



Estimating Country-Level Cybercrime Legal Framework Upgrades Through Markov State Classification of Regulatory Maturity

R. Elavarasi^{1,*}, P. Gajalakshmi², S. Murugesan³, K. Srinivasan⁴

^{1,3} Electrical and Electronics Engineering, AMET Deemed to be University, Tamil Nadu, India

² Electronics and Communication Engineering, University College of Engineering, Tamil Nadu, India

⁴ Electrical and Electronics Engineering, Madras Engineering College, Tamil Nadu, India

ABSTRACT

Cybercrime poses a growing transnational threat, prompting countries to develop regulatory frameworks aimed at strengthening legal enforcement and institutional capacity. However, substantial disparities persist in the maturity of national cybercrime legal systems. This study proposes a predictive modeling framework to estimate country-level cybercrime regulatory transitions using a combined Markov state-transition and supervised machine learning approach. Regulatory maturity scores for 194 countries across 2020 and 2024 are discretized into three states—Low, Medium, and High—and modeled as multi-class transition outcomes. An empirical Markov transition matrix is first constructed to estimate baseline probabilistic movement between regulatory states. Results indicate strong path dependence, with low-maturity countries exhibiting high persistence and medium-level systems demonstrating greater volatility. High-level regulatory systems show partial stability but also evidence of boundary-sensitive transitions. To enhance predictive capability, a Random Forest classifier is trained using baseline maturity score, income level, and regional classification as structural predictors. Cross-validated performance achieves moderate predictive accuracy, suggesting that regulatory transitions are partially but not fully predictable from structural variables. Feature importance analysis reveals that baseline regulatory maturity is the dominant determinant of subsequent state transitions, while income and regional factors provide limited incremental explanatory power. The discrepancy between in-sample fit and cross-validation performance highlights the importance of reproducible evaluation and cautious interpretation of training metrics. The findings suggest that cybercrime legal framework development is primarily driven by institutional inertia and cumulative capacity, rather than rapid structural transformation. By integrating probabilistic transition modeling with machine learning classification, this study advances a governance-oriented perspective on cybersecurity analytics. The proposed framework offers a reproducible foundation for forecasting regulatory upgrades and informing international capacity-building strategies.

Keywords Cybercrime regulation; Markov transition modeling; Machine learning classification; Governance forecasting; Institutional path dependence

Introduction

Cybercrime has emerged as a major transnational security and governance challenge in the digital era. The rapid expansion of digital technologies, online services, and interconnected infrastructures has increased both the scale and complexity of cyber-enabled criminal activities, including fraud, ransomware, data breaches, cyber espionage, and other forms of technology-facilitated crime. These activities generate substantial economic, social, and security

Submitted 22 June 2026
Accepted 3 August 2026
Published 1 September 2026

*Corresponding author
R. Elavarasi,
elavarasir2014@gmail.com

Additional Information and
Declarations can be found on
page 236

DOI: 10.63913/jcl.v1i4.41
© Copyright
2026 Elavarasi, et al.

Distributed under
Creative Commons CC-BY 4.0

How to cite this article: R. Elavarasi, P. Gajalakshmi, S. Murugesan, K. Srinivasan, "Estimating Country-Level Cybercrime Legal Framework Upgrades Through Markov State Classification of Regulatory Maturity," *J. Cyber. Law.*, vol. 2, no. 3, pp. 222-239, 2026.

consequences and increasingly challenge conventional legal and institutional mechanisms that are primarily organized around national jurisdictions [1], [2].

The transnational nature of cybercrime has made the development of effective legal and institutional responses increasingly important. Existing research emphasizes the need for harmonized legislation, stronger law-enforcement capabilities, improved digital investigation mechanisms, and sustained international cooperation to address jurisdictional barriers and differences in national legal systems [3], [4]. International legal instruments have consequently become an important component of cybercrime governance, although their effectiveness depends not only on formal treaty commitments but also on domestic legal capacity and operational enforcement capabilities [5]. This suggests that national cybercrime regulation should not be viewed solely as a static legal attribute, but rather as an evolving institutional process in which countries may strengthen, maintain, or modify their regulatory frameworks over time.

Despite increasing international attention, substantial differences remain in the maturity of national cybercrime legal frameworks. Some countries have developed relatively comprehensive legal provisions, specialized institutional capacities, and stronger alignment with international instruments, whereas others continue to face legislative gaps, limited enforcement capacity, regulatory fragmentation, or insufficient international cooperation [4], [6], [7]. In particular, differences in regulatory capacity can create uneven levels of preparedness and complicate cross-border enforcement, since effective cooperation requires a sufficient degree of compatibility between national legal and institutional systems [8]. Understanding the factors associated with regulatory development is therefore important for assessing the future evolution of global cybercrime governance.

Research on cybersecurity has traditionally placed considerable emphasis on technical dimensions, such as intrusion detection, malware identification, attack prevention, and digital forensics. Although these approaches remain essential, cybercrime governance represents a complementary institutional dimension involving legislation, enforcement capacity, regulatory coordination, and international cooperation [9]. Recent legal scholarship further indicates that cybercrime regulation must continuously adapt to technological change and emerging forms of cyber-enabled criminality, reinforcing the importance of viewing regulatory development as a dynamic process rather than a fixed condition [10]. However, comparatively limited attention has been given to quantitatively forecasting how national cybercrime regulatory systems transition from one level of maturity to another.

