

CHAPTER 5 – CONCLUSION AND RECOMENDATION

5.1. Conclusion

This research thesis has successfully demonstrated that supervised machine learning can be a highly effective tool for predicting Indonesian ISP compliance with negative content filtering mandates. By analyzing key operational metrics, a Random Forest model was developed that can predict an ISP's compliance capability with 97% accuracy. The analysis identified that high DNS query load (QPS), interpreted as a proxy for architectural simplicity, is the single most powerful predictor of non-compliance. This study provides a clear, data-driven pathway for Kominfo to enhance its regulatory oversight and for ISPs to identify and remedy critical deficiencies in their compliance infrastructure.

It is important to acknowledge the limitations of this study. The analysis is based on a static, point-in-time dataset from April 2025. Furthermore, the model predicts an ISP's capability to comply based on its infrastructure, not its real-time behavior in adhering to the dynamic blacklist.

5.2. Rekomendasi

The successful development of a high-performance predictive model is not the end goal of this research, but rather a means to an end. The ultimate value of this analysis lies in its ability to be translated into actionable intelligence that can tangibly improve the state of regulatory compliance within the Indonesian telecommunications landscape. This final section leverages the model's predictive power and the insights gleaned from the feature analysis to provide concrete, data-driven recommendations for both the regulatory body, Kominfo, and the Internet Service Providers themselves. It concludes by summarizing the study's contributions and outlining potential directions for future research.

To demonstrate this capability, the best-performing model from this thesis can be applied to the entire dataset of ISPs. Of particular interest are the ISPs that the model flags as high-risk despite them not meeting the strict definition of non-compliance used

for training. These organizations represent a hidden risk tier that traditional, threshold-based monitoring would miss.

By applying the trained Random Forest model to the full dataset, a list of ISPs predicted to be 'non-compliant' can be generated. This list would include the 130 ISPs with zero RTBH activity, which the model would identify with high accuracy. More importantly, it would also flag a cohort of ISPs that are technically 'compliant' (possess some RTBH) but whose operational metrics (high QPS, low TNG) strongly suggest they are either not using their capabilities effectively or are otherwise at high risk of failure. These are the prime candidates for proactive engagement, targeted audits, and technical assistance from the regulator.

5.2.1. Actionable Recommendations For Regulators (Kominfo)

Adopt Predictive, Risk-Based Oversight Kominfo should integrate this or a similar predictive modeling framework into its standard oversight and monitoring activities. By analyzing regularly submitted operational data, the model can generate a dynamic risk score for every ISP in the country. This would allow Kominfo to allocate its limited audit and enforcement resources more efficiently, focusing on the small subset of ISPs identified as high-risk rather than conducting broad, untargeted spot-checks. This aligns with modern best practices in predictive compliance monitoring.

The this thesis has shown that metrics beyond simple blocking capability are highly predictive. Kominfo should consider refining its data collection requirements. Given the high predictive power of the QPS metric, Kominfo could work towards developing and mandating a standardized set of metrics for DNS filtering performance and reliability. This would provide a more nuanced and accurate picture of an ISP's compliance sophistication than binary blocking checks alone.

5.2.2. Actionable Recommendations For Non-Compliant and At-Risk ISPs

Prioritize Investment in Foundational Technologies for the 130 ISPs identified as non-compliant (having zero RTBH activity), the recommendation is immediate investment

in the implementation of RTBH filtering is the baseline requirement for achieving compliance capability. This is the most fundamental tool for executing network-level blocks and its absence represents a critical infrastructure gap.

Conduct an Architectural Review for High QPS Indicators for ISPs flagged by the model due to a high QPS profile, the recommendation is not to artificially lower the metric. Instead, this should serve as a trigger for an internal network architecture review. These ISPs should investigate their DNS infrastructure for points of over-centralization and bottlenecks. Investing in a more distributed, resilient, and scalable DNS architecture will not only improve their compliance risk profile but will also lead to enhanced service quality and reliability for their end-users.

Enhance DNS Filtering Efficiency for ISPs with low TNG scores, the focus should be on improving the robustness of their DNS-layer filtering systems. This involves ensuring their RPZ implementation (or equivalent technology) is correctly configured, its blacklist is kept up to date with Kominfo directives, and it has sufficient capacity to handle the ISP's query load without performance degradation.

