Strategies for Promoting Sustainability in IT and Analytics

S Lakshmi [1]

[1] SRM Institute of Science and Technology, Chennai, Tamil Nadu
lakshmis10@srmist.edu.in

Abstract. Since we are all in information era especially in the electronic age, it is a foremost requirement to follow the sustainable practices in not only in IT but also in data analytics to reduce the environmental impact. Generally, the organizations long term viability and environmental impact comes from the sustainable practices in IT and analytics. Even though we are in various domains, the natural resources are effectively utilized and efficiently handled by practice. Sustainable practices are the continuous process that should be implemented and monitored properly, if changes are required that should be done then and there for elevating the sustainable practices. Once sustainable practices are identified and initiated for implementation, then reaching the sustainability becomes easy and also need of the hour. The driving force of this work is to explain the sustainable practices of IT and analytics to some extent.

Keywords: Sustainability, IT, Analytics, Environment and practices.

1. Introduction

According to the report of World Commission on Environment and Development, the current sustainable development practices meet the needs of the present without compromising the ability of future generations to meet their own needs who are the backbone of our human community. The ULCA Committee of Sustainability defined the term sustainability as follows: “the integration of environmental health, social equity and economic vitality in order to create thriving, healthy, diverse and resilient communities for this generation and generations to come”. The software organization Deloitte has introduced a three-minute guide of Sustainability Analytics by considering the past, present and future Sustainability related factors such as energy, resource, greenhouse gas emissions can be measured and improved their overall resource efficiently through the practices of sustainability [2]. Sustainable development goals (SDGs) are framed by United Nations by taking the requirements of the future generation and the roles and responsibilities of the current generation. The key areas of the big data analytics such as predictive analytics, borderless supply chain, autonomous vehicles, manufacturing and healthcare industry.

Sustainability can be reached in all aspects of our life by considering the important issues in detail. The sustainable agriculture is the need of the hour due to the lack of resources, insufficient water and greenhouse gas emission and so on. Smart farming can be constructed by integrating the information and communication technologies into machine, sensors, actuators and network devices in agriculture.[3].

M. Rani Nimmagadda et al. (eds.), *Proceedings of the 3rd International Conference on Reinventing Business Practices, Start-ups and Sustainability (ICRBSS 2023)*, Advances in Economics, Business and Management Research 277,
https://doi.org/10.2991/978-94-6463-374-0_74

Ray [4] discussed about IoT based agriculture which consists of six layers. The physical layer contains the various devices like sensors and micro controllers to gather information and pass it other devices. The network layer is considered as the second layer which includes the internet and other technologies for communication. The third layer is an IoT based layer which deals with device management and security related issues. The fourth layer provides storage facility and the fifth layer used to perform data processing for prediction. The sixth layer used to make proper communication with the farmers through the social network. The computer scientists have done research to focus on mainly in social sustainability [5,6,7,8,9,10,11].

Sustainability can be enriched by computing techniques. Sustainable problems have uncertainty in nature, machine learning algorithms, big data analytics, Artificial intelligence, remote sensing, decision making and optimization techniques are used to solve the problems efficiently. [2] stated that the big data analytics has a positive impact on supply chain and operations management. The invention of computer technologies like Internet, the World Wide Web, Global Positioning System (GPS), satellite communications, remote sensing, and smartphones are dramatically accelerating the pace of discovery, engendering globally connected networks of people and devices. The rise of practically relevant artificial intelligence (AI) is also playing an increasing part in this revolution, nurturing e-commerce, social networks, personalized medicine, IBM Watson and AlphaGo, self-driving cars, and other radical transformations [12]. Organizations should pass the information about the initiatives to achieve sustainability to the stakeholders periodically through newsletter or mail [13]. Transparency played a major role to achieve sustainability among the supply chain collaborators [14] and sharing of data that can be used for predictive modelling as well as decision making [15]. The various sub fields of supply chain management actively utilizing the big data analytics effectively. Some of the fields are souring of materials [16], management of risk [17], selection of supplier [18], manufacturing and planning of production [19], inventory management [20] and data warehousing [21].

1.1 Sustainable Practices

It is our foremost responsibility to provide the natural resources to our younger generation so that they can understand and realize the power of nature and they can come up with good ideas to lead their life without harming the environment. The best way of practicing the sustainability through the important 3 R's i.e., Recycle, Reduce and Reuse. The examples for sustainable development of renewing the water resource are rainwater harvesting, avoid the blockage of the water routes and avoiding the mixing of

bad things in water. Energy conservation such as wind mill, solar panel and the environment, economic and social aspects of sustainability are also the examples of sustainable development.

