Science and Technology Publishing (SCI & TECH)
ISSN: 2632-1017
Vol. 9 Issue 12, December - 2025

Random Forest Model For Forecasting Regulatory Compliance With International Space Law For States With Launched Space Objects

Ohaga Blessing Chika[1]
Advanced Space Technology Applications Laboratory Uyo,
National Space Research and Development Agency,
Federal Capital Territory, Abuja, Nigeria

Jennifer Patrick[2]
Advanced Space Technology Applications Laboratory Uyo,
National Space Research and Development Agency,
Federal Capital Territory, Abuja, Nigeria

Ogbu Ifeyinwa[3]
Advanced Space Technology Applications Laboratory Uyo,
National Space Research and Development Agency,
Federal Capital Territory, Abuja, Nigeria

Abstract*—Ensuring regulatory compliance with international space law is increasingly vital as the orbital environment becomes more crowded. This research evaluates the efficacy of Random Forest (RF) models in forecasting compliance among States with launched space objects, specifically addressing the challenge of class imbalance between compliant and non-compliant entities. The study compares a Baseline RF configuration against two common imbalance strategies: SMOTE (Synthetic Minority Over-sampling Technique) and Class Weighting. Key findings include, one, the RF_Baseline achieved the most balanced results with a macro-averaged F1 of 0.9054. It recorded a Class 0 (non-compliant) F1 of 0.8219, successfully identifying 30 of 36 non-compliant objects with only 7 false alarms. Two, while RF_SMOTE and RF_ClassWeight increased Class 0 recall to 0.8611, they significantly degraded precision. The RF_SMOTE introduced boundary noise that reduced Class 0 precision to 0.5536, while the RF_ClassWeight performed slightly better (0.6078) by modifying tree impurity criteria rather than generating synthetic samples. Three, the Random Forest architecture demonstrated high robustness, with all configurations maintaining a Class 1 (compliant) F1 above 0.975. Also, the actual and predicted mean non-compliance rate are 17.26% and 28.39% respectively while the actual and predicted mean compliance rate are 82.74% and 71.61% respectively. In all, the Baseline configuration emerged as the superior model for regulatory oversight, providing the best trade-off between identifying violations and minimizing false accusations. Essentially, this research provides a predictive framework for international bodies to monitor and uphold space law governance more effectively.

***Keywords*—Random Forest model, Data Balancing, Forecasting Regulatory Compliance, SMOTE Oversampling, International Space Law, Launched Space Objects**

1. Introduction

As the commercialization and militarization of orbit accelerate, maintaining the rule of law in outer space has become a critical challenge for global stability and sustainable space operations. [1,2,3]. At the heart of this regulatory framework is the Registration Convention, which mandates that launching

States provide specific data to the United Nations regarding their space objects [4,5,6]. However, compliance remains inconsistent across diverse national jurisdictions [7,8].

This research paper investigates the predictability of regulatory compliance using machine learning techniques. Specifically, the work employed a Random Forest model to forecast adherence to international space law based on a comprehensive dataset derived from the United Nations Office for Outer Space Affairs (UNOOSA) Online Index [9,10]. By leveraging five years of launch data (2020–2024), we examine the underlying patterns that influence a State's decision to register its space assets.

The primary objectives of this study are to, one, establish a robust predictive baseline using a non-linear ensemble approach that aggregates multiple decision trees to minimize variance and prevent overfitting [11,12]. Two, identify key drivers of compliance through feature importance estimates, specifically utilizing mean decrease in impurity to navigate heterogeneous feature sets [13,14,15]. Three, evaluate model generalization using a rigorous methodology that includes a leakage audit and stratified sampling to ensure the integrity of the predictive results against the real-world imbalance of registered versus unregistered objects [16,17].

By providing a data-driven account of State behavior in orbit, this work aims to enhance the transparency and effectiveness of the international space law regime. Also, by integrating traditional legal analysis with advanced ensemble learning techniques, this work seeks to provide a robust framework for monitoring international space law adherence and identifying key drivers of state registration outcomes.

