2. Project Overview

The agriculture industry is experiencing a rapid and significant transformation, as technology plays an increasingly vital role in modernizing farming practices with an impact on the marketing and distribution of agricultural products. Many emerging technologies, such as precision agriculture and smart irrigation systems, are assisting farmers in increasing yields, reducing waste, and preserving resources. Thanks to the advent of e-commerce and online marketplaces, farmers can now sell their products to a global audience, thereby increasing their customer base and enabling them to sell their products at a better price. This is particularly crucial for small-scale farmers, who may find it difficult to reach local markets due to transportation and distribution constraints. Moreover, e-commerce has transformed the agriculture industry by increasing market access, improving efficiency, promoting transparency, and providing better customer experiences. The use of machine learning in the agriculture sector is also on the rise. With machine learning algorithms being used to classify crops based on their appearance, this allows farmers to monitor crop growth and health more efficiently.

The project is driven by the desire to create a more efficient and transparent agricultural supply market. By the fusion of machine learning's image classifications with an ecommerce website, we hope to eliminate the barriers that have historically prevented farmers from selling their crops directly to consumers. The smart crop (harvest) quality detection system will use image classification to analyze various parameters like seed size, shape, weight, color, type, soil, fertilizers used, and climate. By providing a comprehensive quality analysis of each crop, farmers can better market their produce and consumers can make informed purchasing decisions. The e-commerce platform will feature a product catalogue where customers will be able to browse through and select the products they want based on their preferred quality parameters. They can also add the products to the cart, choose a payment method, and complete the transaction in a seamless and secure manner.

This platform has the potential to revolutionize the agriculture industry by empowering farmers to connect directly with customers.

3. Problem Statement

The Farmers' Produce Trade and Commerce Act, of 2020 aimed to give farmers more options for marketplaces and give them more negotiating leverage, enabling them to sell products to anyone even without a license. This would enable them greater flexibility and freedom, but it concerned the farmers as they might get into a poor arrangement with a person causing them to sell harvest at a price less than the minimum support price (MSP).

The recent repeal of farm laws in India has led to a lot of uncertainty and confusion among farmers and traders involved in the agricultural sector. As a result, there is a growing need for a reliable and transparent e-commerce platform that can facilitate the buying and selling of agricultural products in a fair and efficient manner.

Many farmers and traders feel that they are being exploited by middlemen and agents who take a large commission without providing any real value. In addition, there are concerns about the quality and authenticity of the products being sold, as well as the pricing and payment mechanisms.

To address these issues, an e-commerce website for buying and selling agricultural products must be provided as a secure and transparent platform that ensures fair pricing, timely payments, and quality assurance. This can be achieved through a combination of features such as verified seller and buyer profiles, real-time pricing information, secure payment methods, quality checks and certifications. Additionally, the platform should provide relevant information and resources to help farmers and traders navigate the changing regulatory landscape and make informed decisions about their businesses.

The problem statement is:

"To develop a machine learning based e-commerce system to classify the crops based on their quality and predict the price."

4. Need Analysis

To bridge the gap between farmers and consumers, increase transparency and efficiency in the supply chain, and benefit the agriculture industry as a whole, after the disorders created by Farm Laws, we aim to provide a solution that is acceptable to all. We need such a platform because of the following reasons:

- The issue of farmers getting prices lower than the Minimum Support Price, which is calculated by the government of India, can be solved by providing them with a quality & price prediction system for generating newer and better prices.
- The issue of direct communication between the farmers and consumers has been identified from the literature study. Therefore, there is a need to provide quality & price prediction to generate newer and better prices.
- Farmers would benefit as they can now sell the products directly to any buyer depending on the crop quality and current market price, thus removing the need for a middleman and increasing their profit.
- The industry's adoption of quality measuring tools has also encouraged an estimation-based mindset. This, together with machine learning estimations, might make an e-commerce platform which is more effective than traditional selling techniques.
- Use of Machine learning based algorithms will decrease food waste, improve food safety, and help people make better decisions since they would enable farmers to receive better feedback.
- Furthermore, the web application could also offer real-time market data and trends, allowing farmers to make informed decisions regarding their crops and sales. This would enable them to adjust their production based on demand, thereby minimizing waste and maximizing profit.