The electronic waste and carbon emissions in industries are given more important to raise sustainability. Due to exponential increase in the human population throughout the world, there is an increase in need to produce more quantity of food to satisfy basic human needs of the human in the world. The food and agriculture industries try various strategies to combat the challenge of producing sufficient quantity of food. To do so, they started using a greater number of fertilizers and other chemicals to increase the yield, to reduce weeds etc. All these chemicals release hazardous gases into the atmosphere when they are utilized by the plants. These gases contribute to air pollution and are the major air pollutants. Growing more trees and plants can aid in reducing some amount of air pollution thereby helps in raising sustainability. Billions of people can lead the modern life style successfully through the industrial revolution in mass production and cultivation of food. The overexploitation of natural resources and the imbalanced nature of ecological system have been linked directly with the humanity during COVID-19 pandemic [22]. The human-made mass will exceed the all living biomass on the planet Earth [23].

1.1.1 Digital Sustainability

Digital sustainability can be achieved by the development of digital resources to improve the environment like pollution free environment and the well-being of the society [1]. The UN has been constructing the awareness towards sustainability and seventeen sustainable goals are presented from various stakeholders to meet the Sustainability Development Goals. This can be achieved by using information systems research, green computing, cloud computing and digital transformation.

1.1.2 Sustainability in IT

It is also called as Green IT used to minimize the environmental impact of information technology through the optimization of data centres, automation of the process and reduction of the resource usage.

1.1.3 Sustainability in Health Care

Sustainability in health care industry used to raise the benefits of people and also used to minimize the negative impacts from the environment. Healthcare systems used to improve the quality of life.

1.2 Current Sustainable Practices of IT and Analytics

By taking the foremost responsibility of the current generation, especially the following things will be discussed in this paper to reach the sustainability in a good manner. They are green computing, digital transformation, cloud computing and big data analytics.

1.2.1 Green Computing

Green computing can be defined as a broad area of information technology which describes the technical way of designing, manufacturing and disposing of computer equipment and reduce the harmful components from the environment to some extent. The carbon footprint reduction is the important goal of green computing. Now a days, we are using so many devices like computer, video game, mobile phone and variation in all these devices with more advanced options make our life so simple and so easy and also make us to think in different angles for making wise decisions. Some of the important areas of green computing which will lead towards sustainability through hardware (efficient utilization), virtualization and data centre optimization.

1.2.2 Digital Transformation

Organizations make use of digital transformation using the various technologies like block chain, artificial intelligence and internet of things (IoT). Making all the employees with necessary skills is the foremost digital transformation called digital literacy. Innovation can be introduced by ensuring the benefits of digital technology to all the stakeholders.

1.2.3 Cloud Technology

The cloud technology not only starts with the mantra like infrastructure as service, software as service, platform as service but also used to build a more sustainable future. Using cloud nearly 35% of the cost of ownership is reduced through the carbon footprint reduction. It is used to reduce 65% of energy consumption and carbon emission by 80%. Selection of the particular software service can also be used to reduce the energy consumption in cloud environment.

1.2.4 Big Data Analytics

By analyzing huge volume of data, companies can get more details related to resource, supply chain and other things related to business which can help them to take decisions very well and to some extent they can predict the future outcomes based on the analysis of data. This analytical process surely helpful to reach sustainability and can achieve sustainable development goals. In 2020 alone 64.2 zettabytes of data created. Analyzing this exponential growth of data using artificial intelligence leads us to get more information about human behavior and perfection. Now a days, public and private sectors are gathering more details in consumer behaviour, services and predictive analysis by dealing with the large volume of data.

2. Materials and Methods

Sustainable development goals were framed by United Nations in the year 2015 with seventeen subgoals aimed to improve the life of the human by the year 2030. Basically, the goals cover the economic development to overcome poverty and environmental protection to save the planet. The goals one to six speaks about the poverty, the goals from seven to twelve deals with energy conservation and the goals from thirteen to seventeen speaks about the climate changes, natural resources and bio-diversity. Continuous monitoring of natural resources is useful to identify the sudden changes in the environment at the earliest which leads us to move towards sustainable development through better machine learning algorithms.