5.3. Future Work

These limitations point toward several promising avenues for future research. Near future to use latest data and more ISPs up to 80% of the subset to ensure compliance level achieved Komdigi's objectives. Incorporating a wider range of features such as ISP size (number of subscribers), years in operation, financial data, and geographic service area could further enhance the model's predictive power. Exploring more complex algorithms, such as Gradient Boosting Machines or Neural Networks, may yield incremental gains in accuracy. Finally, future work could aim to develop a more nuanced, multi-level compliance score, moving beyond the binary classification to provide a richer assessment of an ISP's regulatory maturity. By building on this foundation, data science can continue to offer powerful solutions for creating a safer and more compliant digital ecosystem in Indonesia.

REFERENCES

- Altaheri, F., Shaalan, K., 2020. Exploring Machine Learning Models to Predict Harmonized System Code, in: Themistocleous, M., Papadaki, M. (Eds.), Information Systems, Lecture Notes in Business Information Processing. Springer International Publishing, Cham, pp. 291–303. https://doi.org/10.1007/978-3-030-44322-1_22
- Amalia, A., Gunawan, D., Lydia, M.S., Wesley, 2019. The Identification of Negative Content in Websites by Using Machine Learning Approaches, in: 2019 5th International Conference on Computing Engineering and Design (ICCED). Presented at the 2019 5th International Conference on Computing Engineering and Design (ICCED), IEEE, Singapore, Singapore, pp. 1–6. <https://doi.org/10.1109/ICCED46541.2019.9161105>
- APJII, 2024. Survei Penetrasi Internet Indonesia Q1 2023-2024.
- Blanquero, R., Carrizosa, E., Ramírez-Cobo, P., Sillero-Denamiel, M.R., 2021. Variable selection for Naïve Bayes classification. *Comput. Oper. Res.* 135, 105456. <https://doi.org/10.1016/j.cor.2021.105456>
- Bukhori, H.A., Widyatmoko, T., Sunarti, S., Wahyono, I.D., Sakti, K.F.L., Pranajaya, B., Ul Muyassaroh, L., Saryono, D., Tiwari, S., 2023. Modification Metric of Class Document on Naïve Bayes for Sentiment Analysis of Online Learning Evaluation, in: 2023 1st International Conference on Advanced Engineering and Technologies (ICONNIC). Presented at the 2023 1st International Conference on Advanced Engineering and Technologies (ICONNIC), IEEE, Kediri, Indonesia, pp. 304–309. <https://doi.org/10.1109/ICONNIC59854.2023.10467513>
- Direktorat Jenderal Aplikasi Informatika, 2021. Rencana Strategis Direktorat Jenderal Aplikasi Informatika 2020 - 2024.
- Direktorat Pengendalian PPI, D.P.A., 2024. Surat Edaran Uji Coba Pemblokiran Konten Negatif melalui Metode RTBH dan DNS Trust-NG.
- Durga, S., Nag, R., Daniel, E., 2019. Survey on Machine Learning and Deep Learning Algorithms used in Internet of Things (IoT) Healthcare, in: 2019 3rd International Conference on Computing Methodologies and Communication

- (ICCMC). Presented at the 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), IEEE, Erode, India, pp. 1018–1022. <https://doi.org/10.1109/ICCMC.2019.8819806>
- Febrianto, T., Soediantono, D., 2022. Enterprise Resource Planning (ERP) and Implementation Suggestion to the Defense Industry: A Literature Review. *J. Ind. Eng.*
- Hidalgo-Fort, E., Gómez-Galán, J.A., González-Carvajal, R., Sánchez-Cárdenas, P., Clemente-Maya, C., 2023. Battery-Less Industrial Wireless Monitoring and Control System for Improved Operational Efficiency. *Sensors* 23, 2517. <https://doi.org/10.3390/s23052517>
- Ichise, H., Jin, Y., Iida, K., 2025. RPZ-Based Suspicious Direct Outbound DNS Traffic Detection Mechanism With Adaptive Policy Updates. *IEEE Access* 13, 30694–30704. <https://doi.org/10.1109/ACCESS.2025.3542234>
- Ichise, H., Jin, Y., Iida, K., 2023. Policy-based Detection and Blocking System against Abnormal Applications by Analyzing DNS Traffic, in: 2023 22nd International Symposium on Communications and Information Technologies (ISCIT). Presented at the 2023 22nd International Symposium on Communications and Information Technologies (ISCIT), IEEE, Sydney, Australia, pp. 1–6. <https://doi.org/10.1109/ISCIT57293.2023.10376042>
- Laayati, O., Bouzi, M., Chebak, A., 2022. Smart Energy Management System: Design of a Monitoring and Peak Load Forecasting System for an Experimental Open-Pit Mine. *Appl. Syst. Innov.* 5, 18. <https://doi.org/10.3390/asi5010018>
- Liaw, A., Wiener, M., 2002. Classification and Regression by RandomForest 2.
- Malathi, K., Ruby, E.D.K., 2024. Evaluating Machine Learning Linear Regression Model for DVT Prediction in Rehabilitation Patients, in: 2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM). Presented at the 2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), IEEE, Chennai, India, pp. 1–5. <https://doi.org/10.1109/ICONSTEM60960.2024.10568850>
- Mallapragada, P.K., Rong Jin, Jain, A.K., Yi Liu, 2009. SemiBoost: Boosting for Semi-Supervised Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 2000–2014. <https://doi.org/10.1109/TPAMI.2008.235>