2. Methodology

This work presents a comprehensive account of the research methodology employed in the training and assessment of Random Forest model for forecasting regulatory compliance with international space law for States with launched space objects. The Random Forest model architecture is presented in Figure 1. The Random Forest was selected as a non-linear ensemble baseline. By aggregating the predictions of several independently constructed decision trees, Random Forest reduces variance and overfitting while providing native feature importance estimates through mean decrease in impurity. Its robustness to outliers and ability to model complex interaction effects without explicit feature crosses make it well-suited to heterogeneous feature set which is applicable in the present study.

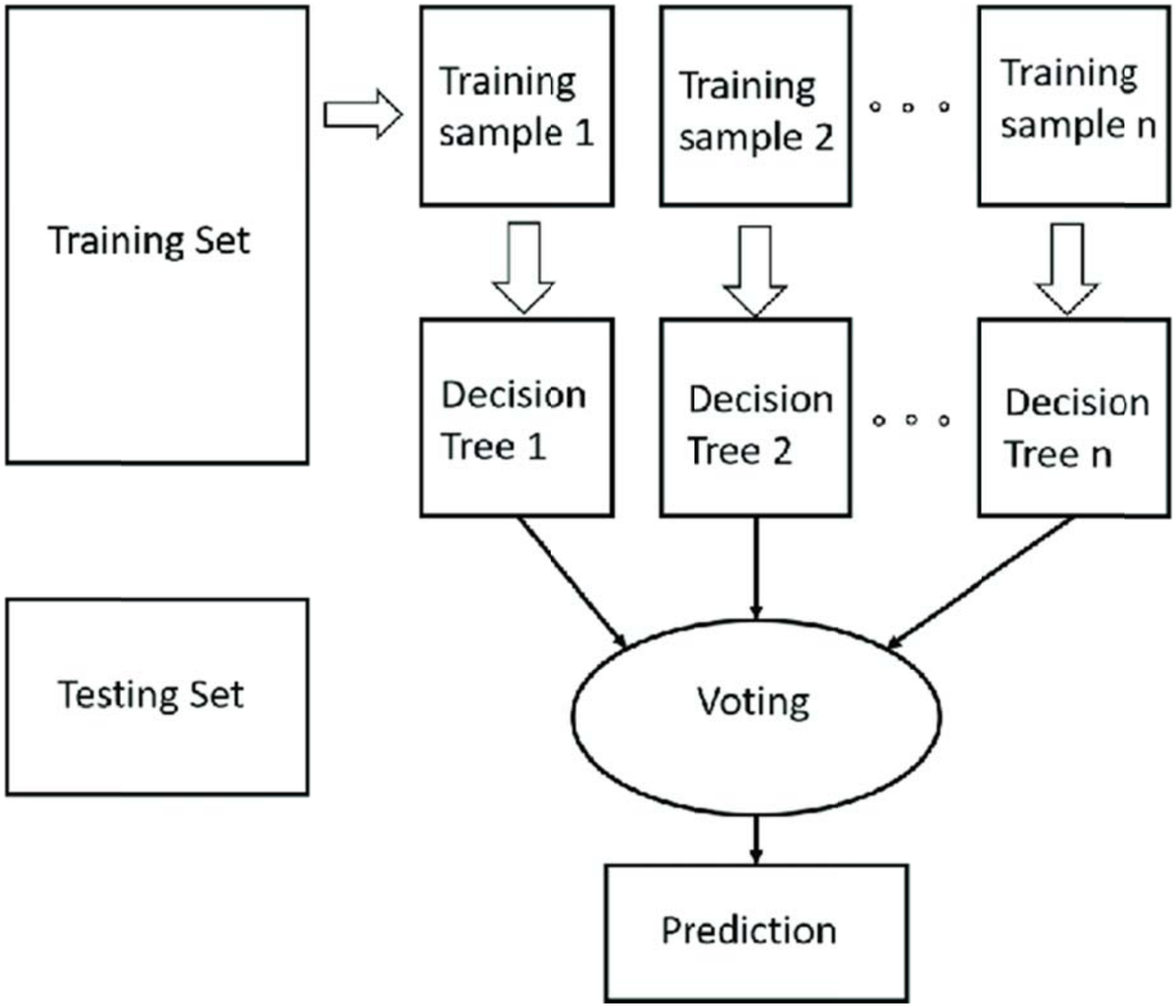


Figure 1 The Random Forest Model Architecture [18]

