Enhancing IoT Security with Explainable AI-Powered Intrusion Detection Systems

A Capstone Project Presentation

19 September, 2024

Presenter:
Sani Abdullahi Sani

Supervisor:
Dr. Ibidun Obagbuwa

Agenda Overview

01 Introduction & Problem Statement

02 Background of the study

03 Related Work

04 Research Questions

05 Methodology

06 Experimental Setup

07 Results

08 Conclusion

09 Acknowledgement

10 Reference

Introduction and Problem Statement

- The Internet of Things (IoT) emerges as a transformative technological paradigm.
- According to Cisco, the global count of IoT devices is anticipated to hit around **30 billion** by 2030.
- However, its heterogeneity and dependency on other technologies like Fog and Cloud make it vulnerable.
- Robust security measures are imperative, but conventional solutions are inadequate due to resource constraints.
- AI-Based Intrusion Detection Systems (IDS) shows great promise but lack transparency.

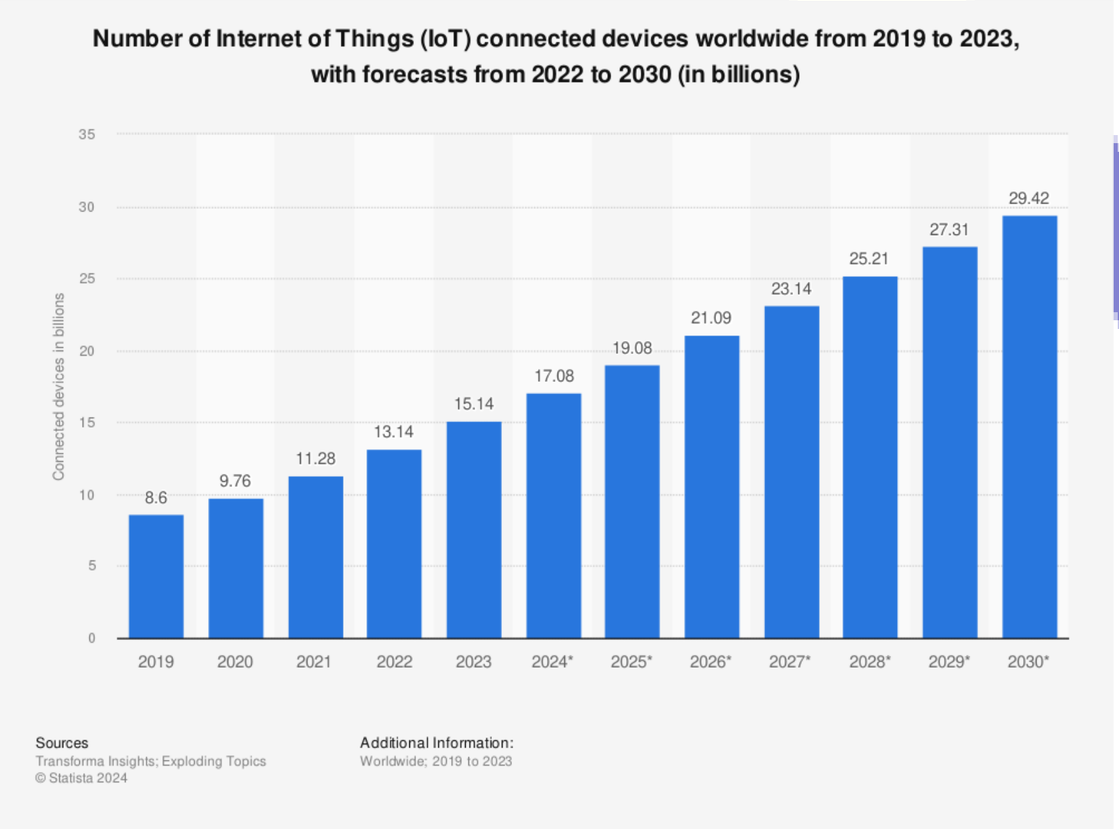


Image Source: IoT Connected Devices Worldwide. Statista 2024. URL: https://www.statista. com/statistics/1183457/iot-connected-devices-worldwide/)

Background of the Study

Category		Key Points
IoT	Devices are categorized into	Sensors & Actuators
	Architecture	Perception, Network & Application
IDS in IoT	Attack types	Internal & External
	Key Elements	Monitoring, Detection, Reporting
	Implementation Strategies	Host-Based, Network-Based
	Methodologies	Signature-Based, Anomaly-Based, Rule-Based
	Placement Strategies	Centralized Approach, Distributed Approach, Hybrid Approach
AI-Driven IDS in IoT	Types	Machine Learning (ML), Deep Learning (DL), Ensemble Methods, Federated Learning (FL)
Explainable AI (XAI)	Scope	Local & Global.
	Model Dependency	Model Specific & Agnostic.

