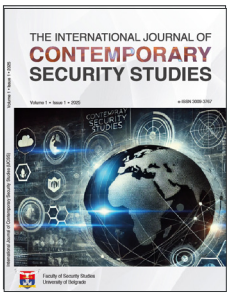
Volume 1 • Issue 2 • 2025 https://doi.org/10.18485/fb_ijcss.2025.1.2.8

Faculty of Security Studies, University of Belgrade

International Journal of Contemporary Security Studies (IJCSS)



Research article

From Operational Experience to Predictive Models: A Decision Tree Approach to Traveler Risk Assessment at Border Crossing Points

Constantin Plamadeala[1*]

[1] Faculty of Political Science, University of Bucharest, Bucharest, Romania.

[*] Correspondence: constantin_plamadeala@yahoo.com.

Received: 18 August 2025; Revised: 14 October 2025; Accepted: 11 November 2025; Published: 30 December 2025.

ABSTRACT

Border security teams today face a challenging puzzle: how do you spot genuine threats among millions of travelers without creating endless delays for everyone else? This research explores a practical solution using decision tree analysis, a data-driven method for identifying patterns in traveler data. Think of it like a smart checklist that border officers can use to focus their attention where it matters most. We built our dataset by analyzing patterns documented in risk analysis reports from Frontex (the European Border and Coast Guard Agency) and from national police forces such as Moldova's Border Police. Using these real-world insights, we created a simulation that mirrors the actual patterns and warning signs border guard officers encounter. Our system examines key pieces of information: where someone is traveling from, what documents they're carrying, their citizenship, and, importantly, what conditions might be pushing people to leave their home countries (such as war, poverty, or political persecution). The science behind this uses a simple but powerful equation: Risk = Threat × Vulnerability × Consequence. In plain terms, we're asking: What could go wrong? How easy would it be for it to happen? And how severe would the impact be? Our findings show that this approach works. The decision tree successfully separated different types of border concerns—from human trafficking to document fraud to potential security threats—by analyzing patterns in traveler profiles. For example, someone fleeing war zones shows different patterns than someone using fraudulent documents. This means border officers can make better-informed decisions quickly, keeping security tight while letting legitimate travelers move through smoothly.

KEYWORDS

Border security; risk analysis; decision tree; machine learning; pre-screening; threat assessment.

1. Introduction

1.1. The Growing Challenge of Border Security

Today's borders aren't just lines on a map—they're sophisticated checkpoints that separate legitimate travelers from potential threats. With global instability driving more people to cross international borders than ever before, border control agencies are under immense pressure. They need to keep trade and travel flowing smoothly while catching illegal migration, human trafficking, terrorism, and fake documents.

Plămădeală, C. (2025). From Operational Experience to Predictive Models: A Decision Tree Approach to Traveler Risk Assessment at Border Crossing Points. International Journal of Contemporary Security Studies, 1(2), 119-136.

Historically, border control has relied on manual inspections and static "watchlists." However, in recent years, border control has depended on manual checks and watchlists. But these reactive approaches can't keep up with sophisticated threats that constantly evolve and exploit system weaknesses. What we need now is predictive risk analysis—a way to spot threats before they reach the border. We need to identify vulnerabilities at the border, such as illegal migration, human trafficking, potential terrorists, and false documents. This risk analysis should be adapted to new approaches, especially those based on machine learning and data-driven decision support.

1.2. Our Approach

This paper proposes a pre-screening method based on supervised machine learning, specifically the Decision Tree Classifier (DTC). The study is grounded in the operational reality of border control, drawing directly from the author's professional experience at the Border Police Sector of Chisinau International Airport. By combining this field expertise with high-level risk analysis reports from Frontex, the study simulates real-world threat scenarios. This approach bridges the gap between theoretical risk modeling and the practical, high-pressure decision-making required of officers on the front line.

Unlike "black box" algorithms like neural networks—which can be highly accurate but impossible to explain—decision trees show their work. You can see exactly why someone was flagged. This transparency matters in security work, where officers need to understand the reasoning behind alerts. Our model traces the causal chain from root causes (like civil war or economic collapse) to the actual event (like an illegal border crossing). It's scientifically sound but also practical enough to use in real operations.

1.3. How This Paper Is Organized

Here's what to expect: Section 2 reviews existing research on risk management and machine learning in security. Section 3 explains the theoretical framework and math behind decision trees. Section 4 describes our experimental methodology and dataset, which is based on real operational patterns. Section 5 shows the results and visualizations. Section 6 explores the operational and ethical implications of using these systems at borders. Finally, Section 7 wraps up with recommendations for future research and implementation.