5. Literature Survey

A quality amount of research has been done on introducing e-commerce in the agricultural field. A few of the works done by some people have been discussed below:

Chen et al. [1] presented their research on the rapid development of agricultural product sales via e-commerce in China in their paper. Their conclusion was that the evaluation of agricultural product suppliers or vendors has become crucial for customers. The paper proposed an evaluation model that uses a case-based distance approach to vendor evaluation by employing the ranking decision in MCDA. The proposed method utilizes the available prior evaluation information as a case set to produce a case-based distance method to handle a multiple-criteria ranking decision more efficiently and effectively. The model is flexible to construct and could simplify and improve the current model of vendor evaluation for agricultural products. A numerical example was used to demonstrate the proposed procedure, and it was proven to be effective.

Huang et al. [2] described that E-commerce has experienced significant growth in almost all fields over the past two decades, driven by computer and internet technology. E-commerce has significantly changed the rules of business, and numerous research institutions and enterprises have made it more intelligent and convenient. The authors propose a novel prototype of a next-generation e-commerce platform with an architecture framework and theoretical models. Each subject, including individuals, enterprises, and administrative departments, has a personalized portal to synchronize subject information, release supplies, satisfy demands, and make social contacts. Using the personalized portal, consumers and suppliers can complete intelligent matching transactions without intermediaries, instead of the traditional trading platform. Moreover, the overall transaction process can be reviewed, making transactions safer, more transparent, and more interesting. The interconnected personalized portals solve the isolated islands of information, and the counterparts support parallel processing, which may improve the operating efficiency of the entire society.

Patel et al. [3] have described that agriculture plays a crucial role in the development of a country's economy and provides numerous opportunities to innovate and apply

innovations. However, trading in agriculture from farmers to consumers goes through a lengthy process where farmers negotiate the yield's price with a middleman. The middleman charges a set fee of brokerage on yield from farms based on quality, which consumes significant time and deteriorates quality, while the farmer gets no return on investment. To overcome these issues, an e-commerce solution can be found, which is transparent enough to ensure that the farmer gets a return on investment with profits and is used without a middleman (broker) to sell the yield. Currently, there is no proper solution that exists for transparent trading between farms and consumers. This paper highlights the importance of e-commerce in agriculture with minimal or no brokerage and emphasizes that it can also help farmers invest in ways that increase the yield.

In their paper, Ali et al. [4] concluded that seed purity is an important indicator of crop seed quality. Moreover, corn is a crucial crop in modern agriculture, accounting for over 40% of worldwide grain production. The objective of this study was to investigate the feasibility of a machine learning (ML) approach for classifying different types of corn seeds. The digital images (DI) of six corn varieties, namely Desi Makkai, Sygenta ST-6142, Kashmiri Makkai, Pioneer P-1429, Neelam Makkai, and ICI 339, were captured using a digital camera in a natural environment without the need for a complicated laboratory system. The acquired DI dataset was converted into a hybrid feature dataset, which combined histogram, texture, and spectral features. For each corn seed image, a total of fifty-five hybrid features were acquired on every non-overlapping region of interest (ROI), including sizes (75 × 75), (100 × 100), (125 × 125), and (150 × 150). Nine optimized features were obtained using the correlation-based feature selection (CFS) technique with the Best First search algorithm. Random Forest (RF), BayesNet (BN), LogitBoost (LB), and Multilayer Perceptron (MLP) were used to build the classification models, employing optimized multi-features using a (10-fold) cross-validation approach. In a comparative analysis of the four ML classifiers, MLP demonstrated an outstanding classification accuracy of 98.93% on ROIs of size (150 × 150). The accuracy values achieved by MLP on the six corn seed varieties, namely Desi Makkai, Sygenta ST-6142, Kashmiri Makkai, Pioneer P-1429, Neelam Makkai, and ICI-339, were 99.8%, 97%, 98.5%, 98.6%, 99.9%, and 99.4%, respectively.

Sobhana et al. [5] presented their research in a paper which concluded that the quality of seeds is crucial for farmers to obtain a good yield. Typically, farmers purchase high-quality seeds from companies, but there is a risk that they might replace good quality seeds with damaged ones while collecting them manually. Therefore, it is necessary to predict the seed quality to prevent this problem. The aim of this study is to provide a solution that eliminates the need for manual seed quality checks in commercial farming. To detect pure and damaged seeds without human intervention, computer vision and deep learning techniques are used. The proposed model uses OpenCV to detect each seed grain in the seed lot and a convolutional neural network to predict the quality of the detected seed grain. The model's output is a prediction of the seed lot's purity percentage. This paper suggests a solution that reduces the amount of manual labor and time required to filter out damaged seeds, thus increasing the efficiency of the farming process.

