

FULL TEXT LINKS



[Comput Intell Neurosci](#). 2022 Aug 26;2022:3131153. doi: 10.1155/2022/3131153. eCollection 2022.

Apache Spark and Deep Learning Models for High-Performance Network Intrusion Detection Using CSE-CIC-IDS2018

[Abdulnaser A Hagar](#)^{1 2}, [Bharti W Gawali](#)²

Affiliations

PMID: 36059395 PMID: [PMC9439899](#) DOI: [10.1155/2022/3131153](#)

Abstract

Keeping computers secure is becoming challenging as networks grow and new network-based technologies emerge. Cybercriminals' attack surface expands with the release of new internet-enabled products. As many cyberattacks affect businesses' confidentiality, availability, and integrity, network intrusion detection systems (NIDS) show an essential role. Network-based intrusion detection uses datasets like CSE-CIC-IDS2018 to train prediction models. With fourteen types of attacks included, the latest big data set for intrusion detection is available to the public. This work proposes three models, two deep learning convolutional neural networks (CNN), long short-term memory (LSTM), and Apache Spark, to improve the detection of all types of attacks. To reduce the dimensionality, random forests (RF) was employed to select the important features; it gave 19 from 84 features. The dataset is imbalanced; thus, oversampling and undersampling techniques reduce the imbalance ratio. The Apache Spark model produced the best results across all 15 classes, with accuracy as high as 100% for all classes, as seen by the experiments' findings. For the *F1*-score, Apache Spark showed the highest results with 1.00 for most classes. The findings of the three models showed outstanding results for multiclassification network intrusion detection.

Copyright © 2022 Abdulnaser A. Hagar and Bharti W. Gawali.

[PubMed Disclaimer](#)

Figures

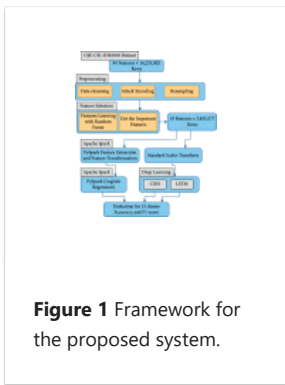


Figure 1 Framework for the proposed system.

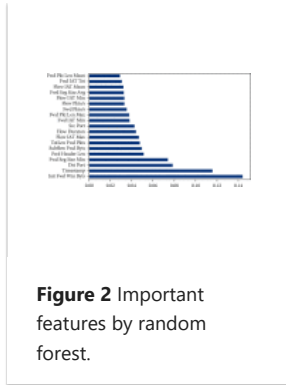


Figure 2 Important features by random forest.

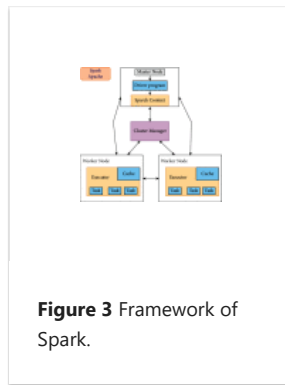


Figure 3 Framework of Spark.

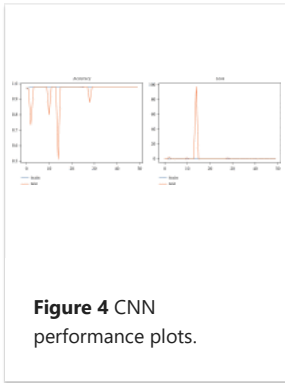


Figure 4 CNN performance plots.

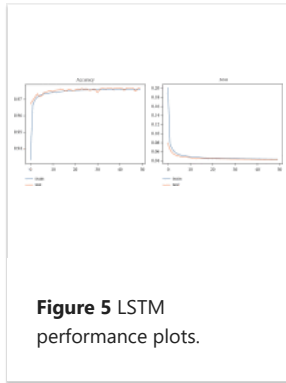


Figure 5 LSTM performance plots.

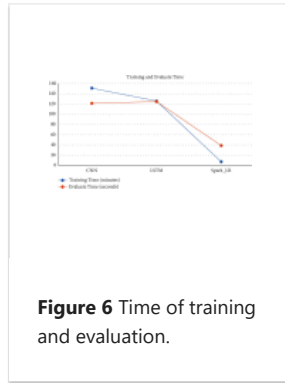


Figure 6 Time of training and evaluation.

All figures (9)

Related information

[PMC images](#)

LinkOut – more resources

Full Text Sources

- [Europe PubMed Central](#)
- [Hindawi Limited](#)
- [PubMed Central](#)

Research Materials

- [NCI CPTC Antibody Characterization Program](#)

Miscellaneous

- [NCI CPTAC Assay Portal](#)