MEKATRONIKA
e-ISSN: 2637-0883
VOL. 2, ISSUE 2, 74 – 80
DOI: https://doi.org/10.15282/mekatronika.v2i2.7022

ORIGINAL ARTICLE

Rain Classification for Autonomous Vehicle Navigation : A Support Vector Machine Approach

Abdul Haleem Habeeb Mohamed[1,2], Muhammad Aizzat Zakaria[1,2*], Mohd Azraai Mohd Razman[1], Anwar P. P. Abdul Majeed[1] and Mohamed Heerwan Peeie[2]

[1] Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, Malaysia.
[2] Autonomous Vehicle Laboratory, Centre for Automotive Engineering, Universiti Malaysia Pahang, Malaysia.

ABSTRACT – The advancement of LIDAR technology used in the autonomous vehicle (AV) system has made it increasingly popular. Despite that, the ability of the sensor to adjust to human behaviour in sensing and perceiving different environments is still unsolved as it significantly impacting the performance of LIDAR, causing the effect of missing points and false positives detection. The immerging of machine learning algorithms that have greatly impacted solving uncertainties and LIDAR's reliability in making judgments has proven a great success. This paper aims to classify different rain rates conditions in a controlled environment with real rain using a LIDAR. Then, the feature extraction using the time-domain method was employed to generate more features with a variation of SVM models in developing classification models. The preliminary observation shows that the Poly-SVM model can achieve a test classification accuracy of 97%. Noting that, the proposed method has the potential to evaluate weather classification.

ARTICLE HISTORY
Received: 10th Nov 2020
Revised: 3rd Dec 2020
Accepted: 20th Dec 2020

KEYWORDS
LIDAR
Autonomous Vehicle
Support Vector Machine

INTRODUCTION

The autonomous vehicle (AV) field is increasingly gaining interest today due to drastic sensor technology changes. Nowadays, car technology includes various driver assistance systems, communications and electronics, automation requirements create a whole new demand [1]. There have been some accomplishments in the AV area of research due to its popularity in recent years showed by research worldwide. The main objective of AV is to reduce traffic congestion and human error that could be fatal through the employment of machine intelligence [2], [3].

The advancement of LIDAR technology used in the autonomous vehicle (AV) system has made it increasingly popular. Despite that, the ability of the sensor to adjust to human behaviour in sensing and perceiving different environments is still unsolved as it significantly impacting the performance of LIDAR, causing the effect of missing points and false positives detection. The immerging of machine learning algorithms that have greatly impacted solving uncertainties and LIDAR's reliability in making judgments has proven a great success. This paper aims to classify different rain rates conditions in a controlled environment with real rain using a LIDAR. Then, the feature extraction using the time-domain method was employed to generate more features with a variation of Support Vector Machine (SVM) models in developing classification models. The preliminary observation shows that the Poly-SVM model can achieve a test classification accuracy (CA) of 97%. Noting that, the proposed method has the potential to evaluate weather classification.

The transition towards the utilization of AVs in Malaysia is somewhat hampered by its weather conditions, particularly its high value of rainfall precipitation (RP) [4]. It has been reported that the RP is between 84.7 and 153.9mmh^{-1} in Peninsular Malaysia and between 81.8 and 143.8mmh^{-1} in Borneo [5], [6]. This problem is not uncommon as it has been reported in the literature that rain droplets and fog are a huge obstacle to the progress of unmanned vehicle technology, especially with Light Detection and Recognition (LIDAR) due to its water droplets which causing false positive and missing detection [7], [8]. It has become a serious issue where the software created to manoeuvre the automated car requires adaptation to various environmental conditions perceived by humans [1].

LIDAR is a promising weather detection technology due to its high definition data and spatial resolution as it scans the environment, generating large-scale point clouds and 3D maps [9]–[11]. An active extrinsic sensor enables the efficient location of weather effects both day and night and over a considerable range [6], [12].

Nowadays, the current trend of large-scale data implementation is always to be used with machine learning due to its large datasets for humans to handle. A study was carried out by using LIDAR and machine learning in classifying different rain weather rates by comparing two classifiers which is *k*-Nearest Neighbors (*k*-NN) and SVM [13].