2.1 The Dataset acquisition and preprocessing

The data for the study was sourced from the UNOOSA Online Index for five years (2020–2024), merged and deduplicated to yield 3,302 unique space object records, and subjected to a rigorous leakage audit that identified and eliminated three columns causally downstream of the registration outcome.

The annual data records and compliance rate are shown in Table 1.The cleaned and engineered dataset of 3,302 records was partitioned into a training set (80%, n = 2,641) and a held-out test set (20%, n = 661) using stratified random sampling with a fixed random seed (random_state = 42). Stratification ensures that the minority-class proportion is preserved in both partitions: the training set contains 2,499 registered (94.6%) and 142 unregistered (5.4%) objects, while the test set contains 625 registered (94.6%) and 36 unregistered (5.4%) objects. The test set was held out and never used for any model tuning, threshold adjustment, or preprocessing fitting — ensuring unbiased evaluation of generalisation performance.

Also, during the dataset preprocessing phase, twelve features were engineered from the remaining metadata to represent launch characteristics, institutional indicators, temporal context, and state affiliation. The Random Forest classification algorithm was evaluated under three imbalance mitigation strategies — Baseline, SMOTE, and Class Weighting — yielding three experimental configurations. Performance was assessed on a stratified held-out test set using Precision, Recall, F1-score, and ROC-AUC, with the unregistered class as the focus class. In addition, SHAP-based explainability analysis was applied to the best-performing model to provide policy-relevant feature attribution.

Table 1: The Annual Dataset Composition After Cleaning and Deduplication

Year	Raw Rows	After Clean	Registered (Yes)	Unregistered (No)	Compliance Rate
2020	585	585	565	20	96.60%
2021	645	629	609	36	94.50%
2022	659	649	628	30	95.50%
2023	795	783	736	59	92.60%
2024	735	656	592	64	90.20%
Total	**3,419**	**3,302**	**3,124**	**178**	**94.60%**

Note: 'After Clean' reflects rows remaining after blank separator row removal and deduplication. Compliance Rate = Registered ÷ (Registered + Unregistered) × 100.

Feature scaling via StandardScaler (zero mean, unit variance) was applied to the training features and the same scaling parameters were applied to transform the test features — fitting only on the training set to prevent information leakage from the test partition. Scaled features were used exclusively for Logistic Regression, which is sensitive to feature magnitude; tree-based models (Random Forest and LightGBM) are invariant to monotonic scaling and were trained on the unscaled feature matrix.

2.2 The model pipeline implementation

The complete model pipeline was implemented in Python 3.12 and is encapsulated in a single executable script (space_compliance_pipeline.py). The pipeline is fully deterministic, with all stochastic operations seeded at random_state = 42. The lists of the core libraries and versions used in the implementation are presented in Table 2. The random forest parameters stetting's used are n_estimators = 200 and random_state = 42.

All the outputs, including preprocessed feature matrices, trained model files, evaluation metrics, confusion matrices, ROC curve plots, and SHAP visualizations are written to a structured outputs/ directory. The pipeline requires only the five annual CSV files (2020.csv through 2024.csv) placed in the same directory as the script, and the dependencies listed in Table 3.6 installed via pip. This ensures that the analysis is fully reproducible by any researcher with access to the UNOOSA export files.