Related Work

Author(s)	Key Focus	Findings
Liu & Pasquale (2018)	Security challenges in IoT	Importance of robust authentication and need for advanced IDS.
Moustafa et al. (2018)	Ensemble-based IDS using statistical flow features	Improved detection rate and minimal false positives compared to conventional techniques.
Kelton da Costa et al. (2019)	Survey of ML for IoT intrusion detection	Effectiveness of ensemble methods, and challenges associated with model interpretability.
Choo et al. (2019)	ML-based IDS for smart homes	efficacy of the ML-based approach in distinguishing between normal and malicious traffic.
Sridhar et al. (2020)	Review of ML techniques for IoT intrusion detection	Growing importance of explainability in ML models for trust and understanding.
Mane & Rao (2021)	Explainable AI in IDS	importance of explainability in enhancing trust and the need for further research on integrating explainable AI into existing systems.
Manickam et al. (2021)	XAIDIDS with explainable AI	System effectiveness in detecting intrusions and the significance of explainability in enhancing transparency.
Vakula Rani et al. (2023)	XAI in IoT network intrusion detection	XAI improves trust and transparency in AI-based IDS solutions.

Research Questions

- How can an IDS be designed and implemented using CNN, XGB, RF, DT, and MLP classifiers to detect intrusions within IoT networks?
- How can XAI techniques be integrated into the system to provide transparent and interpretable explanations?
- What impact does feature engineering technique has in the performance of IDS?
- How does the predictive performance of White box models (DT, RF, XGBoost) compare with that of black box models (MLP and CNN)?

Methodology

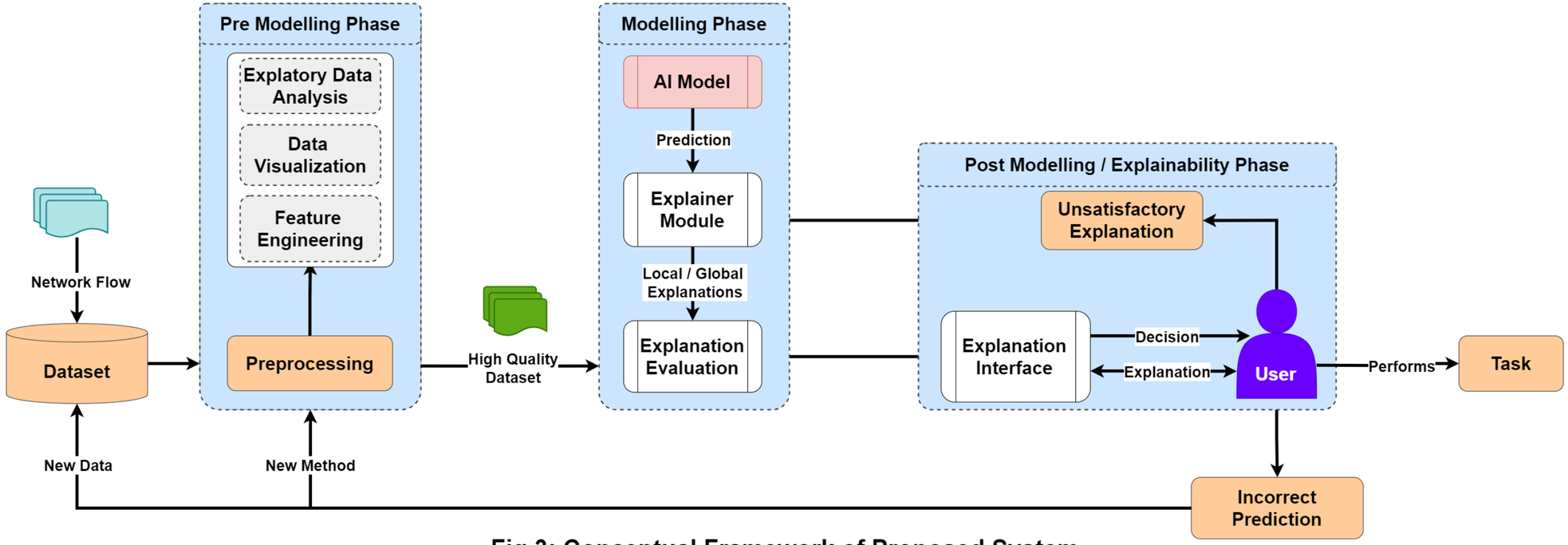
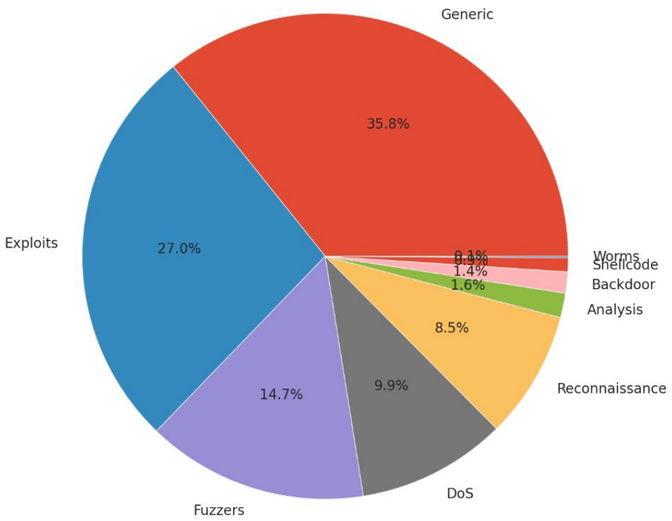
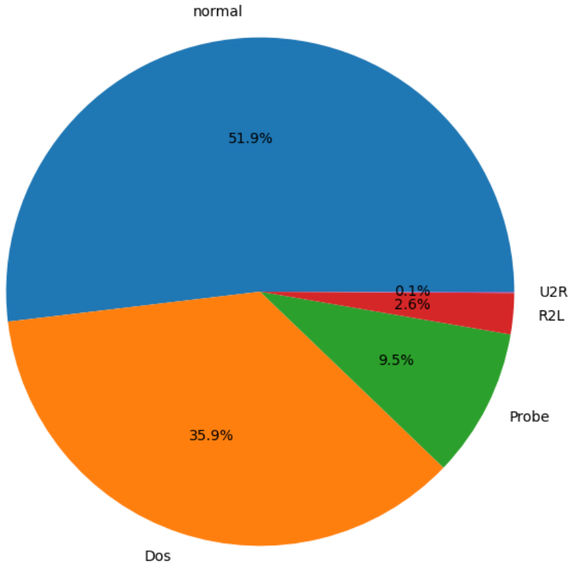


