Optimizing the Visibility of Agricultural activities on the Farm using a sound analytics platform

MSc Research Project
Research in Computing

Manoj Kanthraj
Student ID: 21218315

School of Computing
National College of Ireland

Supervisor: Sean Heeney

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Manoj Kanthraj
Student ID:	21218315
Programme:	Research in Computing
Year:	2023
Module:	MSc Research Project
Supervisor:	Sean Heeney
Submission Due Date:	18/09/2023
Project Title:	optimizing the Visibility of the Agricultural activities on the Farm using a sound analytics platform.
Word Count:	5997
Page Count:	21

I hereby certify that the information contained in this (my submission) is informationpertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at therear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Manoj Kanthraj
Date:	18th September 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placedinto the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Optimizing the Visibility of Agricultural activities on the Farm using a sound analytics platform

Manoj Kanthraj
21218315

Abstract

Norway's agricultural sector is significant culturally and economically, and it thrives despite severe challenges. The production of dairy products is at the top, followed by those of sheep and pigs, with salmon aquaculture also on the rise. Governmentally approved strict criteria for biodiversity, emissions reduction, and sustainability (including livestock welfare) are used to cultivate crops including grains, potatoes, and organic vegetables. Sound-based analytics are used by cutting-edge AI technologies for monitoring farm machinery. The method uses supervised classification to pick out relevant audio data from the noise and then uses one-class classification to identify unique content. Over 20 distinct farming activities may be identified using multi-class classification, with assistance from farm boundary detection. With the Urban Sound dataset, we can see that Random Forest achieves 90% accuracy in just 22.83 seconds, which is significantly faster than XGBoost's time of 326.69 seconds. XGBoost can identify both human voices and garbage in 0.81 seconds, but Random Forest takes 1.27 seconds to do it with just 85% accuracy. XGBoost can identify agricultural tasks with 97% accuracy in about 0.04 seconds, while Decision Trees can only get to 95% accuracy. These changes show how technology is improving life in Norway's rural areas.

Keywords: Agriculture, Machine Learning, Fast Fourier Transform, Mel Frequency Cepstral Coefficients, XGBoost, Ensemble, Space Complexity

1. Introduction:

Even though the land in Norway is hard to work, the agricultural business, which plays a significant role in the history of the nation's culture and economy, is flourishing. The dairy, sheep, and pig industries, as well as the growing fishing business, all make a big difference. Organic farming is becoming more popular, and the most common crops are grains, potatoes, and veggies. The government has strict rules that give priority to wildlife, reducing waste, and living in a sustainable way. AI is changing agriculture in many ways, such as by making it easier to find diseases and pests, controlling weeds, and applying pesticides more accurately. When robots with AI and computer vision are used, the amount of herbicide that goes to waste is cut by up to 80%. AI-driven precision agriculture is important to increase food production in a way that doesn't hurt the environment, keep food quality high, and meet the growing demand for food around the world. This method, which is powered by artificial intelligence, brings in a new age of farming by letting farmers make informed decisions in real-time (Knutsen, H., 2020; Olsen, T.O., 2020; Lundmark Hedman et al., 2021).

1.1 Business Objective

One of the difficulties that farmers encountered in the middle of the 2000s was a lack of efficient field data management, which necessitated the integration of several devices, such as satellite photos and the Raven Receiver for field tracking. The need for a comprehensive system to streamline agricultural processes became apparent in this era. Due to the country's agricultural constraints, such as the short growing season, an automated task management system is crucial for maximizing harvests in Norway. Providing automated guidance and tracking to help farmers properly manage their duties is intended to increase profitability in the agriculture-dependent segment of the Norwegian economy. The ultimate goal is to equip farmers and landowners with these smart technologies so that they may increase their agricultural enterprises' profits.

1.2 Motivation

The use of robotics and automation have been employed in increasing numbers in agriculture to boost output. Agriculture is the second-biggest sector for service robots, where sensors that are vision systems, and artificial intelligence are utilized to detect harvestable produce. By hearing sound records recorded by farmers utilizing their smartphones, the research issue seeks to enhance agricultural operations.

1.3 Impact of the Problem

An apparent view of the types of tasks being completed on the farm is provided to the farm owners through an automated task management system, which takes sound as input and attempts to propose the following activity. The owner of the work now has transparency into what is going on and can optimize the task based on a predefined set of actions.

This has the following impact,

1. Building a unique database using any adaptable smartphone.
2. Activity tracking based on the sound of agricultural activity.
3. Reduction of the number of wasted hours.

1.4 Outcome

The research's output is the creation of a system that can manage a large amount of data flowing in from several farmlands and from various farmers operating in each farmland. The audio recordings must be organized according to each user that has logged into the system. The computation of the algorithms should be optimized by the cloud platform in addition to space optimization. To prevent any false predictions from disrupting the workflow, the performance of the installed algorithms should be excellent. In accordance with GDPR regulations, human voices should not be kept; instead, the system should be able to discern between the human voice, undesirable sounds, and appropriate sounds. The system must be capable to handle and categorize any identified unknown activity to function more effectively. A self-learning module with feedback should be included in the system's deliverables so that inaccurate labels and unidentified files can be taught and labeled for better performance.

1.5 Research Questions

The sound recorded by the smartphones will be used throughout the entire scenario. The entire process is fraught with difficulties, including GDPR concerns with the handling of personal data, memory and storage limitations of the smartphones used to record sound data, and model performance on top of all of that.

Based on this the following are the research questions designed,

RQ1: How can this solution be made automated so that Farmers may use cloud platforms to help them make decisions?

RQ2: Which machine learning algorithm performs best when sound signals are considered, and what are the best techniques for extracting sound signal features?

RQ3: How can the solution be deployed so that cloud handling and computation costs are not impacted?

RQ4: How may this outcome be applied in real time?

A detailed review of various works of literature produced by various researchers will be examined in the section that follows. Based on this, the methodology will go into detail about the approaches and procedures. The planned process flow will be explained in the chapter on implementation, and all the examples will be thoroughly examined in the chapter on results.

2. Related Work:

Abioye et al. (2022) offer an ambient sound categorization (ESC) model that uses STFT spectrograms to work with mono and stereo inputs. Using methods from the picture domain, such as ResNet, Siamese networks, and attention, they overcome challenges associated with leveraging successes from other domains. The study analyses popular datasets, refines the design, and explores the impact of pre-training to advance ESC.

2.1 Artificial Intelligence in Agricultural Sector Enhancement

Abioye et al. (2022) focus on smart irrigation using machine learning models to enhance agricultural water management. With the extensive use of freshwater in agriculture as a backdrop, this project explores mobile and online frameworks for managing irrigation operations. By relieving farmers of the burden of constant on-site monitoring, digital technologies open the door to more efficient water management. Potential problems and research directions for this area are also mentioned. Bao and Xie (2022) examine the applications of AI in the animal husbandry sector in great detail. The analysis of 131 scientific publications centred on the detection and identification of animal behaviour, particularly in swine, cattle, and poultry. Animal behavioural recognition and classification were the main topics of discussion. This research relies heavily on experimental methods of data collecting, which include themes including disease tracking, population projections, and environmental tracking. The challenges of implementing technological solutions into industrial animal production are highlighted. The characteristics of machine learning and the key distinctions between traditional programming and machine learning are outlined by Anagnostis et al.

