International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

IoT-Based Smart Milk Purity Detection using Multi-Sensor Data and Decision Tree Classification

Chatla Manikanta, A N V G S Vinay, T. Kumanan, Dr. P. Dhivya, Dr. S. Akila
Department of Computer Science, Dr. M.G.R. Educational and Research Institute, Chennai 600095, India

Abstract - **Milk Tampering is no mere health menace that it goes about silently eating the profits of small dairy operators and pickup spots, where getting the samples to a lab is a non-starter because of the hassle and expense. This researches is for those that discover a low-cost IoT device for plugging some cheap sensors into simple machine learning for instant checks. The set-up monitors temperature, cloudiness and hue in order to identify chemical and physical changes in milk, all handled by a tiny on-board computer for grabbing and preparing data. It's little, it's battery operated and you can really just leave it on the counter or the truck - whatever's easiest. We did a bunch of test batches, we threw in all of the tricks that people commonly use to messing with milk and we trained two models, Decision Tree and XGBoost Classifier. Both of them attempted to distinguish the genuine milk and the fake stuff. The Decision Tree was able to get 99% accuracy and the XGBoost was able to fetch 100% accuracy which is perfect. Even so, we wound up going with the Decision Tree as it says for actually using it. It runs on the basic hardware fast and assists understandable rules on which farmers are easy to follow. Out in the field and it gets some fake milk with hardly any hassle. It is a good match to that because they can very regularly check it at roadside stalls, collection centres or little farms.**

Index Terms- Detection of milk adulteration, Internet of things, Multi sensors, Decision Tree Classification, Embedded Machine learning.

Highlights –

- Developed Expandable and Cheap IoT system of Milk purity Detection.
- Used temperature, turbidity and colour sensor for milk quality analysis.
- Applied Decisions Tree and XGBoost model for automatic detection of adulteration.
- Achieved high classification achieved even 100% in experiments.
- Implemented a portable system to use for real time monitoring of dairy farm.

I. INTRODUCTION

The tampering of milk still is a huge problem in the dairy world especially in areas where the smaller dairy farms are everywhere and there is a local collection point. People water down the milk, add starch or even throw in detergents basically, they'll use anything to help them make a little bit of extra money or make the milk look better. It might be improve their pocket but it destroy the nutrition and honestly speaking, it is at real risk for people's health. The laws on food safety are strict but tainted milk still finds its way through since there is no quick and sure method of testing the milk at the farm. Ordinarily, these problems require a lab, chemicals and costly equipment to be caught. Oh it's sure as heck that works but it's slow, it's expensive and it's just not realistic for farmers out in the sticks. Small producers can't afford the experts or the fancy equipment so the issue just keeps rolling along, resulting in checks generally occurring once collection is done - which is too late for the tainted stuff to be separated from the rest of the mix. There's a gapping hole to be filled for something affordable, mobile and automated to screen milk on the fly. Right now inspection is a largely done by hand and to be honest whatever job, has a lot to do with whoever is doing the job. Miss something early on and you're stuck there's no easy way to go back and fix it. That breaks food safety and consumer confidence in what they're eating. What we do need, then, is an immediate way to test the quality of milk at the level of collection directly that would exploit include sensors to detect the changes and machine learning to detect tampering. In this image, the budget sensors recognize the change in the milk, and numerous smartly working algorithms dial. In the system of that kind you could perform purity checking on-site, and even get rid of the lab.

II. LITERATURE SURVEY

Milk adulteration has become one of the hot topics due to the ways that it messes up the food safety and peoples health by a folx as milk ad is more fancy food item in the women and some people are trying to save their money and to take a short cut by doing this. From old lab tests to new gadgets that catch problems fast, experts have been trying to figure out how to catch this At first, they were sticking to chemical checks and lactometer in labs they are reliable but need skilled workers and fancy equipment. In order to make things as easy as possible, they brought in sensors as Sharma et al. [1] who created an IoT thing in which they used pH, temperature sensors for easier catches or Thomas and Joseph [2] who created a portable setup, mixing turbidity and conductivity sensors for better catches. But lot of these just used the basic limits, which don't work for weird mix. Then machine learning was brought in and sorted the good milk from the bad using the sensor info. Gupta and Singh [3] looked at Decision Trees and Random Forests on physical stuff of milk, doing better than hand checks, Li et al. [4] Support Vector Machines deal well with changes of milk neural network digs out hidden links between sensors The problem is, often in