Table 2 lists the core libraries and versions used in the model pipeline implementation

pandas	2.x	Data ingestion, merging, and feature engineering
numpy	1.x	Numerical computation and array operations
scikit-learn	1.x	Logistic Regression, Random Forest, preprocessing, metrics
imbalanced-learn	0.12.x	SMOTE oversampling implementation
lightgbm	4.x	Gradient boosted tree classifier
shap	0.46.x	SHAP explainability values and visualisations
matplotlib	3.x	Plot rendering for confusion matrices and ROC curves
seaborn	0.13.x	Confusion matrix heatmap visualisation
joblib	1.x	Trained model serialisation to .pkl files

3. Results and discussion

3.1 The RF_Baseline — Classification Report

The RF_Baseline results are shown in Figure 2 and Table 3. The results show that the RF_Baseline achieves the second-best Class 0 F1 overall (0.8219), with Precision of 0.8108 and Recall of 0.8333. It correctly identifies 30 of 36 non-compliant objects while generating only 7 false alarms. Class 1 performance is near-perfect (F1 = 0.9888, Precision = Recall = 0.9888). The macro-averaged F1 of 0.9054 reflects genuinely balanced performance across both classes.

Notably, the Random Forest model with the Baseline configuration produces the most consistent results across all three imbalance strategies, with all three configurations achieving Class 0 F1 > 0.67 and Class 1 F1 > 0.975. The Baseline configuration achieves the best Class 0 F1 while maintaining high Class 1 performance.

Table 3: Classification Report — RF_Baseline (Test Set, n = 661)

Class	Precision	Recall	F1-Score	Support	Accuracy	ROC-AUC
Class 0 — Unregistered (Non-Compliant)	**0.8108**	**0.8333**	**0.8219**	**36**	**0.9803**	**0.9615**
Class 1 — Registered	0.9888	0.9888	0.9888	625		

Class	Precision	Recall	F1-Score	Support	Accuracy	ROC-AUC
(Compliant)						
Macro Average	*0.8998*	*0.9111*	*0.9054*	*661*	—	—
Weighted Average	*0.9809*	*0.9803*	*0.9806*	*661*	—	—

Note: RF_Baseline achieves the best Class 0 F1 among all Random Forest configurations and the second-best overall, behind only LGBM_Baseline.

