# A Cutting-Edge Hybrid Deep Learning Technique with Low Rank Approximation for Attacks Classification on IoT Traffic Data

Ankita Sharma[1] and Shalli Rani[1]

[1]Chitkara University

August 25, 2024

**Abstract**

Network security is experiencing huge challenges as network attacks on traffic data become more frequent and sophisticated. In this paper, we employ hybrid deep learning models and low-rank approximation to present a novel method for multi-label categorization of network assaults on traffic data. Our suggested solution, LR-CNN-MLP, consists of three models While the CNN and MLP models extract features and categorise data, respectively, the low-rank approximation model reduces the input's dimensionality. Overall, by combining hybrid models and low-rank approximation, our proposed LR-CNN-MLP approach provides a promising solution for multi-label categorization of network attacks on traffic data.

Received <day> <Month>, <year>; | Revised <day> <Month>, <year>; | Accepted <day> <Month>, <year>

DOI: xxx/xxxx

**ARTICLE TYPE**

# A Cutting-Edge Hybrid Deep Learning Technique with Low Rank Approximation for Attacks Classification on IoT Traffic Data†

Ankita Sharma | Shalli Rani*

[1]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

**Correspondence**
*Shalli Rani, Chitkara University,Punjab,India. Email: shallir79@gmail.com

**Abstract**

Network security is experiencing huge challenges as network attacks on traffic data become more frequent and sophisticated. In this paper, we employ hybrid deep learning models and low-rank approximation to present a novel method for multi-label categorization of network assaults on traffic data. Our suggested solution, LR-CNN-MLP, consists of three models While the CNN and MLP models extract features and categorise data, respectively, the low-rank approximation model reduces the input's dimensionality. Overall, by combining hybrid models and low-rank approximation, our proposed LR-CNN-MLP approach provides a promising solution for multi-label categorization of network attacks on traffic data.

**KEYWORDS:**
IoT, Machine Learning, Deep Learning, Security

## 1 | INTRODUCTION

Because of modernity's rising reliance on digital technologies such as mobile phones, Windows PCs, and possibly other linked equipment that make up the Internet of Things, society is more vulnerable to a variety of hostile, targeted attacks [1]. Multi-task learning, which tries to manage multiple distinct computing tasks simultaneously, is a growing subject of study in machine learning applications. This is because various circumstances frequently require the execution of many tasks [2]. To classify something, one must draw inferences about the label or labels that should be ascribed to future events based on previous experience and knowledge. An successful learning approach involves input and output properties. A label, a class, or a collection of labels that are associated with each instance of the dataset are the output features of a classification process in which the input attributes serve as discriminating predictors. Every dataset in the single-label classification method will have the same output attribute [3]. Binary classification also assigns each occurrence to one of two possible categories: if it fits in our category, we code it as "one," and if not, we code it as "zero." Furthermore, multi-class classification produces a s Similar to binary classification, multi-label classification uses an array of outputs to represent each occurrence. Depending on the situation, each output can have the value 0 or 1. Every dataset instance uses an array with the same length, but each one has a different set of active label combinations [4]. Single output attribute per dataset, similar to binary classification. Every instance in a multi-class classification can be assigned any value from a limited set of labels, not simply one or zero. Binary classifiers, on the other hand, limit values to either one or zero. In contrast to binary and multiclass classifications, multi-label classification allows for several output attributes per instance [5]. Similar to binary classification, multi-label classification uses an array of outputs to represent each occurrence. Depending on the situation, each output can have the value 0 or 1. Every dataset instance uses an array with the same length, but each one has a different set of active label combinations [6].Although shallow machine learning approaches

†This is an example for title footnote.

[0]**Abbreviations:** ANA, anti-nuclear antibodies; APC, antigen-presenting cells; IRF, interferon regulatory factor

were applied, the intended results were not obtained. Support vector machines, one of the several classical classifiers in use, performed the best in multi-label classification. The problem with support vector machines is their inability to discriminate between distinct class types that share the same label [7]. This challenge is solved or handled using deep learning classifiers. Deep learning has recently emerged as the most effective technique for representing data [8]. Deep learning methods train a neural network with numerous layers to extract hierarchical patterns from unstructured data and give high-level, abstractive features for learning tasks [9]. In this study, a low-rank technique for multilabel classification on hybrid datasets is presented, taking into account the parameters required to achieve the desired results while using specific pre-processing approaches [10].