(2022). In this study, we focus on the potential of machine learning algorithms to enhance agricultural productivity and product quality, and we examine a number of techniques and their applications. Research on agricultural machine learning algorithms published in academic journals between 2018 and 2020 is summarised in the last portion of the study. The Multimedia-assisted Business Evaluation Model (MBEM) is presented by Zhong et al. (2022). The MBEM is a machine learning-based DSS designed to improve company effectiveness via the evaluation of alternative business models. The MBEM paradigm integrates creative thinking with cutting-edge technology to boost a business's market competitiveness. Issues related to scarce resources are given special attention. The experimental results demonstrate the model's potential to enhance conventional approaches to business evaluation, decision-making, performance, and profitability.

2.2 Application of Artificial Intelligence and Machine Learning in Sound signals analysis

New approaches to real-time tool condition monitoring during machining processes are presented by Pimenov et al. (2022). These methods employ a wide range of sensors and various forms of artificial intelligence. There are many different types of models that fall under this umbrella, including but not limited to: Bayesian networks, support vector machines, ensembles, decision and regression trees, k-nearest neighbours, artificial neural networks, markov models, single spectrum analysis, and genetic algorithms. This study explores the potential for employing these technologies to automate turning, milling, drilling, and grinding processes. The effects of excessive tool wear and precise machining responses are emphasised. Khorasani et al. (2022) look at the potential of deep learning networks in agricultural analysis, namely in the areas of crop image analysis and equipment noise analysis. The major focus of this study is on using deep learning models for sound analysis of harvesters to ensure their operation is safe. CNNs and stacking ensembles are used to classify spectrogram images into safety modes in under a second, guaranteeing the well-being of harvesters in the process. Through a series of votes, the research achieves a perfect classification rate of 100%, demonstrating the need of rigorous voting procedures. Using information gathered from a variety of sensors, Gultekin et al. (2022) investigate the problem of intelligent failure diagnosis in ATVs. Their deep learning and convolutional neural network-based multimodal defect identification system provides more reliable diagnoses than either single or dual sensor approaches. The experiment demonstrates the importance of accurate defect identification as these vehicles gain in popularity throughout the world through the collection of sound and vibration information from two ATV engines. An in-depth analysis of sensor data can reveal hidden problems in industrial processes, leading to better overall safety.

2.3 Feature Extraction from different sound signals

Li et al. (2019) classifies cardiac sound waves as normal or pathological using a one-dimensional convolutional neural network (CNN) model that includes denoising autoencoder (DAE) produced deep features. The model outperforms competing neural network architectures known as backpropagation neural networks (BPNs) by a wide margin (99.01%). In their comprehensive review of the research on audio signal processing, Sharma et al. (2020)

highlight the importance of feature extraction in boosting the efficiency of machine learning algorithms. This study investigates issues in the time and frequency domains and highlights the need to integrate state-of-the-art machine learning strategies with audio signal processing to effectively address these challenges. Chen et al. (2019) looks at how well-stretched convolutional neural networks (CNN) can classify background noise. Dilated CNN is shown to be superior to other methods, including CNN with maximal pooling, according to the study's findings. The research also stresses the need of increasing the dilation rate and expanding the frame coverage for better accuracy. The results also show that the size of the overlay frame significantly affects the feature extraction process and that sound signals are stable for brief periods of time. Deng et al. (2020) provide a unique method for categorising heart sounds that makes use of enhanced Mel-frequency cepstrum coefficient (MFCC) features and convolutional recurrent neural networks (CRNN). Neural networks are used in this process. It is possible to compute the MFCC without splitting the signal, and improved feature extraction may be achieved with CRNN. Using a combination of convolutional neural networks (CNNs) and recurrent neural networks (RNNs), the deep learning system shows great potential for accurate heart sound classification. The success of the framework depends on this mix.

2.4 Sound Classification using different machine learning applications deployedin cloud

Meshram et al. (2021) offer a deep neural network–based strategy for acoustic monitoring of bee activity. Mel-frequency cepstral coefficients (MFCC) are extracted from beehive audio recordings to track bee activity. The assessment of IoT-based systems makes use of both lossless WAV and lossy MP3 formats. The method achieves a 94.09% classification accuracy using uncompressed audio, which is a significant improvement over the prior hidden Markov model approach. However, adopting the MP3 format reduces the precision of deep neural networks by more than 10%. Classifying agricultural equipment based on data gathered from accelerometers and gyroscopes is the focus of Waleed et al. (2021), who provide multi-class supervised machine learning as a solution. Rotavators, levellers, and cultivators, for example, are classified using algorithms like KNN, SVM, DT, Random Forest, and Gradient Boosting. When vibration and tilt data are coupled, the accuracy of categorisation improves, with Random Forest proving to be the most accurate. However, overfitting has been observed in Random Forest and Gradient Boosting. Gradient boosting requires the most time to train, whereas decision trees take the least.

2.5 Summary of the above research

IaaS, PaaS, and SaaS are three forms of cloud computing that increase agricultural efficiency by optimizing resource utilization and crop output. The hybrid cloud and fog computing approach combine in-field real-time data processing with cloud-based data storage, analysis, and decision-making. In a mobile cloud approach, data is gathered via mobile devices, while storage and processing are handled by cloud computing. Heart sound analysis (Li et al., 2019), audio feature extraction (Sharma et al., 2020), categorization of ambient sounds (Chen et al., 2019), and classification of heart sounds (Deng et al., 2020) are just a few examples of the many ways AI may be put to use. Research in this area focuses on improving accuracy and productivity in a range of sectors, from agriculture to industry and beyond. Irrigation (Meshram et al., 2021) and animal farming (2021) are two examples of sectors in which machine learning has been applied, and their potential and challenges have been studied.

3. Methodology and Design Specifications:

3.1 Research Resource – Data Understanding

Dataset 1 – Agricultural Real-Time Dataset: The information is gathered with the use of a smartphone app and then filed away according to the various tasks performed. Online data scraping and collaborations are used to collect agricultural information. While "garbage sounds" and tagged audio could coexist, the use of a startup simplifies feature extraction. The network architecture in use for task detection in real-time is shown in Figure 1. There are in total 23 classes in the dataset while for the data visualization, we have 5 classes for ease of explanation.

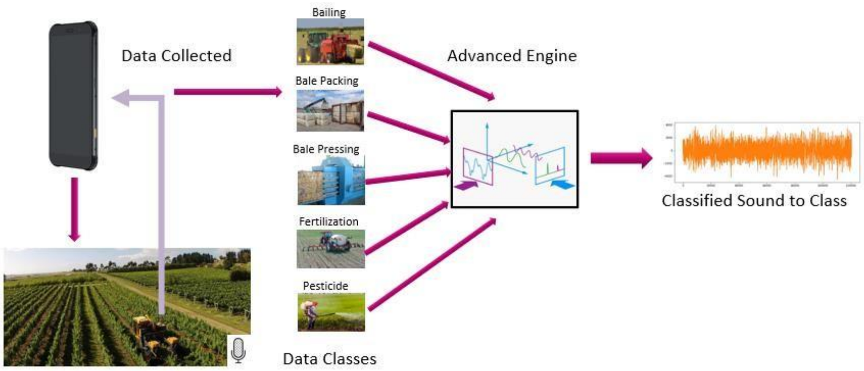

Figure 1: Method of Collecting the Data (Source: Analytic Labs)

GDPR: Classes should be named with keywords like "Human Voice," "Music," and "No Sound," and data training should be performed with and without the usage of human voices. Information that does not meet the criteria will be deleted using a conditional approach. This comprehensive answer addresses both the need for reliable real-time task identification and the need to protect individual confidentiality.