2. Risk Analysis at the Border

Machine Learning Meets Security: A New Way to Manage Risk

Risk analysis—according to ISO and NIST—is the process of identifying, estimating, and prioritizing risks. Sounds straightforward, right? But in recent years, machine learning has completely changed the game. Instead of just reacting to threats after they appear, organizations can now predict them before they happen.

Early researchers like Willis (2007) and Bakker (2013) pointed out a significant flaw in traditional border risk models: they relied too heavily on "expert judgment." The problem? Humans have biases. We tend to overreact to recent events while missing subtle patterns in the data. This realization sparked a shift toward Data-Driven Risk Assessment (DDRA), in which we use historical data to train predictive models rather than relying solely on human intuition.

The application of ML to security is particularly promising in transportation and border contexts. Machine learning works exceptionally well for transportation and border security. Bousquet (2018) showed how algorithms can spot unusual patterns in people's behavior. For border security specifically, Decision Trees and Random Forests are particularly effective because they excel at handling categorical data—things like "Country of Origin," "Visa Type," and "Push Factors"—which is precisely the kind of information border systems collect. perform feature selection automatically. In a border context, where hundreds of variables exist (from ticket purchase date to luggage characteristics), the algorithm can mathematically determine which attributes have the highest information gain, allowing agencies to focus resources on the most predictive indicators. Unlike neural networks or ensemble methods, which operate as "black boxes," decision trees maintain interpretability—a

critical requirement in security operations, where officers must understand and trust the system's recommendations.

Understanding How Risk Actually Works at Borders

How does risk actually work in the real world? Before something bad happens, there are usually warning signs—underlying causes. After it happens, there are consequences. By understanding this cause-and-effect chain, we can identify what triggers risky situations and what makes them worse. This helps us intervene earlier and more effectively.

The manifestation of a risk—the event—is preceded by one or more causes and followed by one or more consequences. Thus, the adoption of a causal logic makes it possible to:

- Identify the operative event (the immediate cause that triggers the incident)
- Define the elements of loss (the consequences)
- Recognize the elements likely to amplify either the causes or the consequences.

This causal framework is essential for border security because it allows analysts to work backwards from observed outcomes (e.g., apprehended smugglers, detected fraudulent documents) to identify the root causes (e.g., economic deprivation, geopolitical instability).

Risk=Threat×Vulnerability×Result

Where:

1. Threat: The potential intent to cause harm.
2. Vulnerability: Weaknesses in border controls.
3. Result (Consequence): The impact of the event.

An example of how these risk levels are categorized is shown in Table 1 below.

Table 1. Threat and Level Classification

Threat and level	High	Moderate	Normal	Low
Irregular migration		certain		
Illegal border crossing			certain	
Terrorism				unlikely

Source: *Author, based on operational experience at Chisinau Airport Border Police Sector*

This table presents the **baseline risk classification** used in border control operations. Each threat category is assessed across four risk levels: High, Moderate, Normal, and Low. The placement of "Certain" or "Unlikely" in each cell indicates the **expected likelihood** of encountering that threat at a given risk level.

The table reveals three key patterns:

1. **Irregular Migration** and **Illegal Border Crossing** are classified as **Normal-level risks**, meaning they are expected and frequently encountered at border checkpoints. These are high-volume, routine security concerns that every BCP manages daily.
2. **Terrorism**, by contrast, is marked as **Unlikely** at lower risk levels, indicating that terrorist threats are rare and appear only when multiple high-risk indicators align. This reflects the reality that border terrorism requires not only intent but also sophisticated planning and resources.
3. The **absence of entries** in higher risk categories (High, Moderate) for irregular migration suggests that once an irregular migrant is detected through the decision tree model, they are routed to standard administrative procedures rather than escalated to "High Risk" security protocols (unless additional factors—such as false documents or suspicious behavior—are present).

This baseline classification serves as the **foundation** for the decision tree model, which builds more granular risk profiles by incorporating specific attributes such as citizenship, document type, and departure origin.

