Agricultural Data Analysis Using Map Reduce

*[1]B. Bharath Kumar, [2]Mary Subaja Christo, [3]N. Deepa

*[1]UG Scholar, [2,3]Associate Professor, Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences, Chennai, India
*[1]bharathbobba@gmail.com,[2]marysubajachristo.sse@saveetha.com, [3]ndeepa.sse@saveetha.com

Abstract

The information mining idea has gotten one of the most fundamental things to process tremendous measures of information and complete the work in different fields. At first, manual calculations are done as a piece of the information investigation. Be that as it may, with the appearance of PCs, the calculation and examination forms have gotten somewhat simpler than before. Alongside the expansion in innovation, the information to be put away and examined has additionally expanded at higher rates. As a rule, the administration authorities make a note of the real zone present in a town/city/state/nation and they'll be allotted the errand of refreshing this information at customary interims of time. To finish this assignment, they for the most part direct studies and make a note of various classifications of land that is available in that specific zone. After the assortment of this information, the information having a place with every one of the regions present in a specific state are joined to get a general thought of the land having a place with various classes in that state. Similarly, the information having a place with all states are united and broke down to comprehend the absolute classes of land present in that specific nation. As we as a whole realize that Agriculture is one of the key areas that is making human life conceivable to get by on this planet, there comes the need to have the insights of this farming area. It is difficult to process this agricultural territory information in ordinary techniques. Information Mining remains as the best decision to finish this undertaking. In this manner, the time arrangement digging approach is followed for farming region location.

1. Introduction

The need for food can be seen everywhere around the world. The food items that the people eat varies from one region to another. This difference in food items is due to the food crops that are grown in those particular countries. In some countries, only a few kinds of crops can be grown on that soil.India is one of the five biggest nations on the planet that delivers an enormous assortment of agricultural items. As indicated by the measurements acquired from the review led in 2018, it is said that agriculture utilized half of the Indian workforce and contributed 17-18% to the nation's Gross Domestic Product. The individuals from practically every one of the pieces of the world consider this as a work that they are firmly joined to, for it makes their endurance conceivable.

Rural terrains and practices can be limitlessly seen in country territories instead of in the towns. While the unfilled grounds in the towns are being utilized to construct a few wellsprings of amusement or safe house, the vacant terrains present in the rustic zones are being changed over and utilized as the rural terrains. Indeed, even fruitless terrains are being changed over into agricultural grounds by following some logical techniques. In this manner there comes the need to have a refreshed database with respect to the rural terrains and the yields that are being collected.

By this procedure of having a refreshed database with respect to the agricultural zone present in the nation and all through the world, the ranchers can benefit from outside intervention in one manner or the other. How the ranchers get profited through this agrarian territory location is that the significance and need of developing

the yields that are generally required by the individuals in a nation can be educated to the ranchers to get the ideal nourishment items. Along these lines, the salary age for ranchers can likewise be expanded. Hence, both the errands i.e., the appetite extinguishing assignment of the individuals present in the nation or the world and the salary age improvement for the ranchers completes.

Information Mining offers a few different ways to examine and utilize the information that we have. In any case, we have to pick explicit techniques from them dependent on our necessities. For example, the estimation of the political outcome is commonly done utilizing the wistful investigation. In the wistful examination, we accumulate the information acquired from the surveys and the posts from the internet-based life locales. The information acquired from these surveys and posts are investigated, positive surveys and positive presents related on a particular ideological group are isolated and their score is gotten. Similarly, the score of different gatherings is acquired and the last score is determined by looking at all these qualities. Probably the best technique to play out the rural territory discovery is the time arrangement mining approach.

2. Literature Survey

R.S. Lunetta et al performed exercises, for example, observing the areas and disseminations of land-spread changes that are significant for building up linkages between strategy choices, administrative activities and ensuing area use exercises. Past investigations joining two-date change recognition utilizing Landsat information have would in general be execution constrained for applications in naturally complex frameworks. This investigation investigated the utilization of 250 m multi-transient MODIS NDVI 16-day composite information to give a robotized change recognition and caution ability on a 1-year time-step for the Albemarle-Pamlico Estuary System (APES) locale of the US. This exploration incorporated the use of a computerized convention to initially channel the MODIS NDVI information to evacuate poor (debased) information esteems and afterward gauge the missing information esteems utilizing a discrete Fourier change procedure to give high caliber continuous information to help the change identification investigation. The techniques and results definite right now just to non-farming zones. Extra confinements credited to the coarse goals of the NDVI information incorporated the overestimation of progress zone that required the utilization of a change zone amendment factor [12].

B.F.T. Rudorff et al proposed the idea which portrays the information on dependable appraisals of regions under sugarcane development is fundamental for the Brazilian agribusiness, since it helps in the advancement of open approaches, in deciding costs by sugar factories to makers and permits building up the coordination of generation removal. The target of this work was to build up an approach for mapping the sugarcane crop territory in the province of Paraná, Brazil, utilizing pictures from the Landsat/TM/OLI and IRS/LISS-3 satellites, for the harvest a long time from 2010/2011 to 2013/2014. The mappings were led through the regulated Maximum probability order (Maxver) accomplishing, by and large, a general exactness of 94.13% and kappa list of 0.82. The connection with the official information of the IBGE ran from moderate to solid ($0.64 \leq rs \leq 0.80$) with normal understanding (dr) of 0.81. There was an expansion of 2.73% (18,630 ha) in the region with sugarcane in Paraná between 2010/2011 and 2013/2014 [11].