A mathematical model approach with maximum and rain rate as the parameter was used to investigate the degradation of LIDAR and incorporate the model into a physical-based simulation using Velodyne VLP-16 by simulating the

*CORRESPONDING AUTHOR | Muhammad Aizzat Zakaria | ✉ maizzat@ump.edu.my

maximum distance of detection throughout three different rain conditions, which is 0 mm/h, 9 mm/h, 17 mm/h and 45 mm/h. In results shows that, reduce the range of detection and number of points [16].

Nonetheless, it is noteworthy to mention that limited studies have been carried out to address the issue regarding the efficacy of LIDAR towards harsh weather environments, especially with regards to the Malaysian climate. This paper investigates the effectiveness of evaluating a variation of SVM models in classifying LIDAR data, particularly for real raining weather conditions with different rain rates.

RELATED WORK

In past work, there have been few studies on the influence of LIDAR sensors in harsh weather conditions. The influence of the weather was analyzed with four weather conditions: clear, fog, rain: less than 5 mm/h, and snow by setting up the LIDAR Velodyne' VLP' static outdoor. Different parameters such as standard deviation, maximum, peak to peak, mean and so forth were obtained to maximize the CA, focusing on *k*-NN F1-score higher than 80%. The parameter, which is the feature, is to be tested to investigate the significant parameters that impact the classification [13].

Furthermore, a machine learning model approach to analyze the influence of weather on the performance of LIDAR by classifying three different conditions: clear, rain and fog in a controlled environment using a Velodyne' VLP16' and Valeo' Scala'. The LIDAR output data gathered consist of cartesian (x, y, z) and spherical (r, θ, φ) coordinates, e for the echo number, I for the intensity and epw for the echo pulse width. The mean and variance of each frame were calculated to generate 16 attributes feature vectors. The weather condition tested were clear;0 mm/h, rain; 55 mm/h and fog; 20-30 m, 30-40 m, 40-50 m and 50-60 m. Two machine learning classifiers were used for the weather classification: *k*-KNN and SVM. Thus, the classification results for VLP16 using mean union over the intersection with KNN and SVM as the classifier is 96.40% and 97.14%, respectively, which is the best [14].

It is noteworthy to mention that limited studies have been carried out regarding the issue of LIDAR towards adverse weather environments. This paper aims to investigate the efficacy of evaluating SVM machine learning classifier models and its hyperparameter tune specifying different kernel variation using time-domain features as parameters in classifying LIDAR data by assessing the CA for clear and other rain rate weather conditions.

METHODOLOGY

Experiment Setup

The RS-LIDAR 16, as in Figure 2 (a), is a 3D LIDAR consisting of 16 channels of light beam with a laser wavelength of 905 nm where a single return data point of 320,000 pts/s can be emitted a maximum of up to 150 m. It is designed especially for the outdoor environment because of its IP67 environmental protection that protects the LIDAR from dust and water immersion for mobile applications such as AV that is mainly use for its capability to perceive the environment with 360 ˚horizontal angle field of view and 30˚vertical angle field of view (15˚up and 15˚down). The rotation range is 5 Hz, 10 Hz and 20 Hz, which is 300 rpm, 600 rpm and 1200 rpm, respectively. The output data contain range, rotation angle and time stamp. The flowchart for rain rate classification can be seen in Figure 1