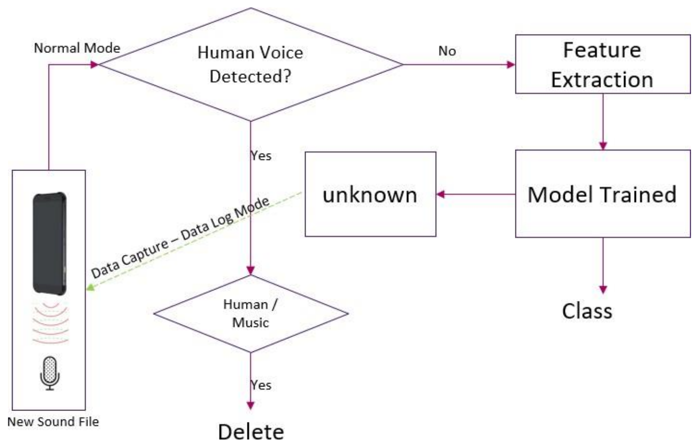

Figure 2: Network Design for the Automated Task Identifier in real time

Standard Dataset – Urban Sound Dataset: The sounds of cities all throughout the world are represented in the urban Sound archive. These recordings provide a plethora of urban sounds, from horns and sirens to street music and beyond. In the realm of audio analysis and classification, the dataset is frequently used as a benchmark for several machine learning approaches. It helps with the development and evaluation of sound recognition algorithms and features many different types of urban noise. Researchers and data scientists frequently use the Urban Sound dataset to test and refine automatic sound classification, environmental monitoring, and urban soundscape analysis methods.

3.2 Research Resource – Features Extraction

Most of the characteristics that are derived from agricultural data are MFCCs and their variations, as well as FFT and their variants. The amplitudes of the frequency plot are used to classify the activities that produce these features. distinct activities result in distinct curves because of the unique combinations of MFCC interactions.

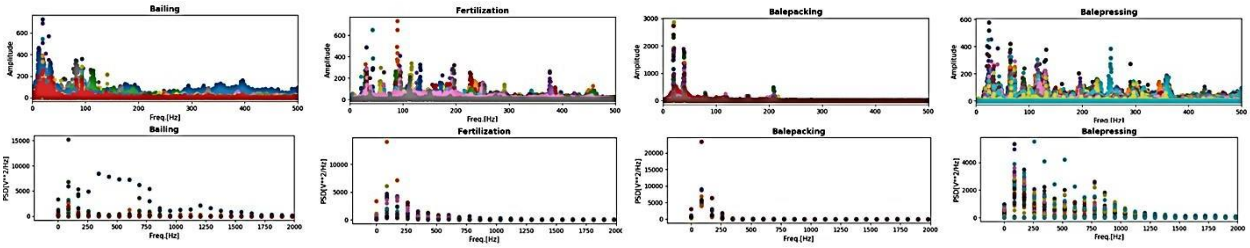

Figure 3: FFT and PSD for 4 different randomly chosen farm activities

Differentiating between the various agricultural activities through the characteristics' amplitude and curve differences improves the likelihood of accurate classification and analysis within the domain.

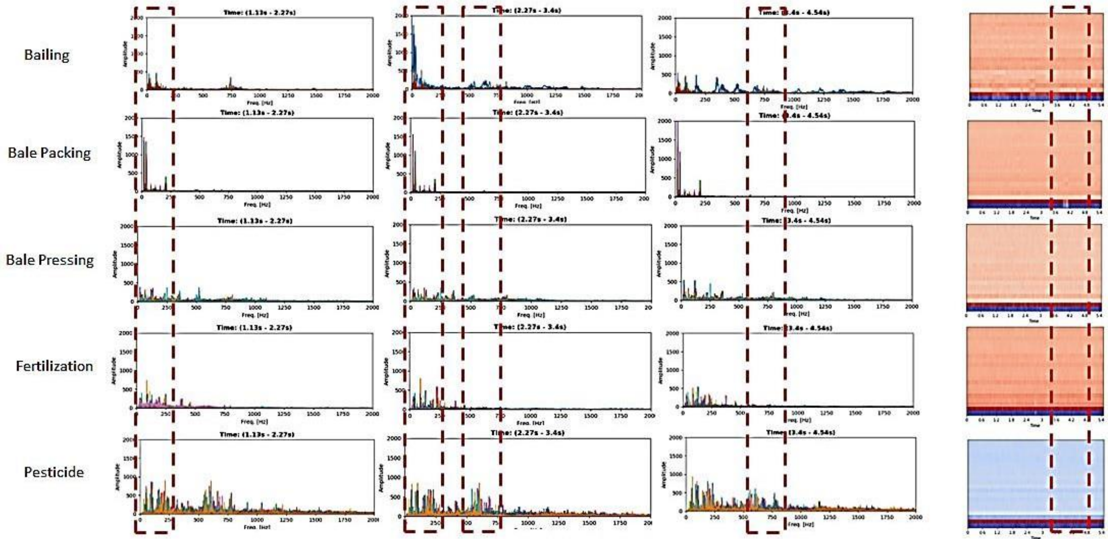

Figure 4: Relationship between different FFT and PSD features for different activities with the MFCC activities

The dataset that is made from the features extraction process has the following feature information,

- Fast Fourier Transform Features: Columns 0 to 19
- Power Spectral Density: Columns 20 to 39
- Auto Correlation Features: Columns 40 to 59
- Mel Frequency Cepstral Coefficients: Columns 60 to 99
- Chroma Features: Columns 100 to 111
- Mel Features: Columns 112 to 239
- Contrast Features: Columns 240 to 246

- Tonnetz Features: Columns 247 to 252

Once these features are extracted, these are passed to the machine learning models to be trained and saved. The saved files can then be deployed on to the application system to perform in the real time environment.

3.3 Research Resource – Machine Learning Modelling

The dataset is samples for training and testing and the training samples are passed to the machine learning models to get trained and perform on the testing samples to check the best models. For this whole pipeline of the model building,

a. Samples are made with both Hold Out and K-Folds with 5 Folds.
b. All the different Combination of Features are used for the detailed analysis with different algorithms.
c. Hyper Parameters are tuned for each mode. The setting of Hyper parameters is as below,
    a. Logistic regression: Penalty with Values as L1 norm and L2 Norm, C
    b. Gaussian Naïve Bayes: The hyper parameters Var Smoothing
    c. Support Vector Classifier: Kernel, Gamma, and C
    d. Decision Trees: Criterion and Max Features
    e. Random Forest: Number of Trees and Max Features
    f. XGBoost: Learning Rate, Gamma, and Maximum Depth
    g. K Nearest Neighbors: Number of Neighbors, Weights, and Metric.

Since different classes can have different number of datasets, the sampling taken here is Stratified Sampling in which the ratio of each of the classes is same to the original ration in both Training and Testing Samples.