Y.L. Everingham proposed propelled satellite symbolism to characterize sugarcane crop attributes. South Africa is the main maker of sugarcane in Africa and one of the biggest sugarcane makers on the planet. Sugarcane is become under a wide scope of climatic, agronomic, and financial conditions in the nation. Stress factors, for example, water what's more, supplement inadequacies, and creepy crawly irritations and sicknesses are among the most significant factors influencing sugarcane generation in the nation. Observing of worry in sugarcane is in this way, basic for evaluating the results on yield and for making a move of their alleviation. The forecast of sugarcane yield, then again is likewise a noteworthy practice for settling on educated choices for viable and sound harvest arranging and the board endeavors in regards to e.g., processing plans, showcasing, estimating, and incomes. In South Africa, the recognition of stress factors, for example, nitrogen (N) lack and sugarcane thrips (FulmekiolaserrataKobus) harm and pervasion are made utilizingcustomary direct techniques whereby leaf tests are gathered from sugarcane fields and the proper lab examination is then performed.

These techniques are viewed as being tedious, work escalated, expensive, and can be one-sided as frequently they are not consistently applied across sugarcane developing territories in the nation. Right now, improvement of efficiently sorted out geo-and time-referenced exact strategies that can recognize sugarcane stress factors and anticipate yields are required. Remote detecting offers close continuous, possibly reasonable, speedy and monotonous information that could be utilized forsugarcane observing. Handling systems of such information have as of late saw more improvement prompting increasingly powerful extraction of data. Right now, point was to investigate the potential utilization of remote detecting to measure worry in and anticipate yield of sugarcane in South Africa [14].

R.K. Dewi et al proposed highlight extraction for recognizable proof of sugarcane rust ailment. The primary abundance of the India is cultivating. In any case, harming pace of that farming item is for the most part through the cataclysmic events like floods and tempest and second factor is the infection or microbes influences the plant. For distinguishing the malady in the plants there are a few specialists despite the fact that it is difficult to move toward without fail. It might be costly

for knowing the infection of the plant. Finding of the influenced once isn't sufficient yet the purpose behind the malady additionally will be valuable to give the pesticides and a portion of the helpful natural materials like manures to expand the obstruction of the plants to yield better quality outcomes. As the outcome the complete exactness of the territory affected. This paper presents recognition of sugarcane plant infections by utilizing DWT (Discrete wavelength change) calculation. By utilizing some well-known strategies like programmed identification methods, the ailments are recognized quick [15].

E.K.Ratnasari et al studied about the sugarcane diseases. About 15% of sugarcane leaf is faulty due to sicknesses, it lessens the amount and nature of sugarcane creation essentially. Early discovery and estimation of plant illness is an approach to control these ailments and limit the serious disease. This paper proposes a model to distinguish the seriousness of certain spot malady which show up on leaves dependent on fragmented spot. The sectioned spot is acquired by thresholding a* part of L*a*b* shading space. Maladies spots are extricated with greatest standard deviation of sectioned recognize that utilization for discovery the kind of ailment utilizing grouping procedures. The classifier is a Support Vector Machine (SVM) which utilizes L*a*b* shading space for its shading highlights and Gray Level Co-Occurrence Matrix (GLCM) as its surface highlights. This proposed model competent to decide the kinds of spot ailments with precision of 80% and 5.73 blunder seriousness estimation normal[16].

J. Zhang et al utilized improved maize developed region estimation over an enormous scale consolidating MODIS–EVI time-arrangement information and harvest phenological data. The precise and auspicious data of yield zone is imperative for crop generation and nourishment security. Right now, Enhanced Vegetation Index (EVI) information from moderate goals Imaging Spectroradiometer (MODIS) coordinated yield phenological data was utilized to evaluate the maize developed zone over a huge scale in Northeast China. The fine spatial goals China's Environment Satellite (HJ-1 satellite) pictures and the help vector machine (SVM) calculation were utilized to separate circulation of maize in the reference region. The mean MODIS–EVI time arrangement bend of maize was removed in the reference zone by utilizing various periods MODIS–EVI information. By examining the transient move of yield schedules from northern to southern parts in Northeast China, the slack worth was gotten from phenological information of twenty-one agro-meteorological stations; here incorporating with the mean MODIS–EVI time arrangement picture of maize, a standard MODIS–EVI time arrangement picture of maize was acquired in the entire investigation region. By computing mean total separations (MAD) map between standard MODIS–EVI picture and mean MODIS–EVI time arrangement pictures, and setting fitting limits in three areas, the maize developed zone was extricated in Northeast China. The

outcomes indicated that the general grouping exactness of maize developed territory was roughly 79%. It showed that MODIS–EVI time arrangement information coordinated with crop phenological data can be utilized to improve the extraction exactness of yield developed territory over a huge scale [17].

J.Li et al performed studies on remote sensing methods. Our capacities for gathering remote detecting pictures have enormously outpaced our capacities to investigate and recover data from the picture databases. This paper shows a disseminated structure for data mining from multi-dimensional remotely detected pictures utilizing Windows High Performance Computing (HPC) Servers and Dryad circulated figuring motor. Land spread and land use types are characterized by Support Vector Machines (SVM) and put away in an item situated database with area quad-tree lists. Dryad LINQ questions, an all-encompassing variant of the LINQ programming model, are created for recovering area spread appropriation data and recognize the progressions of each land spread sort at multi levels. A HPC group with sixteen registering hubs is actualized and the trials are directed on a period arrangement Landsat Thematic Mapper (TM) pictures. The outcomes show the adequacy of the system and its possibilities in other remote detecting applications[19].

