ExpLEA-AIner: Proposition and Development of the Model-Driven Approach to Incorporating Explainable AI in Network Intrusion Detection Systems for Law Enforcement Agencies

Marek Pawlicki[1,2], Aleksandra Pawlicka[1,3], Sebastian Szelest[1], Mikołaj Komisarek[1], Rafał Kozik[1,2], and Michał Choraś[1,2]

[1]**ITTI, Poznań, Poland**
[2]**Bydgoszcz University of Science and Technology, Bydgoszcz, Poland**
[3]**University of Warsaw, Warsaw, Poland**

ABSTRACT

This paper explores the design and implementation of advanced information systems in Law Enforcement Agencies, focusing on how these systems harness artificial intelligence for actionable insights and enhanced operational efficiency, addressing the critical challenges of managing and interpreting vast data sets within Law Enforcement Agencies, where the role of sophisticated information systems has become increasingly central. This paper presents the design, development, and deployment of an advanced Network Intrusion Detection System powered by Machine Learning, integrated with Cyber Threat Intelligence and explainable AI capabilities. The presented system exemplifies a real-world application of Data Science and Machine Learning methodologies, crafted to enhance the operational effectiveness of Law Enforcement. It utilizes a model-driven architecture to process and analyze data from network traffic, effectively identifying and responding to cyberthreats in both real-time and in forensic mode. The development of this system embodies the commitment to pushing the envelope in information system innovation, focusing on practical deployment in high-stakes environments. Through the integration of xAI, transparency and interoperability are provided, supporting the trustworthiness and accountability of Law Enforcement operations. This paper not only explores the technological aspects of the presented system but also highlights its implications for security, operational efficiency, and ethical AI use within Law Enforcement contexts.

Keywords: Network Intrusion Detection Systems; explainable AI; Law Enforcement Agencies; xAI; Cybersecurity in Law Enforcement

INTRODUCTION

Today, in the era of the omnipresent digitization of data, the officers working at Law Enforcement Agencies (LEAs) respond to major challenges in processing vast amounts of information, including crime statistics, information on human resources, as well as call detail records, number plate records, surveillance data, and so on. Not only do they have to uncover the possible patterns and relations emerging from the information but also have to ensure interoperability across a multitude of systems and data formats. If the officers are not able to use the data to their full potential, it directly influences the quality of LEAs decision-making, as well as their capability to combat criminal activities, adapt to evolving threats and challenges, forecast future crimes; eventually, the inability to utilise all the potentially relevant data may affect public safety (3; 27; SAS).

In order to address this concern, artificial intelligence (AI) and machine learning (ML)-powered tools have been successfully employed in the tasks of Law Enforcement, for over a decade now (1), worldwide (20). These advanced technologies give the officers the capabilities to leverage even the largest datasets, actually helping save time and taxpayers' money (1; 6; 20). As Quest et al. (2018) see it, automation of some processes is of great value, for solving the "serious crimes", but foremost for the "mundane crimes" – which normally require repetitive and tedious actions and hundreds of man-hours used. AI has the potential to turn this time into mere minutes (12; 24).

One of the crucial systems leveraging AI in the context of Law Enforcement and data processing are

Network Intrusion Detection Systems (NIDS), which monitor network traffic in real-time to detect and respond to cyberthreats, or allow to investigate previously gathered network traffic in search of nefarious activity. Various network actors have numerous motivations to attack high-stake systems like the ones used by LEAs (21). By integrating ML algorithms, these systems offer an approach to detect anomalies and potential cybercriminal activities effectively. This capability is crucial, not only in safeguarding the data integrity of LEAs but also in protecting the critical infrastructure on which these Agencies rely.

Yet, no matter if the advanced AI algorithms are used to track vehicles in live footage, recognize faces, sift through massive datasets, analyze speech patterns during an interrogation or monitor network traffic for security threats, especially when very sensitive data is at stake, their black-box nature, i.e., the opacity of the AI models becomes a pressing issue (5). In cases when the decision-making process of the AI algorithms is not understandable for human operators, concerns are raised regarding accountability for the algorithmic results. The possible ethical considerations stemming from this problem may go as far as worrying about basic human rights being broken (11; 14). On top of that, the potential challenges resulting from the lack of transparency may lead to an array of unfavourable consequences, such as apprehension of the officers, public trust in AI being undermined, technological progress being slowed down, and the development of new tools being hampered. Ultimately, the possible issues resulting from the lack of transparency may eclipse the possible benefits that AI/ML models may bring to the operations of Law Enforcement.

Consequently, as a solution to this concern, experts have pointed out the need for the AI/ML tools used by LEAs to be made explainable and transparent (6; 11; 14). In this context, explainability relates to uncovering and presenting the reasoning behind algorithmic decisions in such a way that enables humans to understand and interpret them (13). AI explainability (xAI) is not a new concept; it has been deemed "as old as AI itself" (16). Yet, with AI booming, explainability has been in demand like never before; it has been anticipated it will remain a hot and relevant topic for years to come (9; 23).