Machine learning provides a potential methodological approach for addressing this gap because it can identify nonlinear relationships and complex interactions among multiple predictors. Ensemble learning methods, in particular, have demonstrated their ability to improve predictive performance by combining multiple classification models and capturing patterns that may not be adequately represented by conventional linear approaches [11]. Evidence from governance-related prediction problems also indicates that ensemble approaches can capture nonlinear relationships between structural, institutional, and regulatory factors [12]. Moreover, the integration of Markov models with multi-class machine learning has been used to represent transitions between discrete states while simultaneously incorporating explanatory predictors,

providing a methodological basis for combining transition probabilities with supervised classification [13].

Against this background, this study proposes a predictive framework for estimating country-level upgrades in cybercrime legal framework maturity using a state-transition modeling approach. Regulatory maturity is discretized into defined states, allowing changes between states to be modeled as transitions rather than as isolated differences in regulatory scores. This formulation makes it possible to examine both the persistence of regulatory states and the probability that a country will transition toward a higher level of regulatory maturity.

The methodological framework combines a Markov transition matrix with supervised multi-class classification. The Markov component estimates empirical transition probabilities between regulatory states, thereby providing a descriptive representation of regulatory persistence and movement. The machine learning component complements this analysis by incorporating structural predictors, including baseline regulatory maturity, geographic region, and income classification, to estimate the probability of transition across regulatory states. The combination is intended to provide both an interpretable representation of state dynamics and a predictive mechanism for identifying factors associated with regulatory progression.

The empirical analysis uses cybercrime regulatory maturity scores for 194 countries observed in 2020 and 2024. These two observation points provide a globally comprehensive cross-sectional basis for identifying discrete changes in regulatory maturity while allowing the analysis to examine country-level transitions over the four-year interval. Although the limited number of temporal observations constrains the ability to model long-term regulatory dynamics, the global coverage provides substantial cross-country variation across geographic regions and levels of economic development.

The primary contribution of this study is to extend quantitative cybersecurity research from the prediction of cyber threats toward the prediction of cybercrime regulatory development. Rather than forecasting the occurrence of cyber incidents, the proposed framework models the institutional evolution of national cybercrime regulation. By estimating transition probabilities and identifying structural characteristics associated with regulatory upgrades, the study provides a quantitative approach for examining policy convergence, regulatory persistence, and differences in the pace of institutional development. The resulting framework may support comparative cybersecurity governance research and provide evidence that can inform international cooperation and capacity-building strategies.

Literature Review

Cybercrime Detection and Machine Learning

Research on cybercrime has historically focused on the technical detection and mitigation of malicious activities. Machine learning has been extensively applied to intrusion detection, network traffic classification, anomaly detection, and malware identification, where the primary objective is to distinguish legitimate from malicious behavior. For example, supervised learning algorithms such as artificial neural networks, support vector machines, and Naive Bayes have been evaluated for classifying network traffic in intrusion detection systems [14].

Ensemble approaches have subsequently combined multiple classifiers to improve the detection of heterogeneous attack patterns, including malware and network intrusions [15], [16]. These studies demonstrate the value of machine learning for operational cybersecurity, particularly when large and complex digital datasets must be classified automatically.

The technical cybersecurity literature, however, primarily treats cybercrime as a problem of threat identification rather than institutional development. The dependent variables in these studies are generally attack classes, malicious traffic, malware categories, or anomalous behavior, while legal capacity, regulatory maturity, and institutional change are rarely modeled as prediction targets [17]. Consequently, although machine learning has become an established tool for technical cyber defense, its application to the prediction of changes in national cybercrime regulation remains comparatively underdeveloped.

Cybercrime Legislation and International Cooperation

A parallel body of research examines cybercrime from legal and institutional perspectives. This literature emphasizes that the transnational nature of cybercrime creates challenges for conventional criminal law because offenders, victims, digital evidence, and infrastructure may be distributed across multiple jurisdictions. Comparative studies consequently examine differences in criminalization, investigative authority, digital evidence, and mechanisms of international cooperation [18]. A systematic review of cybercrime legislation similarly concludes that comprehensive and regularly updated legislation is an important component of effective responses to cybercrime [19].

International cooperation is particularly important because differences among national legal systems can constrain cross-border investigation and prosecution. Research on international cybercrime law identifies differences in cooperation mechanisms and the exchange of information between law-enforcement agencies as continuing challenges [20]. Similarly, research on EU cybersecurity law shows that regulatory development can involve processes of convergence despite continuing differences between jurisdictions [21]. These findings establish cybercrime regulation as a multidimensional institutional phenomenon shaped not only by domestic legislation but also by international norms, regional integration, and cross-border policy processes.

Regulatory Convergence and Cybersecurity Governance

The governance literature provides further evidence that cyber regulation evolves through processes of institutional adaptation and regulatory convergence. International standards and regional legal frameworks can influence domestic regulatory development, particularly when countries seek closer institutional integration with regional organizations. The evolution of EU cybercrime and cybersecurity rulemaking illustrates how external norms can become incorporated into domestic regulatory structures through multilevel governance processes [22]. Similar dynamics can be observed in the implementation of international cybersecurity standards in national legislation, where regulatory alignment is associated with broader processes of institutional and international integration [23].

Cyber capacity-building research further indicates that regulatory development cannot be separated from institutional and organizational capacity. Capacity-

building initiatives may involve human resources, organizational arrangements, and legal and institutional frameworks, thereby linking cybersecurity development to broader processes of institutional strengthening [24]. Taken together, these studies suggest that regulatory maturity may evolve through identifiable pathways of institutional change and that countries can differ in both their initial regulatory conditions and their subsequent trajectories.