D.G.Stork et al performed studies on pattern classification. Faltering appraisal through the manual characterization of discourse disfluencies is emotional, conflicting, tedious, and inclined to blunder. The point of this paper is to analyze the viability of the 3 discourse include extraction strategies, recurrence cepstral coefficients, direct prescient coding (LPC)- based cepstral parameters, and perceptual straight prescient (PLP) examination, for ordering 2 sorts of discourse disfluencies, reiteration and prolongation, from recorded disfluent discourse tests. Three distinct classifiers, the k-closest neighbor classifier, direct discriminant investigation based classifier, and bolster vector machine, are utilized for the grouping of discourse disfluencies. Discourse tests are taken from the University College London Archive of Stuttered Speech and stammered occasions are recognized through manual division. A 10-crease cross-approval strategy is utilized for testing the unwavering quality of the classifier results. The impact of the 2 parameters (LPC request and edge length) in the LPC-and PLP-put together techniques with respect to the characterization results is likewise researched. The trial results uncover that the proposed strategy can be utilized to help discourse language pathologists in characterizing discourse disfluencies[24].

S.Patra et al studied cluster assumption based active learningRemote detecting pictures are prevalently influenced by the nearness of blended pixels. Delicate classifiers have the bit of leeway to deal with the blended pixels because of the weaknesses of hard classifiers. The fluffy based classifiers have demonstrated to be strong and exact when characterizing land use and land spread

maps. In the writing, the fluffy c-implies classifier has been examined with the Euclidean, Mahalanobis and inclining Mahalanobis standards. Right now, fluffy c-implies classifier has been considered with nine other closeness and divergence measures: Manhattan separation, chessboard separation, Bray-Curtis separation, Canberra, Cosine separation, connection separation, mean supreme contrast, middle outright distinction and standardized squared Euclidean separation.

Cosine standard was seen as the best single standard among all the standards with a general exactness of 75.24%, trailed by the Euclidean standard. These two standards were joined to shape the composite standard which indicated a general exactness of 69.80%. The exactness of the order was likewise estimated on account of an undeveloped class (wheat), which brought about an abatement in the general precision in contrast with the prepared case. To finish up FCM classifier with Cosine standard performed superior to the traditional Euclidean standard. Be that as it may, because of the lack of ability of FCM classifier to deal with commotion appropriately, the arrangement exactness was around 75%[26].

D.Tuia et al performed studies on active learning methods.The legitimacy of preparing tests gathered in field battles are vital for the achievement of land use order models. Notwithstanding, such examples frequently experience the ill effects of an example choice inclination and don't speak to the fluctuation of spectra that can be experienced in the whole picture. In this way, to expand

arrangement execution, one must perform adjustment of the main model to the new information conveyance. Right now, propose to perform adjustment by examining new preparing models in obscure regions of the picture. We will likely choose these pixels in a smart manner that limits their number and expands their data content. Two methodologies dependent on vulnerability and bunching of the information space are considered to perform dynamic determination. Trials on urban and rural pictures show the incredible capability of the proposed technique to perform model adjustment[27].

Problem Statement: The continuous and rapid increase in land inhabitation by people, industries and other organizations has made the area computing task more complicated. The manual survey and computation of area that is inhabited and left over in a particular place is a time-consuming task. However, the survey of this data has become better through the online data updating processes with the help of concerned departments that are assigned the job of collecting the land data. With the use of relational database management system, this process has become a bit easier than before. In the relational database management system, the data can be gathered and used for processing purpose. The use of this method is a wise choice when compared to the traditional area computing methods. But, when this procedure is used to process huge amounts of data, this too fails to deliver the results in a short period of time.
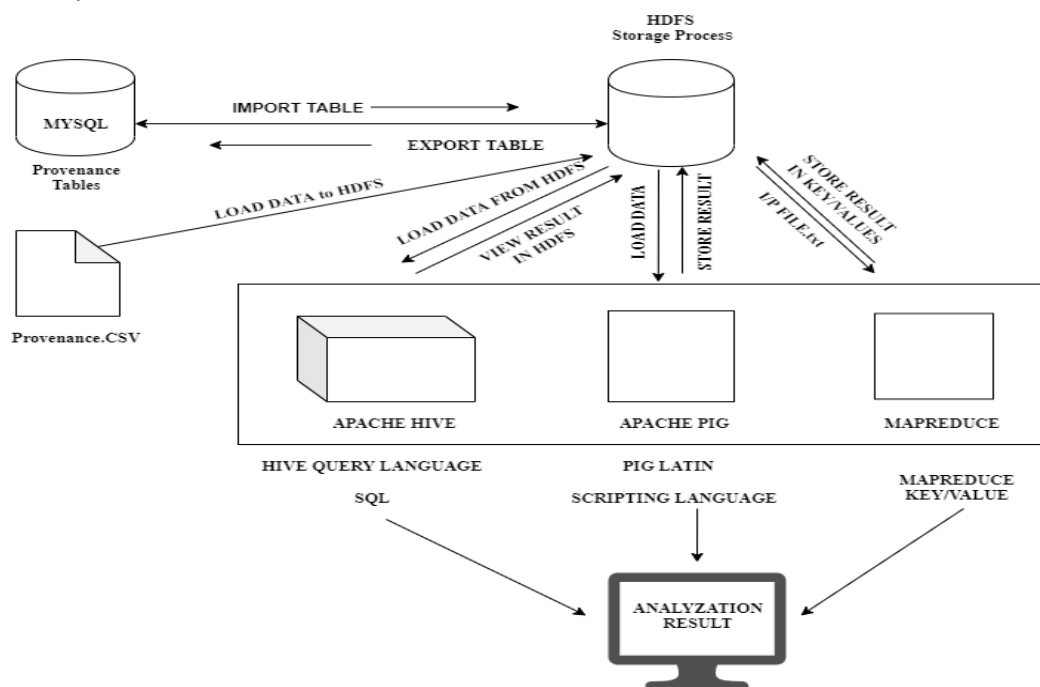
3. Proposed System

