Towards Building a Data-Driven Framework for Climate Neutral Smart Dairy Farming Practices

Mohit Taneja
Dept. of Accounting & Economics
School of Business
Walton Institute for Information and Communication Systems Science
Waterford Institute Of Technology
Waterford, Ireland
mohit.taneja@waltoninstitute.ie

Nikita Jalodia
Dept. of Computing & Mathematics
School of Science & Computing
Walton Institute for Information and Communication Systems Science
Waterford Institute Of Technology
Waterford, Ireland
nikita.jalodia@waltoninstitute.ie

Behnam Dezfouli
Dept. of Computer Science & Engineering
School of Engineering
SCU IoT Research Lab
SIOTLAB
Santa Clara University
California, USA
bdezfouli@scu.edu

2021 IEEE Global Humanitarian Technology Conference (GHTC) | 978-1-6654-3372-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/GHTC53159.2021.9612417

***Abstract*—Climate change is one of the biggest challenges facing agriculture as we must produce food to feed a growing global population. According to Food and Agriculture Organization (FAO) of the United Nations, the world population is expected to be 9.2 billion by 2050, and consequently the food demand will be increased by 70%. To cope with these increasing demands, the food production also needs to increase by at least 50-70% to its current capacity by 2050. Improving production efficiency on farms and minimizing greenhouse gas (GHG) emissions is an ongoing and active research area. The article is a work-in-progress presentation of our ongoing efforts in the direction of building an all-encompassing framework towards less GHG intensive climate-neutral farming practices. The dairy sector within agriculture is being used for validation and experimentation, with cows as the centrepiece, and the farmers as the stakeholders that will be the ultimate actors for sustainable environmental impact from the improved farming practices in place. The work also focuses on measuring the behavioural changes of farmer(s) for adoption of technology.**

The envisioned end goal is to build an overarching toolkit and framework for effective knowledge sharing. The article starts with a brief contextual introduction, followed by key considerations, challenges and identified research questions that need to be addressed and evaluated at a wider level in the research community to build the desired framework at a global scale.

***Index Terms*—data-driven, smart dairy, smart farming, climate neutral, farming practices, framework, smart agriculture**

I. Introduction

The Internet of Things (IoT) ecosystem has expanded to include everything in our environment, from smart homes, smart cities, and manufacturing to environmental sensing. We are in the age of the fourth industrial revolution 'Industry 4.0' which involves sensors and smart data capture systems. Being a multidisciplinary ecosystem, it has penetrated into every sector imaginable, and agriculture is no exception to that.

Today, climate change is a critical challenge and agriculture sector is one of the major contributors for emissions leading to it. The agriculture sector's role in greenhouse gas (GHG) emissions is widely known, but yet digitized and ICT (Information and Communication Technology) enabled solution to address the problem in this sector are lagging compared to other industry sectors. It is estimated that more than one-quarter of the world's GHG emissions come from agriculture. Within the agriculture sector, cattle and dairy cows alone emit enough GHGs to put them on par with the highest-emitting nations list. About 80% of those emissions come from livestock in various forms [1], [2]. Having the food production capacity at the required level to feed to world while fighting climate challenge is not an easy task, and it would not take place automatically. This envisioned and ongoing work is a very small step and contribution in developing a framework for improved farming practices towards sustainable climate-neutral agriculture.

The first step in reducing emissions from agriculture is to enable production (milk, meat, crop, etc.) as efficiently as possible. In its report published in May 2020 [3], McKinsey & Company estimated that GHG-efficient farming technologies and practices could achieve about 20% of the sector's required emission reduction by 2050. One of the measures to be taken to address climate change effectively is to improve animal health monitoring and illness prevention. By improving the health of farm animals, farmers could improve productivity and reduce animal mortality due to disease. The ability to meet the projected production requirement with fewer, healthier animals could reduce direct emissions from livestock. The proposition here is that better management decisions means less GHG emissions on dairy farms, since milk production per cow increases, less cows will be needed. This on-going work and proposed research particularly focuses on the measure of animal welfare and health monitoring as the building block to help address the climate challenge, and aims for providing a data-driven animal health monitoring and welfare framework for better dairy farming practices towards sustainable and climate-neutral agriculture.