With all this in mind, in this paper, the ongoing development of robust information systems designed specifically for LEAs, which integrate ML and xAI to process and analyze vast datasets, transforming them into operational insights are discussed. Specifically, a NIDS is proposed which has explainable AI features integrated directly into the system, by means of ExpLEA-AIner, a solution for ensuring explainability, interpretability, and transparency of AI/ML-based models. This integration is critical as it not only enhances the trust in automated detection systems by making their decisions more interpretable to human operators but also facilitates the broader adaptation of NIDS within Law Enforcement Agencies. The software has been designed with ease of use as a principle so that any user could benefit from all the options it offers.

To date, there have been many explainability methods proposed. Yet, they rarely produce the same results; the need to measure the quality of explanations has been the topic of a heated scientific debate for a couple of years already. As of today, the consensus reached is that it is the user who ultimately decides on the usefulness and applicability of explanations (22). In response, the authors have incorporated a variety of explainability methods into the proposed software and NIDS. This inclusive approach allows users not only to understand the underlying decisions made by the AI but also to select the explanation method that best suits their operational context and personal comprehension needs.

The structure of this paper is as follows. Section 2, the Methodology section, is divided into two parts, describing the software architecture of the network intrusion detection system and the explainability component. This is followed by an exploration of the software functionalities in Section 3, detailed in separate sections for the network intrusion detection system and the explainability component. The Impact Section 4 evaluates the benefits and potential challenges introduced by these systems within Law Enforcement. Finally, the Conclusions summarize the findings and implications of the study.

EXPLAINABLE INTRUSION DETECTION SYSTEM FOR LAW ENFORCEMENT

Architecture of the proposed information system

Architecture of the network intrusion detection system

This section outlines the design and development of a NIDS augmented with Cyber Threat Intelligence and xAI systems, detailing its integration with LEA information systems and its role in enhancing real-time data processing and threat detection capabilities. The proposed NIDS architecture consists of several key components. Network traffic is captured using network probes from devices under surveillance, which is then transmitted to a centralized collector and forwarded to an Apache Kafka message bus. This data serves as input for the AI engine, which examines the traffic for indicators of cyberattacks.

In Fig. 1, the high-level architecture of the component has been presented.

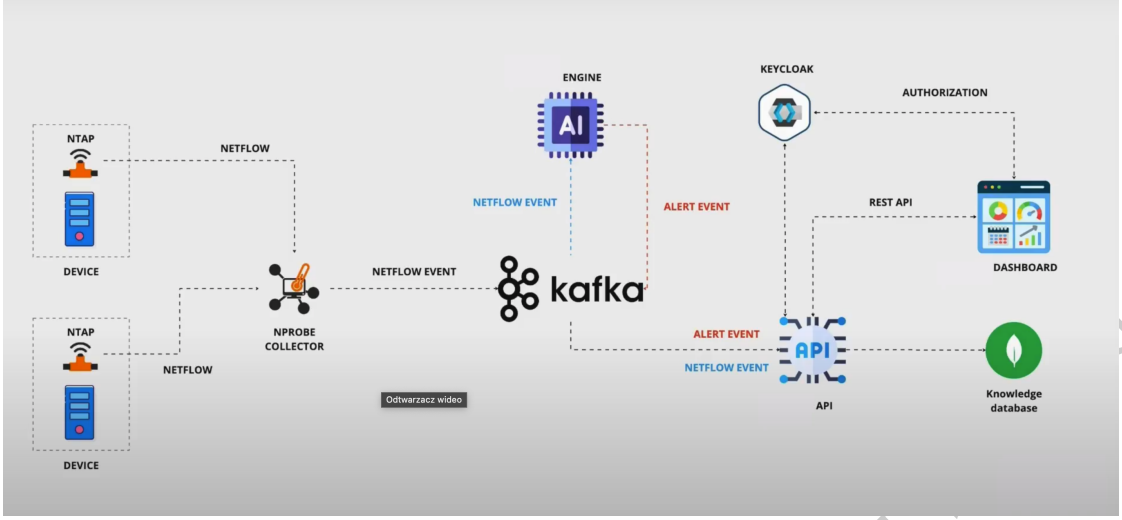

Figure 1. The architecture of the solution presented.

Architecture of the innovative explainability component

The explainability tool presented provides the end user with an informative dashboard which allows them to pick the samples they want to evaluate, choose the AI/ML model the reasoning of which they want to understand, and then select one of the many explainability methods to use. The component then displays the results of the xAI algorithms directly to the user.

In Fig. 2, the high-level architecture of the xAI component has been presented. It starts with the user, who, after authentication can see the dashboard and issue requests. The tool uses a publish-subscribe bus to move samples around the different microservices.