Mathrani, A., Alipour, M., n.d. WEBSITE BLOCKING ACROSS TEN COUNTRIES: A SNAPSHOT.

Menteri Komunikasi Dan Informatika Republik Indonesia, 2021. Perubahan Atas Peraturan Menteri Komunikasi Dan Informatika Nomor 5 Tahun 2020 Tentang Penyelenggara Sistem Elektronik Lingkup Privat.

Menteri Komunikasi Dan Informatika Republik Indonesia, 2020. Penyelenggara Sistem Elektronik Lingkup Privat.

Moreira, D.C., Fachine, J.M., 2018. A Machine Learning-based Forensic Discriminator of Pornographic and Bikini Images, in: 2018 International Joint Conference on Neural Networks (IJCNN). Presented at the 2018 International Joint Conference on Neural Networks (IJCNN), IEEE, Rio de Janeiro, pp. 1–8. <https://doi.org/10.1109/IJCNN.2018.8489100>

Muslim, H.A., 2022. NAIVE BAYES ALGORITHM IN HS CODE CLASSIFICATION FOR OPTIMIZING CUSTOMS REVENUE AND MITIGATION OF POTENTIAL RESTITUTION. J. Inf. Technol. Its Util. 5, 1–9. <https://doi.org/10.56873/jitu.5.1.4740>

Nayak, A.K., Tiwari, D., Pattanayak, S., 2023. IntelliFiber: Optical Network Optimization through Machine Learning Algorithms, in: 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). Presented at the 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, Delhi, India, pp. 1–9. <https://doi.org/10.1109/ICCCNT56998.2023.10308074>

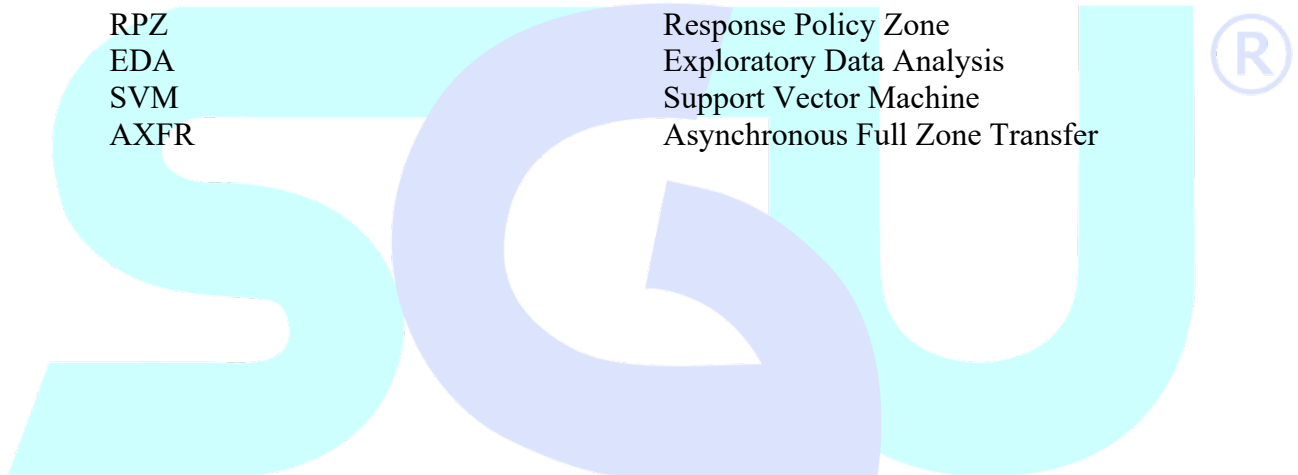
Otsuka, T., Gada, Yamai, N., Okayama, K., Jin, Y., 2015. Design and Implementation of Client IP Notification Feature on DNS for Proactive Firewall System, in: 2015 IEEE 39th Annual Computer Software and Applications Conference. Presented at the 2015 IEEE 39th Annual Computer Software and Applications Conference (COMPSAC), IEEE, Taichung, Taiwan, pp. 127–132. <https://doi.org/10.1109/COMPSAC.2015.220>