Nevertheless, most existing studies of cybercrime regulation remain comparative, doctrinal, or descriptive. They identify differences between legal systems, examine institutional arrangements, or discuss mechanisms of convergence, but rarely estimate the probability that a country will move from one regulatory maturity state to another. This creates a methodological gap between the substantial literature explaining what cybercrime regulatory systems look like and the more limited literature addressing how their regulatory states change over time.

Machine Learning, Explainability, and Governance Prediction

Machine learning provides a potential approach for addressing this gap because institutional and governance outcomes may involve nonlinear relationships among multiple structural factors. Ensemble learning is particularly relevant when the relationship between predictors and outcomes is complex. Importantly, predictive performance alone is insufficient when machine learning is applied to governance-related problems because policymakers and researchers also need to understand which variables contribute to predictions.

This concern has contributed to the development of explainable artificial intelligence (XAI). Research using open government data demonstrates that predictive models such as XGBoost can be combined with SHAP-based explanations to identify and interpret the factors contributing to individual predictions [25]. More broadly, research on explainable AI emphasizes that the increasing use of complex machine learning models in consequential decision-making creates a need for transparent interpretations of model outputs [26]. These developments are relevant to cybersecurity governance because a predictive regulatory model should not only estimate transition probabilities but should also provide interpretable evidence about the structural characteristics associated with those transitions.

Research Gap and Positioning of the Present Study

The literature reviewed above reveals three distinct research streams. The first applies machine learning predominantly to technical cyber-threat detection, including intrusion, malware, and anomaly classification [14]–[17]. The second examines cybercrime legislation and international cooperation, primarily through comparative, doctrinal, and institutional analysis [18]–[20]. The third investigates cybersecurity governance, regulatory convergence, capacity building, and explainable predictive modeling [21]–[26]. These streams have developed substantially but remain only weakly integrated.

In particular, existing cybersecurity machine-learning research generally predicts technical security outcomes, whereas legal and governance research generally explains regulatory differences without developing quantitative transition models. Furthermore, although Markov-based approaches can represent movement between discrete states and machine learning can incorporate multiple explanatory predictors, their combination has received

limited attention in the context of national cybercrime regulatory development. This gap is important because regulatory maturity can be conceptualized not simply as a static score but as a sequence of states through which countries may persist, improve, or potentially regress.

This study addresses this gap by proposing an integrated Markov–classification framework for modeling cybercrime regulatory maturity transitions. The Markov component represents the empirical probability of movement between discrete regulatory states, while supervised multi-class classification incorporates country-level structural predictors to estimate the likelihood of different transition outcomes. The framework therefore connects the descriptive strength of transition modeling with the predictive capacity of machine learning. By applying this approach to cross-national cybercrime regulatory data, the study extends existing research from technical cyber-threat prediction and descriptive legal comparison toward quantitative forecasting of regulatory development.

Methods

Research Design and Modeling Framework

This study employs a quantitative predictive modeling framework to examine the evolution of national cybercrime legal frameworks between 2020 and 2024. The methodological design conceptualizes regulatory development as a discrete state-transition process and combines empirical Markov transition modeling with supervised machine learning.

The framework consists of two sequential components. First, the Markov component estimates the empirical probability of moving between regulatory maturity states. Second, a Random Forest classifier uses country-level structural characteristics to predict the transition class of each country. The integration of these components allows the study to distinguish between the observed baseline dynamics of regulatory change and the conditional prediction of regulatory transitions based on country characteristics.

This conceptualization is consistent with previous applications that combine Markov transition structures with machine learning to model temporal or state-based change and with recent work demonstrating the use of Markov structures alongside Random Forest models for forecasting transitions between discrete states. The methodological workflow is shown in figure 1.

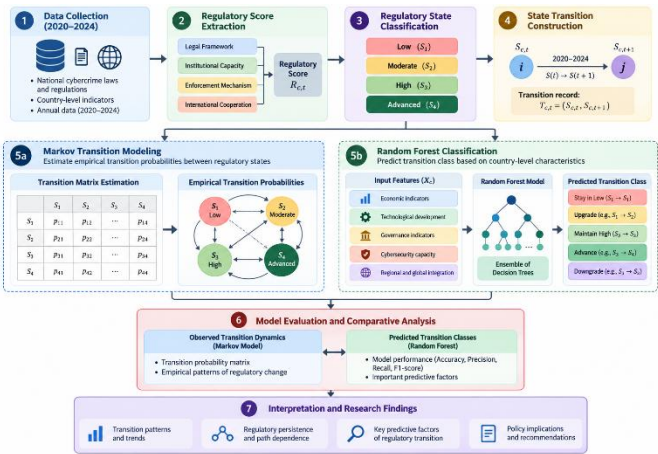


Figure 1 Research Workflow Integrating Markov Transition Modeling and Random

Forest Classification

The central methodological contribution is therefore to treat cybercrime regulatory development as a transition-class prediction problem rather than merely a cross-sectional regulatory scoring problem.

Dataset and Regulatory State Construction

The empirical dataset contains 194 countries with cybercrime regulatory maturity scores observed in 2020 and 2024. Let $R_{i,t}$ denote the regulatory maturity score of country i at time t , where $0 \leq R_{i,t} \leq 1$. The continuous regulatory score is transformed into three discrete regulatory states:

$$S_{i,t} = \begin{cases} L, & R_{i,t} < 0.40, \\ M, & 0.40 \leq R_{i,t} \leq 0.80, \\ H, & R_{i,t} > 0.80. \end{cases} \quad (1)$$

L , M , and H represent Low, Medium, and High regulatory maturity, respectively.

