Phosphate Sulphate, DAP, MAP, etc.

Among these, the widely used fertilizers as of 2017-18 are Urea, DAP, and NPK fertilizers while these fertilizers are also among the most produced fertilizers in India.While India produces large amount of fertilizers, all of these are also imported with ports like Kandla, Mundra, and Vishakhapatnam being the most important ports for importing these fertilizers. Of all the imported fertilizers, Urea and MOP are being the costliest as of 2012-13 annual agricultural year.

Indian sub-continent has majorly 2 agricultural seasons: Kharif and Rabi. The Kharif season roughly lies from June to October which corresponds to the South-West Monsoon winds. Correspondingly follows the Rabi season from November to March which is aided by irrigation systems as well as the rainwater that has percolated through the ground after the monsoon is over. Crops like Rice, Maize, Pulses, Cereals, Cotton, Oilseeds, etc. are preferably grown in the Kharif season while Wheat, Barley, Oats, Mustard, Linseed, etc. are grown in Rabi season. The fertilizers used in each of these periods is also subsequently different. Fertilizers like Urea and DAP (Diammonium Phosphate) mostly available and used in Rabi season while MOP and NPK fertilizers are used in Kharif season. NPK fertilizers consumption is second highest in India after China. All these facts indicate that the use of fertilizers greatly determines the crop yield and in turn allows the farmers to reap maximum profit. Good agricultural yield is necessary to bring prosperity to farmers and also to ensure food supplies to the huge population of India. Geolocation is the co-ordinate of a place on the map. It allows one to distinctively identify position of a place using 2 attributes known as Latitude and Longitude. These attributes can be found using Google APIs for instance. Instead we would be using Chrome Driver and correspondingly Python libraries like Selenium to find the latitude and longitude for any given place or district, saving cost of requesting Google APIs. Given these attributes for any given pair of places, we can calculate the distance between them too. Various APIs exist to perform the distance metric calculation as well. However, we simply prefer Haversine formula which would again be cost effective way instead of requesting the paid APIs.

Application of geolocation and the distance metric calculation is the standard base of this project.

Revised Manuscript Received on June 07, 2019.

Chintan Rajvir, School of computer science and engineering, Vellore Institute of Technology, India.

Rajasekaran Rajkumar, School of computer science and engineering, Vellore Institute of Technology, India.

Jolly Masih, Erasmus School of Economics, Erasmus University, the Netherlands,Mosaic Company India Pvt. Ltd., India.

Paviter Singh Matharu, Erasmus School of Economics, Erasmus University, the Netherlands,Mosaic Company India Pvt. Ltd., India.

II. DATASET

The input data can be tabulated and viewed as: [1]

State Id	State	District	Wholesalers	Retailers
1	Andaman Nicobar	South Andaman	2	2
2	Andhra Pradesh	Anantapur	32	552
2	Andhra Pradesh	Chittoor	33	446
2	Andhra Pradesh	East Godavari	218	1136

…and so on.

The input data is in the ".CSV" format and can be visualized as below:

```
7,Chhattisgarh,Koriya,12,77
7,Chhattisgarh,Mahasamund,24,157
7,Chhattisgarh,Narayanpur,1,8
7,Chhattisgarh,Raigarh,24,202
7,Chhattisgarh,Raipur,88,494
7,Chhattisgarh,Rajnandgaon,28,210
7,Chhattisgarh,Surguja,43,179
7,Chhattisgarh,Uttar Bastar Kanker,28,133
8,Dadra & Nagar Haveli,Dadra & Nagar Haveli,0,1
9,Daman & Diu,Daman,0,1
10,Goa,North Goa,9,53
10,Goa,South Goa,9,43
11,Gujarat,Ahmadabad,65,418
11,Gujarat,Amreli,41,358
11,Gujarat,Anand,91,504
11,Gujarat,Banas Kantha,74,867
11,Gujarat,Bharuch,48,229
11,Gujarat,Bhavnagar,43,586
11,Gujarat,Dohad,27,195
11,Gujarat,Gandhinagar,42,272
11,Gujarat,Jamnagar,50,423
11,Gujarat,Junagadh,56,593
```

Figure 1. Comma-Separated Values format of dataset which includes unique ID of state, name of the state, name of the District, numeber of wholesalers and number of retailers in order

The first value corresponds to the unique ID value assigned to each state in India. Second value corresponds to the name of the state. Third value corresponds to the name of the district in that particular state. The last two integer values represent the number of wholesalers and retailers respectively in the district.