2.1 Case study 1 - Air Quality Index Prediction

Air is the basic necessary thing for the survival of the human being. Air pollution is considered as one of the main reasons for disrupting the human life. For the wellbeing of our human life, it is mandatory to monitor the quality of the air. Generally, air pollution affects our respiratory system easily which leads us to other physiological disorders. Due to the gas emission from industries, air is polluted. It can be avoided to some extent by measuring the air quality index i.e., AQI to know the pollution level and take the necessary steps to reduce the pollution level by continuous monitoring systems.

AQI can be calculated by using the twelve air pollutants such as Nitrogen dioxide (NO_2), Sulfer dioxide (SO_2), Carbon monoxide (CO), ozone(O_3), ammonia (NH_3), PM_{10} (10 microns) and benzene. The high value of AQI indicates the contamination of air. The various machine learning algorithms such as linear regression, decision tree, support vector machine and random forest used to measure the air quality index and the performance of the algorithm is calculated by mean square error and mean absolute error.

2.1.1 Workflow of AIQ Calculation

The implementation process of air quality index involves various steps and the diagrammatic representation of the entire process is explained in the Fig.1. The sample dataset for measuring the air quality index is shown in the Table.1. A very few records are shown in the given table with sixteen features.

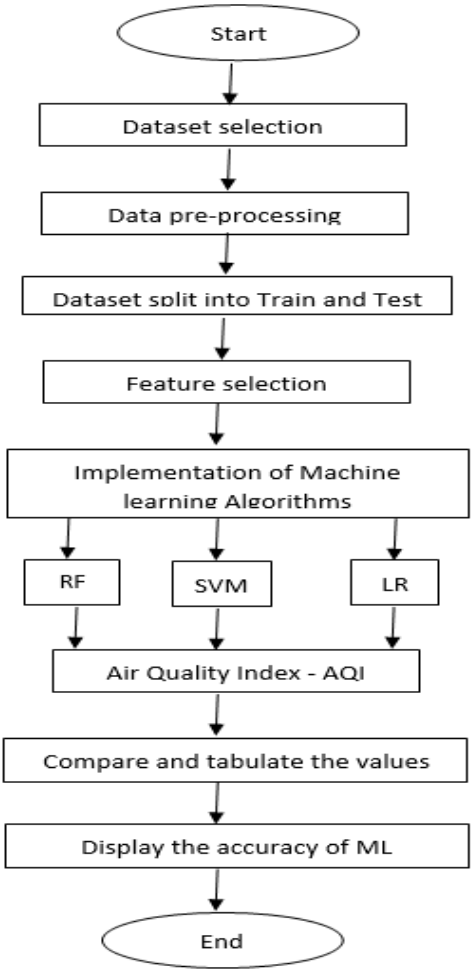

Fig. 1. Workflow of AIQ Prediction

2.1.2 Procedure for predicting the AIQ

The flowchart shows the process of calculating the air quality index by using various machine learning algorithms. The following steps involved in the air quality index checking process:

Step 1: Choose the dataset for analysing the air quality index in csv format.

Step2 : Data preprocessing is considered as an important process and in which the null values, missing values and erroneous information is removed from the dataset so that the dataset is cleaned and ready for further processing.

Step 3: The dataset is splited into training and testing in the ratio of 80:20. Generally, the best results are obtained if 80% of data used for training and the remaining 20% used for testing. Random samples are used for dividing the dataset into training and testing data.

Step 4: In order to make the dataset for doing calculation and modification, normalization is used. The standard Scalar is the better preprocessing technique which is from Scikit library. It is used to normalize the features.
Step 5: The various machine learning algorithms are applied and the air quality index is predicted. The Random Forest supervised learning algorithm is used for classification and regression and it creates decision tree for the samples. The support vector machine and linear regression algorithms are implemented for predicting the air quality index and the results are tabulated. The Mean square error and mean absolute error are used as an evaluation metric and the accuracy of the machine learning algorithm is displayed.