4. Implementation:

The flow diagram of the functioning of the solution is presented below,

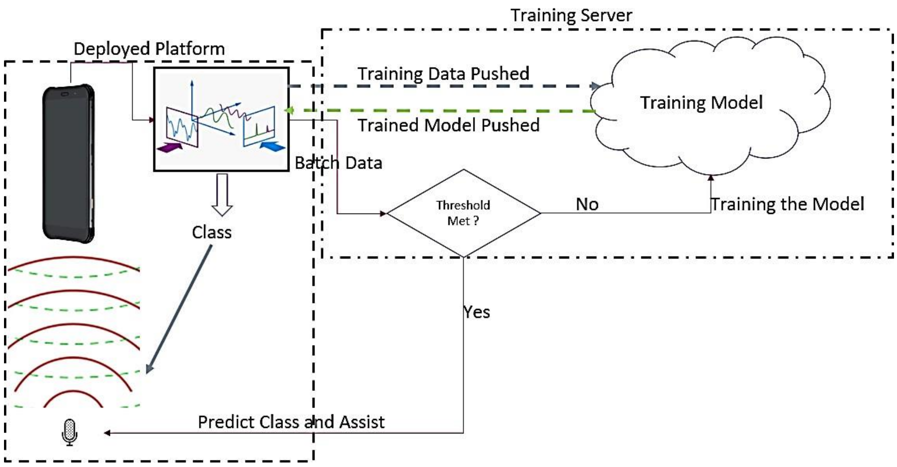

Figure 5: Complete analysis suite, including retraining component and Amazon Web Services (AWS) cloud deployment, required to identify optimal model performance.

Algorithm Flow:

Step 1 – Smart Audio Capture: Audio is gathered for two different sets using a mobile phone app.
Step 2 – Feature Extraction: Several characteristics are extracted at this stage, including the FFT, Auto Correlation, and MFCC. These many features are the food for the machine learning process.

Step 3 – Information Retrieval: The General Data Protection Regulation (GDPR) filtering technique may identify and separate human voices from background noise. Only valid signals are retained, while invalid ones are discarded.
Step 4 – Modelling: Machine learning is able to help in identifying the classes. The data is initially split into training and testing sets so that various machine-learning models may be compared.
Step 5 – Validation: To do real-time validation, a trained feature extraction and ML model must be uploaded to the cloud where it may be subjected to live testing.
Step 6 – Deployment: Based on remote servers Improvements to the Model If the model's performance falls below a threshold, it will be retrained on the cloud, resulting in more recently trained files.

4.1 Cloud Model Deployment

Due to its greater accuracy and durability, XGBoost was chosen as the best method for deployment after a thorough review of numerous machine-learning models where I have mentioned all the performances of the model in the evaluation section. For the cloud deployment, I have used the EC2 instance (54.216.255.188) where I have developed a Flask web application tailored for predicting classes, seamlessly integrating with the machine learning model. This application enables file uploads through Postman, a versatile API testing tool. In the Flask app, a dedicated endpoint /prediction is established to handle incoming POST requests. When a file is uploaded using Postman, it is received and processed within the Flask app. The uploaded file is retrieved from the request using the request. files['sound_data']. Predictions are then generated utilizing the trained machine learning model. The predictions are finally returned as a JSON response using the jsonify function. This comprehensive process ensures efficient class predictions via file uploads, seamlessly bridging the gap between the user interface and the underlying predictive model. Fig 6 prediction result in postman.

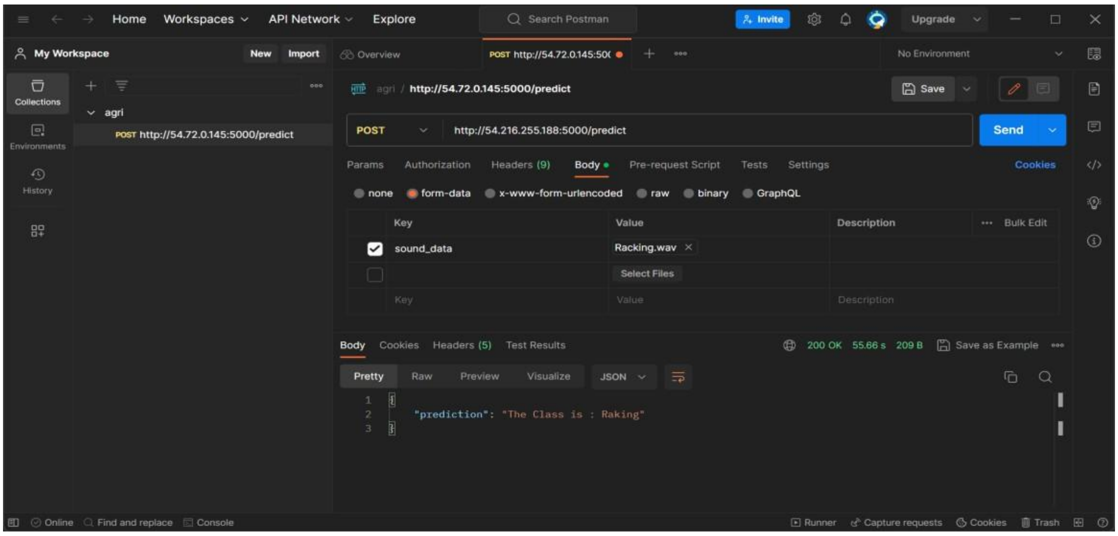

Figure 6: Prediction Result in Postman.

5. Evaluation:

In this section, a complete analysis will be done to check the hypothesis of this pipeline in sound detection. Since the proposal is to make a generic pipeline that will train the model in the cloud and generate a trained file that will be deployed in the smart phone. The different classes to be checked are Aerate gras, Precision chop silaging, Self-loading, wagon, Bailing, Balepacking, Balepressing, Fertilization, Forge Wagon, Mowing, Pesticide, Raking, Sowing, Transportation classes will/can be included as the system is deployed in the real-time. Of these above classes, there are also sounds of Human Voice, No Sound, Garbage Sound, and Music. These can be considered as Not Applicable sounds and need to be treated. But to prove that this methodology works out well, we have considered the urban sound to check on this. Let's start with the analysis of different feature extraction and machine learning models on Urban Sound Data.

Experiment / Case Study 1: Analysis of the Urban Sound Data

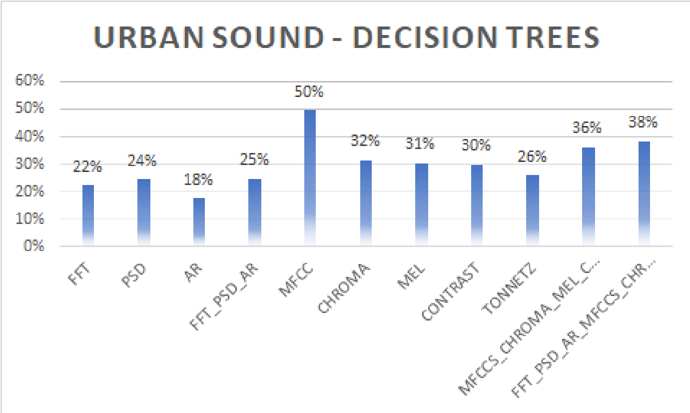

Figure 7: Decision Trees performance on Urban Sound Data

Decision trees didn't perform well on the Urban Sound. The maximum accuracy obtained is 50% and the features used were MFCC. For all the other features the accuracy is less than 40%.

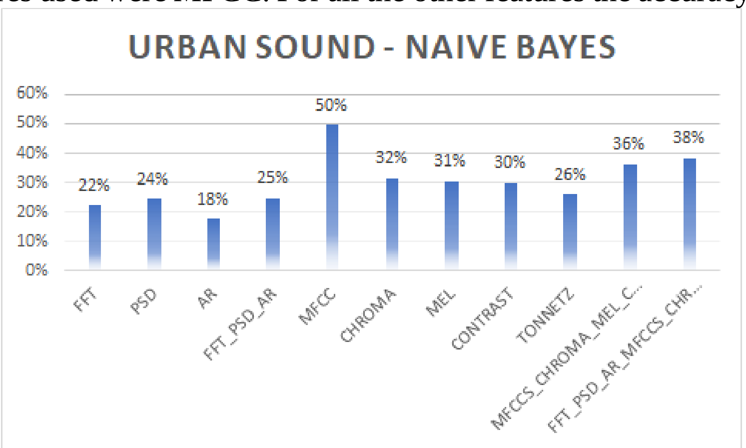

Figure 8: Naïve Bayes performance on Urban Sound Data

Similar kind of performance is also seen with Naive Bayes classifier.