Figure 1: System Architecture

Fig.1 portrays the general procedures that are associated with recognizing the rural region. At first, the information present in the database is brought into the

hdfs stockpiling. From this hdfs stockpiling, the APACHE HIVE and APACHE PIG stacks the qualities present in it. The other usefulness present in the apache

hive is that we can see the outcomes. The other usefulness present in the apache pig is that we can procure the outcomes and store them in the hdfs stockpiling. On account of MapReduce, the info record will be gained from the hdfs stockpiling and stores the acquired outcome in it. At last, the outcomes acquired in the wake of handling the information apache hive, apache pig and guide decrease are store in an alternate module and are utilized for the examination and location of the necessary zone esteem.

4. Modules

Existing Application (MySQL): MySQL is a social database the board framework.The architecture of MySQL can be seen in Fig.2. RDBMS utilizes relations or tables to store Agriculture information as a network of lines by sections with essential key. With MySQL language, Agriculture information in tables can be gathered, put away, prepared, recovered, removed and controlled for the most part for business reason. Existing idea manages giving backend by utilizing MySQL which contains parcel of disadvantages for example information constraint is that preparing time is high when the information is tremendous and once information is lost, we can't recoup so in this way we are proposing idea by utilizing Hadoop instrument.
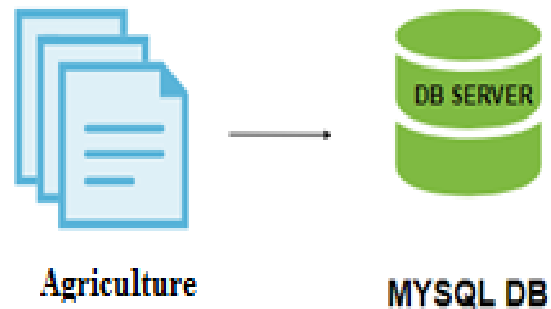

Figure 2: MySQL Architecture

Connector (Sqoop): Sqoop is a medium of data exchange that helps in transferring the agricultural data between the databases. Here in MySQL database having Agriculture information need to import it to HDFS utilizing Sqoop. Agribusiness information can be moved into HDFS/Hive from MySQL and afterward it will create the java classes. In past cases, stream of information was from RDBMs to HDFS. Utilizing "send out" apparatus, we can import information from HDFS to RDBMs. Prior to performing send out,Sqoop brings table metadata from MySQL database. In this way we first need to make a table with required metadata. The architecture of Sqoop can be seen in the Fig.3.
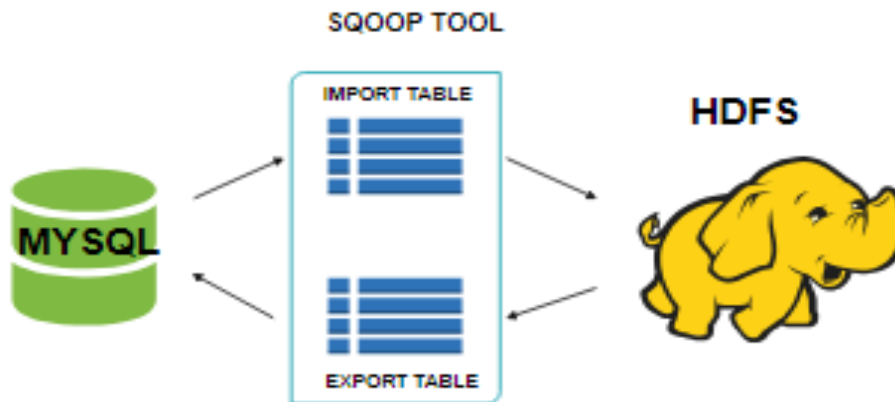

Figure 3: Sqoop Architecture

Analysis Query Language (Hive): Hive is an information product house framework for Hadoop that runs structured query language like queries called hive query languageto handle and make use of data.The architecture of Hive can be seen in the Fig.4. In Hive, Agriculture information tables and databases are made first and afterward information is stacked into these tables. Hive as Agriculture information distribution centre intended for overseeing and questioning just organized