Table 1: Sample Air Quality Index Dataset

StationId	Date	PM2.5	PM10	NO	NO2	NOx	NH3	CO	SO2	O3	Benzene	Toluene	Xylene	AQI	AQI_Bucket
AP001	24-11-2021	71.4	116	1.75	20.65	12.4	12.19	0.1	10.8	109	0.17	5.92	0.1	174	Moderate
AP001	25-11-2021	81.4	125	1.44	20.5	12.1	10.72	0.12	15.2	127	0.2	6.5	0.06	184	Moderate
AP001	26-11-2021	78.3	129	1.26	26	14.9	10.28	0.14	27	117	0.22	7.95	0.08	197	Moderate
AP001	27-11-2021	88.8	135	6.6	30.85	21.8	12.91	0.11	33.6	112	0.29	7.63	0.12	198	Poor
AP001	28-11-2021	64.2	104	2.56	28.07	17	11.42	0.09	19	138	0.17	5.02	0.07	188	Very Poor
AP001	29-11-2021	72.5	115	5.23	23.2	16.6	12.25	0.16	10.6	110	0.21	4.71	0.08	173	Moderate
DL030	30-11-2021	69.8	115	4.69	20.17	14.5	10.95	0.12	14.1	118	0.16	3.52	0.06	165	Moderate
DL030	01-12-2021	74	114	4.58	19.29	14	10.95	0.1	13.9	124	0.17	2.85	0.04	191	Moderate
DL030	02-12-2021	89.9	140	7.71	26.19	19.9	13.12	0.1	19.4	129	0.25	2.79	0.07	191	Moderate
DL030	03-12-2021	87.1	131	0.97	21.31	12.1	14.36	0.15	11.4	115	0.23	3.82	0.04	227	Poor
DL030	04-12-2021	84.6	125	4.02	26.98	17.6	14.41	0.18	9.84	112	0.31	3.53	0.09	168	Moderate
DL030	05-12-2021	88.4	122	3.7	20.23	13.8	13.72	0.12	14	118	0.24	2.92	0.03	198	Moderate
DL030	06-12-2021	96.8	139	1.6	25.65	15	15.12	0.11	16.5	117	0.29	4.45	0.07	201	Poor
CH001	07-12-2021	117	182	4.26	41.1	25.3	17.34	0.13	28.8	94.6	0.36	6.21	0.17	252	Poor
CH001	08-12-2021	123	209	5.56	54.87	33.7	17.96	0.27	23	68.6	0.36	6.28	0.21	310	Very Poor
CH001	09-12-2021	74.3	141	6.1	44.97	28.9	15.73	0.09	21.9	60.6	0.26	4.79	0.16	196	Moderate
CH001	10-12-2021	50.3	103	1.73	33.85	19.4	12.56	0.1	13.7	68.2	0.2	4.29	0.1	132	Severe
CH001	11-12-2021	58.5	115	4.93	41.64	26.2	15.2	0.16	18.4	73.8	0.23	5.51	0.16	147	Severe
CH001	12-12-2021	89.4	131	7.97	42.1	28.9	21.24	0.24	7.42	44.7	0.28	7.01	0.19	179	Very Poor

2.2 Results and Discussions

Sustainability measures is a process of measuring the sustainability attributes like greenhouse gas, climate changes and environmental impacts. Generally, sustainability is impossible to measure and sometimes it is very difficult to measure. Even though it is not measurable, it is in need of hour to measure the sustainability in some way to make effective decisions in industries. It is inevitable to conduct periodic analysis of all aspects of sustainability to measure the progress.

The metrics used in this work with accuracy are R-SQUARE, Mean Square Error (MSE), root mean square error (RMSE) and mean absolute error (MAE). The better model can be identified by considering the higher value of R-SQUARE. The lower

value of MSE used to identify the better model. If the value of MSE is equal to zero, it is considered the model is perfect. The RMSE model has the range to predict the value and the mean absolute error i.e., MAE evaluates the distance between the observed and predicted values. Accuracy is always the best way to predict and evaluate the performance of the best machine learning algorithm. Table 3 - shows the performance metrics of all four machine learning algorithms clearly.

Table 2. Comparison results of Accuracy

Algorithm	Accuracy
Naïve Bayes	86.67
Support vector Machine	92.3
Decision Tree	91.9
Random Forest	99.74

Table 3: Comparison results of various performance metrics

Algorithm	R-square	MSE	RMSE	MAE
Naïve Bayes	0.83	0.054	0.19	0.31
Support vector Machine	0.92	0.09	0.25	0.22
Decision Tree	0.89	0.07	0.21	0.25
Random Forest	0.93	0.078	0.28	0.2