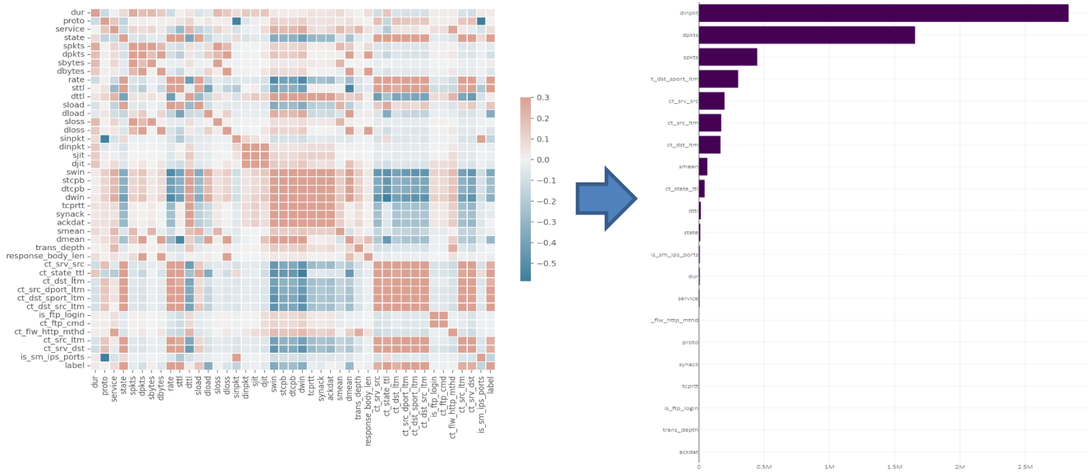
Fig 3: Conceptual Framework of Proposed System