## 1.1 | Our Contributions

1. Using the created hybrid dataset as the basis for a new security hybrid deep learning model for IoT networks.

2. To achieve adaptive multilabel classification, test weight optimisation strategies with deep learning classifiers that use low rank matrix factorization.

3. The findings were generated using the multi-label classification method.

## 2 | RELATED WORK

The authors suggested a low-rank representation method for multi-label classification based on the Inexact Augmented Lagrange Multiplier. When used in conjunction with low-rank representation, the following strategy produced better results. Preprocessing procedures are not used. The approach used was shown to be inadequate for single label occupations and simple data [5]. In a similar vein, the challenge of filling a single label is addressed by filling multiple labels, which also removes noise and outliers without increasing computing costs. the application of the same methodology and data cleansing technique, except in the case of the missing coefficient matrix [6]. Although the parameters lamba 1 and lambda 2 are not systematically established, the coefficient matrix is created by the normalisation pre-processing of the linearized alternating direction method (LADM) [7]. The parameter lambda in [8] is produced by fixing the value using the alternate direction technique. For noise-free data with independent labels, data cleaning and normalisation procedures are used. The matrix for data representation is still required. Despite its low performance, the Wilcoxon signed rank approach in [9] is used to generate the data representation matrix without the need for any pre-processing methods. Nonetheless, the low rank of the acquired samples in low-dimensional space is used to generate a useful graph that may be used to capture local data without requiring pre-processing methods such as SAW, TOPSIS, and MCDM procedures. Global information collecting remains vital [10]. The low rank representation methodology is used without any pre-processing method to capture both local and global data samples; nevertheless, dictionary building is still necessary [11]. The right dictionary is created by using normalisation to combine low rank with sparse matrix decomposition. A single distribution does not completely eliminate noise, outliers, or anomalies [12]. The Manhattan Distance Mixture of the Guassian distribution replaces the single distribution, but with fixed parameters such as lambda and beta equal to 0.1 [13]. In contrast, no pre-processing procedures are utilised while creating a decision matrix with a low rank using the locally enhanced low-rank prior method. It is assumed that the cardinality c and rank value r are unique [14]. A weakly supervised low-rank representation technique is used in conjunction with a normalisation preprocessing technique to discover abnormalities and noise in the data; however, a weakly supervised approach without low-rank representation is inapplicable [15].

## 3 | PROPOSED METHODOLOGY

The following part describes the dataset, preprocessing procedures, low-rank matrix factorization, and multilabel classification. The provided methodology clarifies how to merge two datasets. The findings are obtained utilising the baseline model and a low-rank technique on the combined dataset, which employs hybrid deep learning models such as CNN-MLP. UNSW NB 15 [4] and BoT-IoT [3] security datasets were chosen for their capacity to handle multi-label problems. The first stage of data preprocessing includes merging the two incursion datasets using the column "label." A hybrid dataset, which includes all of the observations

from the combined data sets, contains 182000 observations. When it comes to missing and redundant data, cleaning it up is the first step. to use discretization, the second preparation procedure, on the pooled dataset. Each feature's value range is consistently and linearly mapped within the [0,1] intervals using a normalised processing method, which facilitates arithmetic processing and dimension reduction.