Focusing on Smart Dairy Farming as the IoT vertical, the ongoing and proposed work aims to develop a data-driven framework using artificial intelligence methods for

This work has emanated from research funded by (i) NGI Explorers Fellowship Grant (grant agreement no. 825183), (ii) CISCO Research Gift Fund, (iii) SFI and DAFM on behalf of the Government of Ireland to the VistaMilk SFI Research Centre (16/RC/3835).

improving milk production capacity of cows by building an animal welfare and health monitoring system. The underlying proposition is that - Opting smart dairy farming principles which unify Internet of Things (IoT), data analytics, artificial intelligence (AI), fog computing, and cloud computing can help meet the rising demands and contribute to sustainable growth in the dairy farming. With improved health monitoring, early intervention and illness prevention, the milk production per cow will increase, thus less cows will be needed.

However, to build such a framework, there are challenges such as data collection, data complexity, interoperability between farming systems, varying farm and geographical conditions etc., to name a few. The proposed research and ongoing work aims to address these key challenges, and presents these key challenges in this article.

The end-to-end systematic objectives of the presented and ongoing work are as follows:

- To build a Data-Driven Framework for Climate Neutral Smart Dairy Farming Practices. The premise is to have "improved milk quantity and quality" with fewer animals. The objective is to enable the farmers to meet the projected production requirement with fewer, healthier animals leading to reduced emissions.
- To measure and quantify the behavioural change(s) of farmer(s) for adoption of technology and smart solutions.
- To measure the milk quality and quantity with and without the smart animal welfare framework solution in place and quantify the improvements. The milk quality data will consist of fat, protein, non-soluble solids percentage and somatic cell count at herd level.
- To build the nutritional profile of milk, and analyze the improvements
- To measure or estimate the GHG emissions from the farm in use and take action for carbon offset, if necessary
- One of the end-product of the project is envisioned as a software-based toolkit utilizing analytics approaches built on the data being collected from the farm under trial.

Fig. 1 below gives an info-graphic representation of the various components involved in the ecosystem.

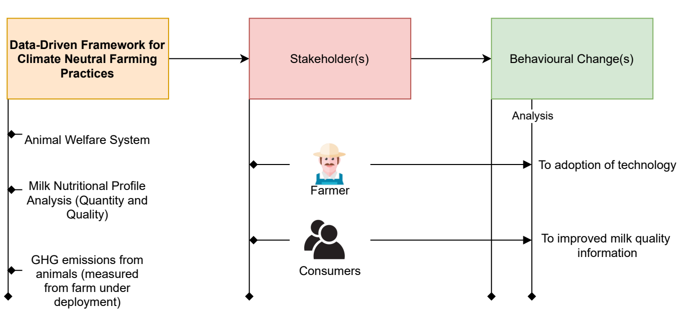

Fig. 1. An info-graphic diagram of the use of the proposed framework and actors involved.

The rest of the article has been organized as follows: §II presents background information and related work outlining the research gaps and motivation behind this work, §III presents the key challenges in the ongoing and envision work, §IV presents different phase involved in the system and solution design, and §V presents concluding remarks.

II. Background, Related Work and Motivation

Data-driven methodologies presents a huge opportunity for verticals like the dairy industry to increase production by gaining actionable insights into farming practices, improving and adjusting them from the learned insights, resulting in increased efficiency and yield. There has been active initiation and movement in the agricultural domain to move towards tech-enabled smart solutions to improve farming practices. The concept of Smart Dairy Farming is no longer just a futuristic concept, and has started to materialize as different fields such as machine learning have found a prosperous application in this domain.

There have been proposed systems in industry [4], [5] as well as in academia [6] for animal health management in dairy farms. A study by authors in [7] gives an overview of the sensor systems available for health monitoring of animals in dairy farms.