Methodology

K-Fold Cross Validation
is used to train our model







Feature Selection Strategy

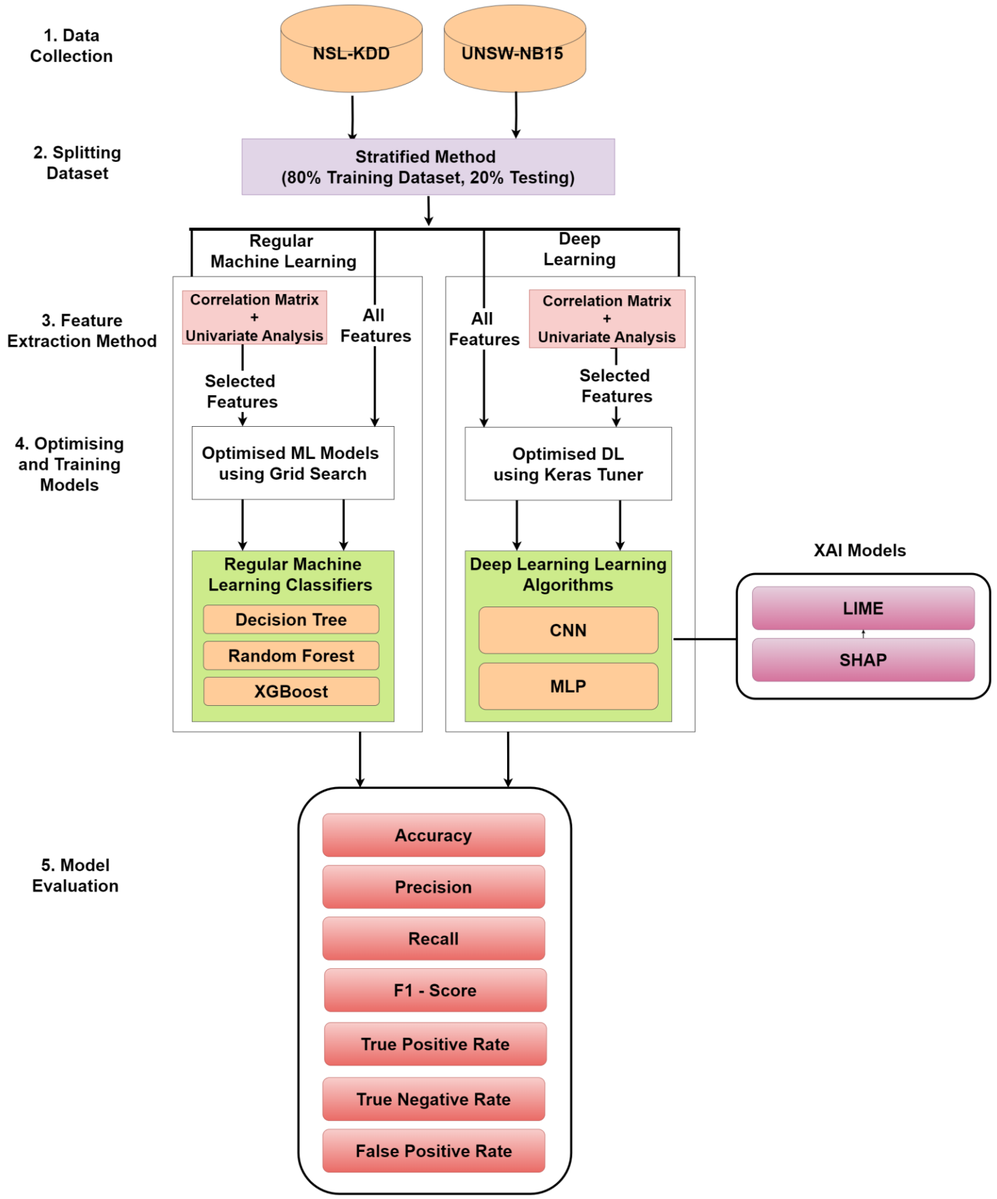
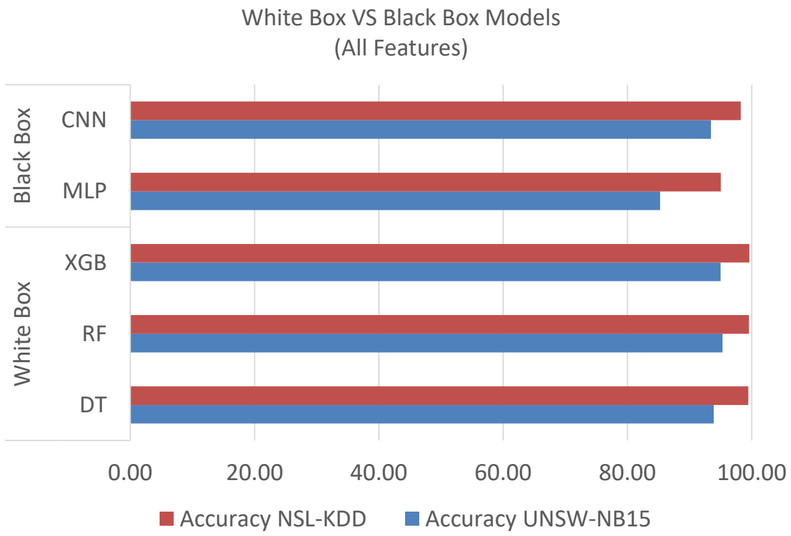
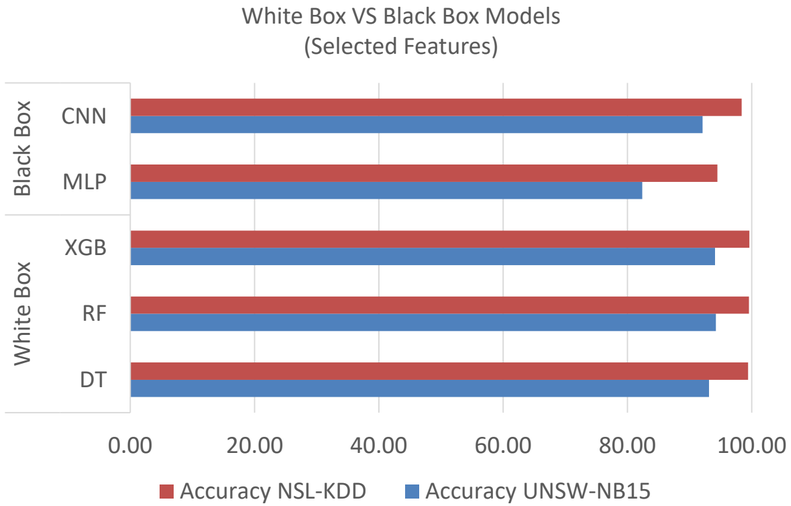
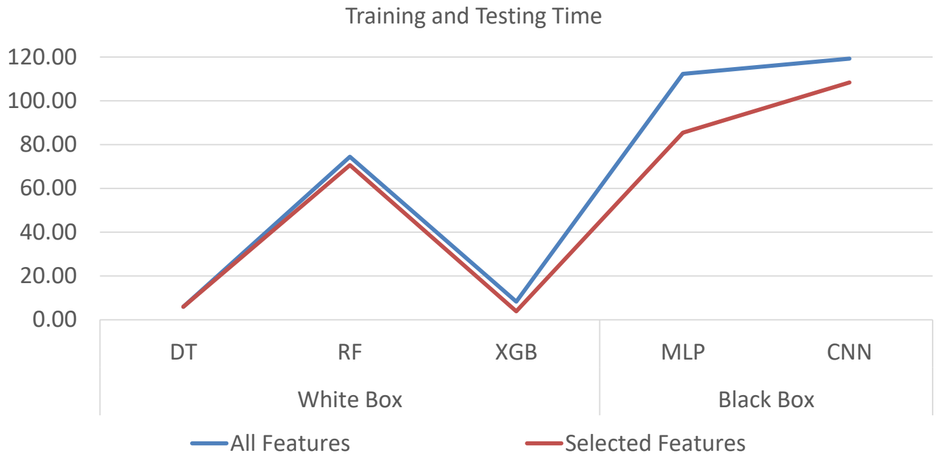