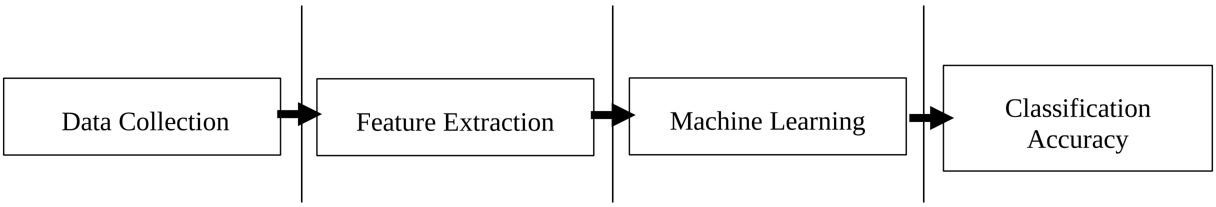
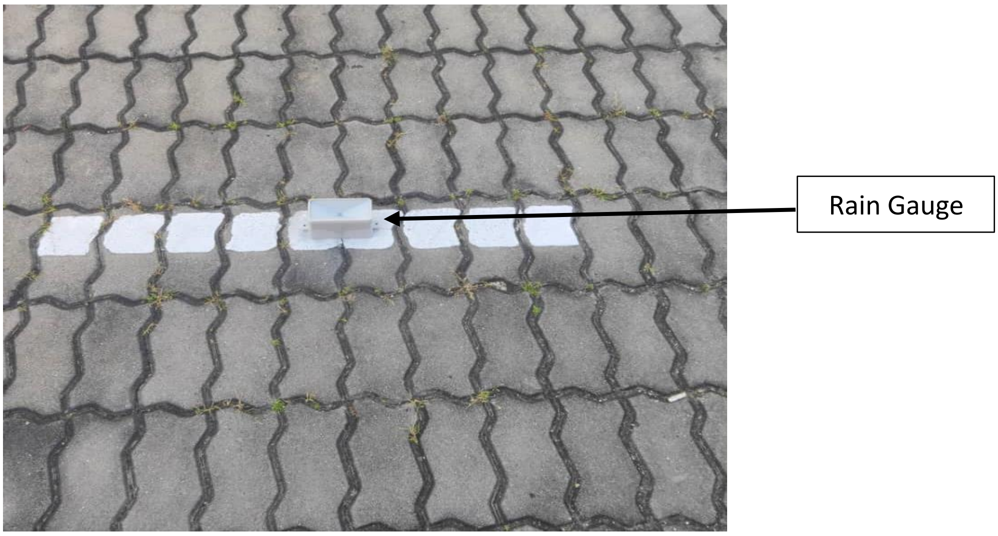


Figure 1. Flowchart of the rain rate classification

The investigation gathers the data for different rain rates, and the selected area was Universiti Malaysia Pahang in Peninsular Malaysia.

(a)



(b)

Figure 2. (a) RS-LIDAR 16 set up (b) Rain gauge position

The experiment was done for three months (September 2020 to December 2020). It was the rainy season, and the statistics obtained from the meteorological station reported that the rainfall precipitation (RP) was between 84.7 mm/h and 153 mm/h during this period. The drop size distribution of the droplets is considered constant along the area of study. As depicted in Figure 2 (a) and (b), LIDAR and the rain gauge were placed respectively in a static position throughout the experiment. The LIDAR rotation range was set to 10 Hz, 600 rpm, and the horizontal angle field of view was set to 120˚. A total of 10 frames consists of approximate 10,700 points per frame, and a total of 107,000 pts/s was collected for each clear weather and 18 mm/h, 36 mm/h, 54 mm/h, 72 mm/h, 90 mm/h and 108 mm/h for rainy weather with a combination of 70 frames. The output LIDAR data consist of cartesian (x, y, z) coordinates were gathered.

In this investigation, time-domain features (Equations 1 to 7) consist of mean, median, variance, standard deviation, skewness kurtosis, standard error mean, minimum and maximum were calculated for every frame and each cartesian coordinates in which will create a total of 27 features a frame. Features extraction is one of the analysis techniques for processing the raw LIDAR data that can be classified as a time-domain analysis method. The benefits of this method are to average the rain noise and become more sensitive to the background noise [15].

$$\text{Mean} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (1)$$

$$\text{Median} = \left(\frac{n+1}{2}\right) \quad (2)$$

$$\text{Variance} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{3}$$

$$\text{Standard deviation} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{4}$$

$$\text{Skewness} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}\right)^3} \tag{5}$$

$$\text{Kurtosis} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^2} \tag{6}$$

$$\text{Standard Mean error} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}}{\sqrt{n}} \tag{7}$$

For this investigation, 400 frame data for each clear and variation of rain rate weather, which is 18 mm/h, 36 mm/h, 54 mm/h, 72 mm/h, 90 mm/h and 108 mm/h, sum up to a total of 2800 frame was evaluated using ML classifier. The classifier used was SVM and was further tuned by hyperparameters using the linear kernel compared to the best ML classifier model. One of the purposes of hyperparameters is to prevent the overfitting of the model [16]. The ML model used was supervised learning and classifying big data such as LIDAR data [11]. The model was fit with the features extracted data using the nine-time domain features for each channel of (x, y, z) that creates 27 features, and the dataset was used as an input to train the ML model [17].