The tool uses AI model integration interfaces, which handle communication with different production-grade AI models. Following this, there are different data preprocessing services, which handle the transformation of data to be fit to use in different algorithms.

The software is packaged with a demonstrator model coming from the domain of network intrusion detection. Having picked the demo model and one of the example datapoints, the user can click 'Explain', and the sample is forwarded to the xAI component, where the particular explanation method can be chosen from a range of listed options.

The proposed NIDS has been designed to optimize real-time threat detection and analysis within stringent time constraints. Utilizing a model-driven architecture that incorporates advanced ML algorithms, the system swiftly processes and analyzes incoming network traffic data. This setup ensures that potential cyber threats are identified and evaluated instantaneously, facilitating immediate response. Key to this capability is the integration of an Apache Kafka message bus, which efficiently manages high-throughput data streams, enabling the system to handle vast amounts of information swiftly and reliably.

Software functionalities — network intrusion detection system

The network intrusion detection system processes and visualizes network traffic in real time via an interactive dashboard, enhancing situational awareness for network operators. In the event of a detected cyber threat, the system suggests appropriate countermeasures, drawing on intelligence from cyber threat platforms and the MITRE ATT&CK database. The dashboard displays various analytical views, such as time-series traffic analysis, volume per source IP, distribution of traffic by application layer, and protocol usage, which are crucial for a comprehensive security posture assessment. The dashboard overview has been presented in Fig. 3. Upon detection of an anomaly, such as a Denial-of-Service (DoS) attack, the system categorizes and displays the threat based on severity, encoded with colour indicators. The alert remains in a pending state until reviewed by an operator (as shown in Fig.4). Detailed information about each detected threat, including the source, target, data volume, and protocol involved, is systematically presented, allowing for an informed response strategy. The screen showing the details of the attack has been shown in Fig.5. The presented network intrusion detection system introduces novel aspects that differentiate it from existing solutions. Primarily, the integration of real-time, machine learning algorithms enables the system to dynamically adapt to new threats, enhancing its ability to detect and respond to anomalies with minimal human intervention. Additionally, the use of the MITRE ATT&CK framework not only enriches the threat detection capabilities but also facilitates a comprehensive understanding of attack vectors, which is often lacking in conventional systems. This approach allows for preemptive action

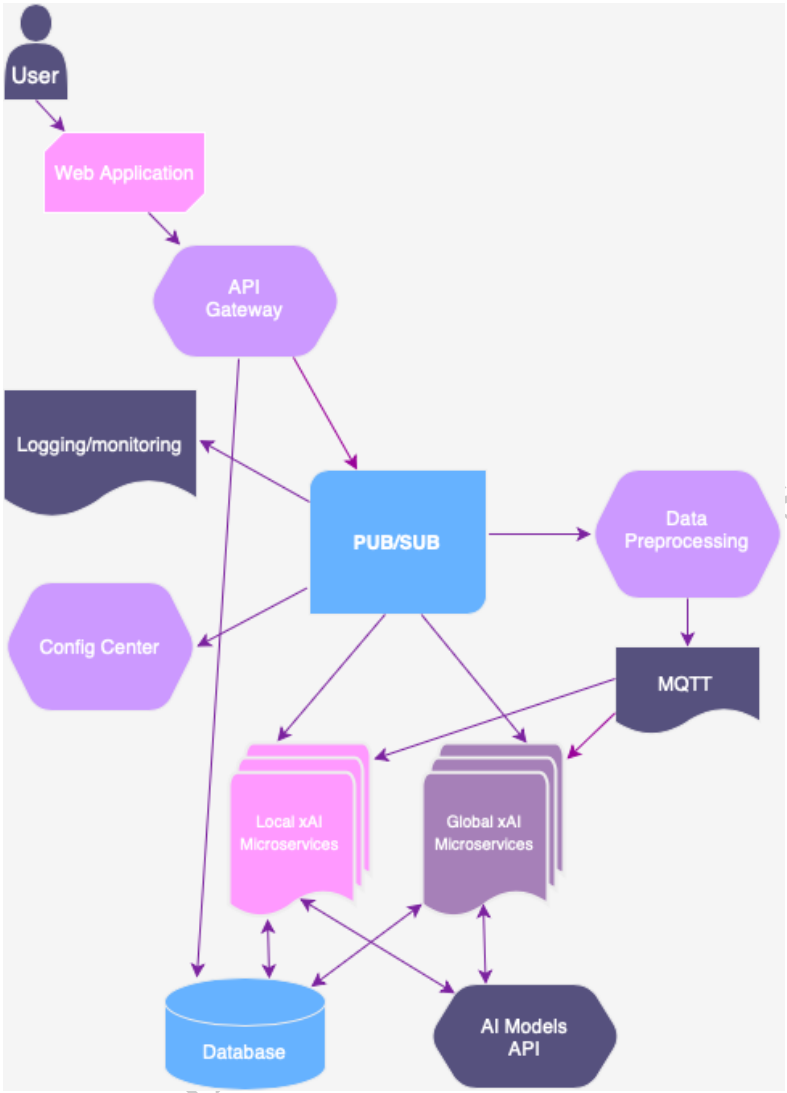

Figure 2. The architecture of the solution presented.