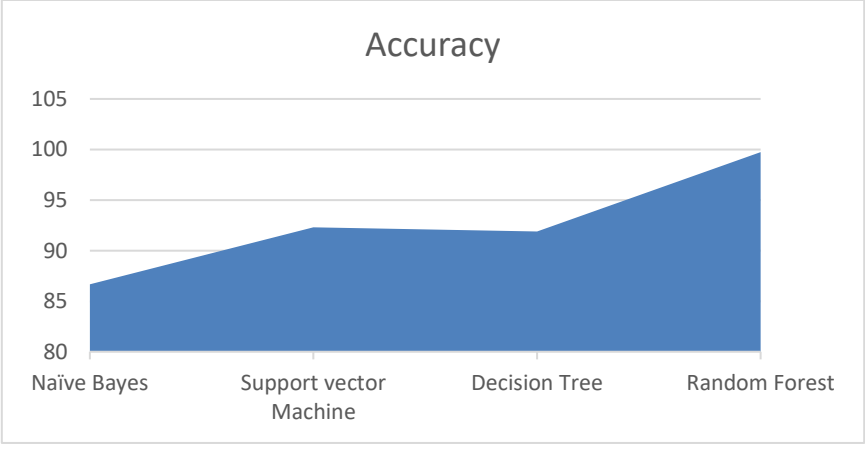

Fig.2. Comparison graph – accuracy

The diagrammatic representation of the accuracy prediction results are shown in the Fig.2.

2.3 Other case Studies

The air quality measurement is the need of the hour to move towards the sustainability in environment which is deeply discussed in 2.2. In the same manner some other case studies are taken and discussed as follows:

2.3.1: Case Study 2 - Textile industry

Let us discuss the sustainability measurement and metrics in a textile industry. The material impacts of the textile industry should include raw material supplier, supply chain, labor practices (waste, water and emissions…), transportation (raw material and products from manufacture to store) and customers. The sustainable metrics may include safety in working environment, health insurance of the employees, workers and labours, economic development and customers (recycling and consignment stores). Since the sustainability measurement is a complex process, it improves the overall production, develop the healthy environment and safe guard the nature for the younger generation. By taking the above said metrics constructing the IT systems and analysing the data surely leads us towards the success of human circulatory system.

2.3.2: Case Study 3 - Upgradation of the Systems

As we are all living in information era, choosing the right tool to perform the right task is the sole responsibility of the organizations. With the advent of IoT and Artificial Intelligence, we are able to get the updated information and the details about the system. In IT industry every day, we come across so many tools to measure the performance and the efficiency of the systems that should be analysed and implemented in a proper way to get more benefits. Suppose we are using the systems which was designed 10 to 15 years back that means, the efficiency of the systems is diminished that should be identified earlier and we need to change the system so that we can safeguard the nature. Hence it is our duty to understand the time to upgrade the system for not only raising the production or but also, we need to take care of our society.

2.3.3: Case Study 4 - Sustainable agriculture

Sustainability agriculture can be achieved through the usage of less chemical, reduce the gas emission and saving the natural resources. Technology oriented applications definitely helps the farmers to use data properly for making right decisions. Sustainable development is required in agriculture for producing food for all. Since the innovative systems will enhance the production, we are in the position to utilize the smart ideas to face the challenges in agriculture.

Designing Smart technologies using IoT enhances the sustainability performance by increasing the production, quality of the crop, resource utilization and improving the quality of the life [24].

3. CONCLUSION

This paper aimed to investigate the sustainability in all aspects of our human life to bring the importance of natural resources and the responsibility of the current generation. The role of IT and Analytics were discussed in detail by considering the case studies with all other sustainability. The importance of analytics in resource utilization, quality of production, optimize the energy consumption and smart monitoring systems in industries as well as agriculture. This is motivated by the question "How can sustainability improve through IT and Analytics?". This paper discussed about sustainability in all aspects of our life. In future planning to consider a single issue in sustainability and try to find solution to maintain the sustainability which relates the growth of an organization.