A variation of SVM models based on the different kernels are evaluated to investigate the effectiveness in classifying the weather [11], [18]. SVM machine learning model is the most popular model which is widely used in nowadays applications for classification. This model works by choosing the dataset features which depend on the least feature value recursively [19]. The different SVM models and hyperparameter tuning was evaluated in Spyder IDE, a Python programming environment. The scikit-learn library was used for the development of the SVM models. The kernels were varied, whilst other hyperparameters were set, as shown in Table 1. The kernels evaluated are linear, polynomial (second-order), radial basis function (RBF) and sigmoid, as referred from Equations 8 until 11, respectively. The evaluation for the classifiers is evaluated based on its CA, and the fivefold cross-validation technique was used to train the models as it has been reported to reduce the overfitting behaviour [20]. The purpose of using hyperparameter tuning is to find the best model and parameters algorithm compared by just using default settings.

1) Linear kernel algorithm:

$$f(x) = \sum_{i=1}^{n} y_i(x_i \cdot w + b) \tag{8}$$

Where *x* is the nearest point to the margin and *w* absolute *b* is the distance between two hyperplanes.

2) Polynomial kernel algorithm:

$$f(x_i, x_j) = (x_i \cdot x_j + 1)^e \tag{9}$$

Where *e* denoted as the degree of polynomial.

3) RBF kernel algorithm:

$$f(x_i, x_j) = exp\left(-r\|x_i - x_j\|^2\right) \tag{10}$$

Where *r* represents the radial kernel parameter.

4) Sigmoid kernel algorithm:

$$f(x_i, x_j) = tanh(rx_i \cdot x_j + c) \tag{11}$$

Where c represents a moving parameter which regulates the mapping threshold, where $r > 0$ and $c < 0$.

The default SVM model was used and compared with tuned SVM hyperparameters to see the capability of the model to adapt the nature of this data.

Table 1. SVM hyperparameters tuning

Kernals	**Parameters**		
	Gamma	**C**	**Degree**
Linear	-	0.01,0.1,1,10,100	-
Polynomial	0.01,0.1,1,10,100	-	2,3
RBF	0.01,0.1,1,10,100	0.01,0.1,1,10,100	-
Sigmoid	0.01,0.1,1,10,100	0.01,0.1,1,10,100	-

EXPERIMENTAL RESULTS

In this study, a real-world environment of the rainy weather and, in this case, was clear and different raining weather conditions were collected. As mention in the previous section, all the features generated from LIDAR data was used to train the SVM models for each different kernel with a total of 69 machine learning model. The data were separated randomly, where 70% of the dataset's data was used to train the SVM models. The remaining 30% was tested with the developed, trained SVM models to observe the accuracy further. As shown in Figure 3, it is the default SVM model and the best hyperparameter tuning for SVM. It can be seen that using the default parameter for SVM achieved a CA of 33 % for train and 32 % for the test, which shows an unreliable machine learning model. After hyperparameter tuning with the parameter set shown in Table 1, the best model achieved by using poly as the kernel was to set parameter degree to 3 and gamma 100. The CA achieved for the train is 100% and 97% for the test, and it shows that the development of optimized ML model has the credibility for different rain rate weather classification