Authors in [8] proposed a fuzzy-logic based model to detect abnormalities in dairy cows based on variability in milk properties. GHG emission being a critical issue in the subject-filed, authors in [9] proposed a regression model to predict methane emissions. Further, multiple linear regression, adaptive-neuro fuzzy inference system has also been proposed to estimate the dairy grassland biomass [10]. In recent times, use of Deep Learning (DL) methods for complex data analytics has shown great potential compared to the ML models mentioned above, although their application in the field of dairy farming is still limited [11]. The review by authors in [11] provides a comprehensive discussion on DL models and their application in various IoT verticals, including Smart Farming. Authors in [12] highlight that it is quintessential to gain more in-depth knowledge about relationship between different parameters to optimize farm management. A microservices-based fog computing–assisted IoT platform towards data-driven smart dairy farming smart agri system termed 'SmartHerd' was proposed in [13] to improve decision making process while reducing latency.

Most of the ML and analytics model have mainly been proposed for various single farm usage, are application-specific, and are limited in terms of scalability and re-usability. Such drawbacks limit their application, re-usability, and in certain instance the explain-ability by such models gets lost when applied to a new use-case or application. Thus, necessary modifications are required to use them cooperatively in large-scale smart farming applications.

While there has an initiated movement towards data-driven agriculture in recent times for sustainable and productive growth, there is still lack of leveraging emerging paradigms such as fog computing and applying innovating and advanced machine learning models to solve a specific problem in the dairy sector while ensuring interoperability and re-usability. Thus, the envisioned data-driven framework needs to be built using existing models available in literature as initial building block, but needs to go beyond to include heterogeneity and other key factors to craft a comprehensive environment agnostic framework for improved smart dairy farming practices.

III. Challenges and Direction - Aims, Objectives and Key Research Questions

The core ground-breaking research question to be addressed as part of ongoing efforts from the research community is that *"how can we feed the world sustainably?"* . The overall research aims to develop an in-depth data-driven framework for providing farm environment agnostic smart dairy farming practices.

We identify and list below key thematic areas, corresponding research questions that need to be addressed and evaluated to achieve specific objectives, and build the solution at a global scale.

- **Data Complexity**
  Question: What are the methodologies and techniques that can be used to reduce the data complexity for resource and connectivity constrained farm environments with minimal information loss while maintaining the explainable feature space?
  Objective: There are a large number of devices in a smart dairy farm setup that continuously monitor various farming processes and generate massive amount of data for e.g., milk quality data, accelerometer data of animal mobility data, rumination or imagery data, etc. These data points coming at different time stamps from distributed set of data sources being aggregated to form a complete dataset(s) are often times complex in nature with noise accumulation and contain redundant and unnecessary data. Thus, require cleansing to avoid inconsistencies while building and selecting machine learning models. Farms are usually located in geographically remote locations facing constrained Internet connectivity. Computing and processing data in resource-constrained farming environment is challenging, and thus effecting methods for cleansing, dimensionality reduction need to be applied prior to applying data analytics to extract insights. In scenarios with low/no Internet connectivity, it becomes ideal to process the data locally as much as possible and send the aggregated or partial outputs over the Internet to the cloud for further enhanced analytical results. Thus, design of novel approaches for explaining complex data, effective data processing methods which can optimally utilize the available resources is of utmost importance.
- **Interoperability**
  One of the major issues with smart farming system is lack of interoperability, thus leading to the following question:
  Question: How can an environment agnostic artificial intelligence-based model be developed that works with a significant accuracy and precision to deliver desirable predictions of milk production and health of an individual cow in the herd?
  Objective: The objective of this research question is to use heterogeneous multi-modal non-invasive monitoring data such as animal mobility data, milk quality data to develop a data analytics predictive model for improved animal welfare and health monitoring. The projected outcome is to generate a framework that can do early detection of a set of disease posed as multi-class classification problem in dairy cows and can quantify the corresponding effect on milk productivity of the animal. The associated challenge here is to develop novel approaches using ensemble learning and federating learning as the baseline methodology to build an environment agnostic framework.
- **Farm Size and Scale**
  Question: What kind of variation(s) need to be considered in proposed practice for a small-scale/smallholder farm and a large-scale farm?
  Objective: Given that the size and scale of a farm can differ significantly in terms of area, herd size etc., the objective here is to see the variation that need to be considered with such changes. One of the primary challenges to address here is change in data distribution or data shift that might occur when the developed framework is deployed on a new farm. Approaches that takes such effects into account, inform of variations and change in practice required need to be developed.