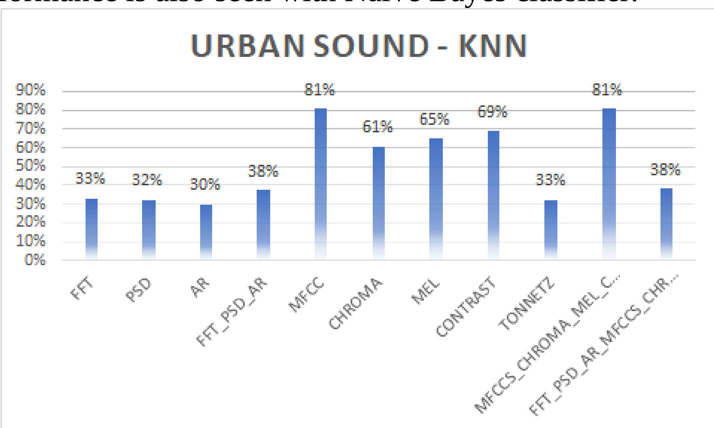

Figure 9: K-Nearest Neighbors performance on Urban Sound Data

K-Nearest Neighbors performed better than Decision Trees and Naïve Bayes. The maximum accuracy was obtained using MFCC and a combination of MFCC, Contrast, Tonnetz, Mel and

Chroma. If the complexity is taken into the picture, the MFCC performance with Naïve Bayes was better. Since with lesser information the model was able to achieve the better results.

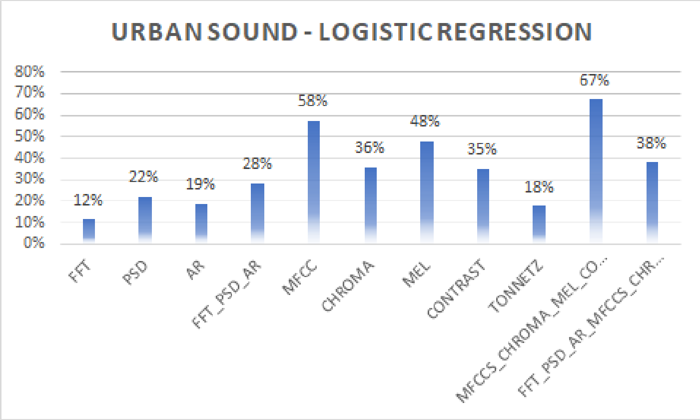

Figure 10: Logistic Regression performance on Urban Sound Data

Logistic regression performed well with a combination of MFCC, Contrast, Tonnetz, Mel and Chroma features with an accuracy of 67%. Using the MFCC alone, the model was able to achieve 58%.

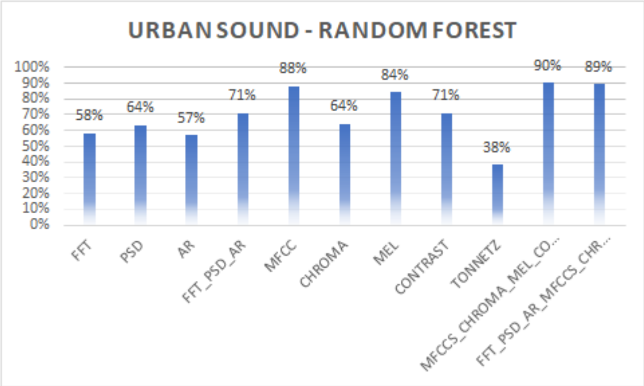

Figure 11: Random Forest performance on Urban Sound Data

Random Forest performed well with a combination of MFCC, Contrast, Tonnetz, Mel and Chroma features achieving an accuracy of 90%. With only MFCC the model performed with 88% accuracy.

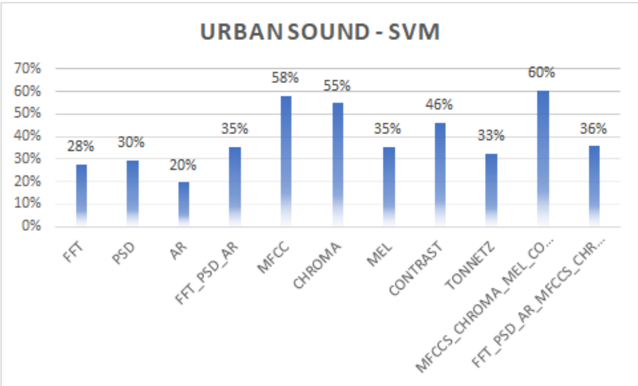

Figure 12: Support Vector Machines performance on Urban Sound Data

Support Vector Machines didn't perform well on Urban Sound data. The maximum accuracy was obtained using the combination of MFCC, Contrast, Tonnetz, Mel and Chroma features with 60% accuracy.

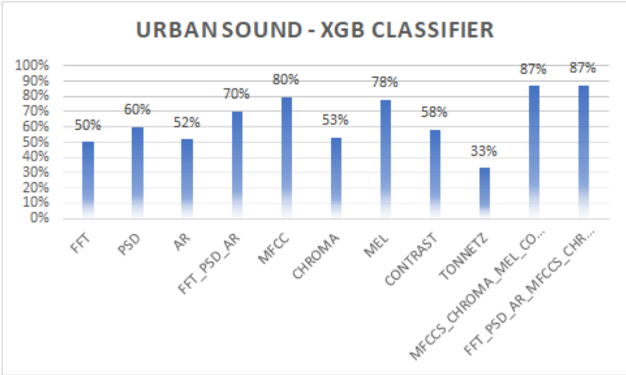

Figure 13: XGB Classifier Performance on Urban Sound Data

XGBoost also performed well but second to Random Forest with 87% accuracy. The features had to be made a lot of complex and full of additional features.

From the above summary report, it is found that Random Forest and XGBoost performed well. The following table discusses the ability of the performance and the computation time taken to run the whole pipeline.

Table 1: Summary report of Random Forest and XGBoost Performance on Urban Sound

Model	Features	Mean(STF_ACC)	Std_dev (STF_ACC)	VAL_ACC	Exec. Time (s)
RANDOM FOREST CLASSIFIER	MFCCS_CHROMA_MEL_CONTRAST_TONNETZ	90%	1%	90%	22.83
XGB CLASSIFIER	FFT_PSD_AR_MFCCS_CHROMA_MEL_CONTRAST_TONNETZ	87%	2%	88%	326.69

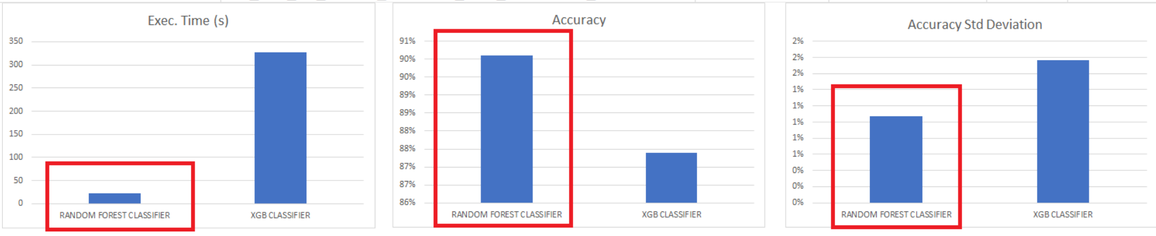

Figure 14: Performance Comparison of Random Forest and XGBoost in terms of (a) Execution Time, (b) Validation Accuracy, and (c) Accuracy Standard Deviation in terms of K-Folds