information that is put away in tables. Hive sorts out Agriculture information tables into allotments. It is a method for isolating a table into related parts dependent on the estimations of divided segments. Utilizing segment, it is anything but difficult to question a part of the given dataset. Tables or parcels are sub-isolated into pails, to give additional structure to the Agriculture information that might be utilized for increasingly effective questioning. Bucketing works dependent on the estimation of hash capacity of some section of a table.
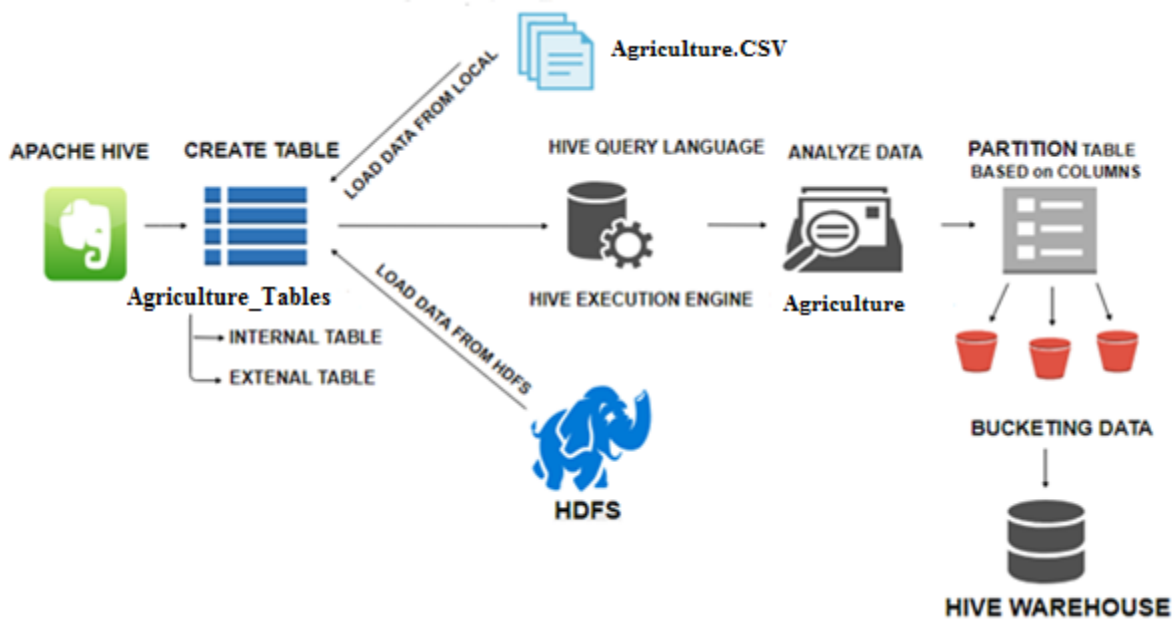
Figure 4: Hive Architecture

5. Results

The agricultural area analysis is initially done using some traditional methods such as manual computing, etc due to the lack of advanced methods to carry out this job. Later, these manual computations are replaced by the method that used the relational database management system. With the use of the relational database management system and the structured query language, the task of computing this agricultural data has become a bit simpler than before. Manual computing may result in some repeated data, whereas there is no chance for such flaws in the relational database management system since the similar data are grouped under the same category here.

The rise in population led to the rise in need for agricultural products which in turn resulted in the need for the rise of the agricultural area. In order to understand and have updated data regarding the agricultural area, there came the need for an advanced method of handling the agricultural data. Thus this method of handling the agricultural data using the MapReduce method is proposed.

As we can see in Fig.5, the accuracy of the data obtained is high in the MapReduce method when compared to the other two methods. Similarly, the time taken by the old methods and MySQL to process the data is high when compared to the MapReduce method.
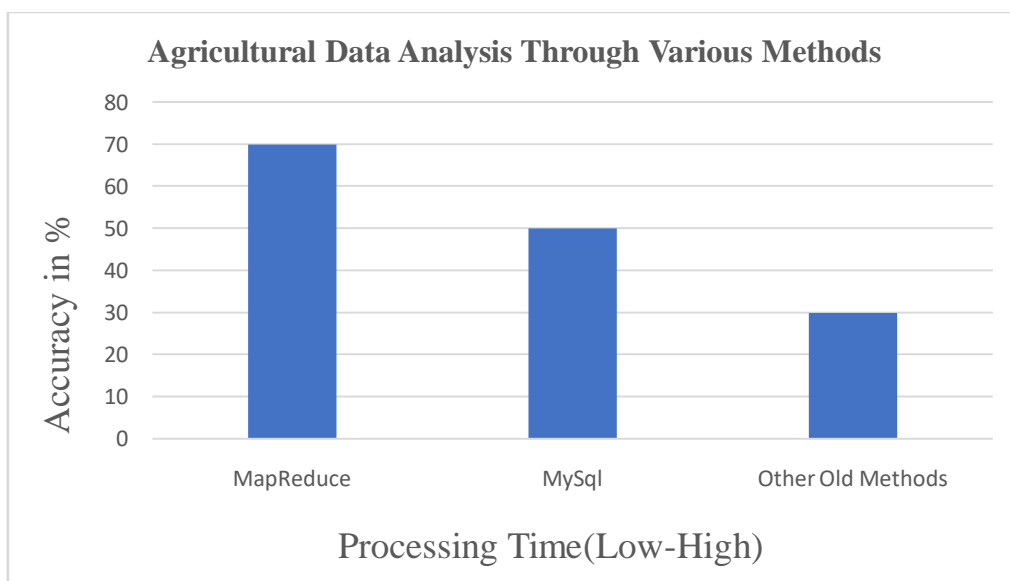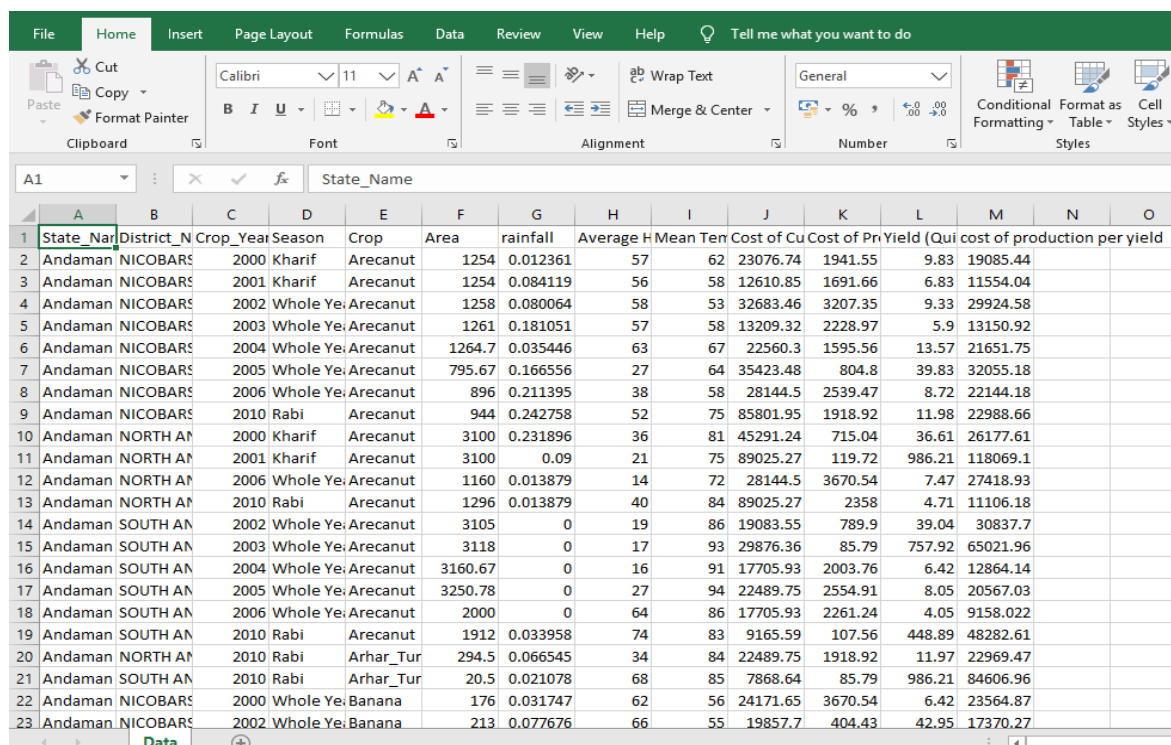