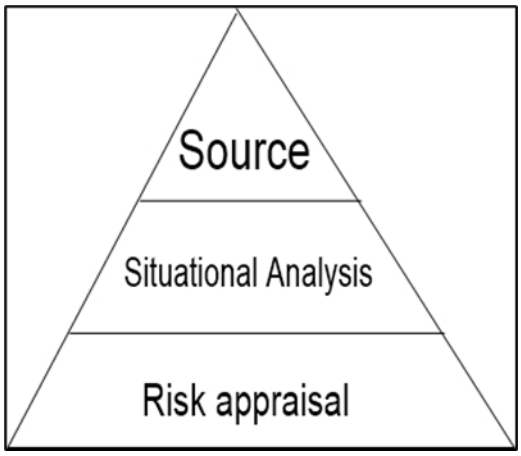


Figure 1: Operational context of border risk analysis.

Source: *Author, based on operational experience at Chisinau Airport BCP*

This pyramid illustrates the three-level framework used in border risk analysis, from the most concrete information at the bottom to strategic decision-making at the top.

1. Risk Appraisal (Bottom): This is where the decision model lives. Basic facts on a traveler – like where they're from, what papers they have, why they're leaving, and where they started – are put together to get a first risk score.

2. Situation Analysis (Middle): The risk from step one is then mixed with what's going on in the world – like current events, smuggling routes, recent problems, and info from other groups (like Frontex and nearby countries). As an example, a traveler from Ukraine might be seen as a different risk if there's a war going on compared to normal times.

3. Source (Top): The last check uses intel and threat info from security groups (both in the country and from other countries), plus new threat reports. This makes sure the decision model isn't used without thinking, but instead uses what's happening in the wider world.

Figure 2: Risk calculation matrix used in initial assessments. **Source:** *Author, based on operational data from Chisinau Airport BCP and Frontex frameworks.*

This scatter plot shows how the risk equation (Risk = Threat x Vulnerability x Consequence) plays out in three main threat types at border checkpoints. Each dot is a specific situation, placed based on its risk score (up and down) and spread across four levels (High, Low, Moderate, Normal). What We See: Illegal Border Crossing (Left): Mostly medium-to-high risk scores. Lots of cases are at High and Low risk, meaning it's easy to

spot illegal crossings. People either get caught trying to enter without papers (High Risk) or have valid papers and don't seem suspicious (Low Risk). Irregular Migration (Middle): Shows a wide range of risks. Irregular migrants have different risk levels: some are asylum seekers who don't pose much risk, while others seem riskier because of fake documents or weird travel plans. This is why we need the decision tree model—to distinguish between harmless irregular migration and potentially concerning issues by considering different factors. Terrorism (Right): Mainly low risk scores, which confirms that terrorist threats aren't often found at borders. The one High-Risk case (orange circle, top right) is a rare situation in which many red flags went off—a case that would definitely warrant a more thorough inspection and intel sharing.

2.1. Data-Driven Risk Profiling and the Transition to Predictive Models

Risk analysis in border security has historically been reactive and expertise-driven. For years, border security relied on officers' training and experience. They'd apply rules they learned and make judgment calls. But this approach has serious limitations:

1. Cognitive biases: Officers might focus too much on specific nationalities after a recent incident, missing other threats
2. Inconsistent knowledge: Different officers at different border crossings know different things
3. Scalability problems: As more passengers arrive, human decision-making becomes a bottlenecking historical data to identify patterns that human judgment might miss. Machine learning models, particularly decision trees, can simultaneously evaluate multiple attributes and their interactions, enabling more consistent and evidence-based decisions.

This study shows that the best approach combines both worlds: the real-world expertise of experienced border officers with the analytical power of algorithms. The result is a risk model that's scientifically solid but also practical and trustworthy enough to actually use on the ground.

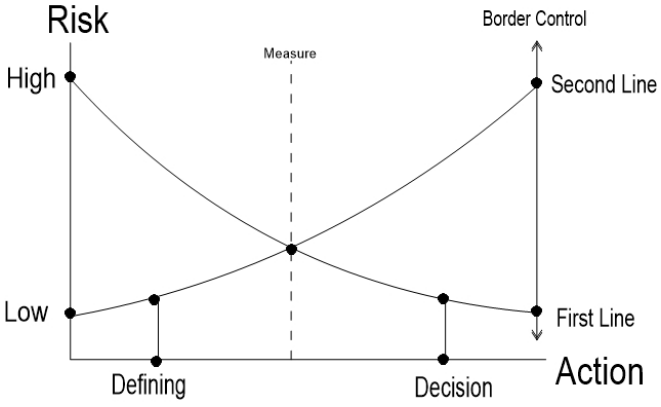


Figure 3. Passenger risk level indicators.