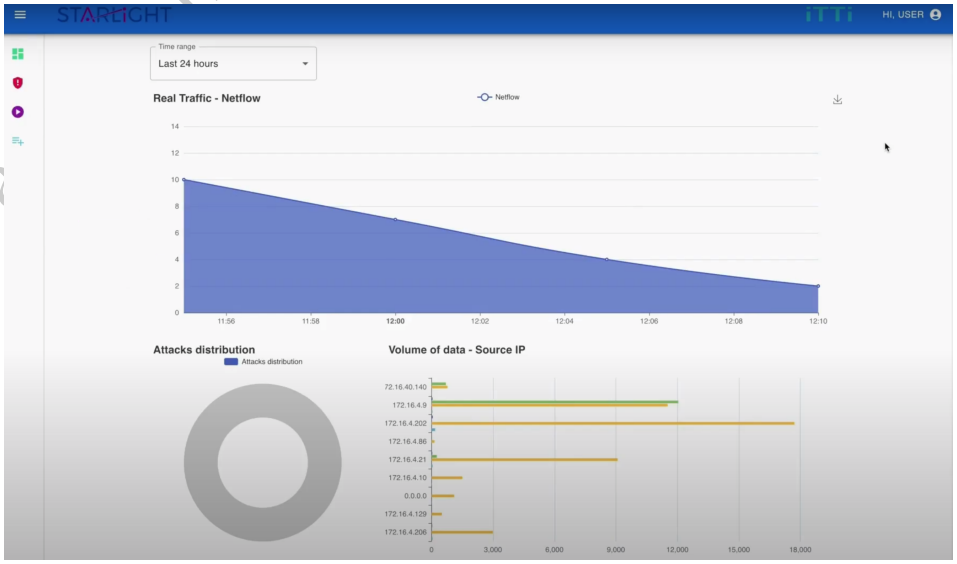

Figure 3. The dashboard of the NIDS solution.

based on predictive analytics and community-shared intelligence, significantly reducing response times and increasing the accuracy of threat classification.

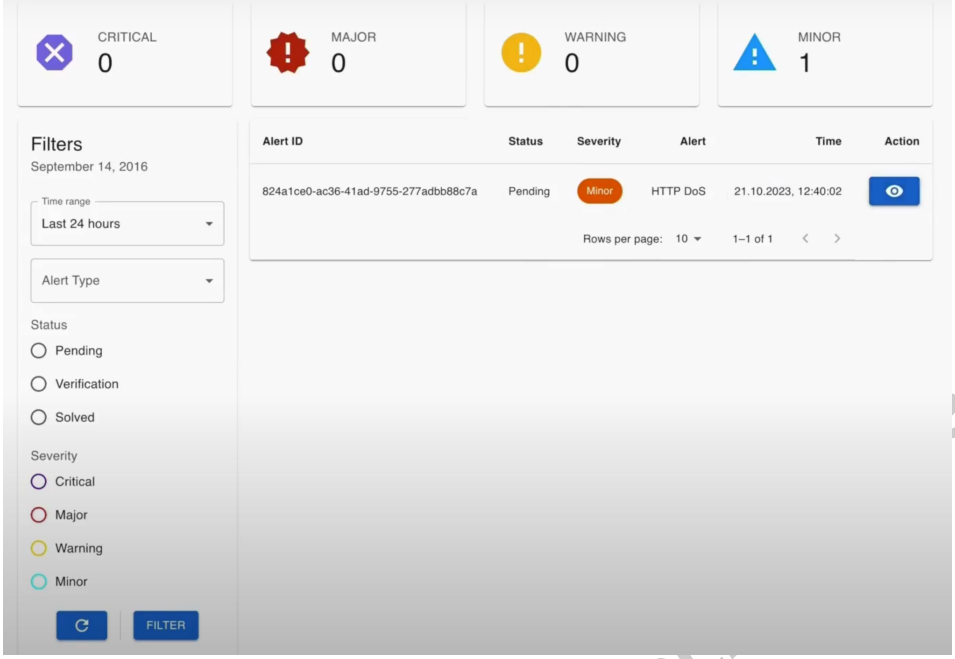

Figure 4. The attack listed in the Alerts view.