Fig 4: The workflow of Machine learning and Deep learning Models.

Results

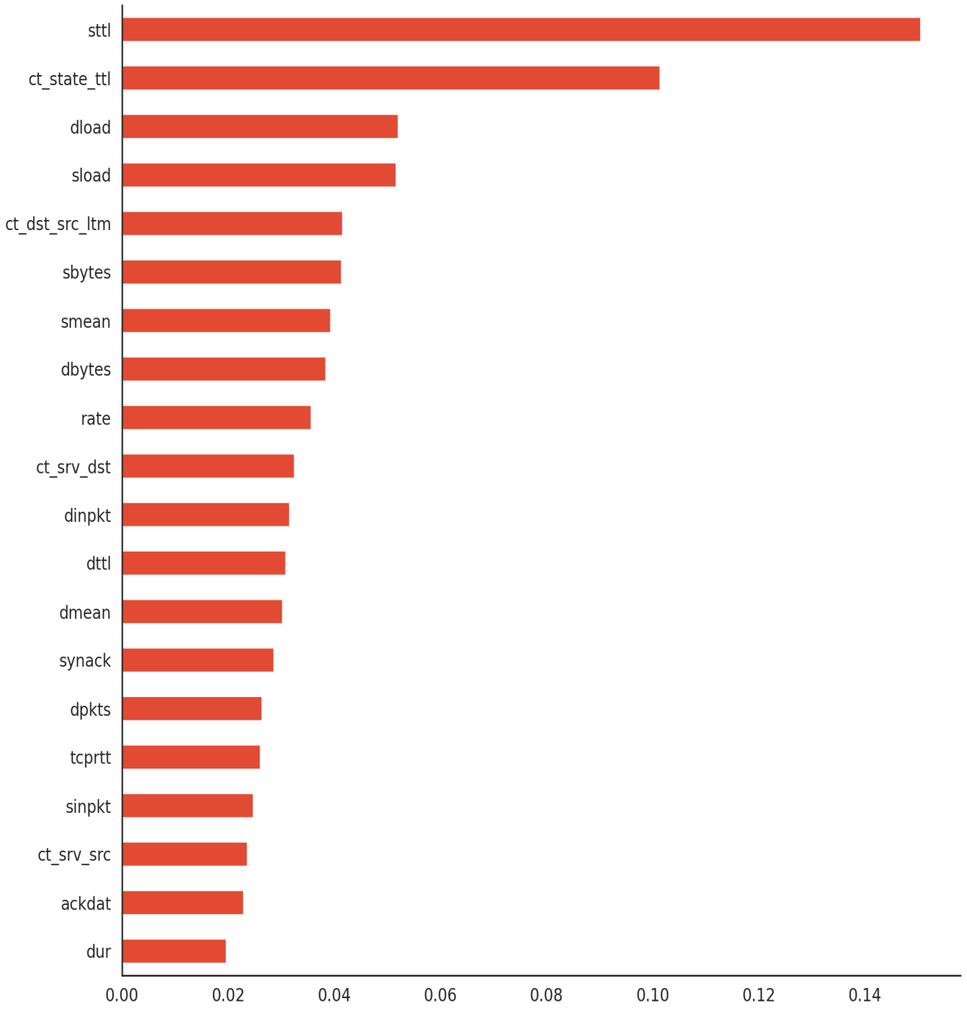






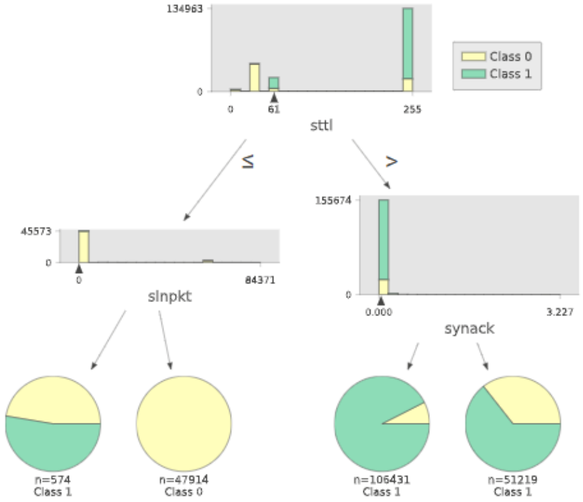
- Feature selection shows improvement in computational efficiency though with a slight drop in performance.
- White box models outperforms black box models, with Random Forest achieving the highest performance.