Challenges:

From the above research papers, we have observed that there is high demand of e-commerce in agricultural field but it comes with its own challenges. We will try to overcome those challenges observed in our project as follows:

1. **Good Prediction Accuracy**: We will make a user-friendly website with a detailed description of the agricultural products and ensure the products are of good quality and provide multiple images of the products.
2. **Partnership with farmers**: We will ensure a partnership with farmers to ensure a constant supply of agricultural products. This will help in maintaining the quality of the products and builds trust with customers.
3. **Competitive pricing**: We will offer competitive pricing to attract customers and stay ahead of the competition. It is also essential to maintain transparency in pricing.
4. **Responsive Application**: We will ensure that the website is mobile-friendly as a significant percentage of user's access websites through mobile devices.
5. **Localization**: Understanding the local market and provide relevant products to customers is also necessary. It also involves providing the products using local currencies.

6. Objectives:

Some potential objectives behind the project are:

- To study the available techniques and approaches of machine learning in crop quality classification.
- To find a relevant dataset (images) for training the model.
- To propose and develop the optimal technique for crop quality classification and price prediction.
- To develop a multi-modular, user-friendly ecommerce application.

7. Methodology:

The methodology would manifest the flow of the processes behind the work of the project.

Part-I

- **Login (as farmer):** Initially the farmer would login to the portal and create a selling inventory/catalogue post. The process would include registering a crop with various meritorious features like seed size, shape, color, type, weight etc.
- **Providing Input:** The next step would include capturing and uploading the seeds images (image acquisition). The images should be of high quality with proper resolution and lighting conditions.
- **Pre-Processing:** This would be followed by pre-processing the images to enhance their quality and eliminate any noise present in them. This step includes operations like image filtering, edge detection, and image normalization.
- **Image Segmentation:** The following stage is for Image Segmentation where the processed image will be then segmented to identify the regions of interest. In this step, the regions of interest like the seed area are separated from the background.
- **Feature Extraction:** The next stage is for Feature Extraction where the features of the segmented seed image will be then extracted. Features can be described as the characteristics of an object that distinguish it from other objects. In seed (harvest) quality detection, features like size, shape, color, and texture are extracted from the seed image.
- **Feature Selection:** This will be followed by Feature Selection where once the features are extracted, the most relevant features will be then selected for classification. The selection of features depends on the type of classification algorithm used.
- **Classification:** The following stage is Classification: In this step, the selected features are fed into a classification algorithm to classify the seed as good or bad. Classification

algorithms like Support Vector Machines, Decision Trees, Random Forest, and Neural Networks can be used for seed quality detection.

- **Quality Prediction:** The following stage will be of Quality Estimation where once the seed is classified, the quality of the seed will be estimated based on the classification result. If the seed is classification result is high, it will be considered of good quality, and if classification result is low, it will be considered of poor quality.

- **Price Prediction:** After the quality of the crop is predicted, it is fed into another predicting algorithm to generate a decent price for the crop – depending on the results generated in the previous step. The price generated is in accordance with the current market prices in the local markets.

- **Deployment:** The prices generated in the previous steps are then deployed/reflected on the webpage for the item being sold and the entirety is closed.

Part –II

If the user is a Customer, he would go through the following path:

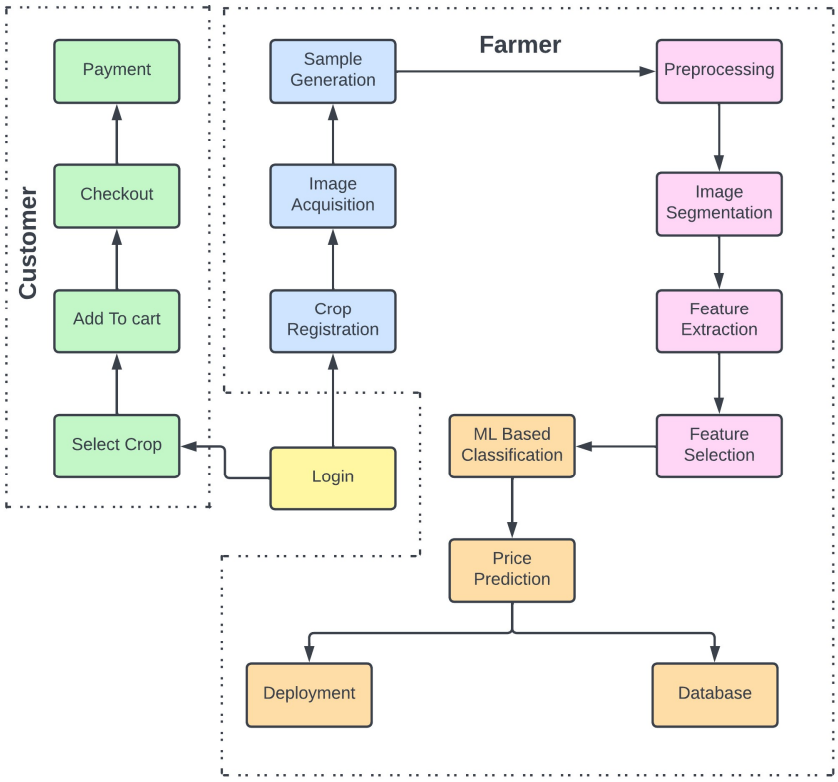
- **Login (as customer):** Initially the customer would login to the portal and will start browsing for a product.

- **Select Product:** Customer will select a product he desires on the basis of price, quality, and availability.

- **Add to Cart:** Then the customer will Add the items to Cart for making a purchase.

- **Payment:** Customer will make a purchase and do the payment.

- **Review (optional):** After purchasing the crops, customers can post a review or a rating for the farmers to give them a feedback.



Dataset:

Datasets with crop features like shape, size, color, type, moisture content, climate would be merged with the results from image processing to generate a final dataset for price prediction. The final outputs of predicted prices would be based on this dataset using regression.

Pre-Processing:

This stage would enable to remove anomalies, outliers and noise from the dataset generated in the preceding stage. If the seller doesn't input the feature data at the time of registration, the prediction could generate poor results. Data smoothing techniques like Z-Score and Interquartile Ranges would be used to remove the anomalies.

Proposed Algorithms:

With the availability of datasets associated with crop diseases and quality categorization, the project would be using algorithms for quality classification using image processing algorithms like – Convolutional Neural Networks, Support Vector Machines, etc.

8. Project Outcomes and Individual Roles:

Outcomes:

The website would be a user-friendly and reliable platform with following outcomes:

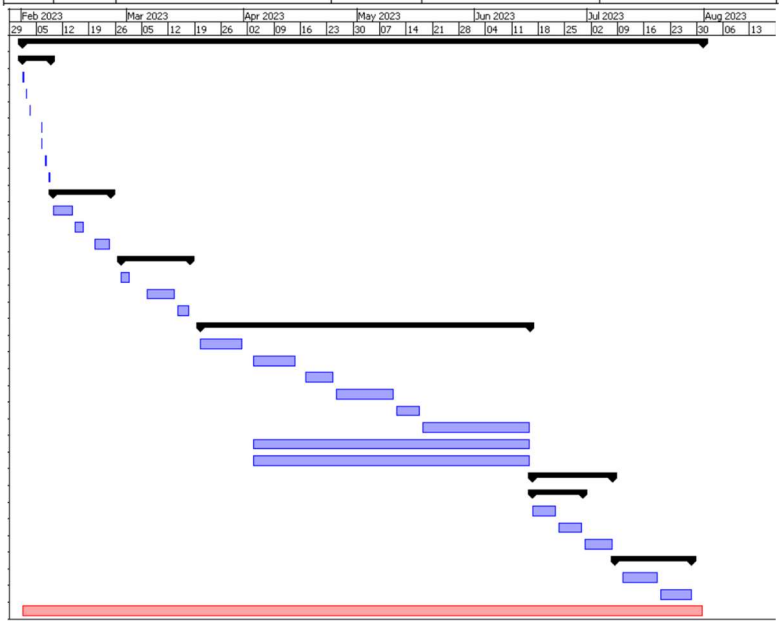
- **Market Outlet Diversification** - The current farming supply market would be supplemented with an online market place that would have genuine crop prices as compared to local markets.
- **Quality Prediction** - The crop qualities would be predicted using machine learning algorithms that can accurately define the standard of the crops.
- **Price Prediction** - The price of the crop would be anticipated upon the quality of the crop predicted by the previous algorithms. The price prediction would also consider the local market prices to predict the new outcomes.
- **Customer Convenience** - Customer who cannot visit local market for purchase would now have a reliable market place, equipped with quality prediction capabilities, to purchase the crops or desired products in a hassle free manner.