This transformation is intended to provide interpretable regulatory regimes and enable transition-based analysis. Each country therefore has an initial state $S_{i,2020}$ and a subsequent state $S_{i,2024}$. The regulatory transition is defined as

$$Y_i = S_{i,2020} \rightarrow S_{i,2024}. \quad (2)$$

Thus, a country moving from Low to Medium is represented as $L \rightarrow M$, while a country remaining in the High state is represented as $H \rightarrow H$.

The resulting transition variable constitutes the target variable for the supervised classification model.

Markov Transition Model

The first analytical component estimates the empirical transition probability between regulatory states. A first-order Markov formulation is adopted, in which the probability of the 2024 regulatory state is conditioned on the observed 2020 state. The transition probability from state i to state j is defined as

$$P_{ij} = P(S_{2024} = j \mid S_{2020} = i). \quad (3)$$

The empirical estimate is calculated as

$$\hat{P}_{ij} = \frac{N_{ij}}{\sum_k N_{ik}}, \quad (4)$$

N_{ij} represents the number of countries that transition from state i in 2020 to state j in 2024. The resulting transition matrix is

$$P = \begin{bmatrix} P_{LL} & P_{LM} & P_{LH} \\ P_{ML} & P_{MM} & P_{MH} \\ P_{HL} & P_{HM} & P_{HH} \end{bmatrix}. \quad (5)$$

The matrix provides the empirical baseline against which the supervised model is interpreted. Diagonal elements represent regulatory persistence, while off-diagonal elements represent regulatory movement. In particular, P_{LM} , P_{LH} , and

P_{MH} represent upward regulatory transitions.

Because only two observation points are available, the Markov model is interpreted as a one-step empirical transition model, rather than as a long-term stochastic process. It therefore describes observed regulatory movement between 2020 and 2024 without assuming unobserved intermediate transitions.

Supervised Transition Classification

The second component extends the Markov framework by incorporating structural country characteristics. The transition class is modeled as a function of the initial regulatory score, geographic region, and income classification. The feature vector for country i is defined as

$$X_i = [R_{i,2020}, G_i, I_i], \quad (6)$$

$R_{i,2020}$ is the baseline regulatory score, G_i represents geographic region, and I_i represents income classification. The conditional transition probability is therefore formulated as

$$P(Y_i = c \mid X_i), \quad (7)$$

c denotes a specific transition class, such as $L \rightarrow M$, $M \rightarrow H$, or $H \rightarrow H$.

This formulation differs from the Markov baseline because the probability of transition is no longer determined solely by the initial regulatory state. Instead, it is conditioned on additional structural characteristics that may distinguish countries sharing the same initial regulatory regime.

Random Forest Classification

A Random Forest classifier is used to estimate the conditional transition probabilities in Eq. (7). The model consists of $B = 300$ decision trees constructed using bootstrap sampling and randomized feature selection. For a given country i , each tree b produces a transition prediction $T_b(X_i)$. The estimated probability of transition class c is calculated as

$$\hat{P}_{RF}(Y_i = c \mid X_i) = \frac{1}{B} \sum_{b=1}^B \mathbb{I}[T_b(X_i) = c], \quad (8)$$

$\mathbb{I}[\cdot]$ is an indicator function. The final predicted transition is obtained from the class with the highest estimated probability:

$$\hat{Y}_i = \arg \max_c \hat{P}_{RF}(Y_i = c \mid X_i). \quad (9)$$

The Random Forest model is selected because the relationship between regulatory development and structural characteristics may be nonlinear and involve interactions between predictors. Previous research has demonstrated the usefulness of Random Forest and Markov-based combinations for predicting discrete transitions in complex systems.

Feature Preprocessing

The predictor set consists of one continuous variable and two categorical variables. The 2020 regulatory score is retained in its continuous form to

preserve information about the country's distance from the regulatory-state boundaries. Geographic region and income classification are transformed using one-hot encoding. For each categorical variable, category membership is represented as a binary indicator.

The complete preprocessing and classification procedure is implemented as a unified pipeline. Consequently, the same transformations are applied during cross-validation, model fitting, and prediction. No missing values were identified in the core variables used in the analysis.

Model Training and Validation

The supervised model is evaluated using stratified five-fold cross-validation. Stratification is applied to preserve the distribution of transition classes across folds. For each fold, the model is trained on four folds and evaluated on the remaining fold. The mean cross-validation performance is calculated across the five evaluation folds.

After cross-validation, the final model is retrained using the complete dataset. A fixed random seed is used throughout the modeling process to ensure reproducibility. Because the dataset contains only 194 countries, the analysis emphasizes stratified validation rather than relying on a single random train-test split. This approach reduces the sensitivity of the reported performance to a particular partition of the relatively small cross-national dataset.

Evaluation and Interpretation

The predictive model is evaluated using accuracy, precision, recall, and F1-score. Because the transition classes may be imbalanced, macro-averaged F1-score is additionally reported to provide equal weight to each observed transition class. A confusion matrix is used to identify systematic classification errors, particularly between regulatory persistence and upgrade classes.