Figure 5: Graph Depicting the Agricultural Data Analysis Through Various Methods

Implementation Screenshots

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	State_Nar	District_N	Crop_Year	Season	Crop	Area	rainfall	Average H	Mean Ten	Cost of Cu	Cost of Pr	Yield (Qui	cost of production per yield		
2	Andaman	NICOBARS	2000	Kharif	Arecanut	1254	0.012361	57	62	23076.74	1941.55	9.83	19085.44		
3	Andaman	NICOBARS	2001	Kharif	Arecanut	1254	0.084119	56	58	12610.85	1691.66	6.83	11554.04		
4	Andaman	NICOBARS	2002	Whole Ye	Arecanut	1258	0.080064	58	53	32683.46	3207.35	9.33	29924.58		
5	Andaman	NICOBARS	2003	Whole Ye	Arecanut	1261	0.181051	57	58	13209.32	2228.97	5.9	13150.92		
6	Andaman	NICOBARS	2004	Whole Ye	Arecanut	1264.7	0.035446	63	67	22560.3	1595.56	13.57	21651.75		
7	Andaman	NICOBARS	2005	Whole Ye	Arecanut	795.67	0.166556	27	64	35423.48	804.8	39.83	32055.18		
8	Andaman	NICOBARS	2006	Whole Ye	Arecanut	896	0.211395	38	58	28144.5	2539.47	8.72	22144.18		
9	Andaman	NICOBARS	2010	Rabi	Arecanut	944	0.242758	52	75	85801.95	1918.92	11.98	22988.66		
10	Andaman	NORTH AN	2000	Kharif	Arecanut	3100	0.231896	36	81	45291.24	715.04	36.61	26177.61		
11	Andaman	NORTH AN	2001	Kharif	Arecanut	3100	0.09	21	75	89025.27	119.72	986.21	118069.1		
12	Andaman	NORTH AN	2006	Whole Ye	Arecanut	1160	0.013879	14	72	28144.5	3670.54	7.47	27418.93		
13	Andaman	NORTH AN	2010	Rabi	Arecanut	1296	0.013879	40	84	89025.27	2358	4.71	11106.18		
14	Andaman	SOUTH AN	2002	Whole Ye	Arecanut	3105	0	19	86	19083.55	789.9	39.04	30837.7		
15	Andaman	SOUTH AN	2003	Whole Ye	Arecanut	3118	0	17	93	29876.36	85.79	757.92	65021.96		
16	Andaman	SOUTH AN	2004	Whole Ye	Arecanut	3160.67	0	16	91	17705.93	2003.76	6.42	12864.14		
17	Andaman	SOUTH AN	2005	Whole Ye	Arecanut	3250.78	0	27	94	22489.75	2554.91	8.05	20567.03		
18	Andaman	SOUTH AN	2006	Whole Ye	Arecanut	2000	0	64	86	17705.93	2261.24	4.05	9158.022		
19	Andaman	SOUTH AN	2010	Rabi	Arecanut	1912	0.033958	74	83	9165.59	107.56	448.89	48282.61		
20	Andaman	NORTH AN	2010	Rabi	Arhar_Tur	294.5	0.066545	34	84	22489.75	1918.92	11.97	22969.47		
21	Andaman	SOUTH AN	2010	Rabi	Arhar_Tur	20.5	0.021078	68	85	7868.64	85.79	986.21	84606.96		
22	Andaman	NICOBARS	2000	Whole Ye	Banana	176	0.031747	62	56	24171.65	3670.54	6.42	23564.87		
23	Andaman	NICOBARS	2002	Whole Ye	Banana	213	0.077676	66	55	19857.7	404.43	42.95	17370.27		

Figure 6: Agricultural Dataset

6. Conclusion

The productive usefulness of this strategy relies upon the qualities that are utilized during the examination procedure. Now and again, the qualities present in the datasets that are accommodated investigation may contain missing qualities. These missing qualities influence the general outcome. To maintain a strategic distance from this, the missing qualities are to be loaded up with some estimated qualities. The general outcome comprises of the territory including the occupied grounds and the fruitless terrains. The general information acquired in the examination is considered and the rural territory information is isolated.