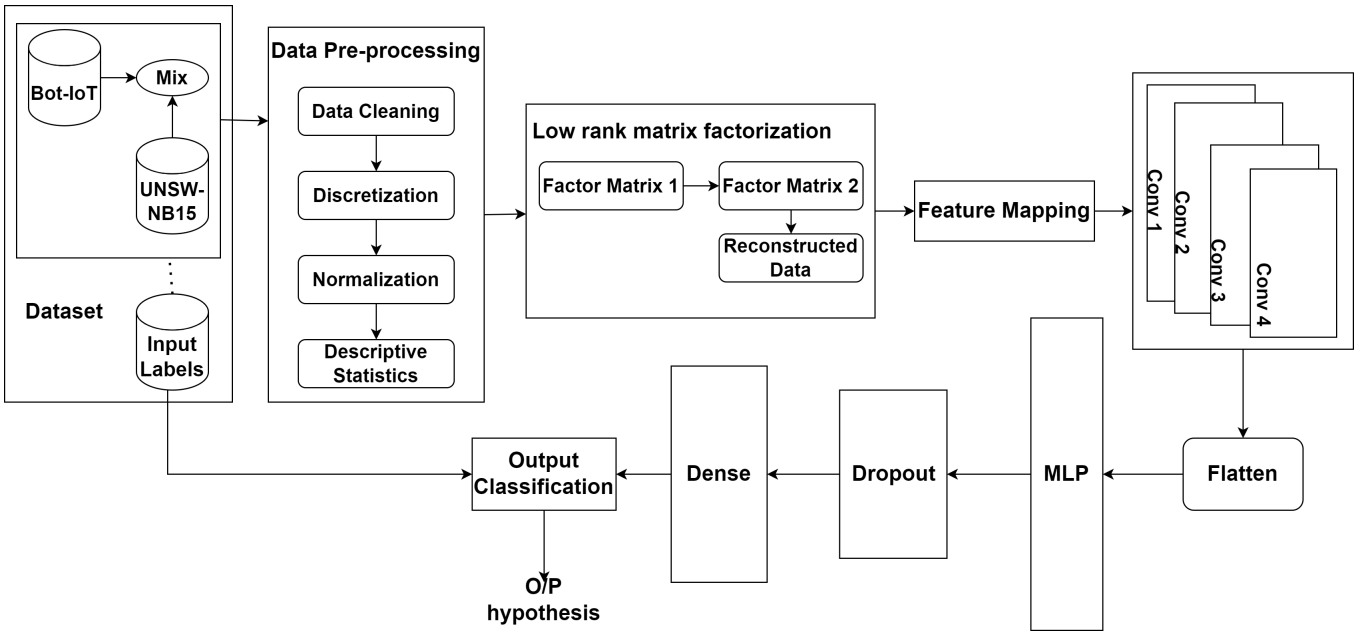


**FIGURE 1** Proposed Methodology LR-CNN-MLP

After creating the hybrid dataset and using the three pre-processing procedures, the features were computed using covolutional layers. The flatten layer transforms two-dimensional feature maps into linear forms. The information was transmitted using backpropagation and a multilayer perceptron. The dropout layer will choose which information to leave behind and which to keep. Based on the dropout layer's output, the dense layer is used to classify the number of classes. As a result, the same label is output and assigned to a variety of classes, providing the desired results.

The TCP and OSI models are combined to form the four layers of the Internet of Things. Because the network layer is vulnerable to several types of attacks, such as DoS, DDoS, MITM, and others, it is where information is transmitted. The hybrid dataset, which includes the BoT-IoT and UNSW NB 15 datasets, is used to identify the classes that originate the shared label in the multi-label classification issue, in which numerous classes share a common label. Pre-processing methods are classified into three categories: normalisation, discretization, and data cleaning. The hybrid dataset, formed by integrating the two datasets, contains redundant data, inconsistent values, and missing values. Missing values are filled in utilising data cleaning procedures, either manually or by using column mean values. The dataset contains nominal and category values. A discretization pre-processing technique employs discrete values, allowing maximum boundaries to be induced on a sorted list. The normalisation procedure is used to eliminate outliers and repetitive data. The next step is to use a correlation function to select attributes that are highly reliable. The dataset is now divided into 80 for training and 20 for testing, following the 80/20 rule. To construct a low rank matrix, 80 training data of strongly correlated characteristics are factorised using k cross fold stratified validation and factor matrices 1 and 2. Reconstructed data is then obtained. Next, feature mapping is used to move input between convolutional layers. When converting multidimensional input to one-dimensional input, the flatten layer is used. The one-dimensional input is acquired and fed into a multilayer perceptron to achieve non-linear data separation. The dropout layer removes some superfluous inputs while leaving the other inputs unaffected.