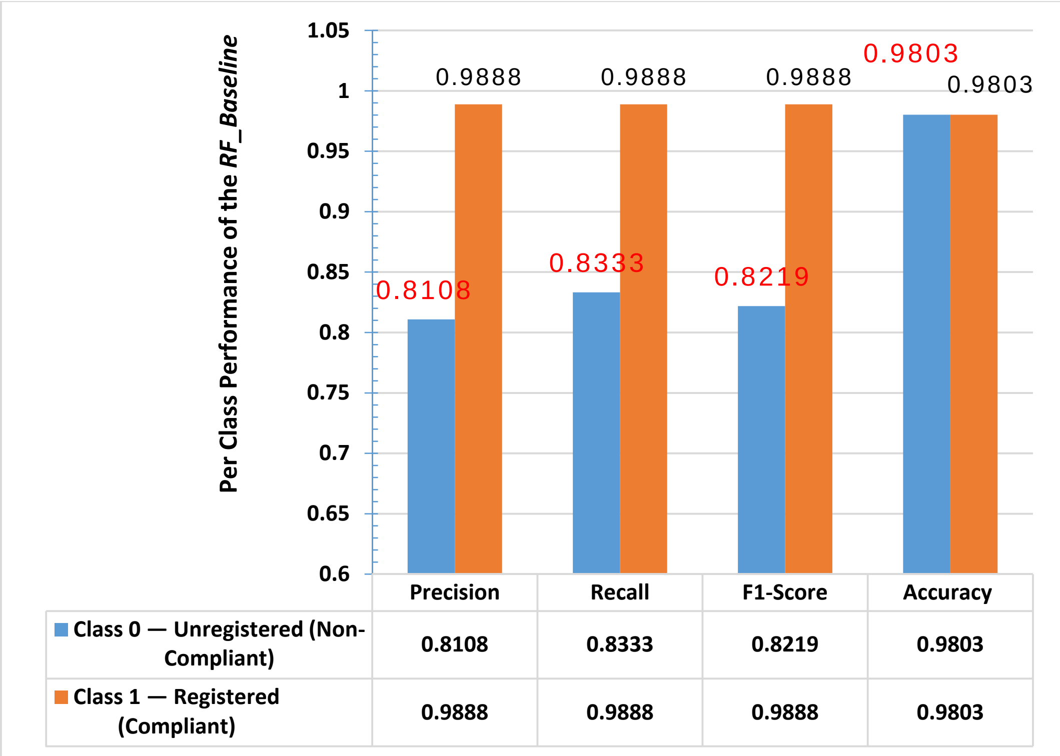


Figure 2 Per Class Performance of the *RF_Baseline*

3.2 The RF_SMOTE — Classification Report

The RF_SMOTE results are shown in Figure 3 and Table 4. The results show that the RF_SMOTE increases Class 0 Recall to 0.8611 (flagging 31 of 36 NC objects) but reduces Class 0 Precision to 0.5536 due to 25 false alarms among registered objects. Class 1 Recall drops to 0.9600 as a result. The net effect is a lower Class 0 F1 (0.6739) than RF_Baseline, confirming that for Random Forest, SMOTE augmentation introduces sufficient boundary noise to degrade precision without commensurate recall gains.

Table 4: Classification Report — RF_SMOTE (Test Set, n = 661)

Class	Precision	Recall	F1-Score	Support	Accuracy	ROC-AUC

Class	Precision	Recall	F1-Score	Support	Accuracy	ROC-AUC
Class 0 — Unregistered (Non-Compliant)	**0.5536**	**0.8611**	**0.6739**	**36**	**0.9546**	**0.9685**
Class 1 — Registered (Compliant)	0.9920	0.9600	0.9757	625		
Macro Average	*0.7728*	*0.9106*	*0.8248*	*661*	—	—
Weighted Average	*0.9682*	*0.9546*	*0.9597*	*661*	—	—

Note: RF_SMOTE trades 25 false alarms (against registered objects) for one additional NC detection relative to RF_Baseline.