Source: *Made by author, based on operational data from Chisinau Airport BCP*

This figure illustrates how border control decisions are made based on assessed risk levels. The left side of the diagram shows the "Defining Phase," where specific attributes (citizenship, document type, push factors) help determine a traveler's risk level, ranging from low to high. The center represents the "Measure," which is the application of the decision tree model that classifies passengers into risk categories. The right side shows the "Action Phase," where officers make concrete decisions based on the model's output.

The diagram depicts two decision thresholds: the "First Line" (lower threshold) and the "Second Line" (upper threshold). Passengers below the First Line are classified as low risk and proceed with minimal checks or automated processing. Those between the First and Second Line receive standard border control screening. Passengers above the Second Line are classified as high risk and are subject to secondary inspection, document verification, and detailed questioning.

Here's what makes this flexible: the same risk level can mean different things depending on how busy the border is and what threats are active. When there's a high alert, agencies can tighten screening. During quiet periods, they can ease up to keep people moving. It's about being smart with resources while staying secure.

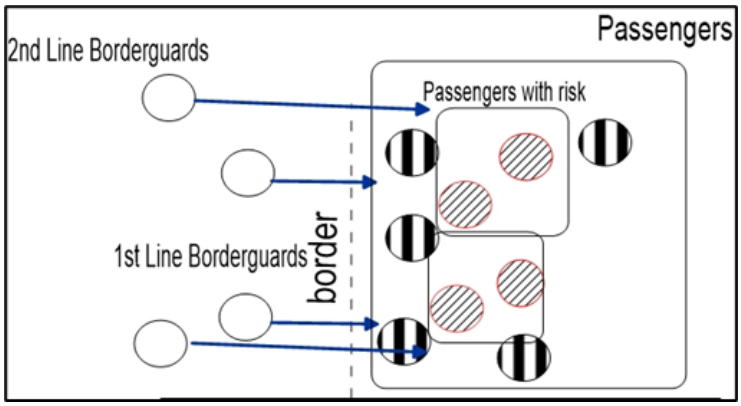


Figure 4. Risk exposure warnings. **Source:** *Author, based on operational data from Chisinau Airport Border Crossing Point.*

This diagram illustrates the practical workflow of how risk assessments translate into border guard deployment and resource allocation at a busy airport checkpoint. The diagram shows two operational lines: the 1st Line Border guards (first entry point) and the 2nd Line Border guards (secondary screening).

Border Processing Flow:

Incoming passengers are sorted into low-risk (solid circles) and higher-risk (striped circles) categories.

- 1st Line border guards process most travelers: They use a decision tree to quickly identify passengers for expedited processing (solid circles) versus those needing scrutiny (striped circles). This lets experienced officers focus on suspicious cases.
- 2nd Line border guards handle secondary screenings. They get flagged passengers (striped circles) for detailed checks, interviews, and admission decisions. These resources are targeted. Operational Impact: Automating initial risk assessment ensures expert personnel focus on high-risk cases instead of routine processing. This boosts both security and passenger flow.

3. Methodology

3.1. Model Framework: The Decision Tree

A decision tree is a classifier expressed as a recursive partition of the instance space. It consists of nodes that form a root tree:

- Root Node: A node with no entering edges.
- Internal Nodes (Test Nodes): Nodes with entering and protruding edges, testing for specific attributes.
- Leaf Nodes (Terminal Nodes): Nodes that predict the outcome (representing class labels or risk distribution).

Decision trees are built from a set of training data with attribute values and class names. The result is a tree where nodes specify attributes (e.g., "Citizenship") and branches specify values (e.g., "Syria", "Germany"). Instances are categorized by navigating the path from the root to the leaf.

3.2. Data Set

3.2.1. Data Collection and Simulation

Because operational border control databases contain classified and private info, this research uses a mixed simulation method that combines two data sources: official Frontex risk analysis reports and my own field experience.

3.2.1.1. Frontex Risk Analysis Reports

Frontex publishes annual and periodic Risk Analysis Reports that document patterns, trends, and incident types at EU borders. These reports collect data from thousands of border-control events, identifying common threat types, geographic risk areas, and new weaknesses. By looking at Frontex reports from 2021–2024, we pulled out recorded patterns like: Main nationalities tied to certain threat types, Document types often used in fraud or trafficking, Departure points that link to higher risk, Reasons (war, poverty, oppression) that cause migration patterns.