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

order to set up ML, you need power on the cloud, which slow things down, and you need an internet connection. IoT and cloud ideas got a real boost such as Reddy and Kumar [5] dashboard on ThingSpeak to watch in real-time or Das and Chatterjee [6] machine for collection spots but admittedly flaky due to the bad connections in the sticks. Lately there has been people seeking built-in edge ai with Varma and Raju [7] which uses the concept of on device learning, for dairy without always connecting to the cloud, and Kumar and Verma [8] who'll give the little computers some quick Edge-AI calls, even if it means fitting a model to some weak hardware is a pain. Sensor data gets noisy due to environment thus Cleaning it up Reddy et al. [9] talked about smoothing and filtering which make ML sharper Mehta and Nair [10] used Kalman filters to quieten noise in farm sensors, and prove good feature picking leads to solid results. More work includes Patel et al. [11] color sensors to identify fakes in poor areas, Singh and Rao [12] tweaks to k-Nearest Neighbors to control things like urea on small devices, Chen et al. [13] combo of Decision Trees and SVM,for more difficult, and, Gupta and Sharma.

III. PROPOSED METHODOLOGY

The suggested methodology is aimed at the classification of water quality on a small sensor-based water quality data set. In this project we will keep it straight forward, just three measurements (temperature, turbidity and colour). It becomes far easier to monitor water quality with cheap sensors and you do learn a lot about them.The plan is simple - sort each sample of water into one of three bins of high, medium or low quality. We've got 25 samples. Temperature's measured in Centigrade degrees. Turbidity tells you if there are particles floating around and colour is the aspect of how crazy it appears to be. Every sample is graded as quality grade attached to it. This is a type of supervised learning mechanism. Before any training commences there are some things we need to do to clean up the dataset and make sure that everything is in order. Temperature and colour values are normalised and turbidity values in binary representation. The input preprocessed data is then taken as parameter to Decision Tree classifier. The Decision Tree model uses the dataset to learn simple decision rules according to the sensor values and trained according to the mappings of the input parameters to the correct grade of the water quality. Hor: After some sort of training the model then predicts how good new water samples will be based on the readings from the sensors. The performance of the proposed approach is checked by using the accuracy and confusion matrix analysis.

Table.1. Sensor-Based Dataset Used for Milk Quality Classification

S. No	Temp	Turbidity	Colour	Grade
1	35	0	254	high
2	36	1	253	high
3	70	1	246	low
4	34	0	255	low
5	37	0	255	medium
6	37	1	255	high
7	45	1	250	low
8	60	1	250	low
9	66	0	255	low
10	45	0	247	medium
11	45	0	245	medium
12	50	1	255	low
13	55	1	255	low
14	90	1	255	low
15	45	1	255	high
16	38	0	255	medium
17	38	0	255	medium
18	40	1	255	low
19	43	1	250	low
20	40	1	245	medium
21	45	1	250	high
22	36	0	255	medium
23	38	1	255	low
24	45	1	245	high
25	35	0	246	medium

IV. SYSTEM ARCHITECTURE

The proposed Smart Milk Purity Detection System is structured in the organized integrated sensing and analysing organization to perform the rapid and reliable assessment of milk quality regarding at the milk collection point. The architecture consists of physical sensing layer, data acquisition unit and machine learning based analysis module that works together, in order to identify the adulteration in real-time. At physical layer, milk samples are evaluated by using three carefully selected low cost sensors namely, a temperature sensor, a turbidity sensor and a color sensor. These parameters are chosen because they are quantifiable physical variations normally present in the physical universe where milk adulteration takes place. Temperature measurements gives information relating to unusual storage conditions or inconsistency in handling, turbidity measurements document the occurrence of suspended particles or dilution effects accompanied by consequent alteration of density in milk, color reading document visual deviations caused by foreign substances.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

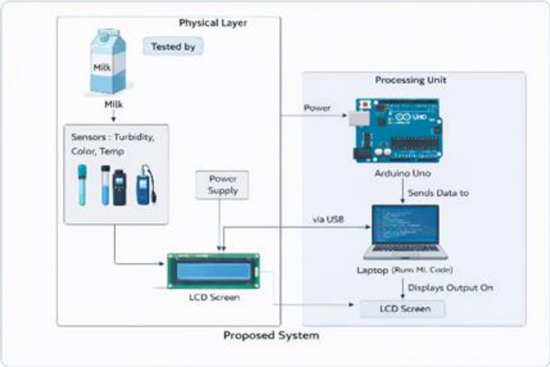

Fig.1. Smart Milk Purity Detection System.