Explainability - Whitebox

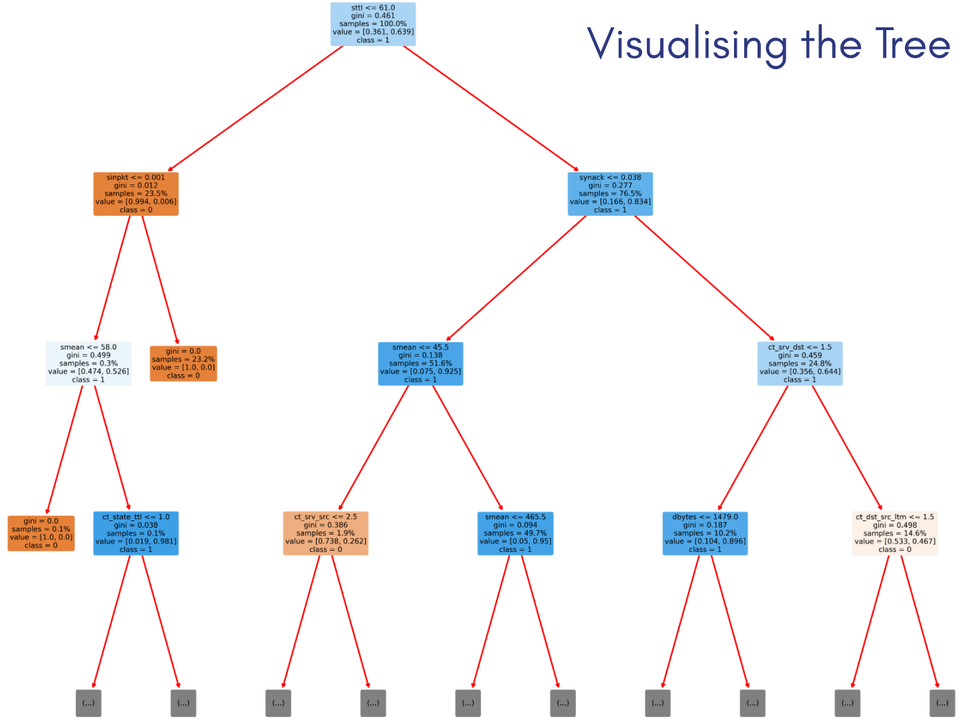


Feature Importance

Visualising the Rules

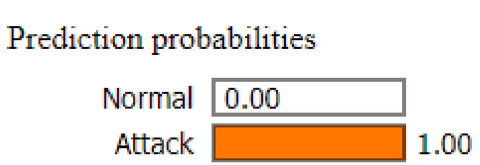


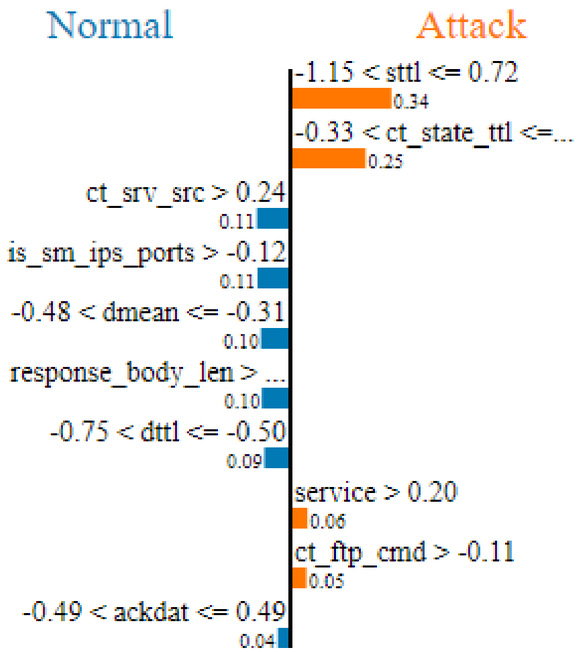
Visualising the Tree



Explainability – Black-box

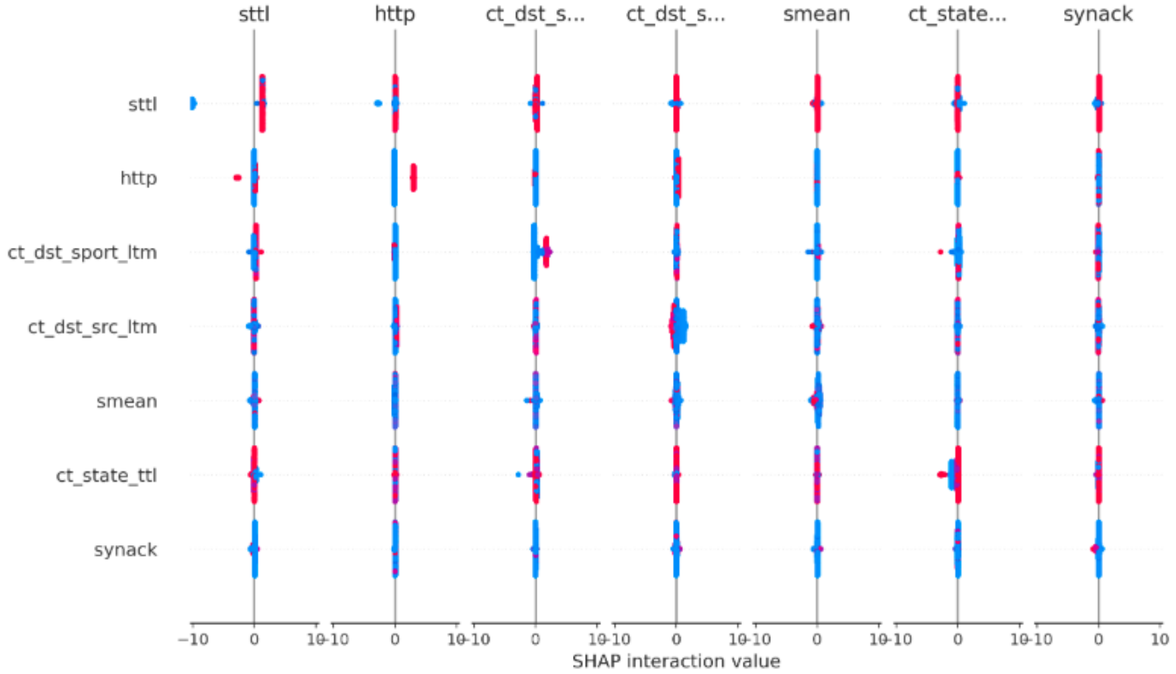
lime



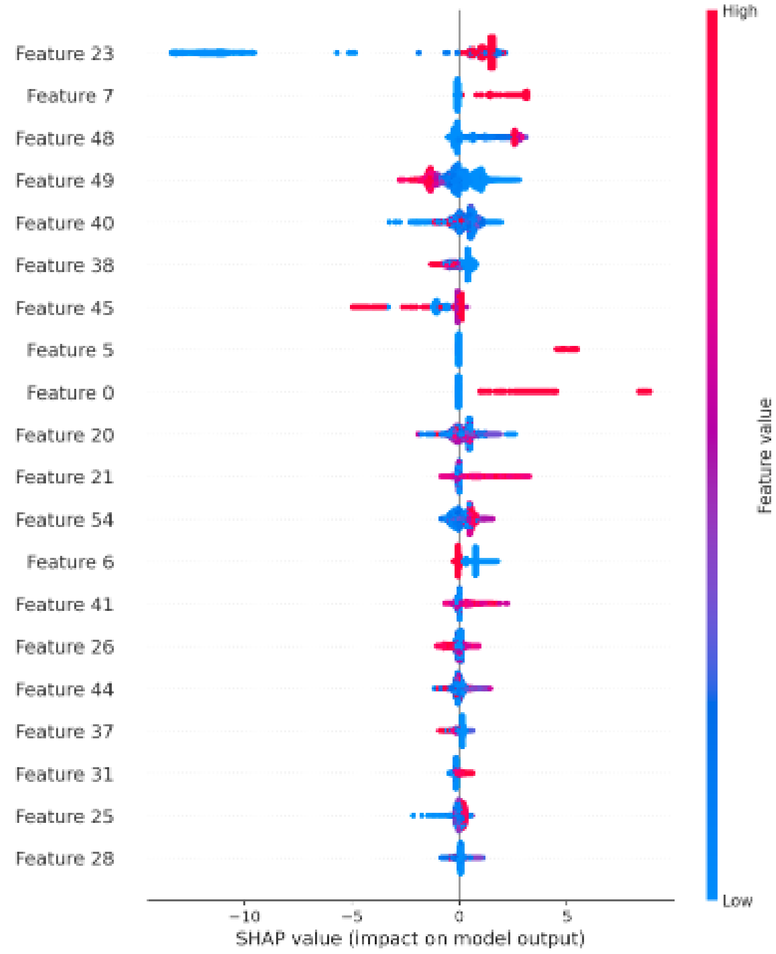


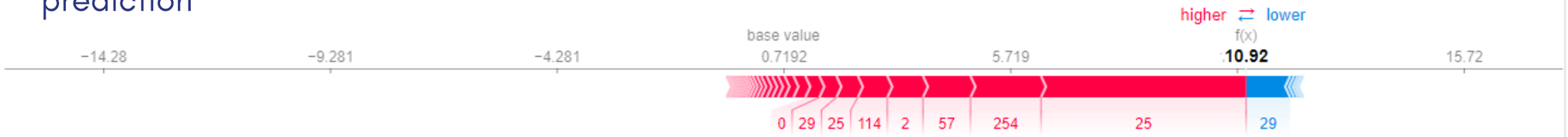
Feature	Value
sttl	0.72
ct_state_ttl	0.68
ct_srv_src	0.33
is_sm_ips_ports	-0.12
dmean	-0.48
response_body_len	-0.04
dttl	-0.75
service	0.20
ct_ftp_cmd	-0.11
ackdat	-0.48

Explainability – Black-box



provide insights into how pairs of features interact to impact the model's prediction





Single prediction quantifying the impact of each feature

Shap

Conclusion

- We developed an IDS with XAI techniques.
- Phases includes pre-modelling, modelling, and post-modelling, focusing on data quality, model performance, and interpretability.
- We aim to enhance cybersecurity capabilities by providing actionable insights for informed decision-making and threat mitigation.