From the above figure, we find that Random Forest performed extremely well in identifying the validation sound samples with an accuracy of 90% at a faster computation time of 22.83 seconds in comparison to XGBoost of 326.69 seconds and with higher stability. Stability is measured with the ranges of accuracies achieved by the model in different Folds. Random Forest had a 1% standard deviation which is lower than XGBoost with 2% standard Deviation inaccuracies. Hence, to conclude the analysis of Urban Sound standard dataset classification, our pipeline model performed well, and Random Forest turned out to be the selected model. This analysis will be carried out in case of the GDPR issues prediction i.e., recognizing and predicting the Human Voice and checking the Agricultural activities.

Experiment / Case Study 2: Analysis of Human Voice detection to take care of GDPR issues.

In this analysis, four different classes will be predicted which are Human Voice, Garbage, Music, and No Sound if any. Till the time these classes are present the sound signals will be discarded and won't be passed to the agricultural sound detection system. Hence, this step is very important in terms of how perfectly the prediction happens and in terms of GDPR issues that no personal information to be passed into the system or saved.

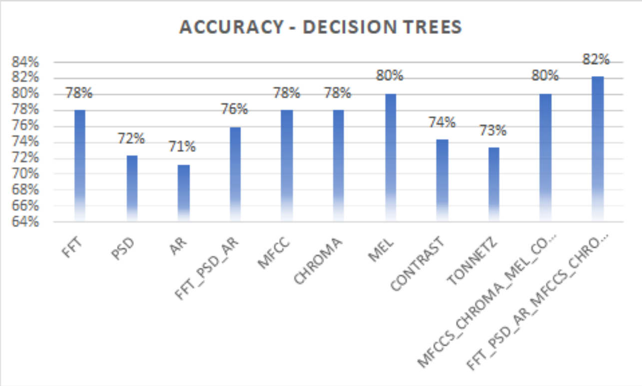

Figure 15: Human Voice and Garbage sound Signals Detection using Decision Trees

Decision trees got an accuracy of 82% with the inclusion of all the features that were extracted from the sound signal. Using Mel features, it achieved an accuracy of 80%.

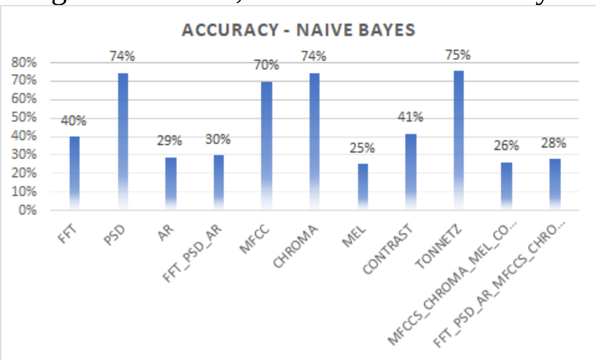

Figure 16: Human Voice and Garbage Sound Signals Detection using Naïve Bayes

Naïve Bayes didn't perform better than Decision Trees. The maximum accuracy of 75% was achieved with the Tonnetz feature.

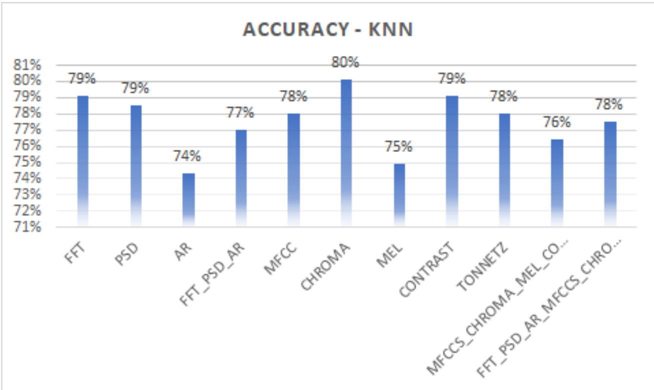

Figure 17: Human Voice and Garbage Sound Signals Detection using K-Nearest Neighbors

K-Nearest Neighbors achieved an accuracy of 80% using Chroma features followed by 79% using FFT and Contrast features.

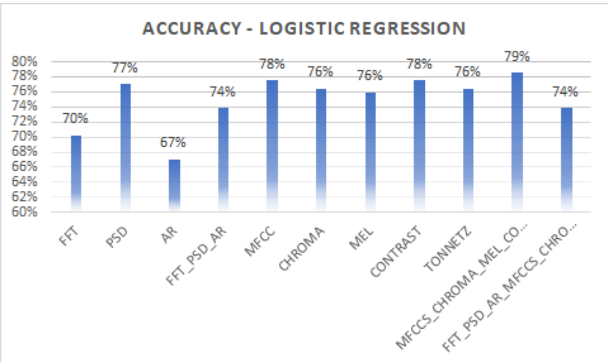

Figure 18: Human Voice and Garbage Sound Signals Detection using Logistic Regression

Logistic Regression also performed with an accuracy of 79% using the combination of Mel, MFCC, Tonnetz, and Contrast features. All the features achieved a similar range of 70-78% accuracy.

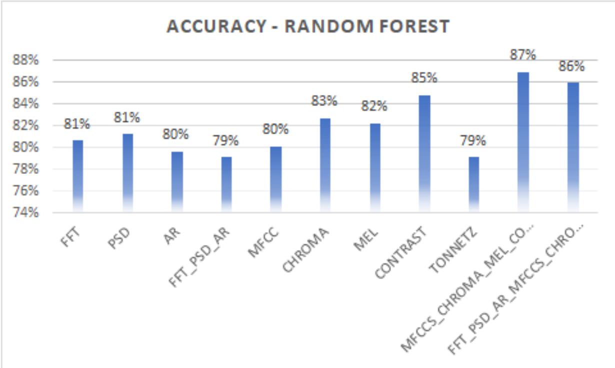

Figure 19: Human Voice and Garbage sound Signals Detection using Random Forest

Random Forest performed well with an accuracy of 87%, 86% and 85% for the features combination of Mel features (MFCC, MEL, Tonnetz, Contrast and Chroma), all the features combination and Contrast features respectively.

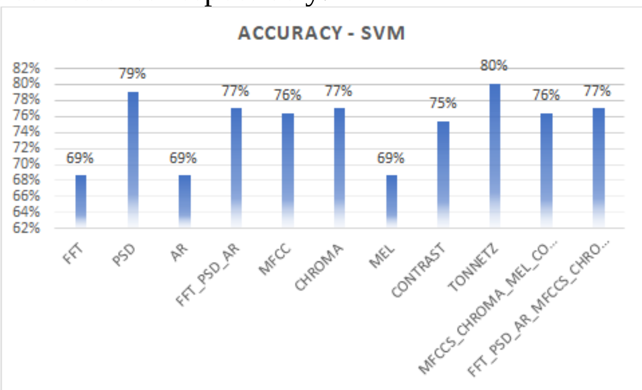

Figure 20: Human Voice and Garbage sound Signals Detection using Support Vector Machines

Using Tonnetz as the feature SVM achieved an accuracy of 80%.