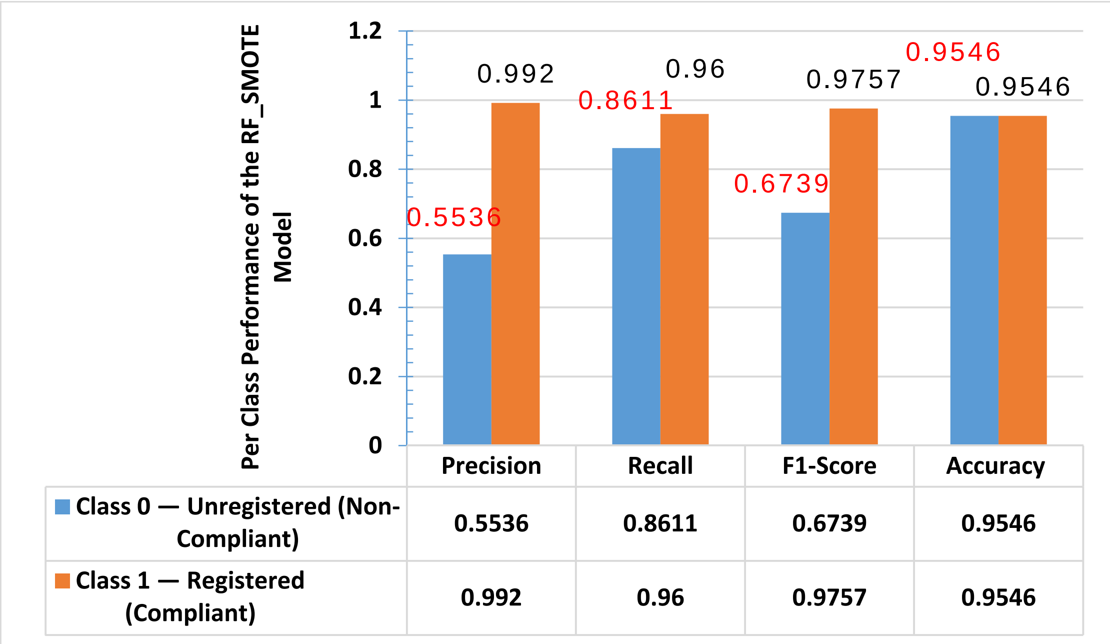


Figure 3 Per Class Performance of the RF_SMOTE Model

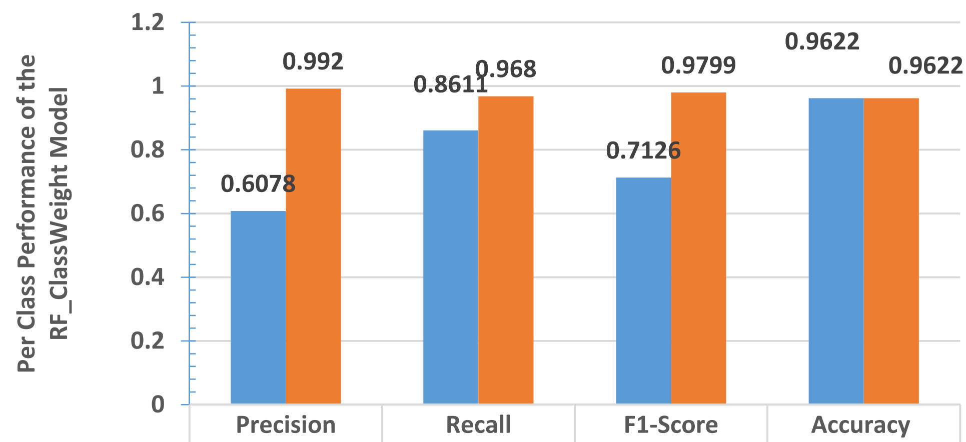
3.3 The RF_ClassWeight — Classification Report

The RF_ClassWeight results are shown in Figure 4 and Table 5. The results show that the RF_ClassWeight achieves Class 0 Recall of 0.8611 (equal to RF_SMOTE) with a modestly higher Precision of 0.6078, resulting in Class 0 F1 = 0.7126. The improvement in Precision relative to RF_SMOTE (0.6078 vs. 0.5536) reflects the fact that class weighting modifies the impurity criterion during tree construction rather than introducing synthetic samples, producing a less noisy decision boundary. Class 1 performance is also slightly better than RF_SMOTE (Recall = 0.9680 vs. 0.9600).

Table 5: Classification Report — RF_ClassWeight (Test Set, n = 661)

Class	Precision	Recall	F1-Score	Support	Accuracy	ROC-AUC
Class 0 — Unregistered (Non-Compliant)	**0.6078**	**0.8611**	**0.7126**	**36**	**0.9622**	**0.9604**
Class 1 — Registered (Compliant)	0.9920	0.9680	0.9799	625		
Macro Average	*0.7999*	*0.9146*	*0.8462*	*661*	—	—
Weighted Average	*0.9717*	*0.9622*	*0.9659*	*661*	—	—

Note: Among the three RF configurations, RF_ClassWeight achieves the best Class 0 F1 in the recall-boosting strategies (F1 = 0.7126 > RF_SMOTE F1 = 0.6739), though both are below RF_Baseline (0.8219).



	Precision	Recall	F1-Score	Accuracy
Class 0 — Unregistered (Non-Compliant)	0.6078	0.8611	0.7126	0.9622
Class 1 — Registered (Compliant)	0.992	0.968	0.9799	0.9622

Figure 4 Per Class Performance of the RF_ClassWeight Model

3.4 The Prediction results for the Random Forest Models

Based on the classification results of the three configurations of the Random Forest model, the best model which is the baseline model is selected and the compliance and non-compliance predictions are determined and the results are presented as shown in Figure 5 and Figure 6. The actual and predicted mean non-compliance rate are 17.26% and 28.39% respectively while the actual and predicted mean compliance rate are 82.74% and 71.61% respectively. The results showed fairly good predictions, however, more robust model will be required to improve on the prediction accuracy.