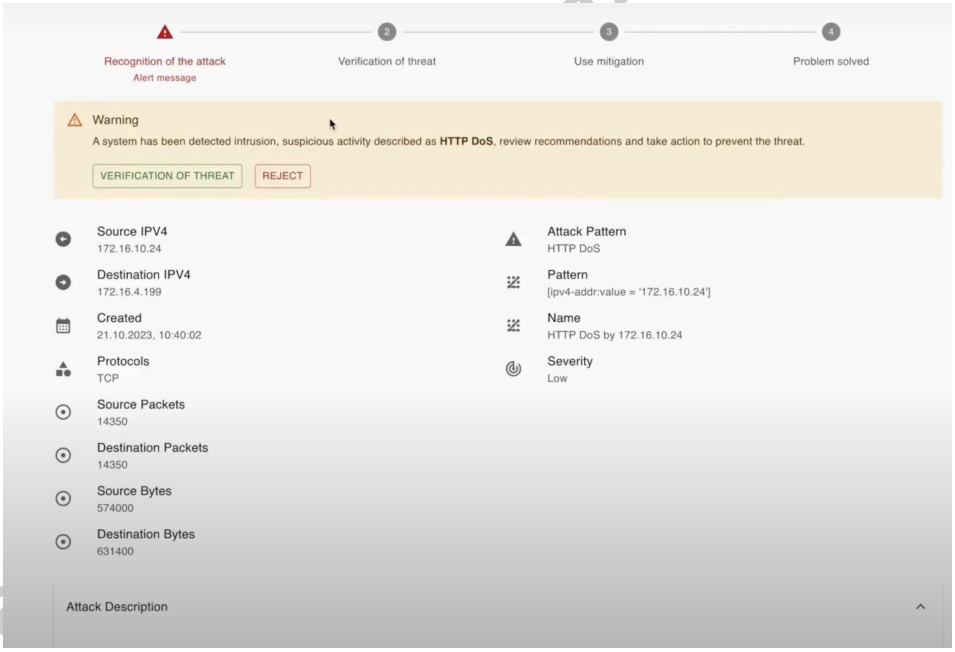

Figure 5. The details of the attack.

The dashboard design is another innovative aspect, presenting a comprehensive view of network traffic and threat intelligence in an intuitive manner. This not only aids in quick threat assessment but also supports detailed forensic analysis, allowing operators to make informed decisions swiftly. The structured alert handling process further ensures that each threat is addressed from initial detection through to resolution, guided by evidence-based strategies. Such end-to-end integration of advanced technologies and methodologies establishes this system as a significant advancement in the field of cybersecurity.

Software functionalities — explainability component

As mentioned in the Introduction, the ultimate choice of the explanation method depends on the operational context and the needs of the user. Thus, it has been decided for the explainability component described to employ a number of various acknowledged explainability techniques. These methods have been briefly presented in this section in the alphabetical order.

Accumulated Local Effects (ALE)

Type of explainer: global, model-agnostic.
This explainer assesses the effects of individual features on model predictions across the data distribution, accumulating local effects to offer a global view of feature influence (2).

Scoped Rules (ANCHORS)

Type of explainer: local, model-agnostic.
The explainer identifies 'anchors', i.e., explanations in the form of rules, which are specific conditions that, when satisfied, predict the same outcome with high probability, and offers insights into what factors are most influential in a model's decision-making process (26).

Fig. 6 shows the explanations provided by means of the ANCHORS algorithm.

Figure 6. The explanations provided by ANCHORS.

Diverse Counterfactual Explanations (DICE)

Type of explainer: local, model-agnostic.
This explainer generates counterfactual explanations, offering alternative scenarios where slightly different input values would lead to a different prediction (19).

Explanation based on Decision Trees

Type of explainer: both local and global, model-dependent.
This explainer uses decision trees to simplify the model's decision-making process into an interpretable form, illustrating how various input features lead to different outcomes (25; 29).

Individual Conditional Expectation (ICE)

Type of explainer: local, model-agnostic.
The explainer visualizes the relationship between a feature and the prediction outcome for individual instances, by plotting how predictions change as a feature varies while other features are held constant (10).

Local Interpretable Model-Agnostic Explanations (LIME)

Type of explainer: local, model-agnostic.
This explainer generates explanations for individual predictions by approximating the underlying model with an interpretable one, such as a linear model or decision tree, based on perturbations of the input data. This approach helps to find out which features significantly influence the output of complex models, making it easier to understand why certain decisions or predictions are made (13; 18).

In Fig. 7, the explanation results given when choosing the LIME method have been presented.

Partial Dependence Plot (PDP)

Type of explainer: global, model-agnostic.
It provides insights into the effect of one or two features on the predicted outcome across the entire dataset, illustrating the average effect of these features, it isolates the relationship between the features and the outcome while averaging out the effects of all other features (7).

Permutation Feature Importance (PFI)

Type of explainer: global, model-agnostic.
This explainer assesses the impact of shuffling each feature on the accuracy of the model to determine its importance. By randomly permuting the values of each feature and observing the resulting decrease in model performance, PFI quantifies the significance of each feature in the model's predictions (4).