The given dataset contains data for 598 districts all over India.

III. METHODOLOGY

We wish to compute the nearest district as mentioned by the user in order to allow him/her to choose the nearest fertiliser store or shop and to save his travel expenses. However, we also wish to focus on real-time scenarios wherein we could return a district as output which is much closer to the user as well as have a good number of dealers available (as specified by the user: whether he/she wants to purchase the fertilizers in wholesale or retail). In case the user is searching for a district which has dealers (as specified) already present, we simply return the same district as the output. In simple words, we check if there is a certain district in the 50 kilometre radius of the nearest district which could provide sufficiently large number of dealers (either wholesalers or retailers, as specified by the user) in comparison to the nearest district. This shall allow the user to benefit in terms of availability of dealers at time of their visit as well as price of the fertilizers could be bargained more where there is high competition. The 50 kilometres is set as the radius because further travel would increase the cost to the user and thus, travel overhead would overcome the benefits reaped from optimizing the fertilizer prizes. In this way the model aims at providing the most fruitful way for the farmers to buy the fertilizers. Firstly, we convert the given dataset to ".JSON" format with additional properties. For each row we will find the latitude and longitude of the district specified. The latitude and longitude are found using the chromedriver.exe app. As we are using Python 3, we could make use of the selenium library which facilitates in the running the webdriver functions. The driver allows us to open a new Google Chrome tab. We put the google maps URL with the given district and state names as the query in the URL. The motivation behind using this approach is that the URL loaded onto the new window of the browser changes as the search is completely loaded. The new URL now contains the latitude and longitude value for the search query (city and state) specified.

The initial URL looks like: https://www.google.com/maps/place/City,State/

The city and state could be searched as needed by placing it in the above URL. The new updated URL would look like:

https://www.google.com/maps/place/Vellore,+Tamil+Nadu/@12.8992994,79.0483018,12z/data=!3m1!4b1!4m5!3m4!1s0x3bad38e61fa68ffb:0xbedda6917d262b5e!8m2!3d12.9165167!4d79.1324986

As we can see the text highlighted in the new URL is the co-ordinates of the place searched. The selenium library's webdriver function now allows us to fetch this

co-ordinates. This allows us to find the latitude and longitude for each district already present in the dataset. Pre-processing and storing these values will allow us to directly compute the output for the user specified query.

These will be the 2 additional properties for the district. The processed data will be JSON file having a list of dictionary objects. Each dictionary object corresponds to one of the districts present in the dataset. We could visualize the JSON file as:

…. 453}, {"ID": 11, "City": "Rajkot", "State": "Gujarat", "Wholesalers": 85, "Retailers": 851, "Latitude": 22.2734269, "Longitude": 70.6812104}, {"ID": 1 ….

The set of all latitudes and longitudes of each district can be treated as a "Network of Chemical Fertilizer Dealers".

After the pre-processing phase, we are ready to process the user queries. The user can enter their choice of district and corresponding state (within India) from where they wish to start their travel. In any case, if user fails to correctly enter the name of the district or state, the driver tries to use its auto-correction ability while loading the initial URL.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Yet, if the browser fails to search for the specified location, the driver displays the default location of the server which shall host the API. Correspondingly, the code checks if the co-ordinates retrieved from the webdriver are same as that of server's exact location, then it returns an error message asking the user to retry his/her search query. If the query is successfully processed, we shall have now the co-ordinates of the location as specified by the user. More specifically, we can also allow the user to turn on their device GPS to automatically detect their location while we publish the application for public use.