IV. Smart Farming Landscape - Research Design and Methodologies

Smart Farming (SF) is a farm management concept that is used to employ sustainable farm management that aims to create greater production and profit with minimum environmental impact and waste. In doing so, modern IoT devices such as sensors, mobile phones, robots, and various ICT infrastructures will collectively monitor the spatial and temporal variability (e.g., grass growth, soil fertility, weather and animal well-being) during the farming process. The collected data is then processed and converted into actionable insights that help stakeholders (farmers, producers) to make meaningful and timely informed decisions to optimize various farm management processes.

Building such IoT solution and frameworks is an intricate process involving end-to-end components, each of which is adapted to the use case being addressed. Generally speaking, an end-to-end IoT solution towards a smart scenario involves the following steps:

- **Connecting the Unconnected:** This step involves the installation of sensors on physical entities such as objects (both static or in motion), remote infrastructure or living entities towards achieving a specified objective such as monitoring. This phase looks at how data will be handled and also other requirements such as frequency, resolution, format and channel of data sources.
- **Data Acquisition:** This involves attaining the sensor data and transferring it to the data analytics platform(s) to achieve actionable insights for better decision making.

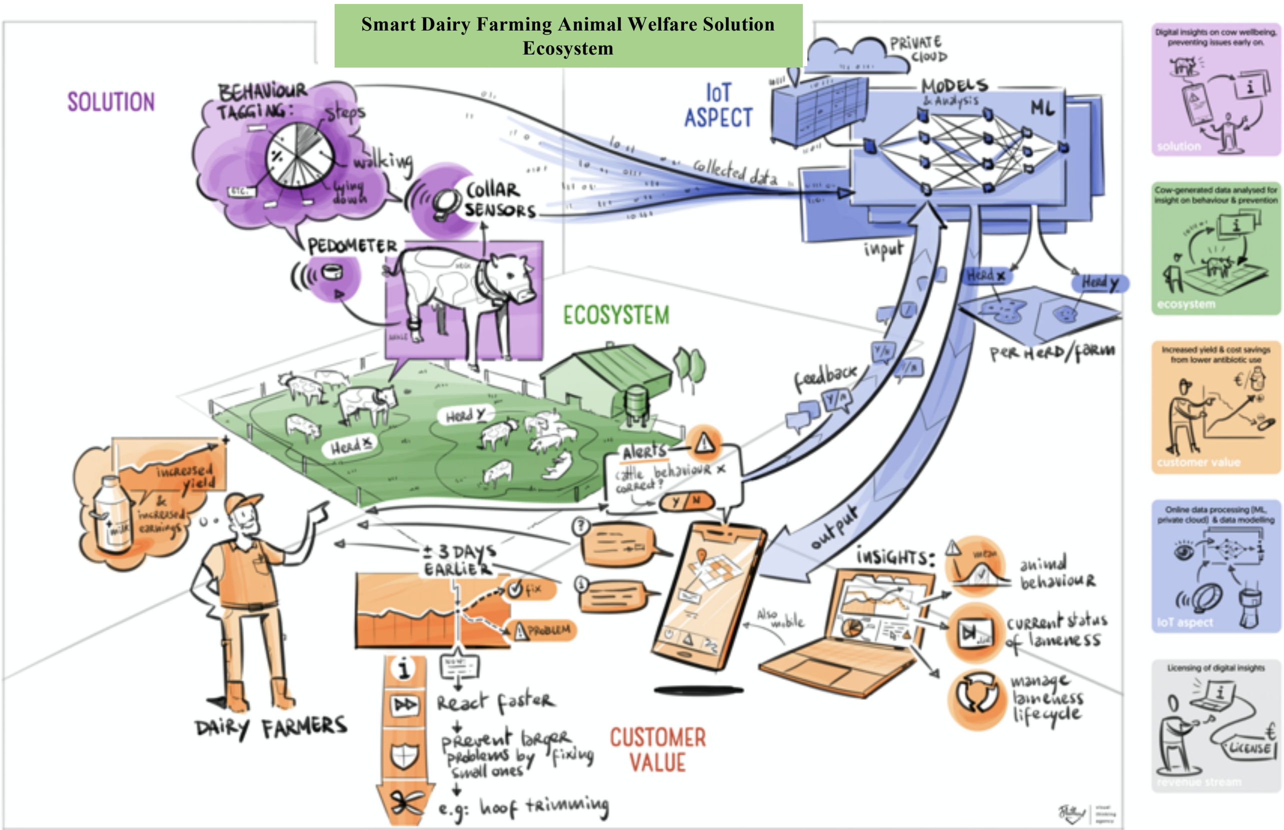

Fig. 2. Visual representation of smart dairy farming ecosystem, IoT aspect and customer value chain.

This becomes a critical problem in scenarios such as ours, wherein a farm location has little or no Internet connectivity.

- **Architecting, Integrating and Management:** This crucial phase involves key decisions on the software architecture and design principles to be used during development of the system. Once finalized, the next step is to integrate, optimize and manage the computing system thus built, which is usually an ongoing process.
- **Data Analytics:** Once the data is at the desired platform (be it fog or cloud), this part involves figuring out how to analyze the data to get the desired information to achieve the specified objective, given the constraints.