9. Work Plan

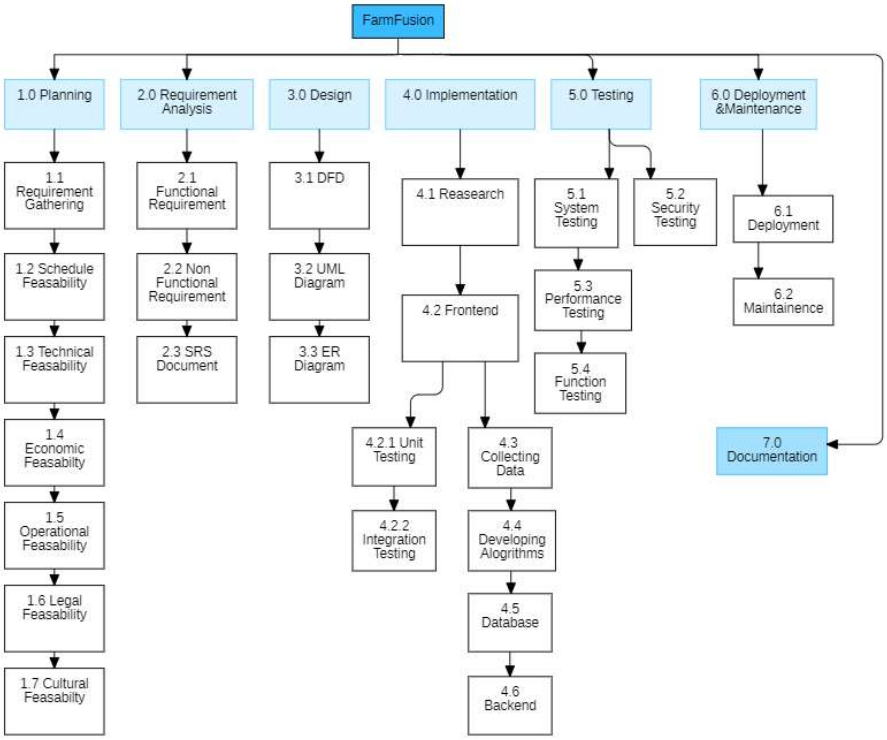
After considerate planning we have come up with the following plan in the form of a Gantt Chart along with a Work Breakdown Structure.

Gantt Chart:

		Name	Duration	Start	Finish
1		**FarmFusion**	**129 days?**	**2/1/23 8:00 AM**	**7/31/23 5:00 PM**
2		**Planning**	**6 days?**	**2/1/23 8:00 AM**	**2/8/23 5:00 PM**
3		Requirement Gathering	1 day?	2/1/23 8:00 AM	2/1/23 5:00 PM
4		Schedule Feasibility	1 day?	2/2/23 8:00 AM	2/2/23 5:00 PM
5		Techinical Feasibility	1 day?	2/3/23 8:00 AM	2/3/23 5:00 PM
6		Economic Feasibility	1 day?	2/4/23 8:00 AM	2/6/23 5:00 PM
7		Operatonal Feasibility	1 day?	2/6/23 8:00 AM	2/6/23 5:00 PM
8		Legal Feasibility	1 day?	2/7/23 8:00 AM	2/7/23 5:00 PM
9		Cultural Feasibility	1 day?	2/8/23 8:00 AM	2/8/23 5:00 PM
10		**Requirement Analysis**	**12 days?**	**2/9/23 8:00 AM**	**2/24/23 5:00 PM**
11		Functional Requirements	4 days?	2/9/23 8:00 AM	2/14/23 5:00 PM
12		Non-Functional Requirements	3 days?	2/15/23 8:00 AM	2/17/23 5:00 PM
13		SRS Document	5 days?	2/20/23 8:00 AM	2/24/23 5:00 PM
14		**Design**	**15 days?**	**2/27/23 8:00 AM**	**3/17/23 5:00 PM**
15		DFD	3 days?	2/27/23 8:00 AM	3/1/23 5:00 PM
16		UML Diagrams	6 days?	3/4/23 8:00 AM	3/13/23 5:00 PM
17		ER Diagrams	4 days?	3/14/23 8:00 AM	3/17/23 5:00 PM
18		**Implementation**	**64 days?**	**3/20/23 8:00 AM**	**6/15/23 5:00 PM**
19		Research	10 days?	3/20/23 8:00 AM	3/31/23 5:00 PM
20		Frontend	10 days?	4/1/23 8:00 AM	4/14/23 5:00 PM
21		Collecting Data	6 days?	4/15/23 8:00 AM	4/24/23 5:00 PM
22		Developing Algorithms	12 days?	4/25/23 8:00 AM	5/10/23 5:00 PM
23		Database	5 days?	5/11/23 8:00 AM	5/17/23 5:00 PM
24		Backend	21 days?	5/18/23 8:00 AM	6/15/23 5:00 PM
25		Unit Testing	54 days?	4/1/23 8:00 AM	6/15/23 5:00 PM
26		Integration Testing	54 days?	4/1/23 8:00 AM	6/15/23 5:00 PM
27		**Testing**	**16 days?**	**6/16/23 8:00 AM**	**7/7/23 5:00 PM**
28		**System Testing**	**10 days?**	**6/16/23 8:00 AM**	**6/29/23 5:00 PM**
29		Peformance Testing	5 days?	6/16/23 8:00 AM	6/22/23 5:00 PM
30		Function Tsting	5 days?	6/23/23 8:00 AM	6/29/23 5:00 PM
31		Security Testing	6 days?	6/30/23 8:00 AM	7/7/23 5:00 PM
32		**Deployment and Maintena...**	**15 days?**	**7/8/23 8:00 AM**	**7/28/23 5:00 PM**
33		Deployment	8 days?	7/8/23 8:00 AM	7/19/23 5:00 PM
34		Maintenance	7 days?	7/20/23 8:00 AM	7/28/23 5:00 PM
35		Documentation	129 days?	2/1/23 8:00 AM	7/31/23 5:00 PM



Work Breakdown Structure:



10. Course Subjects:

- Machine Learning - UML501
- Software Engineering - UCS503
- Artificial Intelligence - UCS411
- Database Management System - UCS310
- Innovation and Entrepreneurship - UTA012
- Source Code Management - UCS537
- Fundamentals of Data Science - UCS538

11. References

1. Wangquan Sun, Y. Chen and HaiYan Xu, “A case-based distance approach to agricultural product's vendor evaluation in e-commerce”,2015 IEEE International Conference on Grey Systems and Intelligent Services (GSIS), Leicester, UK, 2015, pp. 439-445, doi: 10.1109/GSIS.2015.7301897.
2. Y. Huang, Y. Chai, Y. Liu and J. Shen, “Architecture of next-generation e-commerce platform,” in Tsinghua Science and Technology, vol. 24, no. 1, pp. 18-29, Feb. 2019, doi: 10.26599/TST.2018.9010067.
3. Patel, A.N., Ahmed, A., Rohini, B.R. (2022). “An Approach Towards E-Commerce for Agriculture with Modern Technologies.” In: Kumar, A., Senatore, S., Gunjan, V.K. (eds) ICDSMLA 2020. Lecture Notes in Electrical Engineering, vol 783. Springer, Singapore. doi: 10.1007/978-981-16-3690-5_92.
4. Aqib Ali, Salman Qadri, Wali Khan Mashwani, Samir Brahim Belhaouari, Samreen Naeem, Sidra Rafique, Farrukh Jamal, Christophe Chesneau & Sania Anam (2020) “Machine learning approach for the classification of corn seed using hybrid features”, International Journal of Food Properties, 23:1, 1110-1124, doi: 10.1080/10942912.2020.1778724.
5. M. Sobhana, P. Dabbara, G. Ravulapalli and K. S. Kakunuri, “Seed Quality Prediction using Computer Vision and Convolutional Neural Networks,” 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2022, pp. 1134-1138, doi: 10.1109/ICIRCA54612.2022.9985705.