Meanwhile, the received co-ordinates are then used to find distance with each of the co-ordinates of the districts as specified in the processed dataset. We use the Haversine formula to compute the distance between each pair of co-ordinates, correspondingly storing the distance value. The formula can be seen as:

Let radius of earth be (R): 6371 km

We assume the co-ordinates to be (22.2734719,70.7512552) which is Rajkot, Gujarat and (23.0201818,72.4396539)as Ahmedabad, Gujarat. Next, we write each latitude value as radians.

$$\varphi_1 = \frac{\text{Latitude}_1 \times \pi}{180} = \frac{22.2734719 \times 3.14}{180} = 0.3887454 \text{ radians}$$

$$\varphi_2 = \frac{\text{Latitude}_2 \times \pi}{180} = \frac{23.0201818 \times 3.14}{180} = 0.401778 \text{ radians}$$

Followed by, we compute:

$$\Delta\varphi = \frac{(\text{Latitude}_1 - \text{Latitude}_2) \times \pi}{180} = \frac{0.74671 \times 3.14}{180} = 0.013 \text{ radians}$$

$$\Delta\omega = \frac{(\text{Longitude}_1 - \text{Longitude}_2) \times \pi}{180} = \frac{1.6884 \times 3.14}{180} = 0.02947 \text{ radians}$$

$$a = \sin^2\left(\frac{\Delta\varphi}{2}\right) + \left[\cos(\varphi_1) \times \cos(\varphi_2) \times \sin^2\left(\frac{\Delta\omega}{2}\right)\right]$$

$$a = 4.225 \times 10^{-5} + 1.8491 \times 10^{-4} = 2.271567 \times 10^{-4}$$

$$c = 2 \times \text{atan2}\left(\frac{\sqrt{a}}{\sqrt{1-a}}\right) = 2 \times 0.0151 = 0.0302$$

Finally,

The distance between 2 points is (d): $R \times c$ km = $6371 \times 0.0302 = 192.4$ km

This value is computed between the user specified district and each of the districts in the given dataset. Then result is given as output using the approach as specified in the beginning. The user is able to see the place of travel, the number of wholesalers or retailers available (as mentioned by him/her) and the distance to be travelled.

IV IMPLEMENTATION RESULTS

We considered the output for one of the test cases: Bhuj, Gujarat. Subsequently, the output describes the nearest district to be visited computed optimally with respect to distance and type of dealers, whether wholesalers or retailers along with its number and approximate distance of travel in kilometres.

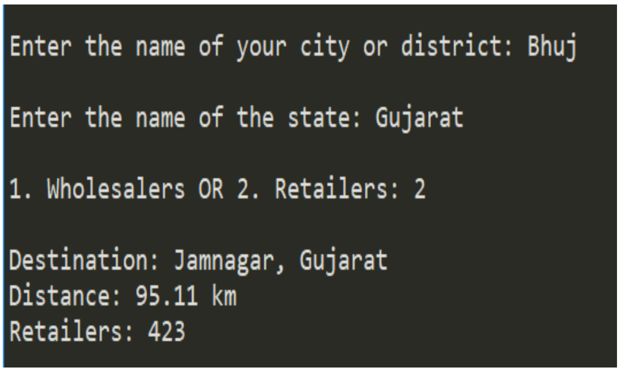

Figure 2. Query regarding a district named Bhuj in the state of Gujarat

V. CONCLUSIONS

Our Paper is aimed at benefitting farmers who do not have proper contacts or whose contact do not have proper affiliation with the Government of India, Agricultural division and are being duped of good fertilizers and seeds. Through our project, farmers can be made aware about the nearest Government certified Fertilizer centres from their location with proper latitude and longitude and they can buy proper fertilizers for their crops. The project yields the nearest fertilizer centre in a 50km radius of their location. Moreover, our project provides options of searching for Wholesale or Retail. The project has potential of boosting our agricultural output since India is primarily an agriculture based country with over 70 percent of the population relying on agriculture as means of income. For future scope, the project can be established as an app, so that farmers can access the data anywhere, anytime hence featuring concept of ubiquitous computing. It will benefit farmers to a very large extent.