References

1. Digital Sustainability in Information Systems Research: Conceptual Foundations and Future Directions, "Digital Sustainability in Is Research Future Directions", Julia Kotlarsky, Ilan Oshri, Nevena Sekulic
2. Kyle Tanger, Deloitte Consulting, "Sustainability Analytics- The three-minute guide"
3. Pivoto, D.; Waquil, P.D.; Talamini, E.; Spanhol, C.; Corte, V.F.D.; Mores, G.D.V. Scientific development of smart farming technologies and their application in Brazil. Inf. Process. Agric. 2018, 5, 21–32. [CrossRef]
4. Ray, P.P. Internet of things for smart agriculture: Technologies, practices and future direction. J. Ambient Intell. Smart Environ. 2017, 9, 395–420. [CrossRef]
5. Faghmous, J. and Kumar, V. A big data guide to understanding climate change: The case for theory-guided data science. Big Data 2, 3 (2014), 155–163.
6. Fang, F., Tambe, M., Dilkina, B., and Plumptre, A (eds.). Artificial Intelligence and Conservation. Cambridge University Press, 2018.
7. Fisher, D.H. Recent advances in AI for computational sustainability. IEEE Intelligent Systems 31, 4 (2016), 56–61; https://doi.org/10.1109/MIS.2016.61
8. Lassig, J., Kersting, K., and Morik, K (Eds.). Computational Sustainability. Studies in Computational Intelligence 645 (2016). Springer.
9. Rudin, C. and Wagstaff, K. Machine learning for science and society, Machine Learning 95, 1 (2014),1–9; https://doi.org/10.1007/s10994-013-5425-9
10. Russell, S.J. et al. Letter to the editor: Research priorities for robust and beneficial artificial intelligence: An open letter. AI Magazine 36, 4 (2015).
11. Tambe, M. and Rice, E (eds.). Artificial Intelligence and Social Work. Cambridge University Press, 2018.
12. Carla Gomes, Thomas Dietterich, Christopher Barrett, Jon Conrad, Bistra Dilkina, Stefano Ermon, Fei Fang, Andrew Farnsworth, Alan Fern, Xiaoli Fern, Daniel Fink, Douglas Fisher, Alexander Flecker, Daniel Freund, Angela Fuller, John Gregoire, John Hopcroft, Steve Kelling, Zico Kolter, Warren Powell, Nicole Sintov, John Selker, Bart Selman, Daniel Sheldon,

David Shmoys, Milind Tambe, Weng-Keen Wong, Christopher Wood, Xiaojian Wu, Yexiang Xue, Amulya Yadav, Abdul-Aziz Yakubu, Mary Lou Zeeman[2019], "Computational Sustainability: Computing for a Better World and a Sustainable Future", Communications of the ACM, September 2019, Vol. 62 No. 9, Pages 56-65,doi:10.1145/3339399
13. Giri, C.B.; Dey, S. Game theoretic models for a closed-loop supply chain with stochastic demand and backup supplier under dual channel recycling. Decis. Mak. Appl. Manag. Eng. 2020, 3, 162–189.
14. Carter, C.R.; Rogers, D.S. A framework of sustainable supply chain management: Moving toward new theory. Int. J. Phys. Distrib. Logist. Manag. 2008, 38, 360–387. [CrossRef]
15. Mageto, J.; Prinsloo, G.; Luke, R. Determinants of logistics outsourcing performance among small and medium enterprises. Int. J. Logist. Syst. Manag. 2020, 35, 541–565. [CrossRef]
16. Mani, V.; Delgado, C.; Hazen, B.T.; Patel, P. Mitigating supply chain risk via sustainability using big data analytics: Evidence from the manufacturing supply chain. Sustainability 2017, 9, 608. [CrossRef]
17. Huang, T.; Lan, L.; Fang, X.; An, P.; Min, J.; Wang, F. Promises and challenges of big data computing in health sciences. Big Data Res. 2015, 2, 2–11. [CrossRef]
18. Choi, Y.; Lee, H.; Irani, Z. Big data-driven fuzzy cognitive map for prioritising IT service procurement in the public sector. Ann. Oper. Res. 2018, 270, 75–104. [CrossRef]
19. Krumeich, J.; Werth, D.; Loos, P. Prescriptive control of business processes. Bus. Inf. Syst. Eng. 2016, 58, 261–280. [CrossRef]
20. Lee, K.H. A GA-based optimisation model for big data analytics supporting anticipatory shipping in Retail 4.0. Int. J. Prod. Res. 2017, 55, 593–605. [CrossRef]
21. Rozados, I.V.; Tjahjono, B. Big data analytics in supply chain management: Trends and related research. In Proceedings of the 6th International Conference on Operations and Supply Chain Management, Bali, Indonesia, 10–12 December 2014.
22. Blaauw OJ (2021) Reevaluating value. Mater Circ Econ 3:11
23. Elhacham E, Ben-Uri L, Grozovski J, Bar-On YM, Milo R (2020) Global human-made mass exceeds all living biomass. Nature 588:442–444
24. Salinee Santiteerakul, Apichat Sopadang, Korrakot Yaibuathet Tippayawong and Krisana Tamvimol, "The Role of Smart Technology in Sustainable Agriculture: A Case Study of Wangree Plant Factory", Sustainability, June 2020,12,4640; doi:10.3390/su12114640