A visualisation of one such ecosystem in smart dairy farming illustrating animal welfare solution, the IoT aspect and customer value chain is shown in Fig. 2 [14].

Some key areas of research and improvement in such ecosystem are listed below:

- Designing Inter-operable Plug and Play Components: With farm management system being heterogeneous in nature, there is need to build plug and play services and solution that can be used by the stakeholders, such as farmers depending on their need and requirement compared to existing lock-in and singular solutions.
- Data fusion and data management: There is need to develop effective methods of defining entities of interest and associating a range of networked data sources to them. Context and use-case aware data fusion algorithms need to be designed and developed to relate the impact of various data sources to an entity.
- Data predictive analytics: A granular focus on extracting information from identified data sets and new fused data sets (entities) in order to determine patterns and predict future outcomes and trends also needs to be a key element of ongoing and future research efforts.

A. Data

The ongoing work uses the animal welfare system developed as part of an EU project termed MELD[1] in the large scale Horizon 2020 EU project named IoF2020[2]. The real-world farm trials under MELD project included five farms in Ireland, Israel and Portugal of approximately 1400 cattle combined. In the ongoing and presented work, the Irish trial is being used for all the analysis and validation. The Ireland based trial in MELD is deployed on a farm with 130 cattle utilises leg mounted sensors and uses Machine Learning based animal welfare analysis system to alert the farmer for anomalies detected in animal behaviour. The solution leverages behavioral

[1]MELD stands for **M**achine Learning for **E**arly **L**ameness **D**etection Beef and Dairy Cattle

[2]Internet of Food & Farm 2020, https://www.iof2020.eu/

analytics to generate early alerts toward the animals' well-being, thus assisting the farmer in livestock monitoring. The MELD project specializes in detecting lameness in dairy cattle at an early stage, before visible signs appear to the farmer or an animal expert. The system involves a mobile application for the farmer to use and be notified of the animal welfare alerts from the developed system.