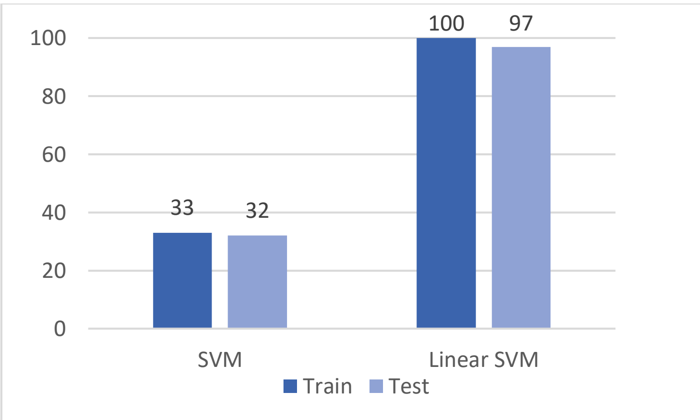


Figure 3. SVM model and hyperparameters tuning

Further observation was carried out on the confusion matrix of both default SVM and polynomial SVM models that are shown in Figure 4 and Figure 5. In Figure 4, shows the misclassification for defaults SVM model came from rain rate of 0 mm/h was misclassified as 36 mm/h, 18 mm/h were misclassified as 36 mm/h and 90 mm/h, 36 mm/h was misclassified as 90 mm/h, 54 mm/h were misclassified as 36 mm/h and 90 mm/h, 72 mm/h rain rate were misclassified as 36 mm/h and 90 mm/h, 90 mm/h was misclassified as 36 mm/h and lastly, 108 mm/h were misclassified as 36 mm/h and 90 mm/h. The misclassification of the tuned SVM model can be seen in Figure 5, where rain rates for 0 mm/h and 18 mm/h were classified correctly. However, there was a misclassification for 36 mm/h that was classified as 54 mm/h, 54 mm/h were misclassified as 36 mm/h and 72 mm/h, 72 mm/h were misclassified as 54 mm/h, and 108 mm/h, 90 mm/h was misclassified as 72 mm/h, and 108 mm/h was misclassified as 90 mm/h. The misclassification for the tuned SVM model has less misclassification compared to the default SVM model. This shows the instability of the model to classify causing a massive error and with hyperparameter tune had shown in better model predictability in classifying rain rates.

Actual Class	0	18	36	54	72	90	108
0	120	0	0	0	0	0	0
18	0	120	0	0	0	0	0
36	0	0	117	3	0	0	0
54	0	0	8	103	9	0	0
72	0	0	0	1	118	0	1
90	0	0	0	0	4	116	0
108	0	0	0	0	0	3	117

Predicted Class

Misclassification
True Classification

Figure 4. SVM default parameter Confusion Matrix.

Actual Class	0	18	36	54	72	90	108
0	75	0	45	0	0	0	0
18	0	0	100	0	0	20	0
36	0	0	108	0	0	12	0
54	0	0	93	0	0	27	0
72	0	0	35	0	0	85	0
90	0	0	36	0	0	84	0
108	0	0	73	0	0	47	0

Predicted Class

Figure 5. SVM-Poly Confusion Matrix.

CONCLUSION

This preliminary study shows that default SVM and Poly-SVM models developed using the selected features from LIDAR data can provide a reasonable CA to evaluate the weather changes. Future studies will be carried out by including more disturbance to the environment, engineering different features, and optimizing the hyperparameter on the best SVM model. The preliminary results obtained encourage the objective-based judgement when using LIDAR that measures the reliability of the sensor in the presence of different rain rates.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Higher Education for providing financial support under Fundamental Research Grant Scheme (FRGS) No. FRGS/1/2018/TK08/UMP/02/1 (University reference RDU190104) and Universiti Malaysia Pahang for laboratory facilities as well as additional financial support under Internal Research grant RDU1903139. The authors also thank the research team from Autonomous Vehicle Laboratory AEC, Innovative Manufacturing, Mechatronics and Sport Laboratory (iMAMS), who provided insight and expertise that greatly assisted the present research work.

REFERENCES

[1] M. Kutila, P. Pyykönen, W. Ritter, O. Sawade, and B. Schäufele, “Automotive LIDAR sensor development scenarios for harsh weather conditions,” *IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC*, pp. 265–270, 2016, doi: 10.1109/ITSC.2016.7795565.

[2] K. Bimbraw, "Autonomous cars: Past, present and future: A review of the developments in the last century, the present scenario and the expected future of autonomous vehicle technology," *ICINCO 2015 - 12th Int. Conf. Informatics Control. Autom. Robot. Proc.*, vol. 1, no. August, pp. 191–198, 2015, doi: 10.5220/0005540501910198.

[3] U. Z. A. Hamid *et al.*, "Autonomous emergency braking system with potential field risk assessment for frontal collision mitigation," *2017 IEEE Conf. Syst. Process Control*, no. December, pp. 71–76, 2017, doi: 10.1109/SPC.2017.8313024.

[4] M. S. Kwan, F. T. Tanggang, and L. Juneng, "Projected changes of future climate extremes in Malaysia. National Symposium on Climate Change Adaptation," vol. 42, no. 8, pp. 1051–1058, 2011.

