International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 08 | August - 2022 Impact Factor: 7.185 ISSN: 2582-3930

A Machine Learning Approach to 5G Infrastructure Market Optimization

#1 K.Ramya , Asst Prof , Dept of MCA, Mohamed Sathak Engineering College
#2.V.Karan Final MCA , Mohamed Sathak Engineering College
#3 Dr.S.Sajitha Banu, Asst Pro, Dept of MCA, Mohamed Sathak Engineering College
#4 G.Bala Murugan , Asst Prof , Mohamed Sathak Engineering College
#5 N.Balasubramanian, Asst Prof , Mohamed Sathak Engineering College

ABSTRACT

Traffic Classification (TC) systems allow inferring the application that is generating the traffic being analyzed. Stateof-the-art TC algorithms are based on Deep Learning (DL) and have outperformed traditional methods in complex and modern scenarios, even if traffic is encrypted. Most of the works on TC assume the traffic flows on a wired network under the same network management domain. This assumption limits the capabilities of TC systems in wireless networks since users' traffic on one network domain can be negatively impacted by undetected traffic transmissions from users in other network domains or detected ones but with no traffic context in a shared spectrum. To solve this problem, we introduce a novel framework to achieve TC at any layer on the radio network stack. We propose a spectrum-based procedure that uses a DL-based classifier to realize this framework. We design two DL-based classifiers, a novel Convolutional Neural Network (CNN) spectrum-based TC and a Recurrent Neural Networks (RNN) as baseline architecture, and benchmark their performance on three TC tasks at different radio stack layers. The datasets were generated by combining packet traces from real transmissions with a standard-compliant waveform generator for 802.11 radio technologies. Performance evaluations show that the best model can achieve an accuracy above 92% in the most demanding TC task, a drop of only 4.37% in accuracy compared to a byte-based DL approach, with microsecond per-packet prediction time, which is very promising for delivering real-time spectrum-based traffic analyzers.

KEYWORDS

RNN- Recurrent Neural Networks
CNN- Convolutional Neural Network

I.INTRODUCTION

Nowadays, wireless technologies are omnipresent and provide access to millions of users and machines to the Internet. This access is mostly offered by complementary technologies, e.g., 4G/5G Mobile Networks and Wireless Local Area Networks (WLANs), with an increased network capacity to support the ever-increasing number of devices and applications. As a result, managing and optimizing the wireless network capacity to provide Quality of Service (QoS) becomes even more challenging [1], [2]. Traditionally, the Network Monitoring Service (NMS) performs a set of tasks to analyze the behavior of the networks and services throughout their traffic. The information provided by this system can be used to determine which applications affect network the most in terms of total bandwidth usage or identify the most critical links so decision-making engines for network management can ensure fast troubleshooting and securing high QoS to the users. More recently, this service has been enhanced with Machine Learning (ML) techniques to perform automatic network traffic analysis such as network state predictions, anomaly detection, malware detection, and TC [3]. Focusing on TC, this network management task allows inferring the application that is generating the traffic [2], [3]. Knowing the traffic class provides a mechanism to enforce specific security and QoS policies on the analyzed traffic. Over the years, several approaches have been designed and developed to follow the evolution of the technologies driving the development of user's application and communication protocols.

II.PROPOSED SYSTEM

In proposed system, we have The RNN architecture, designed as baseline, is inspired on [8] with a fine-tuning and optimization steps based on the dataset provided in this paper. Following a similar approach as with the CNN, we varied the number of recurrent layers and the type and number of recurrent units to find the model that performs the best. The number of recurrent layers varied from 1 to 4, achieving the best performance with 3 recurrent layers, result that is aligned with [8]. We also varied the type of recurrent units between GRU and LSTM and we found that GRU had similar or outperformed the LSTM in both execution time and accuracy in all the experiments we run. This result is also aligned with previous findings in other comparative of plots

ADVANTAGES:

- It performs the good Prediction.
- It enhance the performances.

EXISTING SYSTEM

Finally, transmitting L1 packets over the air will result in receiving packets with a large variety of SNR values compared to the values used in the generated dataset (20 to 30dB). Changes in the SNR values will modify the original signal and, in the case of low SNR values, it will negatively impact the performance of the DL classifier as shown in previous work [37], [45]. For this, we plan to augment the provided dataset with packets generated with a larger set of SNR values (e.g., in the range between -20 to 20 dB) so a SNR sensitivity analysis can be performed. With the resulting dataset, researchers would be able to investigate mechanisms to mitigate the negative impact on the classifier' performance when packets are received with low SNR. For instance, the use of denoising AEs as feature extractors may improve the performance in the presence of high levels of noise. Additionally, the augmented dataset could foster the development and implementation of novel algorithms, closing the performance gap when a classifier is trained with synthetic, but standard compliant dataset, deployed in a real environment

DISADVANTAGES

- Training model prediction on Time is High

III. MODULES DESCRIPTION

DATA SELECTION:

- The dataset was collected from the dataset repository.
- In this step, we have to load the data with the help of panda's packages.