REFERENCES

1. https://data.gov.in/ogpl_apis?sort=desc&order=Updated&org_type=Select&sector=Agriculture&org=Select&org_l1=Select&org_l2=Select&page_size=15&ogpl_apis_title=dealers

AUTHOR DETAILS

Dr. Rajkumar Rajasekaran graduated in Electrical and Electronics Engineering from Madras University in 2000 and received Master's degree in Computer Science & Engineering from VIT University, Vellore. He has completed his Ph.D. from Vellore Institute of Technology University, Vellore in Computer Science and Engineering. Her research area includes Data Analytics, Data Mining, Big Data, Machine learning and Data Visualization. He has been teaching from last 17 years with emphasis on Data Science .He is a Life time member of Computer Society of India and American Association for Advancement Science (AAAS). He has published and communicated more than 50 research papers in SCI, SCIE, and Scopus indexed Journals and Conducted many workshops, Faculty Development Programmes and Seminars at various universities and Colleges. He has filed a Patent on Holographic keyboard for smart Devices. Presently, working at Vellore Institute of Technology Vellore and supervising Ph.D.

candidates from VIT University Vellore and from Anna University Chennai. He is Main Editor of the Book entitled "Contemporary Applications on Mobile Computing in Healthcare Settings", "Artificial Intelligence for Social good in Healthcare" published by IGI Global. "Smart Healthcare" by Bentham Science. He is the Deputy Editor Asia for Public Health in Africa (Springer). Editorial Board Member for International Journal of IOT and Cyber Assurance, Associate Editor for International Journal of Social Computing and Cyber Physical Systems (Inderscience). Special Issue Editor for International of Pharma and Bio Science (Scopus) on Smart Healthcare, Scalable Computing Practices and Challenges (Scopus SCI) on Mobile Cloud Computing.He is working on various projects assigned from American Association of Advancement of Sciences (AAAS).He was the Head of Department for Data Analytics, University Coordinator for Apple University Development Program.

Chintan Rajvir Education:
3rd year B. Tech. CSE, VIT Vellore, Tamil Nadu.

Achievements:
Regionals in ACM ICPC for the year 2017 (Kanpur-Kolkata site) and 2018 (Kharagpur site and Amritapuri site).
Software Developer Intern at Microsoft India, Hyderabad. (2019)

Dr. Jolly Masih, Post Doc Scientist Erasmus School of Economics, Rotterdam, The Netherlands

Dr. Jolly Masih is working on Epidemological research in health care and lifestyles including consumer behaviour studies at Erasmus School of Economics in collaboration with Erasmus Medical College, Rotterdam, the Netherlands. She has acute inclination in mood disorders and depression studies caused due to changes in lifestyle example food habits, change in job, shifting to a new place or any noticeable change in normal routine. She completed her PhD from Drexel University, USA in collaboration with Institution of Agribusiness Management, Bikaner, India.She has written over 35 research papers and is author of book titled "Marketing strategies and distribution mix for gluten-free foods". She also has contributed in field of big data analytics, agribusiness management and auto immune disorders like celiac disease and gluten sensitivity. She is two times winner of RULA Best Scientist International award. She was awarded Best Emerging Young Researcher at Institution of Agribusiness Management, Bikaner. She is Editor of Journal of Depression and Anxiety and has handled special issue titled "Lifestyle changes and depression". Her research goal and to link epidemiological studies to life sciences for betterment of humanity.

Paviter Singh Matharu, Manager-Sales & Marketing The Mosaic Company, Studied at Institute of Agri Business Management, Bikaner

A qualified Agribusiness Manager from Institute of Agribusiness Management, SK Rajasthan Agricultural University (2010-12) with two certified MANAGEMENT DEVELOPMENT PROGRAM from Management Development Institute, Gurgaon (2016) and Indian Institute of Management, Kolkatta (2016). Backed by seven years of experience as Manager-S&M in the Agrochemical and Fertilizer Industry including 5 years in Fortune 500 Crop Nutrition Company (Mosaic India Private Limited, THE MOSAIC COMPANY USA). Always seeking a challenging role where he can excel by contributing in the strategic sales, marketing and logistics by utilizing my proven considered management and interpersonal skills employing professional and ethical business practices. Having excellent professional presentation and communication ability for better framing out the execution of Channel Sales involving strong and efficient negotiations.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication