**CCT College Dublin**

MSc in Data Analytics

Integrated CA2 – Individual

Dublin City Council Parking Meter Data Analysis

**Lecturer Names:**  Marina Iantorno/Vladimir Milosavljevic

Muhammad Iqbal & David McQuaid

**Student:** Claudia Oliveira da Silva

**Student Number:** sbs22026

**Assessment Due Date:** 11th April 2022

**Date of Submission:** 11th April 2022

**Declaration**

By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

**Table**

|  |  |  |
| --- | --- | --- |
| Table | …………………………………………………………… | 2 |
| 1. Abstract | …………………………………………………………… | 3 |
| 2. Introduction | …………………………………………………………… | 3 |
|  | …………………………………………………………… | 4 |
|  | …………………………………………………………… | 4 |
|  | …………………………………………………………… | 4 |
|  | …………………………………………………………… | 5 |
|  | …………………………………………………………… | 6 |
|  | …………………………………………………………… | 8 |
|  | …………………………………………………………… | 8 |
|  | …………………………………………………………… | 9 |
|  | …………………………………………………………… | 10 |
|  | …………………………………………………………… | 10 |
|  | …………………………………………………………… | 11 |
|  | …………………………………………………………… | 13 |
|  | …………………………………………………………… | 14 |
|  | …………………………………………………………… | 14 |
| 7. References | …………………………………………………………… | 15 |

1. **Abstract**

In general technological advancement alludes to the disclosure of better methods of producing goods. Changes in innovation lead to an expansion in efficiency of work, capital, and different variables of production. The use of software algorithms created an avalanche of modernization in Agriculture business. As the global population increase is decisive to leverage productivity, the profitability of businesses and food production and the well-being of rural and urban society.

Buying rural areas is a very profitable investment than investments in foreign currency, fixed income, and gold specially in the period between 2008 and 2012. In this research paper, a comparative study of the price and size of the land designated for farming and the price the relations with the increase of productivity versus quantity of fertilise used in the last years.

All data was collected on government and open websites, which we understand as a reliable source. As Agricultural Production is a very specific topic and not linked to our daily routine, it is easy to understand that vast majority of people do not talk about agricultural topics, such as fertilizer price, type, and price of land or even the type of pasture of milk that we consume. Because of that, in this research was not use web scrapping, instead was made a programmatic request for sentiment analysis.

Performance was compared using evaluation machine learning linear regression prediction . Upon successful completion of experiments, results show that regression-based machine learning models generally showed better results for modelling with lengthier historical data (more than three years). This is especially so AdaBoost, which is the best result overall.

1. **Introduction**

Since the Farmers began to cultivate the land with the support of technology in agriculture, farming has never been the same. Instead, the open data available shows that, from time to time, technological innovations in agriculture provide a paradigm shift in the agricultural production model.

This means that the innovations in agriculture that characterize technologies such as Big Data on internet mobility, soil analysis, biofortification, GPS become part of life in the countryside, revolutionizing modern agriculture.

For us the majority for people that living in the big cities, it may seem obvious but the wireless and high-speed connection has allowed the introduction of numerous new technologies in agriculture. Now farmers can monitor the farm in real time, follow machines during work, manage staff and access data, KPIs and plantation information on their smartphone screen. All business details can be seen through the cloud, making it much easier to manage the production.

The Big Data have revolutionized the agricultural sector, the use of any technology from Big Data made possible results never before imagined. With these innovations, farmers are managing to increase productivity, but without harming the environment, as well as find it difficult to keep up to date and find new professionals adapted to the new market reality.

Now, agricultural producers, rural producers and agroindustry managers need to be aware of the new technological demands they qualify, as advances and modernization in the field do not stop. Every year are new technologies, software(weather forecast, grain price, price of livestock markets), and of machines(GPS, tractors), soil analysis( fertilizers ) to accelerate the production process, new management procedures have been developed to face the sector's competition. The use of technology to management agriculture industry for change the production vision, reduce costs, increase productivity and reduce environmental damage.

Despite not being a very discussed topic within our society, as result became a very challenging subject to collect data for Machine Learning Sentiment Analysis and also to separate in Sentiment analysis between Producers and consumers.

In this research we choose try to find a correlation between, price of the land for farming, use of fertilizers and increase the amount in agricultural production in the same period.

In view of the huge variety of agricultural products, was defined in our research analysis of a single product cow milk and than choose others to products work on. The reason for that was to create a challenge for my personal learning process. Instead, look for some elements (products) that apparently have no correlation with each other. I would have a much more guaranteed correlation if for example, cattle and pasture land price only, or looking for a date on cereal production. I prefer to work with products are a from a view of part without immediate correlation for a person with no knowledge in the agricultural sector.

So, in this paper, was analysed Price of Land , Price of Fertiliser and The main challenge was to shows that is most important thing is not the amount of information, but what is done with it.

*RQ: “Today, big data is ubiquitous, machine learning applications are thriving, artificial intelligence appears in everyday conversations, and the internet of things is present even in household appliances. Businesses and organizations are increasingly managed through cloud computing and high-performance computing is progressively accessible as a service…More effective operations, reduced uncertainties, and real time decision-support could revolutionize agriculture to a great extent . Food could be produced more efficiently, of higher nutritional quality, in more stable supplies, with less environmental damage, and likely with additional economic, social, and ecological benefits.”(Sjoukje A. Osinga, Dilli Paudel, Spiros A. Mouzakitis, Ioannis N. Athanasiadis (2022)).*

*Sub RQ: “What Machine <earning approaches are best suited to modeling different types and lengths of data (in terms of model accuracy and processing time), and how to analysis the correlation between apparently different topics.”*

In addressing the above research questions, our proposed framework makes the following contribution to show how the Machine Learning, Statistic Analysis can be useful and improve the efficiency of Agriculture sector.

The following sections of the paper are organized as follows. Section 3. presents the review of relevant literature in the field of Agriculture Industry. In section 4. the research methodology is described all process of data manipulation with corresponding outputs in section 5. Machine Learning experiment results and evaluation are presented in section 6, while section 7 presents further discussion of results. Section 8 concludes the research with references.

1. **Literature Review**

Many believe that Big Data is a large volume of information that circulates in the digital environment. Big Data, however, also exists in the data analysis part, in extracting relevant information and applying the data transforming it in facts and information.

Currently, the Big Data is stored in the cloud, the data can be consulted from anywhere, from the farm's seating room or in the middle of the field , and transfer to wherever the business is. In short, Big Data serves the purposes of traceability, prediction, and management of production software systems. With the ability to handle any type of digital record, some subjects are indispensable, and Big Data can record helping Farmers and Agriculture producers, such as geolocation videos, soil characteristics, and the history of water consumption and crop rotation.

Based on 5 principles (of speed, volume, veracity, variety, and value), Big Data uses diverse information collected in real-time to generate expertise. Applied to Machine Learning, data analysis provides information to farmers and machines ensuring the optimization of fertilizers, seeds, and others. It is also possible to identify threats and vulnerabilities and planning preventive actions that avoid losses.

Some of the benefits that agricultural technology provides are:

1. Productivity: with improved advances, it detects production bottlenecks and increases the number of plants per hectare, or the livestock thus increasing the productivity of the farm.
2. Reducing consumption of water, fertilizers, and pesticides, which, in addition to increasing profit, reduces the value of the product.
3. The action of environmental impacts on the environment and the decrease in the production of chemicals in rivers and groundwater, make the business more sustainable.
4. Increased efficiency in the application of Fertiliser technology.
5. Detects soil quantity of nutrients that need to be added to the use as pasture for livestock.

One of the aspects that interfere in a good agricultural productivity is the soil, consequently, its manipulation and capacity of production through the correct application of nutrients and fertilization is of very important. In this context, fertilizers stand out as a technology whose function is to replace and provide the soil with the main nutrients essential for plant growth in order to expand the productive potential of crops and for grazing cattle, sheep, swine, as well as guaranteeing an appreciation of the price of the farm.

Historically, the largest consumption of fertilizers in the world has been through products originating from mineral raw materials, which are also known as synthetic fertilizers. The Have several types and quality of fertilise. In an evolutionary process from these mineral raw materials, the fertilizer industry began to develop technologies that could granulate and group various nutrients into a single product. This wide variety of components and combinations of nutrients and minerals with different fertilizers called compounds.

This technology proposes the production of formulas – compounds is based on a simple mixture of chemical fertilizers to form the so-called formulations with granule mixture. This new fertiliser mixture with the new Big Data technology of Soil Analysis helps to protect the environment, save the water consumption, increase the productivity, and increase the price of cultivated land.

In our process of Data Preparation and Visualization was tried different types of charts and graphics in order to locate the most common type (compound) of Fertiliser and from there to test the relation between Price of Land, Productivity and Fertiliser.

|  |  |
| --- | --- |
| In the graphics you can see there are many input attributes, it makes difficult to visualize the data. Basically, it creates a concern that with this dataset with a big numbers of attributes will negatively affects the accuracy and training time of the machine learning model.  After a research the Principal Component Analysis (PCA) shows to be one way to solve this problem and is used for better data visualization. |  |

Here we stipulated two main type of Fertiliser to be used in the Machine Learning Models and the Statistics and Data Visualization process together with others datasets regarding Livestock, Milk production and Land prices.

Also, in the figure above we can see the correlation between several types of Fertilisers (compounds) that shows the majority of the producers in the same period using the same Fertilisers. It is not for a feeling or tradition, it is Big Data been using effectively in improving soil management and consequently increasing agricultural production.

In relation to the feeding different types of herd, great advances occurred from the improvement of existing pastures by Machine Learning algorithms to analysis soil, seeds, weather forecast, and fertilisers. For example, the adoption of selected grasses developed through scientific research, raises the capacity growth and also the animal performance. In periods of confinement to livestock with the use of management software is capable of accurately calculating the amount and rations and supplements offered in the cold seasons, and the technology is present all seasons. Even a small dairy farms, through modern cow milking equipment are using Big Data every day. Whether or not to adopt its use within the property will be increasingly difficult for farmers in general.

The inclusion of Big Data into Agriculture Industry using cloud data to improve production in agriculture, we can also ask ourselves how this technological integration should be managed, and the advantages it offers. In this research shows that is necessary to care with quantity and more important quality over the data collected through different devices and technologies that make up the software and codes used in modern agriculture. Because, it will make a difference for those who seek better results, which can achieve more effective productions and greater profitability for the sector.

This research process presented us with the challenge of working data in a very specific market but at the same time full of information that for a person with no knowledge of the agricultural market has become a "Pandora's box". The necessity to use mor than one Dataset and the lack of knowledge in the subject, became a big task in terms of data cleaning and data mining. Using PCA numerals, ‘drop’, ‘index’, ‘insa’ functions more than we anticipated. But gave us an opportunity to work with different Machine Learning Models and realize that um Machine Learning Model is not always the best for everything.

In Resume, in the research literature we learned that that when we discuss about Agriculture management system, we are relating data referring to production costs that range from the inputs themselves, labour, machinery, charges, taxes, commercial contracts, stock, transport, land price and much more. All this data can be related to data obtained in the field in your area, or in our country in even in your world, to seek better productivity linked to greater operational efficiency.