The model's feature importance is also examined to identify which structural characteristics contribute most strongly to transition classification. These importance measures are interpreted as predictive associations rather than causal effects. The Markov transition matrix and Random Forest predictions are interpreted jointly. The Markov component provides the baseline probability of regulatory movement, while the Random Forest component determines whether additional country-level characteristics improve the discrimination of individual transition classes.

Reproducibility

All stochastic processes use a fixed random seed. The complete preprocessing and Random Forest pipeline is stored as a serialized checkpoint, while model configuration and evaluation results are recorded in structured metadata. This procedure allows the experiment to be reproduced without manually reconstructing the preprocessing or retraining the model when a valid checkpoint is available. The complete analytical process can therefore be summarized as:

$$R_{2020}, R_{2024} \rightarrow S_{2020}, S_{2024} \rightarrow Y \rightarrow \begin{cases} P_{Markov} \\ \hat{P}_{RF}(Y | X) \end{cases} \rightarrow \hat{Y}. \quad (10)$$

Equation (10) summarizes the proposed methodological architecture: observed

regulatory scores are converted into discrete states, state transitions are identified, empirical transition probabilities are estimated using the Markov model, and country-level transition probabilities are subsequently predicted using Random Forest.

Result

Descriptive Transition Dynamics

The Markov transition matrix (figure 2) reveals strong state dependence in cybercrime regulatory maturity between 2020 and 2024. Countries classified in the Low category exhibit extremely high persistence, with approximately 94% remaining in the Low state. Only a small proportion transitioned upward, and virtually none moved into intermediate states. This indicates that low regulatory capacity tends to be structurally entrenched over the observed period.

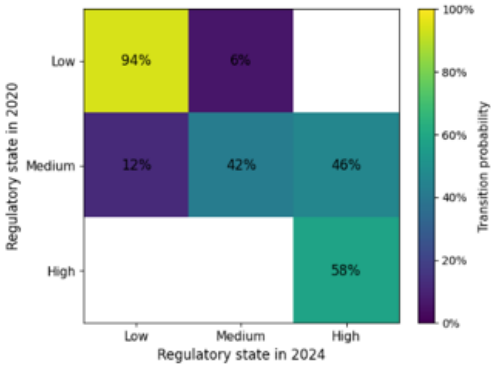


Figure 2 Markov Transition Probability Heatmap (2020–2024)

The Medium category demonstrates more dynamic behavior. Approximately 46% of Medium countries transitioned upward to High, while 42% remained in Medium and 12% regressed to Low. This suggests that Medium-level regulatory systems represent transitional regimes with substantial volatility. Countries at this stage appear to face both upward mobility potential and downward risk.

The High category exhibits partial stability but also notable downgrade behavior. Around 58% of High countries remained in the High state, while a substantial portion transitioned downward. Although this pattern may reflect sensitivity to discretization thresholds, it indicates that regulatory maturity is not strictly monotonic and that institutional maintenance plays an important role in sustaining high-level frameworks.

Overall, the transition matrix highlights that cybercrime regulatory development is strongly path-dependent. Initial maturity levels significantly constrain future movement, and the probability of change varies considerably across starting states. These findings establish a probabilistic baseline against which machine learning predictions can be interpreted. Figure 3 show flow diagram illustrates the observed movement of countries across Low, Medium, and High regulatory maturity states. The dominant transitions are Low→Low (94%), Medium→Medium (42%), Medium→High (46%), and High→High (58%). High→Low and High→Medium probabilities are not shown because their values were not specified in the reported results.

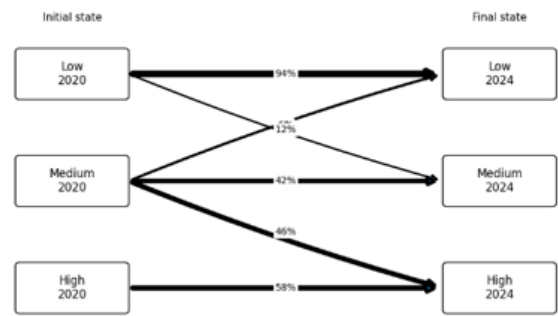


Figure 3 Country-level cybercrime regulatory state transitions between 2020 and 2024.

Predictive Performance of the Classification Model

The supervised classification model achieved a mean cross-validation accuracy of approximately 0.67 across five stratified folds (see figure 4). This indicates moderate predictive capability when evaluated under out-of-sample conditions. Given the multi-class structure and limited feature set, this performance suggests that structural predictors partially explain regulatory transitions but do not fully determine them.

In contrast, the in-sample training accuracy reached 0.94, accompanied by high precision and recall across most transition classes. The confusion matrix demonstrates strong diagonal dominance, particularly for persistent transitions such as High→High and Low→Low. However, this performance reflects model fit to the training data rather than generalizable predictive strength.

The discrepancy between cross-validation accuracy and training accuracy indicates model overfitting and structural memorization of transition patterns. Given the relatively small dataset size and the presence of strong state persistence, ensemble models can easily capture deterministic boundaries. Cross-validation results therefore provide a more realistic assessment of predictive power.

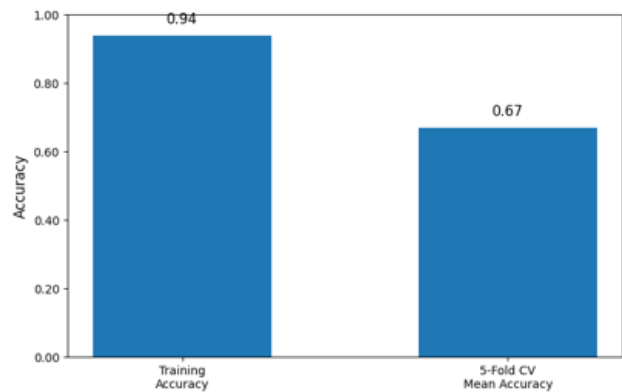


Figure 4 Random Forest classification performance based on training and cross-validation accuracy.