The proposed and on-going work uses the system built in MELD and combines it with the milk analysis data. A pictorial representation of various entities involved at farm level analytics in building the proposed framework, along with the interactions involved has been presented in Fig. 3

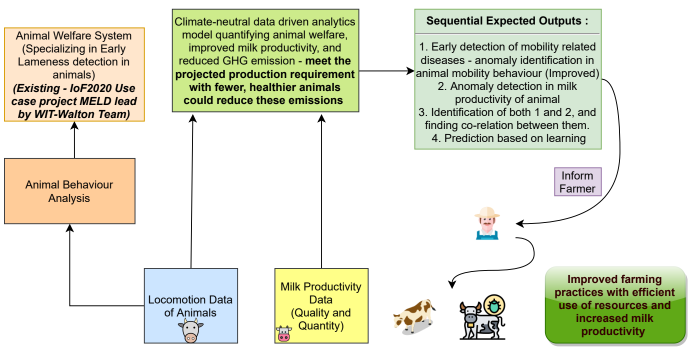

Fig. 3. A very high-level illustrative diagram of the proposed work displaying the various entities involved in farm level data analytics to build the proposed climate-neutral smart farming practices framework.

B. Expected Outcome/End Product from Project:

The end product of the project is envisioned as a software based toolkit utilizing analytics approaches built on the data being collected from the farm under trial. This toolkit can be utilized by researchers in the subject field, especially during stakeholder engagement and surveys. The variable inputs required by the toolkit will be:

- dairy farming milking practice (classical milking or robotic milking),
- daily average milking frequency of cows (i.e. how many times cows are milked in a day? In classical milking practice, its usually twice, but in robotic milking station setup, where cows come to get milked on voluntary basis, it can be between 2 to 4 times a day) type of farm (small-holder or big scale farm)
- current milk production (per year)
- current GHG emission per year (if exact figure available, else estimation based on herd size)
- size of the herd
- it will also use the geo-location based weather API to input data on weather behaviour for the specific location

The toolkit will produce outputs of increase in milk productivity and decrease in GHG emission on an annual basis if the smart animal welfare system is to be implemented on a particular farm. It will be utilizing the prescriptive and predictive analytics model as the block for the calculation built on real data collected from Irish farm trial, which will be wrapped under a multi-variate regression problem for the presented output. The vision is to build an inter-operable and agnostic analytical system to generate personalized farm based result.

This will enable farmers to make an informed decision so as to whether to invest in such smart systems or not, and will put an ease for technology adoption. It should also be noted here that there exist business models now a days that do not require farmer(s) to put a capital investment for implementing smart technology solutions on their farm, rather these solution can be either leased out to them or given to them as a subscription service. With proper data ownership agreements in place, where farmers still own the data and the Ag-Tech service provider acts as user of data, these business models appear as a good option for farmers and are gaining popularity in the farming community [15]. Fig. 4 below gives a clear representation of the complete cycle planned for the project.

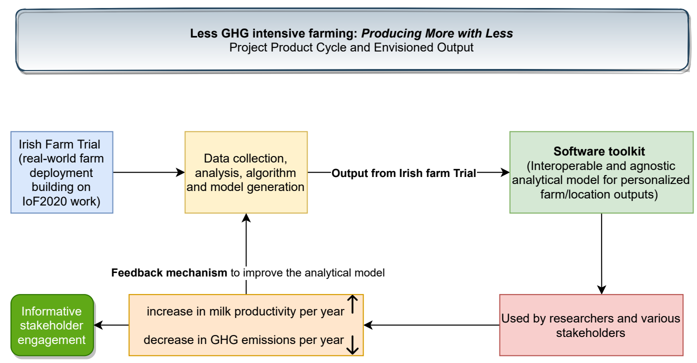

Fig. 4. Project and experiment cycle illustrating the involvements of different stakeholders and output from farm trial headquartered in Ireland.

V. Concluding Remarks

It is becoming increasingly evident that agriculture plays a pivotal role in the socio-economic development of a country. However, many factors impact the sustainability of agriculture in a region. The research focuses on Climate Change as a factor that affects the sustainability of agricultural systems, and the ongoing efforts aim to build a digitized AI based solution to help address these challenges. In this research, Information Communication Technology (ICT) tools can be used for mitigation, adaptation, preparedness, risk management, spreading correct information and as communication empowerment tools.