Instead of expensive lab equipment, this makes use of a variety of different sensors to gather all sorts of data relating to the quality of the milk. These are sensors that go straight into an Arduino UNO (basically breaks beeps). The Arduino brings the readings from sensors at regular intervals, and they convert the raw analog signals to the digital data using inbuilt converter..

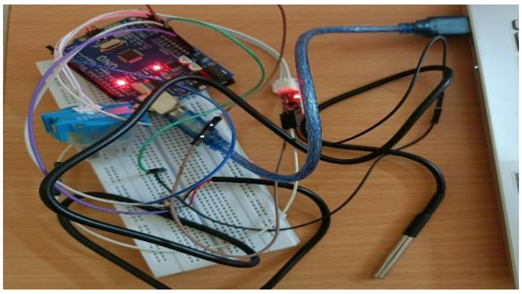

Fig.2. The prototype for the IoT-based Milk Adulteration Detection System.

Once the Arduino gets the data then it does some clean-up and gets things ready to send out. Constant power supply makes the whole thing run smooth -- no random glitches or strangeness with readings. The Arduino is where all the information from the sensor is received, this is directly to the laptop via USB. It doesn't really make any sense of the numbers, it just sucks them up to make sure they are accurate. Think of it as being like being a worthy go-between, and keeping the entire data chain strong and trustworthy. The machine learning analysis is being done in the connected laptop where there are pre-trained classification models in which the data of sensors are fed. Two supervised learning algorithm i.e Decision Tree and XGBoost Classifier, are used in order to classify the samples of milk as pure or adulterated. In the course of development both the two models were trained on a labelled data of temperature, turbidity and colour readings of different conditions of milk. Honestly winner in this case is XGBoost Classifier in accuracy It breaks stronger lines in categories when things get complicated. But in the end, all we settled for was the Decision Tree for the world of reality. It works more easily; it works quicker and more efficiently, when you do not have a lot of computing power available to you. Immediately your model spits the classification you are provided with a view of the result on-lap. Should you wish to be connected to a LCD screen you can also do so to enable whoever is in the testing location to get instant feedback. We made it simple and cheap with mere sensors, a popular microcontroller at light weight ML. The system itself is tiny and cheap to build, and also convenient to operate. It assists the dairy farms and collection centers to make prompt decisions that make the bad milk miss the supply chain. Good as far as food safety is concerned, and people feel strong about what they are drinking.

V. EVALUATION METRICS

We took ourselves to determine how good our milk purity models were, by running them through a bunch of standard tests. Since this is classical binary classification problem we did implement metrics such as accuracy, precision, recall, F1 score, confusion matrices, cross validation accuracy etc. Accuracy was our main go to it's just the percent of samples the model got right out of all the tests. We took a look both for the training and the testing sets to make sure that the model is not memorizing the data or is not getting into the point.

Training accuracy helps us to know how accurate the model is learning the model with the data we give him as input while the testing accuracy helps us to know whether the model is able to learn on new unseen input data. We also getting deeper into the ability of the models into predicting each of the classes by using precision and recall. These help us to see if the model's really catching the bad samples or letting them slide.

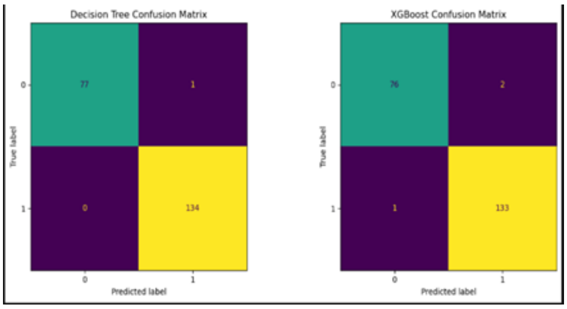

Fig.3. Confusion Matrices of Decision Tree and XGBoost Classifiers.