Importantly, even with moderate cross-validated performance, the model substantially outperforms naive random classification. This confirms that regulatory transitions are not entirely stochastic. Instead, they exhibit structured regularities that can be partially captured using baseline maturity and contextual indicators.

Class-Level Prediction Patterns

The classification report shows near-perfect performance for dominant classes such as High→High and Medium→Medium under in-sample evaluation. These classes are characterized by strong structural persistence and are therefore easier for the model to predict. The high recall for these transitions reinforces the interpretation that institutional inertia is a defining feature of regulatory evolution.

Upgrade transitions such as Medium→High and Low→Medium also demonstrate relatively strong predictive performance in training evaluation. This suggests that baseline maturity provides meaningful information about upward movement probability. However, due to smaller class sizes, caution must be exercised in interpreting performance metrics for rare transitions.

Rare transition categories, particularly those with minimal support, display perfect or near-perfect scores. These results are statistically fragile and likely reflect memorization rather than robust generalization. Such outcomes underscore the importance of cross-validation and careful interpretation of minority classes.

Overall, the model performs best for persistent and structurally predictable transitions. More volatile or less frequent transitions are harder to generalize, indicating that additional explanatory variables may be required to capture complex institutional shifts.

Determinants of Regulatory Transitions

Feature importance analysis indicates that the baseline regulatory score (score_2020) accounts for approximately 72% of total model importance (see [figure 5](#)). This dominance suggests that initial maturity is the primary determinant of subsequent transition outcomes. In other words, cybercrime regulatory evolution is heavily conditioned by prior institutional status.

Income level and regional classification contribute marginally to predictive performance. While these variables provide contextual refinement, they do not substantially alter transition probabilities once baseline maturity is known. This implies that economic development and geographic grouping play secondary roles relative to path dependence.

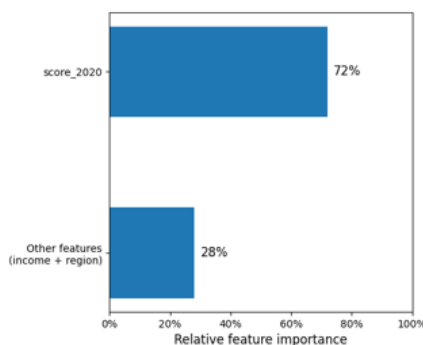


Figure 5 Relative feature importance in the Random Forest classification model.

The strong influence of score_2020 reflects threshold-based dynamics inherent in the discretization framework. Countries near state boundaries are more likely to transition, whereas those deeply embedded within a state exhibit stability.

This pattern aligns with institutional theory emphasizing incremental reform rather than abrupt structural change.

Taken together, the results indicate that regulatory maturity evolves primarily through state-dependent processes. Structural context provides limited additional predictive power beyond baseline maturity, reinforcing the centrality of institutional inertia in cybercrime governance development.

Discussion

The findings provide empirical evidence of path dependence in national cybercrime regulatory development. Countries with low regulatory maturity tend to remain in low-capacity states, while countries with high regulatory maturity are more likely to maintain advanced regulatory frameworks. This persistence indicates that cybercrime governance develops cumulatively and that institutional capabilities are not easily transformed over short periods. The strong importance of baseline regulatory maturity further supports this interpretation, as established legal frameworks, administrative capacity, technical expertise, and institutional experience can provide structural advantages for maintaining and developing regulatory capacity [23], [24]. Conversely, countries with lower maturity may face resource constraints, administrative fragmentation, limited technical expertise, and competing governance priorities that restrict the pace of regulatory modernization [24]. These findings suggest that cybercrime regulatory development is better understood as an incremental institutional process in which existing conditions influence subsequent trajectories.

The transition patterns also reveal greater volatility among countries in the Medium regulatory state compared with those in the Low and High states. Medium-level countries show meaningful probabilities of upgrading, remaining stable, or experiencing regression, indicating that this state represents a transitional phase in the development of cybercrime governance. Such volatility may reflect ongoing institutional reforms, including the introduction of new legislation, strengthening of enforcement mechanisms, development of digital investigation capabilities, and alignment with international standards [18], [19], [21]. These processes can generate opportunities for regulatory improvement while simultaneously creating temporary institutional uncertainty. Countries in the Medium state may therefore possess sufficient foundations for further development but lack the institutional stability characteristic of more mature regulatory systems. The importance of capacity building in strengthening human resources, organizational arrangements, and legal and institutional frameworks further supports the interpretation of regulatory development as an evolving process rather than a static condition [24].

From a forecasting perspective, the moderate predictive performance of the model indicates that regulatory transitions are partially predictable but cannot be fully explained by the structural variables included in the analysis. Baseline regulatory maturity provides an important signal for estimating future regulatory states, but the difference between training and cross-validation performance suggests limitations in generalization. This finding implies that regulatory transitions are influenced by additional factors that are not captured by the current feature set. Political and institutional stability, economic conditions, digital infrastructure, international cooperation, government capacity, and the implementation of international legal instruments may contribute to changes in

cybercrime regulatory maturity [18], [20], [21], [24]. The literature on cybersecurity governance also indicates that national regulatory development is shaped by interactions among domestic institutions, international norms, and cross-border cooperation [21], [22]. Incorporating these variables into future models could therefore provide a more comprehensive representation of the factors associated with regulatory progression and improve the ability to distinguish between countries with similar initial maturity but different subsequent trajectories.