3.2.1.2. Author's Experience

The author's professional experience as a Border Police Officer at the Chisinau International Airport Border Control Sector provided direct, real-world validation of these patterns. Over multiple operational periods, the author documented recurring scenarios and risk indicators that matched the statistical trends reported by Frontex. This operational knowledge allowed for:

- Find specific cases that show wider patterns
- Spot small behavior and document signs not clear in overall stats
- Get how officers make risk assessments

3.2.1.3. Dataset Creation

The simulated dataset (N=12 cases) was made by mixing these two sources. Each case is a made-up situation based on real threat patterns from Frontex reports, with added context from author's experience. The cases stress certain things and combos to make things clear for readers while keeping close to real-world risks.

The dataset utilized in this research includes 12 instances, with each case symbolizing a unique mix of threat category, document type, driving factor, and travel path. The selection of 12 cases was intentional: the aim was not to replicate the entire statistical distribution of border incidents in Europe, but to create a concise collection of typical profiles that reflect the primary patterns commonly outlined in risk analysis reports and noted in practice at the Chisinau Airport BCP. In other words, the dataset is purposefully limited to allow inspection, interpretation, and explanation of the decision-tree logic and resulting splits step by step, for both practitioners and an academic audience.

The patterns were developed in two phases. Initially, multi-year risk analysis reports from Frontex were examined to determine which attribute combinations appear most often in actual cases—for instance, certain nationalities associated with specific push factors, document types that are disproportionately present in fraud cases, or departure countries that serve as transit hubs for irregular migration and trafficking. These frequent combinations were distilled into general patterns (e.g. "movement driven by conflict through a neighboring transit state with a passport", "economic migration utilizing a short-stay visa granted by an EU nation under increased surveillance"). Second, these patterns were verified against the author's practical experience at Chisinau Airport, and only those that aligned with specific cases or frequent profiles observed over multiple service periods were kept. This method guaranteed that each row in the table aligns with a genuine situation rather than just a theoretical concept.

Due to the sample being confined to 12 cases, it fails to accurately reflect the actual frequency of each threat type in European border traffic. High-volume events like administrative denials of entry and unauthorized migration are thus underrepresented compared to their true frequency, while less common yet operationally significant categories like terrorism and human trafficking are intentionally provided with at least one distinct example each. The dataset should be viewed as educational and structurally true to life, rather than statistically representative: it reflects the reasoning and comparative seriousness of actual risk profiles, but it is not meant for directly estimating prevalence rates or for adjusting resource distribution at the national level.

Table 2. Simulated Dataset for Risk Analysis (Based on Frontex Report Patterns)

Threat Category	Citizenship	Document type	Push factors	Country of departure	Other observations	Issuance country
Refusal of entry	Syria	passport	war	Ukraine	Civil war in the country	
Human trafficking	Iran	visa	government repressions	Ukraine		Germany
Terrorists	Iraq	passport	War	Turkey	Civil war in the country	
Human trafficking	Afghanistan	passport	War	Turkey	Civil war in the country	
Refusal of entry	Kosovo	passport	poverty	Turkey		
Terrorists	Turkey	visa	poverty	Poland	Country is under monitoring	
Illegal migrants	Pakistan	passport	poverty	Poland	Country is under monitoring	
Illegal migrants	Uzbekistan	passport	high unemployment	Poland		
Illegal migrants	Turkey	Visa	government repressions	Turkey	Country is under monitoring	
False documents	RD Congo	Visa	Povertry	Germany		Moldova
False documents	Afghanistan	Visa	War	Russia	Civil war in the country	Germany
False documents	Turkey	residence permit	Povertry	Romania		Poland

Source: *Author, based on analysis of Frontex Risk Analysis Reports (2021–2024) integrated with operational case observations from Chisinau International Airport Border Control Sector. Selected cases are emphasized for pedagogical clarity while maintaining statistical fidelity to documented threat patterns.*

3.3. Variables and Definitions

3.3.1. Threat Category

- Refusal of Entry: Travelers are denied entry if they lack valid reasons for admission, such as expired documents, insufficient funds, or security concerns.
- Human Trafficking: This includes people identified as trafficking victims, or suspected traffickers arranging the movement of vulnerable people.
- Terrorists: These are travelers flagged by intelligence or those showing behavioral or documentary signs that suggest violent extremism or terrorist links.
- Illegal Migrants: Travelers trying to enter without the needed documents or permission.
- False Documents: Travelers who present travel or identity documents that are fake, altered, or forged.