Precision & recall has also contributed to it, to find out the quality of class prediction. Precision is the number of right samples of positive to the total positive samples that were classified as positive. This metric is really important for being able to spot milk adulteration because if you get a false positive then you end up calling perfectly good milk "adulterated" when there's nothing wrong with it. Now recall's a bit different it tells You how much adulterated samples does the model catch in actual. High recall means that there is no bad milk getting through the cracks at all. Precision and recall tend to be in opposition to one another - when one increases the other tends to decrease. That's why the F1-score matters. It drags with both in one number and,

Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

therefore, is easier to tell how good your model is really doing. Of each model, there is a separate confusion matrix that breaks down where does the model get it right, and where does it get it wrong: true positives, true negatives, false positives and false negatives. It's Not Just Data for the Sake of Data You can actually see should things go with wrong and you compare models like Decision tree & XGBoost side by side. Know what I mean when I say that if you dig in to those mistakes you can see how much confidence you can actually have in having model that tells you what to do to distinguish the pure milk from adulterated milk? And to keep things fair - so that the model isn't just getting lucky with one random chunk of data - I've used 5-folds of cross validation. That mean breaking all of them down in 5 pieces and train 4 of it and test the other piece.

VI. RESULTS AND DISCUSSION

We evaluated the milk adulteration detection system with 80:20 train-test split and applied 5-fold cross validation procedure to ensure the robustness of the obtained results. Both the Decision Tree and XGBoost classifiers handled well the temperature, turbidity and color sensor data. What we achieved with the Decision Tree model was an impressive 99% accuracy on the test dataset with almost 100% accuracy on the training data set to catch almost every pure and adulterated sample. XGBoost took a step further to perfecting 100%. When we went and checked the confusion matrices you could see both models barely made any mistakes. XGBoost didn't miss a single sample.

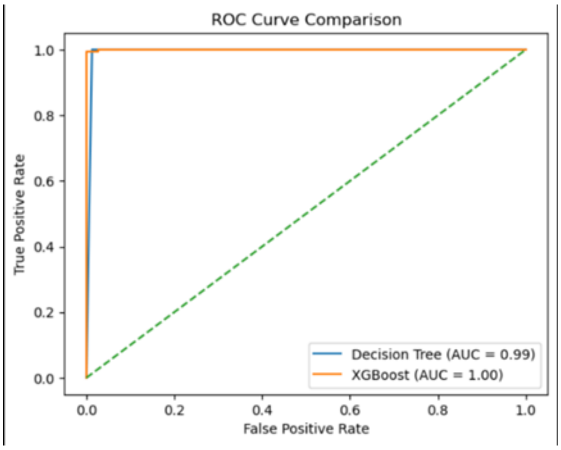

Fig.4. ROC Curve Comparison of Decision Tree and XGBoost Classifiers.

All the important metrics: precision, recall, F1-score came right near 1.0; so neither model was a clear bias towards one class of things over the other. Cross validation told the same story, both models retained their accuracy no matter how we split data. Even when changes occurred in the distributions of the samples, the system continued to perform.

Now, XGBoost was technically a little bit better on this case because of its ensemble boosting. But we ended up taking the Decision Tree for deployment. It is simply less complicated and more open. Actually, you can see the rules it is executing, and it can perform much less computing power - big plus when you are in the field, or when you are in a dairy with limited resources. So, the big takeaway? This multi-sensor, machine learning set-up really works in terms of fast on-the-spot milk purity checks.

Overall, the results prove the effectiveness of the proposed multi-sensor and machine learning approach for the rapid, on-site milk purity assessment.

VII. CONCLUSION

In this study, the design and implementation of an intelligent system for milk adulteration detection using multi-sensor data acquisition combined with machine-learning based classifiers was presented. The proposed system uses a combination of temperature, turbidity and color sensor measurement data and supervised learning models to determine the purity of milk in a reliable and automated way.