At first, the agrarian zone information are viewed as together. Further, the classes of yields are considered and afterward the insights identified with those classifications are isolated. The fundamental goal of this task is satisfied just subsequent to social event the farming region information. In any case, the sub-undertakings engaged with this investigation procedure incorporates finding the kind of agrarian practices that are being followed in a specific nation and the mainland. These outcomes will likewise help oversee import and fare exercises all through the world. The horticultural items that are available which are beyond what required can be sent out to the regions that require them.

The more seasoned strategy used to complete this agrarian zone identification process is utilizing the Hadoop. With the progression in innovation, the Hadoop framework has gotten less proficient in handling this information and in yielding productive outcomes. Accordingly, the productive flash instrument is utilized to process the information with respect to the territory inhabitance and utilization to get the best outcomes. The results obtained by using the MySQL are moderately accurate and useful. Thus there comes the need for the proposed process, which requires less processing time and gives more accurate results.

References

[1] União das Indústrias de Cana-de-Açúcar, "Dados da produção canavieira", UNICADATA, https://bit.ly/2wBqSNa. 2016.

[2] G.M.S. Câmara, and E.A.M. Oliveira, Produção de cana-de-açúcar. Piracicaba, São Paulo: FEALQ, pp. 1-242, 1993.

[3] J.A. Marengo, Mudanças climáticas globais e seus efeitos sobre a biodiversidade: caracterização do clima atual e definição das alterações climáticas para o território brasileiro ao longo do século XXI. Brasília, Federal District: MMA, pp. 1-212, 2007.

[4] H.S. Pinto and E.D. Assad, Aquecimento global e cenários futuros da agricultura brasileira. Campinas, São Paulo: EMBRAPA, pp. 1-82, 2008.

[5] F.R. Marin and D.S.P. Nassif, "Mudanças climáticas e a cana-deaçúcar no Brasil: Fisiologia, conjuntura e cenário futuro," Revista Brasileira de Engenharia Agrícola e Ambiental,

vol. 17, no. 2, pp. 232239, avaliable at https://bit.ly/2PZlD2J, Jan. 2013.

[6] S.J. Wand, G. Midgley, M.H. Jones, and P.S. Curtis, "Responses of wild C4 and C3 grass (Poaceae) species to elevated atmospheric CO2 concentration: a meta-analytic test of current theories and perceptions," Global Change Biology, vol. 5, n. 6, pp. 723-741, avaliable at https://bit.ly/2LRUAmK, Aug. 1999.

[7] A.P. Souza, M. Gaspar, E.A. Silva, E.C. Ulian, A.J. Waclawovsky, R.V. Santos, M.M. Teixeira, G.M. Souza, and S. Buckeridge, "Elevated CO2 increases photosynthesis, biomass and productivity, and modifies gene expression in sugarcane," Plant, Cell & environment, vol. 31, no. 8, pp. 1116-1127, available at https://bit.ly/2PBjysR, Aug. 2008.

[8] D.B. Leakey, M. Uribelarrea, E.A. Ainsworth, S.L. Naidu, A. Rogers, D.R. Ort, and S.P. Long, "Photosynthesis, productivity, and yield of maize are not affected by open-air elevation of CO2 concentration in the absence of drought," Plant Physiology, vol. 140, no. 2, pp. 779-790, avaliable at https://bit.ly/2oB7hJi, Feb. 2006.

[9] G.W. Wall, T.J. Brooks, N.R. Adam, A.B. Cousins, B.A. Kimball, P.J. Pinter Jr., R.L. Lamorte, J. Triggs, M.J. Ottman, S.W. Leavitt, and A.D. Matthias, "Elevated atmospheric CO2 improved sorghum plant water status by ameliorating the adverse effects of drought," New Phytologist, vol. 152, no. 2, pp. 231-248, available at https://bit.ly/2CbPCS2, Nov. 2001.

[10] R.R.V. Gonçalves, P.P. Coltri, A.M.H. Avila, L.A.S. Romani, J. Zullo Jr., and H.S. Pinto, "Análise comparativa do clima atual e futuro para avaliar a expansão da cana-de-açúcar em São Paulo," Proc. 17th Congresso Brasileiro de Agrometeorologia (CBAGRO '11), pp. 1-5, Jul. 2011.

[11] B.F.T. Rudorff, D.A. Aguiar, W.F. Silva, L.M. Sugawara, M. Adami, and M.A. Moreira, "Studies on the rapid expansion of sugarcane for ethanol production in São Paulo State (Brazil) using Landsat data," Remote Sensing, vol. 2, no. 4, pp. 1057-1076, available at https://bit.ly/2L4cG4X, Apr. 2010.

[12] R.S. Lunetta, J.F. Knight, J. Ediriwickrema, J.G. Lyon, and L.D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," Remote Sensing of Environment, vol. 105, no. 2, pp. 142-154, available at https://bit.ly/2oCDL6a, Nov. 2006.