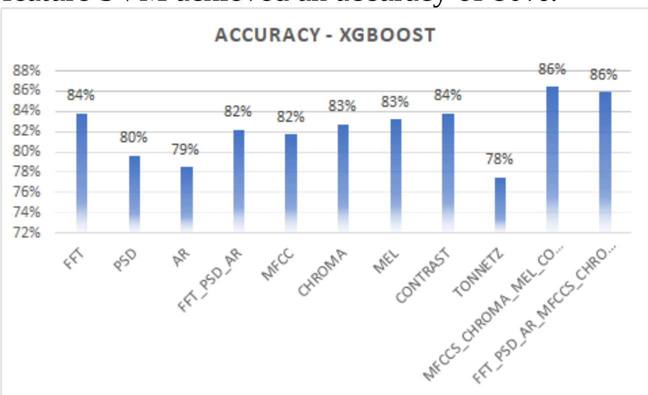

Figure 21: Human Voice and Garbage sound Signals Detection using XGBoost

XGBoost performed like Random Forest but the accuracy was lesser than Random Forest.

In this analysis, Random Forest performed the best while if the complexity of the problem is taken into the picture, then the algorithm should be fast as well. Hence, using Contrast as feature, random forest achieved an accuracy of 85% which is like XGBoost. When the execution time is checked,

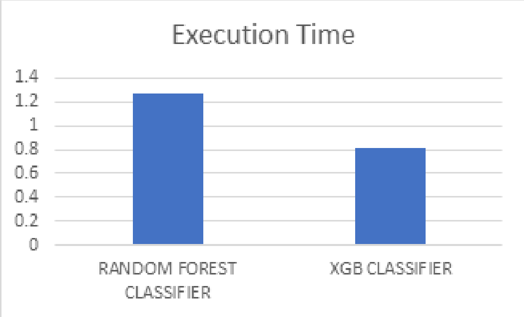

Figure 22: Execution Time comparison of Random Forest and XGBoost for Garbage Sound Detection

XGBoost is fast with a computation time less than one second i.e., with 0.81 second in comparison to Random Forest with 1.27 seconds. Hence as a conclusion, XGBoost will be deployed for canceling the garbage sound from the data.

Experiment / Case Study 3: Analysis of the correct activity detection from the Agricultural Data

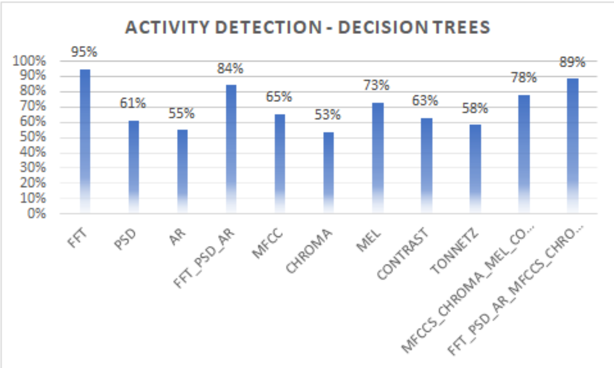

Figure 23: Activity detection from Agricultural data using Decision Trees

Using FFT feature as the input, decision trees achieved an accuracy of 95% followed by 895 using all the features.

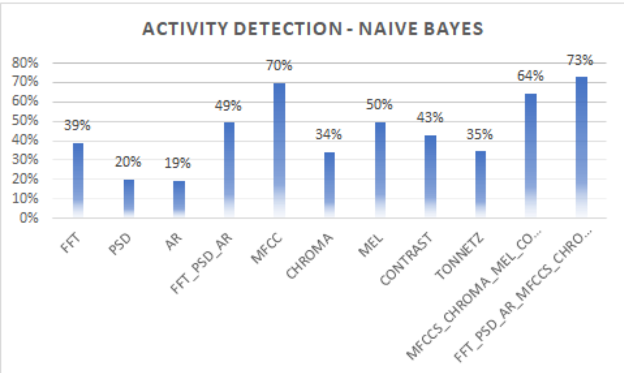

Figure 24: Activity detection from Agricultural data using Naïve Bayes

Naïve Bayes didn't perform good here. The highest accuracy was 73% with all the features.

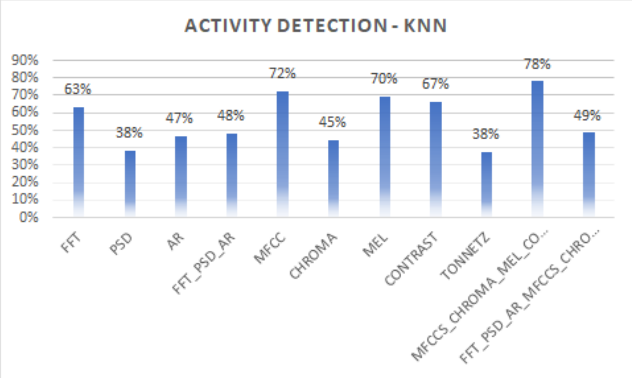

Figure 25: Activity detection from Agricultural data using K-Nearest Neighbors

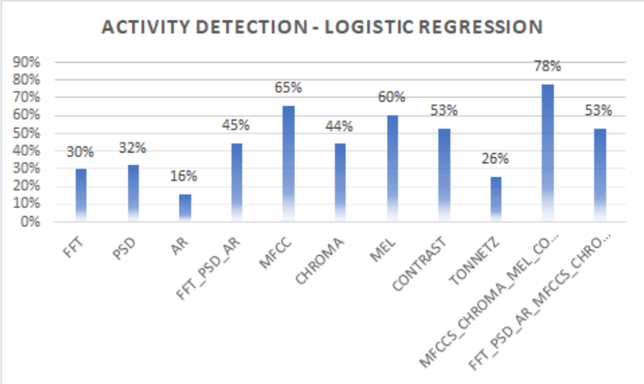

Figure 26: Activity detection from Agricultural data using Logistic Regression

Like the above, KNN and Logistic regression didn't perform well. Both achieved the highest accuracy was obtained using combined Mel features of 78%.

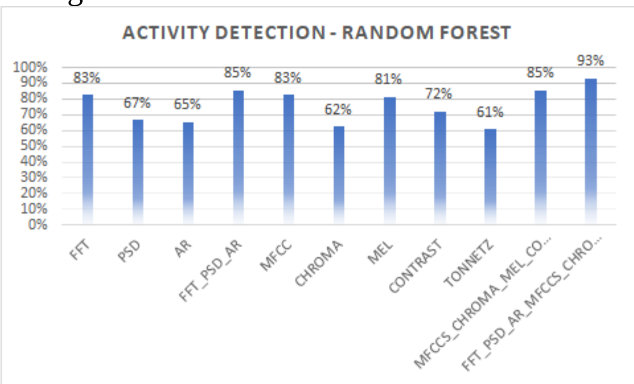

Figure 27: Activity detection from Agricultural data using Random Forest

Random Forest performed well by achieving an accuracy of 93% using all the features as input. This is lesser than Decision trees whose accuracy is found to be 95%. In the next experiment, we find that SVM didn't perform well, and the performance graph is represented below.