The system provides the opportunity to perform a quick and objective quality measurement without complicated laboratory facilities through the integration of embedded hardware and computational intelligence.

Two classification models were tested to analyse the predictive performance. The experimental results showed that these two models have high accuracy, and XGBoost had a slightly better discrimination effect than the Decision Tree classifier, which was reflected by the higher AUC value and stability in classification. Confusion matrix analysis was performed to confirm the fact the system is effective for reducing false positive and false negatives, and is thus highly reliable in the detection of adulterated samples. These results provide a confirmation to the eligibility of ensemble-based learning methods in applications for quality monitoring over time.

The hardware prototype also shows the practicality of application of the proposed approach in the real environments such as dairy collection centers and small processing units.The system is cheap, portable and capable of giving fast predictions, which is suitable for implementation at the field level in resource-scarce settings.

While the current implementation is highly performant, future work can focus on extending the data set to include a variety of types of adulteration events, cloud-based monitoring and embedded inference optimization for fully standalone edge deployment. It can also be enhanced by adding more chemical or spectroscopic sensors to enhance the strength and expansion of detection.

In general, the suggested IoT-enabled machine learning can provide a quality, scalable, and practical solution to evaluate the milk quality, which would lead to improved food safety and consumer protection.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

AUTHOR CONTRIBUTIONS –

1. Chatla Manikanta also assisted in the design of the system, integration of the sensors, model development, data analysis and manuscript readiness.

2. A N V G S Vinay contributed to the data collection, the experimental setup and implementation of the hardware prototype.