[13] P.R. Meneses, T.D. Almeida, A.N.D.C.S. Rosa, E.E. Sano, E.B.D. Souza, G.M.D.M. Baptista, and R.S. Brites, Introdução ao processamento de imagens de sensoriamento remoto. Federal District: Brasília: UNB/CNPq, pp. 1-266, 2012.

[14] Y.L. Everingham, K.H. Lowe, D.A. Donald, D.H. Coomans, and J. Markley, "Advanced satellite imagery to classify sugarcane crop characteristics," Agronomy for Sustainable Development, vol. 27, no. 2, pp. 111-117, available at https://bit.ly/2wClL04, Jun. 2007.

[15] R.K. Dewi, and R.V.H. Ginardi, "Feature extraction for identification of sugarcane rust disease," Proc. Int'l Conf. Information, Communication Technology and System (ICTS '14), pp. 99-104, Sep. 2014, doi: 10.1109/ICTS.2014.7010565.

[16] E.K. Ratnasari, M. Mentari, R.K. Dewi, and R.V.H. Ginardi, "Sugarcane leaf disease detection and severity estimation based on segmented spots image," Proc. Int'l Conf. Information, Communication Technology and System (ICTS '14), pp. 93-98, Sep. 2014, doi: 10.1109/ICTS.2014.7010564.

[17] J. Zhang, L. Feng, and F. Yao, "Improved maize cultivated area estimation over a large scale combining MODIS–EVI time-series data and crop phenological information," Journal of Photogrammetry and Remote Sensing, vol. 50, no. 9, pp. 102-113, available at https://bit.ly/2wDh59Z, Aug. 2012.

[18] D.Xie, P. Sun, J. Zhang, X. Zhu, W. Wang, and Z. Yuan," Autumn crop identification using high-spatial-temporal resolution timeseries data generated by MODIS and Landsat remote sensing images," Proc. IEEE Geoscience and Remote Sensing Symposium (IGARRS '14), pp. 2118-2121, Jul. 2014, doi: 10.1109/IGARSS.2014.6946884.

[19] J. Li, "Remote sensing image information mining with HPC cluster and DryadLINQ," Proc. 49th Annual Southeast Regional Conf. (ACM-SE '11), pp. 227-232, Mar. 2011, doi: 10.1145/2016039.2016099.

[20] S. Niazmardi, S. Homayouni, H. Mc Nairn, J. Shang, and A. Safari, "Multiple kernels learning for classification of agricultural time-series data," Proc. 03rd Int'l Conf. Agro-Geoinformatics, pp. 14, Aug. 2014, doi: 10.1109/Agro-Geoinformatics.2014.6910640.

[21] J. Gruijter, D.J. Brus, M.F. Bierkens, and M. Knotters, Sampling for natural resource monitoring. Berlin, Berlin: Springer Science & Business Media, pp. 1-334, 2006.

[22] W.G. Müller. Collecting spatial data: optimum design of experiments for random fields. Linz, Upper Austria: Springer Science & Business Media, pp. 1-242, 2007.

[23] M.M. Crawford, D. Tuia, and H.L. Yang, "Active Learning: Any value for classification of remotely sensed data?," Proc. of the IEEE, vol. 101, no. 3, pp. 593-608, available at https://bit.ly/2CdRIB5, Mar. 2013.

[24] R.O. Duda, P.E. Hart, and D.G. Stork, "Pattern Classification," New York, New York: Wiley-Interscience, pp. 1-680, 2001.

[25] D. Tuia, F. Ratle, F. Pacifici, M.F. Kanevski, and W.J. Emery, "Active Learning methods for remote sensing image classification," Geoscience and Remote Sensing, vol. 47, no. 7, pp. 2218-2232, available at https://bit.ly/2wES0e5, Jul. 2009.

[26] S. Patra and L.A. Bruzzone, "A fast cluster-assumption based active-learning technique for classification of remote sensing images," IEEE Trans. on Geoscience and Remote Sensing, vol. 49, no. 5, pp. 1617-1626, available at https://bit.ly/2Cndi6m, May 2011.

[27] D. Tuia, E. Pasolli, and W.J. Emery, "Using active learning to adapt remote sensing image classifiers," Remote Sensing of Environment, vol. 115, no. 9, pp. 2232-2242, avaliable at https://bit.ly/2LZCT4Q, Sep. 2011.

[28] D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Munoz-mari, "A survey of active learning algorithms for supervised remote sensing image classification," IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 3, pp.606-617, available at https://bit.ly/2Q3aoWT, Jun. 2011.

[29] J. Li, J.M. Bioucas-dias, and A. Plaza, "Spectral–spatial classification of hyperspectral data using loopy belief propagation and active learning," IEEE Trans. on Geoscience and Remote Sensing, vol. 51, no. 2, pp. 844-856, available at https://bit.ly/2LSyXCI, Jul. 2013.

[30] E. Pasolli, F. Melgani, D. Tuia, F. Pacifici, and W.J. Emery, "SVM active learning approach for image classification using spatial information," IEEE Trans. on Geoscience and Remote Sensing, vol. 52, no. 4, pp. 2217-223, available at https://bit.ly/2oyGaia, Apr. 2014.

[31] Christo, Mary Subaja, "A Decision Based Asymetrically Trimmed Modified Winsorized edian Filter for the Removal of Salt and Pepper Noise in Images and Videos." Multimedia Tools and Applications (2019) Impact Factor:2.4 (WOS/SPRINGER)" 1-18.

[32] Mary Subaja Christo "E-Commerce in Agriculture and Product Recommendation using Structural Balance Theory" International Journal of Emerging Technology in Computer Science & Electronics ISSN:0976-1353 Volume 25 Issue5-APRIL-2018.