[5] A. K. Ab Rahman, R. Abdullah, N. Balu, and F. Shariff, "The Impact of La Niña and El Niño Events on Crude Palm Oil Prices : An Econometric Analysis," *Oil Palm Ind. Econ. J.*, vol. 13, no. September, pp. 38–51, 2013.

[6] T. V Omotosho, J. S. Mandeep, M. Abdullah, and A. T. Adediji, "Distribution of one-minute rain rate in Malaysia derived from TRMM satellite data," pp. 2013–2022, 2013, doi: 10.5194/angeo-31-2013-2013.

[7] C. Goodin, D. Carruth, M. Doude, and C. Hudson, "Predicting the influence of rain on LIDAR in ADAS," *Electron.*, vol. 8, no. 1, 2019, doi: 10.3390/electronics8010089.

[8] R. Heinzler, F. Piewak, P. Schindler, and W. Stork, “CNN-Based LIDAR Point Cloud De-Noising in Adverse Weather,” vol. 5, no. 2, pp. 2514–2521, 2020.

[9] A. Filgueira, H. González-Jorge, S. Lagüela, L. Díaz-Vilariño, and P. Arias, “Quantifying the influence of rain in LIDAR performance,” *Meas. J. Int. Meas. Confed.*, vol. 95, pp. 143–148, 2017, doi: 10.1016/j.measurement.2016.10.009.

[10] C. Goodin, D. Carruth, M. Doude, and C. Hudson, "Predicting the Influence of Rain on LIDAR in ADAS," *Electronics*, vol. 8, no. 1, p. 89, 2019, doi: 10.3390/electronics8010089.

[11] S. S. Teri and I. A. Musliman, "MACHINE LEARNING in BIG LIDAR DATA: A REVIEW," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 4/W16, pp. 641–644, 2019, doi: 10.5194/isprs-archives-XLII-4-W16-641-2019.

[12] G. Péter, B. Kiss, and V. Tihanyi, "Vision and odometry based autonomous vehicle lane changing," *ICT Express*, vol. 5, no. 4, pp. 219–226, 2019, doi: 10.1016/j.icte.2019.09.005.

[13] J. Chen, "Weather Classification Using an Automotive LIDAR Sensor Based on Detections on Asphalt and Atmosphere."

[14] R. Heinzler, P. Schindler, J. Seekircher, W. Ritter, and W. Stork, “Weather influence and classification with automotive LIDAR sensors,” *IEEE Intell. Veh. Symp. Proc.*, vol. 2019-June, pp. 1527–1534, 2019, doi: 10.1109/IVS.2019.8814205.

[15] B. Sreejith, A. K. Verma, and A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks," no. 1, pp. 1–6, 2008.

[16] N. Rtayli and N. Enneya, "Journal of Information Security and Applications Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization," *J. Inf. Secur. Appl.*, vol. 55, no. September, p. 102596, 2020, doi: 10.1016/j.jisa.2020.102596.

[17] J. R. Vergara and P. A. Este, "A review of feature selection methods based on mutual information," pp. 175–186, 2014, doi: 10.1007/s00521-013-1368-0.

[18] G. B. Sai Tarun, J. V. Sriram, K. Sairam, K. T. Sreenivas, and M. V. B. T. Santhi, "Rainfall prediction using machine learning techniques," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 7, pp. 957–963, 2019, doi: 10.13140/RG.2.2.26691.04648.

[19] A. R. Vaka, B. Soni, and S. R. K., "Breast cancer detection by leveraging Machine Learning," *ICT Express*, no. xxxx, pp. 0–4, 2020, doi: 10.1016/j.icte.2020.04.009.

[20] C. C. Sern, A. F. A. Nasir, A. P. P. Abdul Majeed, M. A. Zakaria, M. A. Mohd Razman, and A. Azmi, "Comparison of Support Vector Machine and Friis Equation for Identification of Pallet-Level Tagging Using RFID Signal," *ISCAIE 2020 - IEEE 10th Symp. Comput. Appl. Ind. Electron.*, pp. 215–219, 2020, doi: 10.1109/ISCAIE47305.2020.9108834.