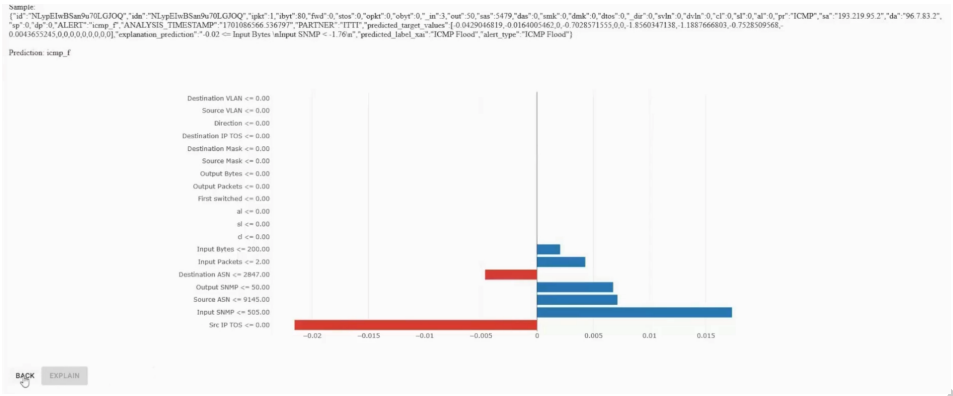

Figure 7. The explanations provided by LIME.

RuleFit method

Type of explainer: global, model-dependent.

This method employs decision rules generated from decision trees along with original features to construct a linear model that predicts the outcome. This approach combines the interpretability of rules with the predictive power of linear models, revealing the influence of individual features and rule conditions on the overall prediction (8).

Shapley Additive Explanations (SHAP)

Type of explainer: local and global, model-agnostic.

This method utilizes Shapley values from cooperative game theory to attribute the contribution of each feature in a prediction, providing a detailed and fair explanation of the model's output. This method allows for both individual explanations (local) and overall model behaviour insights (global), highlighting how different features impact the model's decision-making process (13; 15; 17).

Lastly, in Fig. 8, the way of presenting the explanations given by SHAP was illustrated.

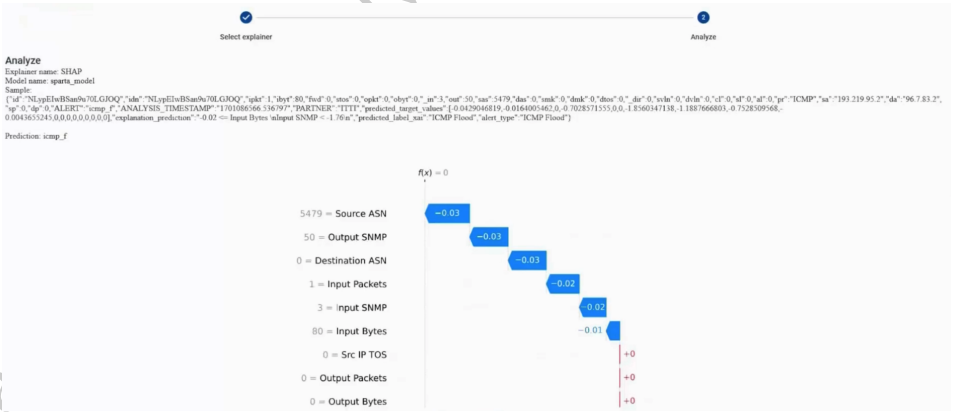

Figure 8. The explanations provided by SHAP.

IMPACT FOR LAW ENFORCEMENT AGENCIES AND SOCIETY

This section evaluates the significant impact of deploying these information systems in LEAs, emphasizing how their development addresses specific operational challenges faced by law enforcement officials, leading to more informed decision-making and efficient resource allocation. The deployment of this advanced NIDS significantly enhances the capabilities of LEAs in their cybersecurity efforts. By providing real-time detection and analysis of potential cyber threats, the system empowers Agencies to preemptively identify and mitigate attacks that could compromise critical infrastructure or sensitive data.

Furthermore, the integration of the MITRE ATT&CK framework offers Law Enforcement a standardized, actionable intelligence format, facilitating more effective coordination and communication across different jurisdictions and units. This is especially crucial in combating sophisticated cybercrime networks that operate across borders. The system's comprehensive logging and reporting features also aid in forensic investigations, enabling Agencies to trace the source of attacks, gather evidence, and prosecute offenders

with higher accuracy and efficiency. Consequently, the NIDS not only strengthens the cybersecurity posture of Law Enforcement Agencies but also enhances their investigative processes, ultimately contributing to more robust national and international cyber defense strategies.

Recognizing the challenges LEAs face when applying AI/ML tools, there is a growing consensus among experts regarding the importance of the methods inherently explainable and transparent. Achieving explainability involves uncovering and presenting the reasoning behind algorithmic decisions in a manner that humans can understand and interpret.