Our proposed strategy, shown in Figure 5, improves multi-label classification performance on CSV datasets by combining the advantages of CNNs with low-rank matrix factorization. The MF-CNN architecture has two components: the convolutional neural network layer and the matrix factorization layer. The matrix factorization layer decomposes the original input matrix X into two low-rank matrices, M and N, using either singular value decomposition (SVD) or non-negative matrix factorization

(NMF). The decomposed matrices M and N represent the latent qualities of the input data, and can be used to train the CNN layer. After receiving the deconstructed matrices M and N, the CNN layer employs pooling and convolutional techniques to extract high-level features from the input data. The acquired characteristics are then processed by a fully connected layer using softmax activation to produce the final multi-label classification results. Different convolution kernels can be applied to each layer of a CNN to create unique feature maps. On the next layer's feature map, each group of neighbouring neurons is associated to a neuron. After the feature map has been constructed, each spatial position in the input has the same kernel. Following a few convolutional and pooling layers, classification is performed using one or more totally linked layers. CNN manages these commonalities by using lower-level information to build unique higher-level traits. Pooling layers aggregate similar feature values computed at numerous places using a fixed value. The pooling layer may handle anomalies with different distributions.The localization and pooling capabilities of the CNN can improve anomaly detection performance. The locality features of convolutional layers protect the model from the effects of noise on data. Convolution layers' recovered features are thus noise-resistant. The difficulty of traditional shallow learning approaches to distinguish between benign and deviant data is due to the lower level equivalence of their created properties.The deep CNN retrieves significant properties from input data by leveraging its weight sharing property, which delivers speed advantages. To avoid overfitting, we utilise a dropout approach to randomly remove neurons from the model's training. An MLP is a feedforward neural network with many layers of nodes. After receiving input from the previous layer, each node in a layer uses an activation function to generate an output, which is then passed on to the next layer. The final layer generates the network's forecast.

## 4 | RESULTS

This section discusses the experimental findings on the two datasets. On hybrid datasets, CNN and MLP classifiers were used to achieve the necessary accuracy for multi-label classification. The experiment was conducted with different learning rate parameters. The experiment at first has been done on LR-CNN. In optimisation, the learning rate function determines the step size at each loop with the lowest loss function. When using low rank CNN with the provided methodology, 150 epochs with a learning rate of 0.020 yield the best results with 94.26% accuracy.

**TABLE 1** Multi-label Classification Results on proposed methodology (LR-CNN)

| *LR* | *Precision* | *Recall* | *F1-Score* | *Accuracy* |
|---|---|---|---|---|
| LR1=0.0010 | .888 | .920 | .904 | 92.45 |
| LR2=0.0015 | .902 | .916 | .901 | 91.44 |
| **LR3=0.0020** | **.891** | **.926** | **.909** | **94.26** |
| LR4=0.0025 | .889 | .924 | .907 | 93.21 |

**TABLE 2** Multi-label Classification Results on proposed methodology (LR-CNN-MLP)

| *LR* | *Precision* | *Recall* | *F1-Score* | *Accuracy* |
|---|---|---|---|---|
| LR1=0.0010 | 0.943 | 0.975 | 0.959 | 96.12 |
| LR2=0.0015 | 0.956 | 0.97 | 0.963 | 96.84 |
| **LR3=0.0020** | **0.944** | **0.979** | **0.961** | **98.17** |
| LR4=0.0025 | 0.942 | 0.978 | 0.960 | 97.91 |

Table 1 displays the results of the suggested multilabel classification methods (LR-CNN). Various learning rates were utilised in the experiment. The learning rate of 0.0020 produces the best results, with precision of.891, recall of.926, f1-score of.909, and maximum accuracy of 94.26%. Table 2 displays the results of the suggested multilabel classification methods (LR-CNN-MLP). The experiment was repeated with varying learning rates. The learning rate of 0.0020 produces the greatest results: precision

of.944, recall of.979, f1-score of.961, and maximum accuracy of 98.17%. LR-CNN and LR-CNN-MLP produce the greatest outcomes with a learning rate of 0.0020. As a result, no effect from different learning values has been seen. For comparison the