In this research, the core idea is that AI technology based farm management and animal welfare solutions can assist dairy farmers to achieve efficiency, thus leading to improved food security, which is in line with Sustainable Development Goal 2 (SDG 2). Technology can potentially make agriculture and dairy farming more sustainable, which would be a step towards increasing the productivity to address the rising demands. This ongoing research investigates the potential contribution that technologies can have as support mechanisms for climate change management and awareness of dairy farmers, and their openness to adapt farming practices and new frameworks.

REFERENCES

[1] F. N. Tubiello, C. Rosenzweig, G. Conchedda, K. Karl, J. Gütschow, P. Xueyao, G. Obli-Laryea, N. Wanner, S. Y. Qiu, J. De Barros, *et al.*, "Greenhouse gas emissions from food systems: building the evidence base," *Environmental Research Letters*, vol. 16, no. 6, p. 065007, 2021.

[2] O. A. Ikhuoso, M. Adegbeye, M. Elghandour, M. Mellado, S. Al-Dobaib, and A. Salem, "Climate change and agriculture: The competition for limited resources amidst crop farmers-livestock herding conflict in nigeria-a review," *Journal of Cleaner Production*, p. 123104, 2020.

[3] McKinsey&Company, "Reducing agriculture emissions through improved farming practices — McKinsey." https://mck.co/3jE0HQ8, May 2020.

[4] "Quantified ag® — the best cattle ear tag to identify cattle illness." https://quantifiedag.com/.

[5] "Dairymaster." https://www.dairymaster.com/.

[6] M. Cockburn, "Review: Application and prospective discussion of machine learning for the management of dairy farms," *Animals*, vol. 10, no. 9, 2020.

[7] N. Khan, B. N. Siddiqui, N. Khan, and S. Ismail, "The internet of thing in sustainable agriculture," *Artech J. Res. Stud. Agric. Sci*, vol. 2, pp. 12–15, 2020.

[8] K. Hempstalk, S. McParland, and D. Berry, "Machine learning algorithms for the prediction of conception success to a given insemination in lactating dairy cows," *Journal of Dairy Science*, vol. 98, no. 8, pp. 5262–5273, 2015.

[9] G. Jaurena, J. Cantet, J. Arroquy, R. Palladino, M. Wawrzkiewicz, and D. Colombatto, "Prediction of the ym factor for livestock from on-farm accessible data," *Livestock Science*, vol. 177, pp. 52–62, 2015.

[10] I. Ali, F. Cawkwell, E. Dwyer, and S. Green, "Modeling managed grassland biomass estimation by using multitemporal remote sensing data—a machine learning approach," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 7, pp. 3254–3264, 2017.

[11] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for iot big data and streaming analytics: A survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 2923–2960, 2018.

[12] M. J. O'Grady and G. M. O'Hare, "Modelling the smart farm," *Information Processing in Agriculture*, vol. 4, no. 3, pp. 179–187, 2017.

[13] M. Taneja, N. Jalodia, J. Byabazaire, A. Davy, and C. Olariu, "Smartherd management: A microservices-based fog computing–assisted iot platform towards data-driven smart dairy farming," *Software: Practice and Experience*, vol. 49, no. 7, pp. 1055–1078, 2019.

[14] M. Taneja, N. Jalodia, P. Malone, J. Byabazaire, A. Davy, and C. Olariu, "Connected cows: Utilizing fog and cloud analytics toward data-driven decisions for smart dairy farming," *IEEE Internet of Things Magazine*, vol. 2, no. 4, pp. 32–37, 2019.

[15] "Farmers turn to leasing to keep up with tech - Future Farming." https://www.futurefarming.com/Machinery/Articles/2019/4/Farmers-turn-to-leasing-to-keep-up-with-tech-414885E/.