Paramartha, I.G.Y., Ardiyanto, I., Hidayat, R., 2021. Developing Machine Learning Framework to Classify Harmonized System Code. Case Study: Indonesian Customs, in: 2021 3rd East Indonesia Conference on Computer and Information Technology (EIconCIT). Presented at the 2021 3rd East Indonesia Conference

- on Computer and Information Technology (EIconCIT), IEEE, Surabaya, Indonesia, pp. 254–259.
<https://doi.org/10.1109/EIconCIT50028.2021.9431888>
- Rahul, Gupta, A., Raj, A., Arora, M., 2021. IP Traffic Classification of 4G Network using Machine Learning Techniques, in: 2021 5th International Conference on Computing Methodologies and Communication (ICCMC). Presented at the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), IEEE, Erode, India, pp. 127–132.
<https://doi.org/10.1109/ICCMC51019.2021.9418397>
- Sekar Asmoro Gati, Khoirunurrofik, 2023. THE EFFECT OF TELECOMMUNICATIONS INFRASTRUCTURE ON INCLUSIVE ECONOMIC GROWTH IN INDONESIA 2011–2021. Int. J. Innov. Technol. Soc. Sci. https://doi.org/10.31435/rsglobal_ijitss/30062023/8007
- Utamima, A., Aditya, F.I., Ashaari, F.W.G., Nugraha, A.K., 2023. Applying Machine Learning Algorithms for Opinion Mining on a Digital Internet Service Provider, in: 2023 11th International Symposium on Digital Forensics and Security (ISDFS). Presented at the 2023 11th International Symposium on Digital Forensics and Security (ISDFS), IEEE, Chattanooga, TN, USA, pp. 1–5.
<https://doi.org/10.1109/ISDFS58141.2023.10131886>
- Varshney, G., Shah, N., 2021. A DNS Security Policy for Timely Detection of Malicious Modification on Webpages, in: 2021 28th International Conference on Telecommunications (ICT). Presented at the 2021 28th International Conference on Telecommunications (ICT), IEEE, London, United Kingdom, pp. 1–5. <https://doi.org/10.1109/ICT52184.2021.9511514>

GLOSSARY

Term	Definition
DNS	Domain Name Server
RTBH	Remotely Triggered Black Hole
QPS	Queries Per Second
TNG	Trustworthiness of Next Generation network
TPS	Transactions Per Second
ISP	Internet Service Provider
NAP	Network Access Providers
TCP RST	Transmission Control Protocol Reset
IP	Internet Protocol
URL	Uniform Resource Locator
RPZ	Response Policy Zone
EDA	Exploratory Data Analysis
SVM	Support Vector Machine
AXFR	Asynchronous Full Zone Transfer



Curriculum Vitae - Author

Name : Dika Aghniyasyak Priyo Faqih
Address : East Jakarta
Education : 2012, Bachelor of Informatics Engineering at STT-PLN Jakarta
2024, Profesional Engineering (Ir.) at Trisakti University, Jakarta

Work Experience :

Present Supervisor Log Analyst Command Center Ministry of Communcation
and Digital Republic of Indonesia
2020-2025 IT Governance, at PT. Mubaligh Teknologi Indonesia
2019-2021 System Administrator, at Diskominfotik Pemprov DKI Jakarta
2018-2019 Senior IT Aplication & Networking, at PT. Sumber Rejeki Abadi
Makmur
2016-2017 Engineer Developer Online Monitoring Electrical at PT. Matriks
Megah Sejahtera
2012-2016 Supervisor IT, at PT. Kawan Lama Group, PT. Ace Hardware
Indonesia, Tbk