In response to these concerns, ExpLEA-AIner, the presented solution for ensuring explainability, interpretability, and transparency of AI/ML-based models, has been developed. ExpLEA-AIner enables Law Enforcement professionals to benefit from a range of explainability methods. This approach empowers users to evaluate and compare different explanations, ultimately allowing them to make informed decisions based on the most suitable interpretation method for their specific needs. The major benefit of the module is that through the suite of integration microservices, preprocessing microservices and explanation microservices it can offer xAI to the end-user seamlessly and in near-real-time. The user can choose a datapoint present on a Kafka topic, or, with further integration, have a button allowing them to explain the decision of an AI model with state-of-the-art methods present directly in the dashboard of their AI-based system of choice.

While AI and ML technologies offer tremendous potential for enhancing Law Enforcement capabilities, addressing the challenges of transparency and explainability is essential to ensure accountability, maintain public trust, and maximize the benefits of these innovative tools in promoting public safety. Through initiatives like the proposed ExpLEA-AIner, the path toward achieving transparency and accountability in AI-driven Law Enforcement is within reach, contributing to the more responsible and effective use of these technologies in the future.

The development and implementation of the ExpLEA-AIner solution introduces several avenues for new research, particularly in the field of xAI within the context of Law Enforcement. This software opens the door to exploring how different explainability methods can be optimized for various types of data and scenarios encountered by Law Enforcement Agencies. Research can be directed towards understanding the impact of explainable AI on decision-making processes in high-stakes environments and how these insights can further refine AI/ML models for better accuracy and transparency.

ExpLEA-AIner significantly enhances the pursuit of existing research questions by providing a practical framework to assess and compare the decisions provided by AI systems through the lenses of different xAI approaches. It allows for empirical studies on the effectiveness of explainability in improving user trust and comprehension of AI-supported decisions, thereby contributing to the broader discourse on ethical use of algorithms in sensitive sectors.

In terms of changing daily practice for its users, ExpLEA-AIner democratizes access to complex AI/ML-based insights, enabling officers with varying degrees of technical expertise to understand and leverage AI tools confidently. This shift not only improves operational efficiency but also fosters a culture of transparency and accountability in the use of technology within Law Enforcement. As Law Enforcement Agencies increasingly turn to AI/ML solutions to enhance their capabilities, tools like ExpLEA-AIner will become indispensable for ensuring these technologies are used responsibly and effectively.

Lastly, the ExpLEA-AIner's approach to AI explainability has the commercial potential to influence the development of new products and services within the tech industry, particularly in areas requiring transparent AI solutions. By demonstrating the feasibility and value of explainable AI, it will encourage the emergence of spin-off companies focused on creating accessible and understandable AI technologies for various sectors beyond Law Enforcement, including healthcare, finance, and cybersecurity.

Ethical and Privacy Concerns

The deployment of AI in Law Enforcement is bound to raise significant ethical and privacy concerns. There is the common worry that the use of AI for surveillance and data analysis can lead to unintended biases and invasion of privacy if not properly managed. To mitigate these issues, the approach presented incorporates strict data handling protocols and algorithmic transparency; i.e., the explainable AI component ensures that each decision made by the system can be audited and understood by human operators, promoting accountability. Once deployed in a real-world scenario, regular ethical reviews and adherence to legal standards concerning data privacy will be integral to the deployment strategy, ensuring that the system upholds the rights and freedoms of individuals while enhancing public safety.

Future work

Future enhancements for the proposed Network Intrusion Detection System focus on expanding its capabilities to better handle encrypted traffic and zero-day attacks. Plans include integrating advanced adaptive algorithms that can learn from emerging threats in real-time. Moreover, expanding the explainability

component to include more diverse AI methodologies as well as xAI evaluation metrics will enhance the system's transparency and utility across different operational scenarios. It would also be very beneficial to collaborate with academia and industry in order to explore new frontiers in cybersecurity and AI-driven Law Enforcement technologies.

CONCLUSIONS

The integration of the innovative Network Intrusion Detection System presented in this paper enables real-time threat detection and analysis, significantly boosting the ability of Law Enforcement to respond to cyber threats effectively.

However, to ensure accountability and maintain public trust, addressing the challenges of transparency and explainability in these AI-driven systems is crucial. Initiatives such as ExpLEA-AIner guide the way towards achieving transparency and accountability in AI-driven Law Enforcement tools. By promoting the explainability of AI actions and decisions, these initiatives contribute to a more responsible and effective use of AI technologies, ultimately enhancing public safety while adhering to ethical standards. This balanced approach ensures that the benefits of these innovative tools are maximized, fostering a safer and more secure environment.

This innovative approach and the tool have already been showcased to end-users at the hands-on sessions organized by the project, and will be further adjusted to their specific needs.

ACKNOWLEDGEMENTS

The work described in this paper is performed in the H2020 project STARLIGHT ("Sustainable Autonomy and Resilience for LEAs using AI against High priority Threats"). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101021797.