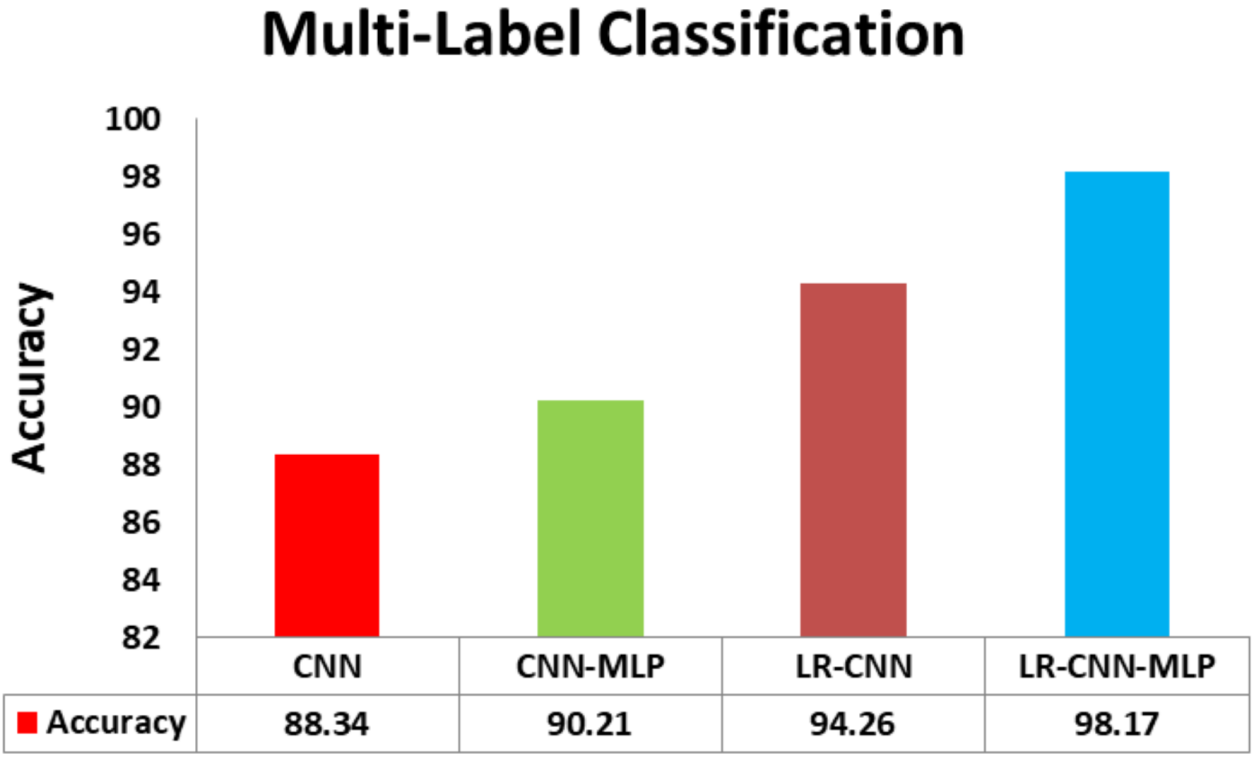


**FIGURE 2** Comparison of Accuracy of Different Deep Learning Classifiers

experiment has been performed on CNN, CNN-MLP alone along with LR-CNN, LR-CNN-MLP.