PREPROCESSING:

- Data pre-processing is the process of removing the unwanted or unnecessary data from the input dataset.
- **Missing data removal**: In this process, the null values such as missing values and Nan values are replaced by 0.
- Missing and duplicate values were removed and data was cleaned of any abnormalities.
- **Encoding Categorical data**: That categorical data is defined as variables with a finite set of label values.

DEEP LEARNING:

- In classification process, we are using ANN algorithm for better performance.
- An **CNN and RNN** is configured for a specific application, such as pattern recognition or data classification, through a learning process.
- Learning largely involves adjustments to the synaptic connections that exist between the neurons.

DATA SPLITTING:

- Data splitting is the act of partitioning available data into two portions, usually for cross-validator purposes.
- One Portion of the data is used to develop a predictive model and the other to evaluate the model's performance.
- Separating data into training and testing sets is an important part of evaluating data mining models.
- Typically, when you separate a data set into a training set and testing set, most of the data is used for training, and a smaller portion of the data is used for testing.

RESULT GENERATION:

- While using collaborative filtering algorithm, the recommend Movie for particular user id.
- The Final Result will get generated based on the overall classification and prediction.

- The performance of this proposed approach is evaluated using some measures like,
- **Accuracy**

Accuracy of classifier refers to the ability of classifier. It predicts the class label correctly and the accuracy of the predictor

- **Precision**

Precision is defined as the number of true positives divided by the number of true positives plus the number of false positives.

Precision=TP/ (TP+FP)

- **Recall**

Recall is the number of correct results divided by the number of results that should have been returned. In binary classification, recall is called sensitivity.

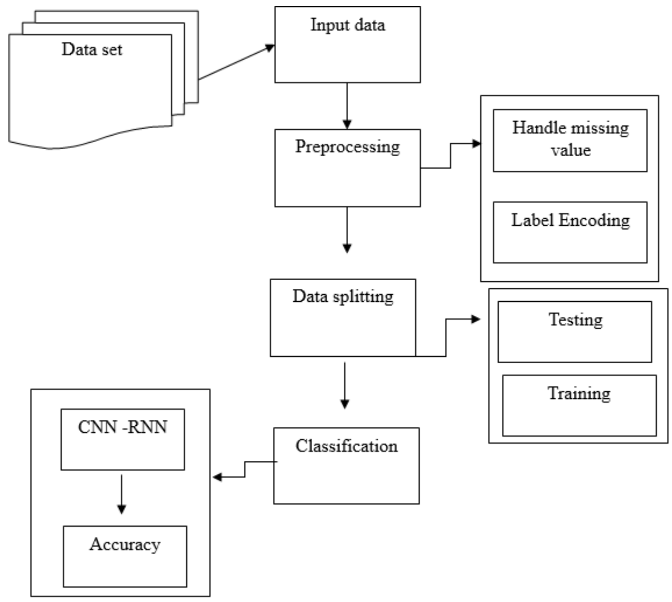
Recall=TP/ (TP+FN)

- **F-Measure**

F measure (F1 score or F score) is a measure of a test's accuracy and is defined as the weighted harmonic mean of the precision and recall of the test.

F-measure=2TP/ (2TP+FP+FN)

SYSTEM ARCHITECTURE :-



SYSTEM TESTING:-

System testing is the stage of implementation, which aimed at ensuring that system works accurately and efficiently before the live operation commence. Testing is the process of executing a program with the intent of finding an error. A good test case is one that has a high probability of finding an error. A successful test is one that answers a yet undiscovered error.

UNIT TESTING

Unit testing is the testing of each module and the integration of the overall system is done. Unit testing becomes verification efforts on the smallest unit of software design in the module. This is also known as 'module testing'.

INTEGRATION TESTING

Data can be lost across an interface, one module can have an adverse effect on the other sub function, when combined, may not produce the desired major function. Integrated testing is systematic testing that can be done with sample data. The need for the integrated test is to find the overall system performance. There are two types of integration testing.

WHITE BOX TESTING

White Box testing is a test case design method that uses the control structure of the procedural design to drive cases. Using the white box testing methods

We Derived test cases that guarantee that all independent paths within a module have been exercised at least once.