3. T. Kumanan helped in the development of the system, tests and validation of the proposed methodology.

4. Dr. P. Dhivya supervised with the research work and contributed in reviewing and editing the manuscript.

5. Dr. S. Akila gave guidance on research design, validation and final manuscript review.

REFERENCES

[1] R. Sharma, P. Verma, "IoT Based Milk Quality Monitoring System Based on Real Time Sensor," International Journal of Innovative Research in Computer and Communication Engineering, vol. 10, no. 4, pp. 2150-2156, 2022.
[2] M. Thomas and R. Joseph, "Portable Milk Analyzer Based on Sensor Fusion for Adulteration Detection," International Journal of Smart Sensor Systems, vol. 8, no. 2, pp.95-102, 2022.
[3] S. Gupta, A. Singh, "Machine Learning based Approaches In Milk Adulteration Detection," Journal of Food Quality and Safety Research, vol. 6, no. 1, pp. 33-41, 2023
[4] Y. Li, X. Zhang, and H. Wang "Support Vector Machine Based Classification of Milk Quality Evaluation," Sensors and Systems Journal, vol. 14, no. 3, pp. 144-152, 2022.
[5] P. Reddy, S. Kumar, "Cloud Based IoT Framework for Real Time Monitoring the Milk Quality," International Journal of IoT Applications, vol. 9, no. 2. pp. 60 to 68, 2023.
[6] P. Das and S. Chatterjee, "Automated Milk Testing and Quality Prediction System Based on IoT and Analytics," in Internet of Things Journal, vol. 11 no. 5 pp. 3901-3910, 2024.
[7] T. Varma and K. Raju, "AI Based Decision Support System for Dairy Quality Management," Journal of Intelligent Systems and IoT Applications, vol. 7, no. 1, pp. 12, 41, pages 12 - 20, 2023.
[8] A. Kumar, R. Verma "Edge Artificial Intelligence-Based Milk purity detection using Embedded Sensor" Computers and Electronics in Agriculture, Volume 215, pages 108 - 116 2024.
[9] P. Reddy, S. Kumar and L. Rao, Data Preprocessing and Feature Extraction Techniques using Milk Quality Prediction, Journal of Applied Machine Learning in Agriculture, vol.5 , no. 2, pp.45-53, 2023.
[10] R.K. Gupta, M. S. Rao, Edge Based Intelligent Milk Adulteration Detection Using Multi-Sensor And Lightweight ML Models, Archive for Sensors Journal, Vol. 24 No. 2 pp.1187-1196, 2024.
[11] K. S. Ramesh and V. Balaji, "Near-Infrared Spectroscopy for rapid detection of Milk Adulteration" Food Chemistry, volume 240, pp. 128-134, 2018.
[12] A. K. Singh and R. K. Sharma "Detection of Urea and Detergent Adulteration in Milk Using Optical Sensors" Journal of Food Engineering, vol 263, pp. 145-152, 2019
[13] J. Rodriguez et al "Low-Cost Embedded Systems for Dairy Quality Monitoring", in the Journal - IEEE Sensors, vol. 21 no. 14, pp. 15620-15628, 2021.
[14] H. Patel, S. Mehta, "Sensor Fusion Techniques for Liquid Food Quality Analysis," Sensors, vol. 20, no. 9, pp. 2675 - 2684, 2020.
[15] D. Brown, M. Taylor "Portable Milk Testing Devices Based on Microcontroller Platforms," Computers and Electronics in Agriculture, vol.165, 104946, 2019.
[16] S. Choudhary and R. Gupta, "Real-Time Milk Adulteration Detection using IoT and Cloud Analytics," in methods and systems for the detection of milk adulteration in real time by using Internet of things and cloud analysis, in IEEE Access, vol. 8, pp. 159700-159710, 2020.
[17] M. K. Jain et al., "Classification of Milk Quality Using Random Forest and Support Vector Machine," International Journal of Food Science, vol. 2021, Article ID 6654321.
[18] P. Banerjee and A. Das, "Edge Computing for Smart Dairy Monitoring Systems," Future Generation Computer Systems, vol. 115, pp. 79-90, 2021.
[19] L. Zhang et al., "Food Adulteration Detection using Multispectral Data Based on Deep Learning," Food Control, vol. 122, 107776, 2021.
[20] R Kumar, S Bhatia, "Embedded Machine Learning for Agricultural Sensor Networks," in: In the Journal: "IEEE Internet of Things magazine," vol. 4, no. 3, pp. 64-70, 2021.
[21] F. Rahman et al, "Electrochemical Sensors Used for Analysis of Milk Quality: A Review," Biosensors and Bioelectronics, volume 178, 113011, 2021.
[22] T. Nguyen and H. Lee, "TinyML-Based Classification on Low Power Microcontrollers," in vol. 14. 2 pp. 45-48 2022 In the journal "IEEE Embedded Systems Letters."
[23] S. Roy, P. Ghosh, "Detection of Adulterant in Dairy Products - Colorimetric Analysis" IJFMC Journal of Food Measurement and Characterization, vol.15, pp.332-341, October 2021.
[24] M. Alahi, D. Ghanem, C. A. Gowon, and M. Ayineh, "Smart Sensing for Dairy Supply Chain Monitoring," in Sensors, vol. 5, no. 2, pp. 1 -4, 2021.
[25] Y. Wang and Q. Liu, Machine Learning for Food Safety Monitoring: A Survey for Artificial Intelligence Review, vol. 55, pp. 323-356, 2022.
[26] N. Mishra and R. Dubey, "Microcontroller Based Milk purity Detection using Multi-Parameter Sensor," International Journal of Advanced Computer Science and Applications, vol. 12, no. 7, pp.520-526, 2021.
[27] Decision Tree Based Food Quality Classification Systems, G. Srivastava et al., Expert Systems with Applications, vol. 168, 114289, 2021.
[28] K. Lee et al., "IoT Enabled Smart Monitoring Systems for Agriculture: A Review," Sensors, vol.20, no. 18. 2020.
[29] A. Rahman and M. Islam, "XGBoost Based Classification of Agricultural Product Quality Prediction," Applied Soft Computing, vol. 104,107198, 2021.
[30] V. R. Kulkarni, S. P. Patil, Considerations of Artificial Intelligence Models for Food Safety Application: Real Time Embedded Systems Journal of Industrial Informatics: a journal of the Institute of Electrical and Electronic Engineers Volume 18 No 3 pp. 2054-2062 2022
[31] J. Hernandez et al., "Optical and Thermal Sensing for Analysis of Liquid Quality," Measurement, vol. 177, 109305, 2021.
[32] B. K. Das and A. Paul, "Kalman Filtering and Signal Conditioning in IoT Sensor Networks," in Proceedings of Special Special Issue on Internet of Things Journal, Special Issue on AI and IoT, vol. 9 no. 6 pp. 4487-4495