## 5 | CONCLUSION

The low rank approximation and hybrid deep learning models used in the LR-CNN-MLP technique provided a revolutionary way to multi-label categorization of network attacks on traffic data. The technique combines the benefits of three distinct models: the low rank approximation for data reduction, the multilayer perceptron for classification, and the hierarchical convolutional neural network for feature extraction. The experimental results showed that the suggested LR-CNN-MLP strategy outperformed numerous cutting-edge techniques currently used for multi-label categorization of network assaults on traffic data. Furthermore, the results revealed that hybrid deep learning models greatly increased classification accuracy, particularly on difficult and large datasets. Overall, the LR-CNN-MLP approach offers a promising solution for multi-label categorization of network attacks on traffic data. This method is predicted to increase classification accuracy in other fields as well. Future aims include applying hybrid deep learning models to additional security-related traffic datasets and deploying the following strategy in real-world scenarios when detecting a specific attack type is necessary.

## ACKNOWLEDGEMENTS

## References

1. Hwang, R.H.; Peng, M.C.; Huang, C.W.; Lin, P.C.; Nguyen, V.L. An unsupervised deep learning model for early network traffic anomaly detection. IEEE Access 2020, 8, 30387–30399.
2. Xiao, Y.; Xing, C.; Zhang, T.; Zhao, Z. An intrusion detection model based on feature reduction and convolutional neural networks. IEEE Access 2019, 7, 42210–42219.
3. Koroniotis, N.; Moustafa, N.; Sitnikova, E.; Turnbull, B. Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset. Future Gener. Comput. Syst. 2019, 100, 779–796
4. Moustafa, N.; Slay, J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS), Canberra, Australia, 10–12 November 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–6
5. Sumarsono, A., Du, Q. and Younan, N., 2015, IEEE June. Hyperspectral image segmentation with low-rank representation and spectral clustering. In 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) (pp. 1-4).
6. Sumarsono, A. and Du, Q., 2016. Low-rank subspace representation for supervised and unsupervised classification of hyperspectral imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(9), pp.4188-4195.
7. Wang, J., Shi, D., Cheng, D., Zhang, Y. and Gao, J., 2016. LRSR: low-rank-sparse representation for subspace clustering. Neurocomputing, 214, pp.1026-1037.
8. Wang, J., Wang, X., Tian, F., Liu, C.H. and Yu, H., 2016. Constrained low-rank representation for robust subspace clustering. IEEE transactions on cybernetics, 47(12), pp.4534-4546.
9. Babbar, H., Bouachir, O., Rani, S. and Aloqaily, M., 2022, April. Evaluation of deep learning models in its software-defined intrusion detection systems. In NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium (pp. 1-6).
10. Sethi, M., Ahuja, S. and Kukreja, V., 2021, August. An empirical study for the deep learning models. In Journal of Physics: Conference Series Vol. 1950, No. 1, p. 012071.
11. Rani, S., Singh, A., Elkamchouchi, D.H. and Noya, I.D., 2022. Lightweight Hybrid Deep Learning Architecture and Model for Security in IIOT. Applied Sciences, 12(13), p.6442.
12. Rani, S. and Bashir, A.K., 2022, October. Analysis of Machine Learning and Deep Learning Intrusion Detection System in Internet of Things Network. In 2022 International Conference on Data Analytics for Business and Industry (ICDABI) pp. 1-9.
13. Sethi, M., Ahuja, S. and Bawa, P., 2023. Deep Learning Techniques Using Transfer Learning for Classification of Alzheimer's Disease. Machine Intelligence, Big Data Analytics, and IoT in Image Processing: Practical Applications, p.1.
14. Lopez-Martin, M., Carro, B., Sanchez-Esguevillas, A. and Lloret, J., 2019. Shallow neural network with kernel approximation for prediction problems in highly demanding data networks. Expert Systems with Applications, 124, pp.196-208.
15. Akinsola, J.E.T., Awodele, O., Kuyoro, S.O. and Kasali, F.A., 2019. Performance evaluation of supervised machine learning algorithms using multi-criteria decision making techniques. In Proceedings of the International Conference on Information Technology in Education and Development (ITED) (pp. 17-34).

**How to cite this article:** Williams K., B. Hoskins, R. Lee, G. Masato, and T. Woollings (2016), A regime analysis of Atlantic winter jet variability applied to evaluate HadGEM3-GC2, *Q.J.R. Meteorol. Soc.*, *2017;00:1–6*.