The integration of Markov transition modeling and supervised machine learning provides a complementary analytical framework for understanding these dynamics. The Markov component captures the empirical probability of movement between regulatory states and provides an interpretable representation of persistence and transition patterns, while the classification component incorporates country-level predictors to estimate the probability of specific transition outcomes. This combination allows the analysis to move beyond descriptive comparisons of regulatory maturity toward a quantitative examination of how regulatory states may evolve. The use of machine learning for cybersecurity prediction demonstrates the ability of data-driven models to identify complex patterns and relationships that may be difficult to capture using conventional approaches [11], [16]. At the same time, explainable machine-learning research emphasizes the importance of interpreting model outputs when predictive models are applied to governance-related or consequential decision-making contexts [25], [26]. From a governance perspective, the proposed framework therefore provides a basis for examining patterns of regulatory persistence and transition while maintaining an interpretable connection between structural predictors and predicted outcomes.

Several limitations should nevertheless be considered when interpreting the findings. The analysis is based on only two observation points, which limits the temporal depth of the study and prevents the estimation of longer-term or higher-order transition dynamics. The classification of countries into Low, Medium, and High regulatory states may also affect the resulting transition probabilities, particularly for countries located near the boundaries between states. In addition, the relatively limited feature set means that important institutional, economic, technological, and political determinants may not be represented in the model. Previous research on cyber capacity building emphasizes that regulatory development involves multiple dimensions, including human resources, organizational arrangements, and legal and institutional frameworks [24], while research on international cybercrime governance highlights the importance of legal harmonization and international cooperation [18], [20]. Future research should therefore employ longer longitudinal datasets, conduct sensitivity analyses using alternative maturity thresholds, and incorporate broader governance and socioeconomic indicators. Expanding the temporal and explanatory dimensions of the dataset would enable more robust modeling of regulatory trajectories and provide deeper insight into the structural factors underlying the evolution of national cybercrime governance.

Conclusion

This study demonstrates that national cybercrime regulatory development exhibits strong path dependence and partial predictability. The Markov transition analysis reveals that countries in low and high regulatory states tend to remain

in their respective categories, while medium-level systems display greater volatility. Supervised machine learning modeling confirms that baseline regulatory maturity is the dominant determinant of subsequent transitions, with income level and regional context contributing only marginally once initial state is accounted for. Although in-sample classification performance is high, cross-validation results indicate moderate generalizable predictive capability, suggesting that structural inertia explains much—but not all—of regulatory evolution.

These findings imply that cybercrime legal framework development is shaped primarily by institutional continuity rather than rapid structural transformation. While machine learning can partially forecast regulatory upgrades, significant unexplained variation remains, pointing to the influence of political, economic, and international dynamics beyond baseline maturity. The integration of Markov transition modeling with supervised classification provides a transparent and reproducible analytical framework for governance forecasting. Future research incorporating richer longitudinal data and expanded structural predictors may enhance predictive accuracy and deepen understanding of the institutional mechanisms driving cybercrime regulatory progression.

Declarations

Author Contributions

Conceptualization: R.E., S.M.; Methodology: R.E., P.G.; Software: S.M., K.S.; Validation: P.G., K.S.; Formal Analysis: R.E.; Investigation: S.M., K.S.; Resources: P.G., K.S.; Data Curation: S.M.; Writing – Original Draft Preparation: R.E.; Writing – Review and Editing: P.G., K.S.; Visualization: K.S.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] O. Sviatun, O. Goncharuk, R. Chernysh, O. Kuzmenko, and I. Kozych, “Combating cybercrime: Economic and legal aspects,” *WSEAS Trans. Bus. Econ.*, vol. 18, no. 1, pp. 1–10, 2021, doi: 10.37394/23207.2021.18.72.