Acknowledgement

- I also express my heartfelt gratitude to my family, especially my mother.
- Special thanks to my Supervisor Dr Ibidun Obagbuwa for invaluable guidance in this research and also to Dr. Helen Robertson for coordinating efforts.

References

- [1] Jayavardhana Gubbi et al. “Internet of Things (IoT): A vision, architec tural elements, and future directions”. In: Future Generation Computer Systems 29 (2013), pp. 1645-1660. DOI: 10.1016/j.future.2013.01.010.
- [2] D. Evans. The Internet of Things: How the next evolution of the Internet is changing everything. White Paper. San Jose, CA, USA: Cisco Internet Business Solutions Group (IBSG), Cisco, 2011.
- [3] IoT Connected Devices Worldwide. 2022. URL: https://www.statista. com/statistics/1183457/iot-connected-devices-worldwide/ (vis ited on 07/27/2023).
- [4] Sarhad Arisdakessian et al. “A Survey on IoT Intrusion Detection: Fed erated Learning, Game Theory, Social Psychology, and Explainable AI as Future Directions”. In: IEEE Internet of Things Journal 10 (5 Mar. 2023), pp. 4059-4092. ISSN: 23274662. DOI: 10.1109/JIOT.2022.3203249.
- [5] OT/IoTSecurityReportFebruary2021.2021. URL:https://www.nozominetworks. com/landing/ot-iot-security-report-february-2021 (visited on 05/01/2021).
- [12] D. Gunning and D. Aha. “DARPA’s explainable artificial intelligence (XAI) program”. In: AI Mag. 40.2 (June 2019), pp. 44-58.
- [18] A.Tabassum,A.Erbad,andM.Guizani.“Asurveyonrecentapproaches in intrusion detection system in IoTs”. In: Proc. 15th Int. Wireless Com mun. Mobile Comput. Conf. (IWCMC). 2019, pp. 1190-1197.
- [37] M. Wang et al. “An explainable machine learning framework for intru sion detection systems”. In: IEEE Access 8 (2020), pp. 73127-73141.
- [41] A. Adadi and M. Berrada. “Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)”. In: IEEE Access 6 (2018), pp. 52138 52160.
- [48] Y. Liu and L. Pasquale. “A survey on security for Internet of Things”. In: IEEE Internet of Things Journal 5.5 (2018), pp. 3018-3028.

References

- [49] K.K.R.Choo,N.Dlodlo, and R. Choo. “An IoT intrusion detection sys tem for smart home environments using machine learning techniques”. In: IEEE Internet of Things Journal 6.1 (2019), pp. 630–641.
- [50] Kelton AP Kelton da Costa et al. “Internet of Things: A survey on ma chinelearning-basedintrusiondetectionapproaches”.In:JournalofComputer Networks 151 (2019), pp. 147–157.
- [51] NourMoustafaetal.“AnEnsembleIntrusionDetectionTechniquebased on proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things”. In: IEEE Internet of Things Journal (2018). 45
- [52] ShraddhaManeandDattarajRao.ExplainingNetworkIntrusionDetection System Using Explainable AI Framework.
- [53] S. Manickam, B. Shanthi, and S. Sangeetha. “XAIDIDS: Explainable ar tificial intelligence-based intrusion detection system for IoT networks”. In: Computers & Security 105 (2021), p. 102267.
- [54] A.Sridhar, P. B. Raj, and B. B. Gupta. “Explainable AI techniques for in trusion detection in IoT networks: A review”. In: IEEE Internet of Things Journal 8.13 (2020), pp. 10810–10823.
- [55] C3 AI. LIME: Local Interpretable Model-Agnostic Explanations. https:// c3.ai/glossary/data-science/lime-local-interpretable-model agnostic-explanations. 2020.
- [56] J. Vakula Rani, Haifa Ali Saeed Ali, and Aishwarya Jakka. “IoT Net work Intrusion Detection: An Explainable AI Approach in Cybersecurity”. In: Institute of Electrical and Electronics Engineers Inc., 2023. ISBN: 9798350327328. DOI: 10.1109/C21659362.2023.10430601.
- [57] Q. Deng and P. Li. “Explainable intrusion detection for IoT networks based on deep learning”. In: IEEE Internet of Things Journal 9.6 (2021), pp. 4835–4847.
- [60] N.MoustafaandJ. Slay. “Unsw-nb15: a comprehensive data set for net work intrusion detection systems (unsw-nb15 network data set)”. In: 2015 Military Communications and Information Systems Conference (Mil CIS). 2015, pp. 1–6. [61] N. Moustafa. “A new distributed architecture for evaluating ai-based security systemsattheedge:networkton_iotdatasets”.In:Sustain.Cities Soc. 72 (2021), p. 102994.

Thank You