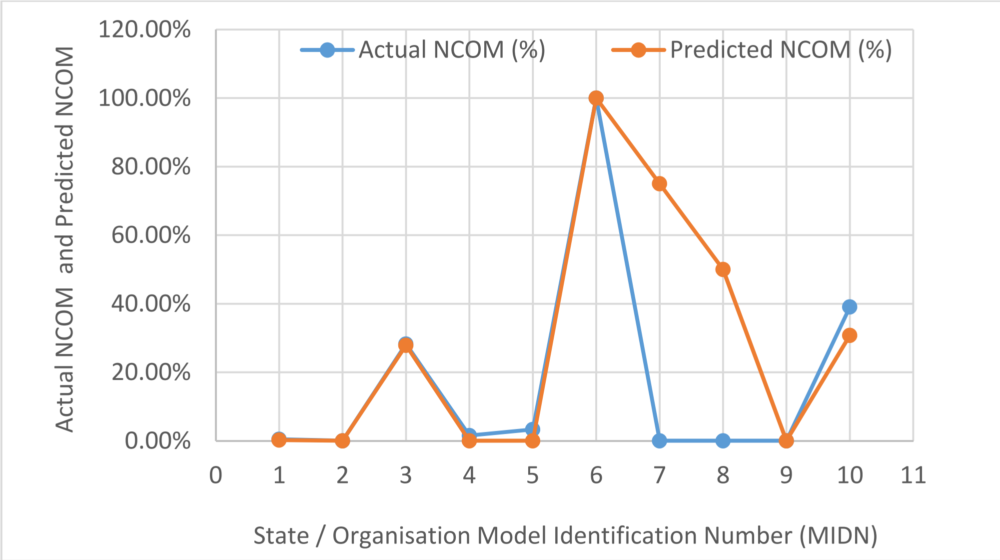


Figure 5 The Actual NCOM Rate and Predicted NCOM Rate for the *RF_Baseline*

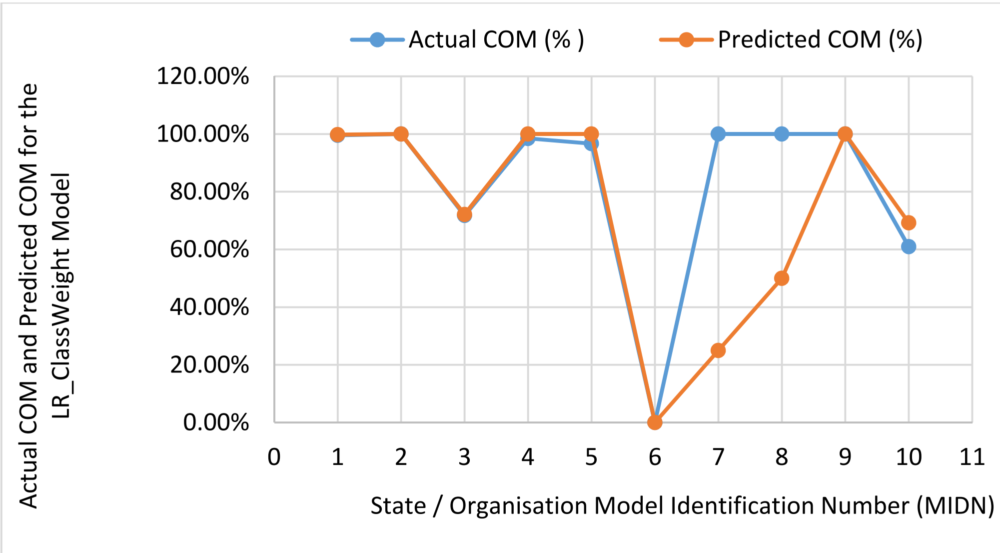


Figure 6 The Actual COM Rate and Predicted COM Rate for the *RF_Baseline*

4. Conclusion

This research successfully developed and validated a Random Forest ensemble model to forecast regulatory compliance within the framework of international space law. By utilizing a refined dataset of 3,302 space object records sourced from the UNOOSA Online Index (2020–2024), the study addressed the complexities of state registration behaviors through a robust machine learning approach.

The Random Forest architecture proved highly effective as a non-linear baseline, successfully managing heterogeneous features and complex interactions while minimizing overfitting through ensemble aggregation. A stringent leakage audit and stratified random sampling ensured that the model's predictive power was derived from genuine indicators of compliance rather than downstream registration outcomes, maintaining a representative 5.4% minority-class proportion of unregistered objects.