REFERENCES

[1] Alzou'bi, S., Alshibly, H., and Al-ma'aitah, M. (2014). Artificial Intelligence in Law Enforcement, A Review. *International Journal of Advanced Information Technology (IJAIT)*, 4(4).

[2] Apley, D. W. and Zhu, J. (2016). Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models.

[3] Boehmer, R. (2017). 5 Ways to Improve the Use of Data by Police Departments. *Hillard Heintze*.

[4] Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.

[5] Choraś, M., Pawlicki, M., Puchalski, D., and Kozik, R. (2020). Machine learning–the results are not the only thing that matters! what about security, explainability and fairness? In *Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part IV 20*, pages 615–628. Springer.

[6] Dees, T. (2019). 3 ways artificial intelligence can work for your agency. *Police 1*.

[7] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5).

[8] Friedman, J. H. and Popescu, B. E. (2008). Predictive learning via rule ensembles.

[9] Gill, S. S., Xu, M., Ottaviani, C., Patros, P., Bahsoon, R., Shaghaghi, A., Golec, M., Stankovski, V., Wu, H., Abraham, A., Singh, M., Mehta, H., Ghosh, S. K., Baker, T., Parlikad, A. K., Lutfiyya, H., Kanhere, S. S., Sakellariou, R., Dustdar, S., Rana, O., Brandic, I., and Uhlig, S. (2022). AI for next generation computing: Emerging trends and future directions. *Internet of Things*, 19:100514.

[10] Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E. (2013). Peeking Inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation.

[11] Hisham, S. (2019). AI will be used for all forms of policing in future. *Geospatial World*.

[12] Kozik, R., Choras, M., Pawlicki, M., Hołubowicz, W., Pallmer, D., Mueller, W., Behmer, E.-J., Loumiotis, I., Demestichas, K., Horincar, R., Laudy, C., and Faure, D. (2019). The Identification and Creation of Ontologies for the Use in Law Enforcement AI Solutions – MAGNETO Platform Use Case. pages 335–345.

[13] Kurek, W., Pawlicki, M., Pawlicka, A., Kozik, R., and Choraś, M. (2023). Explainable artificial intelligence 101: Techniques, applications and challenges. In *International Conference on Intelligent Computing*, pages 310–318.

[14] Lalley, A. Z. (2019). *INTRODUCING ARTIFICIAL INTELLIGENCE INTO THE UNITED STATES LAW ENFORCEMENT COMMUNITY: LEARNING FROM FOREIGN LAW ENFORCEMENT AGENCIES*. PhD thesis.

[15] Lundberg, S. M. and Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

[16] Meske, C., Bunde, E., Schneider, J., and Gersch, M. (2022). Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. *Information Systems Management*, 39(1):53–63.

[17] Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.

[18] Molnar, C. (2022). *Interpretable Machine Learning (Second Edition) A Guide for Making Black Box Models Explainable*. Leanpub.

[19] Mothilal, R. K., Sharma, A., and Tan, C. (2020). Explaining machine learning classifiers through diverse counterfactual explanations. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 607–617, New York, NY, USA. ACM.

[20] OSCE (2019). Law enforcement agencies should embrace Artificial Intelligence to enhance their efficiency and effectiveness, say police experts at OSCE meeting.

[21] Pawlicka, A., Choraś, M., and Pawlicki, M. (2021). The stray sheep of cyberspace aka the actors who claim they break the law for the greater good. *Personal and Ubiquitous Computing*, 25(5):843–852.

[22] Pawlicka, A., Pawlicki, M., Kozik, R., Kurek, W., and Choraś, M. (2024). How Explainable Is Explainability? Towards Better Metrics for Explainable AI. pages 685–695.

[23] Pawlicki, M., Pawlicka, A., Kozik, R., and Choraś, M. (2024). Advanced insights through systematic analysis: Mapping future research directions and opportunities for xAI in deep learning and artificial intelligence used in cybersecurity. *Neurocomputing*, page 127759.

[24] Quest, L., Charrie, A., and Roy, S. (2018). THE RISKS AND BENEFITS OF USING AI TO DETECT CRIME. *Harvard Business Review*.

[25] Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1):81–106.

[26] Ribeiro, M., Sing, S., and Guestrin, C. (2018). Anchors: High-Precision Model-Agnostic Explanations. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAAI-18)*, New Orleans, Louisiana.

[27] Santos, A., Jenkins, I., Mariani, J., Gelles, M., and Mirkow, A. (2019). Investigative analytics Leveraging data for law enforcement insights. *Deloitte Insights*.

[SAS] SAS. How big data analytics can be the difference for law enforcement.

[29] Szczepański, M., Choraś, M., Pawlicki, M., and Kozik, R. (2020). Achieving explainability of intrusion detection system by hybrid oracle-explainer approach. In *2020 International Joint Conference on neural networks (IJCNN)*, pages 1–8. IEEE.