BLACK BOX TESTING

- Black box testing is done to find incorrect or missing function
- Interface error
- Errors in external database access
- Performance errors.
- Initialization and termination errors

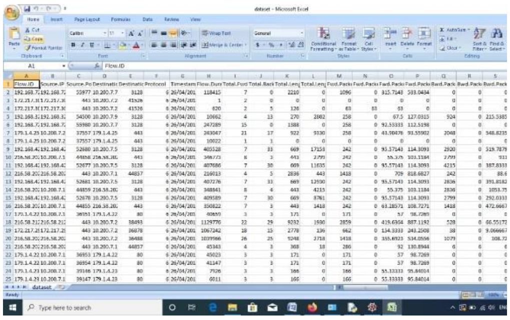
In 'functional testing', is performed to validate an application conforms to its specifications of correctly performs all its required functions. So this testing is also called 'black box testing'. It tests the external behaviour of the system. Here the engineered product can be tested knowing the specified function that a product has been designed to perform, tests can be conducted to demonstrate that each function is fully operational.

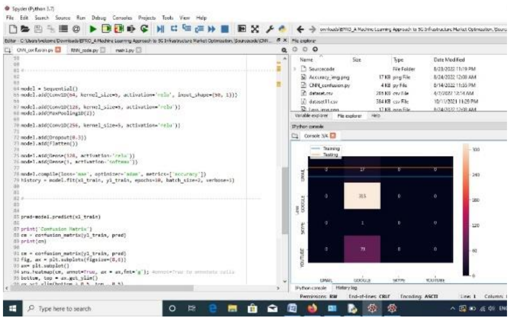
VALIDATION TESTING

After the culmination of black box testing, software is completed assembly as a package, interfacing errors have been uncovered and corrected and final series of software validation tests begin validation testing can be defined as many

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 08 | August - 2022 Impact Factor: 7.185 ISSN: 2582-3930

WORKING:-





CONCLUSION

Performance evaluations showed that the DL model based on CNN could achieve the best performance on the three proposed tasks, achieving above 99.9% in task accuracy discriminating among classes in task 1, 97.8% in task 2, and 92% in task 3. These results are very promising if we compared them to byte-based DL models, where spectrum based achieved similar accuracy on task 1, a drop of 1.38% in task 2, and 4.37% in task 3. Finally, the proposed DL architecture could perform the prediction of a given class in the order of microseconds, prediction times that are compelling for integrating into spectrum-based real-time traffic analysers. As future work, there are several challenges to be addressed with the proposed framework and the spectrum-based procedure for TC.

REFERENCES

[1] M. Roughan, S. Sen, O. Spatscheck et al., “Class-of-service mapping for qos: A statistical signature-based approach to ip traffic classification,” in Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement, ser. IMC ’04. ACM, 2004, p. 135–148.

[2] S. Valenti, D. Rossi, A. Dainotti et al., Reviewing Traffic Classification. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 123–147.

[3] G. Aceto, D. Ciuonzo, A. Montieri et al., “Toward effective mobile encrypted traffic classification through deep learning,” Neurocomputing, vol. 409, pp. 306 – 315, 2020.

[4] F. Pacheco, E. Exposito, M. Gineste et al., “Towards the deployment of machine learning solutions in network traffic classification: A systematic survey,” IEEE Communications Surveys Tutorials, vol. 21, no. 2, pp. 1988–2014, 2019.

[5] C. Zhang, P. Patras, and H. Haddadi, “Deep learning in mobile and wireless networking: A survey,” IEEE Communications Surveys Tutorials, vol. 21, no. 3, pp. 2224–2287, 2019.

[6] M. Lotfollahi, M. J. Siavoshani, R. S. H. Zade et al., “Deep packet: A novel approach for encrypted traffic classification using deep learning,” Soft Computing, vol. 24, no. 3, pp. 1999–2012, 2020.

[7] P. Soto, M. Camelo, J. Fontaine et al., “Augmented wi-fi: An ai-based wi-fi management framework for wi-fi/lte coexistence,” in 2020 16th International Conference on Network and Service Management (CNSM), 2020, pp. 1–9.

[8] T. J. O’Shea, S. Hitefield, and J. Corgan, “End-to-end radio traffic sequence recognition with recurrent neural networks,” in 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2016, pp. 277–281.

[9] S. Riyaz, K. Sankhe, S. Ioannidis et al., “Deep learning convolutional neural networks for radio identification,” IEEE Communications Magazine, vol. 56, no. 9, pp. 146–152, 2018.

[10] M. Camelo, T. De Schepper, P. Soto et al.,

"Detection of traffic patterns in the radio spectrum for cognitive wireless network management," in 2020 IEEE International Conference on Communications (ICC), Jun 2020.