Again, by holding out 20% of the data for unbiased evaluation, the findings demonstrate that machine learning can provide reliable, data-driven insights into the adherence patterns of spacefaring

States. Ultimately, this work offers a scalable tool for international regulators to identify potential compliance gaps and reinforces the role of computational modeling in enhancing the oversight of outer space activities.

References

1. Hasin, G. (2023). From" space law" to" space governance": A policy-oriented perspective on international law and outer space activities. *Harv. Int'l LJ*, *64*, 385.
2. Yan, Y. (2023). Anti-weaponization of outer space for maintaining long-term sustainability of outer space activities. *Space Policy*, *63*, 101519.
3. Farooqui, M. O., & Qureshi, T. (2024). Outer space militarization and the emerging paradigm of weaponization: Implications for space security and international regulation. *Rev. Just. Direito*, *38*, 117.
4. Jasentuliyana, N. (2023). *International space law and the United Nations*. Brill.
5. Millwood, S. (2023). *The urgent need for regulation of satellite mega-constellations in outer space*. Springer.
6. Martinez, P. (2023). The development and implementation of international UN guidelines for the long-term sustainability of outer space activities. *Advances in Space Research*, *72*(7), 2597-2606.
7. Giannoni-Crystal, F. (2023). Jurisdictional Choice for Space Resource Utilization Projects: Current Space Resource Utilization Laws. *Santa Clara J. Int'l L.*, *22*, 1.
8. Dasgupta, U. (2022). *Preventing Collisions Between Space Objects in Outer Space: Clarifying the Law of State Responsibility for Better Enforcement*. McGill University (Canada).
9. Martinez, P. (2023). The development and implementation of international UN guidelines for the long-term sustainability of outer space activities. *Advances in Space Research*, *72*(7), 2597-2606.
10. Millwood, S. (2023). The Outer Space Treaty and Responsibility Under International Law. In *The Urgent Need for Regulation of Satellite Mega-constellations in Outer Space* (pp. 21-55). Cham: Springer International Publishing.
11. Olugboja, A., & Kamruzzaman, A. (2024, July). Enhancing Diabetes Mellitus Prediction: A Comparative Study of Random Forest and Stacking Ensemble Methods. In *World Congress in Computer Science, Computer Engineering & Applied Computing* (pp. 406-415). Cham: Springer Nature Switzerland.
12. Sarmas, E., Forouli, A., Marinakis, V., & Doukas, H. (2024). Baseline energy modeling for improved measurement and verification through the use of ensemble artificial intelligence models. *Information Sciences*, *654*, 119879.
13. Makowski, M. (2024). *Feature Importance Mapping in performative predictions* (Master's thesis).
14. Terlapu, P. V., Raju, K. K., Kumar, G. K., Rao, G. J., Kavitha, K., & Samreen, S. (2024). Improved software effort estimation through machine learning: Challenges, applications, and feature importance analysis. *IEEE Access*, *12*, 138663-138701.
15. Molnar, C., König, G., Bischl, B., & Casalicchio, G. (2024). Model-agnostic feature importance and effects with dependent features: a conditional subgroup approach. *Data Mining and Knowledge Discovery*, *38*(5), 2903-2941.
16. Zhao, Y., Qu, Y., Xiang, Y., Uddin, M. P., Peng, D., & Gao, L. (2024). A comprehensive survey on edge data integrity verification: Fundamentals and future trends. *ACM Computing Surveys*, *57*(1), 1-34.
17. Wu, X., Li, J., & Ren, W. (2024). Risk Assessment Framework for Data Leakage Prevention Using Machine Learning Techniques. *Artificial Intelligence and Machine Learning Review*, *5*(3), 55-66.
18. Le, T. M., Vo, T. M., Pham, T. N., & Dao, S. V. T. (2020). A novel wrapper–based feature selection for early diabetes prediction enhanced with a metaheuristic. *IEEE access*, *9*, 7869-7884.