- [2] I. Elegbe, "Cybercrime legislation: A comparative analysis of legal frameworks, policy responses and recommendations," *Int. J. Educ. Soc. Sci. Res.*, vol. 7, no. 2, pp. 1–15, 2024, doi: 10.37500/ijessr.2024.7211.
- [3] S. Khan, T. Saleh, M. Dorasamy, and N. Khan, "A systematic literature review on cybercrime legislation," *F1000Research*, vol. 11, no. 1, pp. 1–20, 2022, doi: 10.12688/f1000research.123098.1.
- [4] T. M. Klyukanova, A. Vinichenko, and K. Y. Khristinchenko, "Analysis of the legislation of the EEU countries in the field of combating cybercrime," *Eurasian Integration: Economics, Law, Politics*, vol. 2024, no. 3, pp. 83–90, Sep. 2024, doi: 10.22394/2073-2929-2024-03-83-90.
- [5] V. Blikhar, M. Tsymbaliuk, H. Hreshchuk, R. Dostdar, T. Kokhaniuk, and I. Krykavska, "Current state and development trends of international law in the context of economic and legal analysis of financial measures to combat cybercrime in the global environment," *Financial Credit Activity: Problems Theory Pract.*, vol. 2022, no. 6, pp. 1–15, Dec. 2022, doi: 10.55643/fcaptp.6.47.2022.3936.
- [6] S. Snail ka Mtuze and M. Musoni, "An overview of cybercrime law in South Africa," *Int. Cybersecurity Law Rev.*, vol. 4, no. 1, pp. 1–25, 2023, doi: 10.1365/s43439-023-00089-8.
- [7] A. Fuady and F. Hasibuan, "Strengthening cybersecurity and data protection legal framework in Indonesia: A normative analysis of current challenges and future directions," *Law Justice Res. J.*, vol. 1, no. 3, pp. 1–15, 2025, doi: 10.70062/ljrj.v1i3.87.
- [8] B. Šturc, T. Gurova, and S. Chernov, "The specifics and patterns of cybercrime in the field of payment processing," *Int. J. Criminol. Sociol.*, vol. 9, no. 1, pp. 1–10, 2020, doi: 10.6000/1929-4409.2020.09.237.
- [9] M. Mphatheni and W. Maluleke, "Cybersecurity as a response to combating cybercrime," *Int. J. Res. Bus. Soc. Sci.*, vol. 11, no. 4, pp. 1–10, 2022, doi: 10.20525/ijrbs.v11i4.1714.
- [10] M. Alkrisheh, "Criminal protection of corporate websites: An analytical study," *J. Governance Regul.*, vol. 11, no. 3, pp. 1–10, 2022, doi: 10.22495/jgrv11i3art12.
- [11] N. Borjalilu, F. Jolai, and M. Tavakoli, "Cockpit crew safety performance prediction based on the integrated machine learning multi-class classification models and Markov chain," *Aviation*, vol. 2023, no. 1, pp. 1–15, 2023, doi: 10.3846/aviation.2023.19739.
- [12] Q. Jiang, F. Cioffi, F. Conticello, M. Giannini, V. Telesca, and J. Wang, "A stacked ensemble learning and non-homogeneous hidden Markov model for daily precipitation downscaling and projection," *Hydrol. Processes*, vol. 37, no. 1, pp. 1–15, 2023, doi: 10.1002/hyp.14992.
- [13] M. H. Lone, A. Mahindrakar, and K. Kumar, "Evaluation of CA-Markov and MLP-based machine learning approach for long-term land use land cover prediction in Himalayan mountain basins," *Discover Sustainability*, vol. 7, no. 1, pp. 1–15, 2026, doi: 10.1007/s43621-026-03211-y.
- [14] A. Gerka, "Searching for optimal machine learning algorithm for network traffic classification in intrusion detection system," *ITM Web Conf.*, vol. 21, no. 1, pp. 1–10, 2018, doi: 10.1051/itmconf/20182100027.
- [15] A. Venčkauskas, J. Toldinas, and Š. Grigaliūnas, "Ensemble-based classification using neural networks and machine learning models for Windows PE malware detection," *Electronics*, vol. 10, no. 4, pp. 1–15, Feb. 2021, doi: 10.3390/electronics10040485.

- [16] R. Gangula, M. V., and R. Kumar M., "Network intrusion detection system for Internet of Things based on enhanced flower pollination algorithm and ensemble classifier," *Concurrency Comput.: Pract. Exper.*, vol. 34, no. 1, pp. 1–15, 2022, doi: 10.1002/cpe.7103.
- [17] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "A novel ensemble of hybrid intrusion detection system for detecting attacks in networks," *Electronics*, vol. 2019, no. 1, pp. 1–15, 2019.
- [18] M. Dumchykov, M. Utkina, and O. Bondarenko, "Cybercrime as a threat to the national security of the Baltic States and Ukraine: The comparative analysis," *Int. J. Safety Security Eng.*, vol. 12, no. 4, pp. 1–10, Aug. 2022, doi: 10.18280/ijssse.120409.
- [19] S. Khan, T. Saleh, M. Dorasamy, and N. Khan, "A systematic literature review on cybercrime legislation," *F1000Research*, vol. 11, no. 1, pp. 1–20, 2022, doi: 10.12688/f1000research.123098.1.
- [20] M. Sayyad, Tofiq oğlu, and In Law, "Prosecuting cybercrimes under international legal frameworks: Challenges and innovations," *Futurity Economics Law*, vol. 2023, no. Sep., pp. 1–15, Sep. 2023, doi: 10.57125/fel.2023.09.25.05.
- [21] E. Fahey, "The evolution of EU–US cybersecurity law and policy: On drivers of convergence," *J. Eur. Integration*, vol. 46, no. 7, pp. 1073–1088, 2024, doi: 10.1080/07036337.2024.2411240.
- [22] E. Fahey, "The EU's cybercrime and cybersecurity rulemaking: Mapping the internal and external dimensions," *European Union Law*, vol. 2024, no. 1, pp. 1–15, 2024.
- [23] V. Derkachenko *et al.*, "Implementation of international cyber security standards in national legislation," *Cybersecurity Law Policy*, vol. 2022, no. 1, pp. 1–15, 2022.
- [24] P. Pawlak, "Capacity building in cyberspace as an instrument of international cooperation," *Cybersecurity Cooperation*, vol. 2016, no. 1, pp. 1–15, 2016.
- [25] E. Kalampokis, E. Karamanou, and *et al.*, "Applying explainable artificial intelligence to open government data," *Open Government Data*, vol. 2021, no. 1, pp. 1–15, 2021.
- [26] M. Dazeley, P. Vamplew, C. Foale, C. Young, S. Aryal, and F. H. M. Cruz, "Levels of explainable artificial intelligence for human-aligned conversational agents," *Artif. Intell.*, vol. 2021, no. 1, pp. 1–15, 2021.