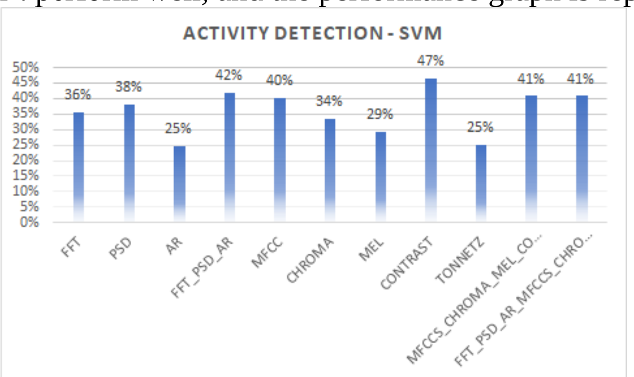

Figure 28: Activity detection from Agricultural data using SVM

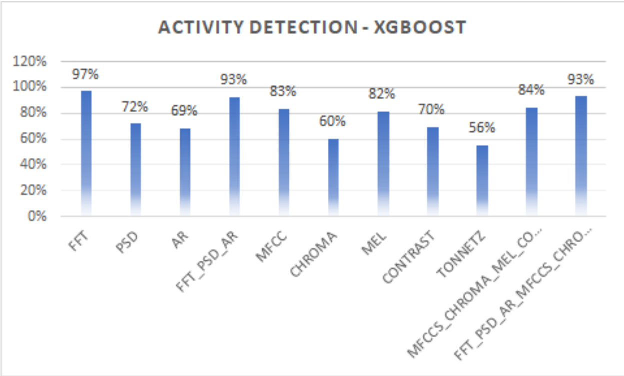

Figure 29: Activity detection from Agricultural data using XGBoost

XGBoost performed the best with an accuracy of 97% using the FFT as the feature. Now, if we compare all the information for the best model selection.

Table 2: Comparison of Decision Trees and XGBoost for the agricultural activity prediction

Model	Mean (STF_ACC)	Std_dev (STF_ACC)	VAL_ACC	Total Exec. Time (s)
DECISION TREE CLASSIFIER	0.916	**0.02546**	0.948	**0.04**
XGB CLASSIFIER	**0.928**	**0.02799**	0.969	4.66

If we check the Validation accuracy and the mean of 5-Folds accuracy, XGBoost performed the best but Decision Trees also achieved very similar results. Hence if the Execution time is taken into the picture, Decision Trees produced the result in only 0.04 seconds. Hence agricultural activities have very different sounds for each different activity and Decision Trees can be used for identifying the different activities in real-time. This can be seen above and if the sound signals are listened.

6. Conclusion and Future Work:

Seed sowing, smart watering, fertilization, and crop harvesting may all benefit from the use of AI-based solutions in the management of agricultural operations. Using just the mobile phones of farmers, we propose to develop an AI solution based on sound to track agricultural operations in real time and track the use of farmers' tractors, tillers, and other machinery components. To keep the data privacy intact, the feature's extracted data was shared with a limited number of rows to cover all the field activities along with the data ambiguity classes. Random Forest performed extremely well in identifying the validation sound samples in the Urban Sound Dataset with an accuracy of 90% at a faster computation time of 22.83 seconds in comparison to XGBoost of 326.69 seconds and with a higher stability. Stability is measured with the ranges of accuracies achieved by the model in different Folds. Random Forest had 1% standard deviation which is lower than XGBoost with 2% standard Deviation in accuracies. Hence, to conclude the analysis of Urban Sound standard dataset classification, our pipeline model performed well, and Random Forest turned out to be the selected model. In case of detecting the Human voice and garbage sound signals, XGBoost is fast with the computation time less than one second i.e., with 0.81 second in comparison to Random Forest with 1.27 seconds with 85% accuracy. Hence as a conclusion, XGBoost will be deployed for cancelling the garbage sound from the data. In, case of the Agricultural Activity of sound detection, XGBoost performed the best but Decision Trees also achieved very similar results. Hence if the Execution time is taken into the picture, Decision Trees produced the result in only 0.04 seconds. The agricultural activities having very different sounds (as can be checked with our analysis) for each different activity, Decision Trees can be used for the detection of different activities in real time.

6.1 Future Work

In future work, I will try to implement LSTM to check how the performance is coming over all the different classes. Since there are many classes in agriculture and identifying the new classes will be an important one. This can be handled when the whole system is built and in the other part, I will try using the multi-cloud approach to mitigate the optimized performance and distribute the workloads for better latency.

References

Abioye, E.A., Hensel, O., Esau, T.J., Elijah, O., Abidin, M.S.Z., Ayobami, A.S., Yerima, O. and Nasirahmadi, A., 2022. Precision Irrigation Management Using Machine Learning and Digital Farming Solutions. AgriEngineering, 4(1), pp.70-103.

Akbal, E., 2020. An automated environmental sound classification method based on statistical and textural feature. Applied Acoustics, 167, p.107413.

Anagnostis, A., Asiminari, G., Benos, L. and Bochtis, D.D., 2022. Machine Learning Technology and Its Current Implementation in Agriculture. In Information and Communication Technologies for Agriculture—Theme II: Data (pp. 41-73). Springer, Cham.

Bao, J. and Xie, Q., 2022. Artificial intelligence in animal farming: A systematic literature review. Journal of Cleaner Production, 331, p.129956.

Chen, Y., Guo, Q., Liang, X., Wang, J. and Qian, Y., 2019. Environmental sound classification with dilated convolutions. Applied Acoustics, 148, pp.123-132.

Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J. and Fan, H., 2020. Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. Neural Networks, 130, pp.22-32.

Gao, L., Campbell, H.A., Bidder, O.R. and Hunter, J., 2013. A Web-based semantic tagging and activity recognition system for species' accelerometry data. Ecological Informatics, 13, pp.47-56.

Gültekin, Ö., Cinar, E., Özkan, K. and Yazıcı, A., 2022. Multisensory Data Fusion-Based Deep Learning Approach for Fault Diagnosis of an Industrial Autonomous Transfer Vehicle. Expert Systems with Applications, p.117055.

Guzhov, A., Raue, F., Hees, J. and Dengel, A., 2021, January. Esresnet: Environmental sound classification based on visual domain models. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 4933-4940). IEEE.

Khorasani, N.E., Thomas, G., Balocco, S. and Mann, D., 2022. Agricultural Harvester Sound Classification using Convolutional Neural Networks and Spectrograms. Applied Engineering in Agriculture, 38(2), pp.455-459.

Li, F., Liu, M., Zhao, Y., Kong, L., Dong, L., Liu, X. and Hui, M., 2019. Feature extraction and classification of heart sound using 1D convolutional neural networks. EURASIP Journal on Advances

in Signal Processing, 2019(1), pp.1-11.
Meshram, V., Patil, K., Meshram, V., Hanchate, D. and Ramkteke, S.D., 2021. Machine learning in agriculture domain: A state-of-art survey. Artificial Intelligence in the Life Sciences, 1, p.100010.

Pimenov, D.Y., Bustillo, A., Wojciechowski, S., Sharma, V.S., Gupta, M.K. and Kuntoğlu, M., 2022. Artificial intelligence systems for tool condition monitoring in machining: Analysis and critical review. Journal of Intelligent Manufacturing, pp.1-43.

Sharma, G., Umapathy, K. and Krishnan, S., 2020. Trends in audio signal feature extraction methods. Applied Acoustics, 158, p.107020.

Waleed, M., Um, T.W., Kamal, T. and Usman, S.M., 2021. Classification of agriculture farm machinery using machine learning and internet of things. Symmetry, 13(3), p.403.

Zhong, X., Tian, X., Thilak, K.D. and Anbarasan, M., 2022. Machine learning-based multimedia services for business model evaluation. Computers & Electrical Engineering, 97, p.107605.