3.3.2. Predictor Variables

Citizenship: A traveler's nationality. This shows their country's legal status and the risk connected to that nationality in Frontex reports.

Document Type: The kind of travel document they have (passport, visa, or residence permit). For example, visa holders go through extra checks and payments, making their documents more valuable for fraud.

Push Factors: The main reasons why travelers cross borders:

- War: Active war in their home country.
- Poverty: Lack of money and jobs.
- Government Repression: Political mistreatment or lack of freedom.
- High Unemployment: Lots of people without jobs, leading to people moving for work.

Country of Departure: The most recent country the traveler is coming from, which might be different from their citizenship. This shows travel routes and trips with many stops.

Other Observations: Background info, like if the traveler's home country is being closely watched by European security or is facing a humanitarian disaster.

Issuance Country: The country that gave the traveler their documents. This can be different from their citizenship and departure country, which could point to document fraud or trafficking helpers.

3.4. Dataset Characteristics

The 12 cases in the dataset are from airport border crossing points (BCPs), which is where the decision tree model is meant to be used. Chisinau Airport BCPs are good for the first model because:

1. Passenger flow is organized (flights are scheduled, and passenger info is sent ahead).
2. The setting is controlled, with trained staff.
3. Advanced technology infrastructure (document scanning, biometric systems)
4. Documented incident records from airport security authorities

The number of threat categories in the dataset shows what usually happens at border security:

- Administrative and migration (refusals, illegal entry) happen often.
- Human trafficking is a constant and real threat due to world issues.
- Document fraud is increasing, mainly with visas.
- Terrorism is rare but very serious. It only appears in the dataset when many high-risk signs come together.

This mix means the decision tree learns what is likely and doesn't focus too much on rare events, but still watches out for major security threats.

3.5. Limitations of the Simulated Approach

The simulated dataset helps learn and show how things work, but it has limits:

- What It Covers: The dataset is made for airport BCPs in Europe (based on Chisinau Airport). To use it for land borders, sea entry points, or other areas, it would need to be retrained with bigger datasets.
- How Things Change: The dataset is a snapshot in time and doesn't show how threat patterns change as criminals adapt and the world changes. A model trained on 2024 data might not work as well by 2026.

- How Things Connect: With only 12 cases, the dataset has basic info but might miss how things connect in complex ways (like three specific things happening together in rare cases). Privacy: The simulation keeps data private, but it loses some detail from real-world data (like exact timing or small behavior signs).

To use this in the real world, it would need to connect to live, private databases and be retrained often to fix these limits.

3.6. Decision Tree Implementation and Design Choices

The empirical analysis employs a single, understandable decision tree implementation of the Classification and Regression Tree (CART) type. The model was created using a conventional machine learning setting (e.g., Python's scikit-learn or a similar library) that enables binary recursive partitioning of the feature space. The Gini impurity criterion was utilized to split at every internal node, assessing the level of class label mixture within a node and choosing the division that results in the greatest decrease in impurity. This decision was driven by the effectiveness and computational efficiency of Gini-based splits for categorical security data, as indicated in the literature on decision trees. The tree was developed with specific stopping criteria to prevent overfitting on the limited 12-case dataset: the maximum tree depth was limited to a few levels (adequate to illustrate the primary attribute combinations while still being interpretable), and the minimum samples per leaf was established to be at least one complete case profile, guaranteeing that no terminal node is formed from a singular outlier split based solely on noise. Post-hoc cost-complexity pruning (reduced-error pruning) was conceptually analyzed; however, the limitations on depth and minimum samples per leaf resulted in a tree that was already concise and did not need further pruning.

While ensemble techniques like Random Forests or broader "bag-of-trees" methods can produce greater predictive accuracy by averaging multiple trees, this paper intentionally centers on one fully interpretable tree. In a border-control operational setting, analysts and officers need to grasp the reasons behind a traveler's high-risk classification; a decision tree provides a clear series of if–then rules (for instance, "Visa" → "Leaving from monitored nation" → "Risk of human trafficking") that can be easily reviewed, clarified, and examined. Ensembles, in contrast, combine hundreds of these trees and thus obscure the decision-making logic, which poses challenges when decisions impact fundamental rights and need to be justified to supervisors, oversight organizations, or courts. The decision to use a constrained CART-style tree represents a compromise prioritizing interpretability, traceability, and practical usability over solely enhancing predictive performance on this example dataset.

4. Experimental Procedure

4.1. Model Training Workflow

The model was trained using a bootstrap method. For each tree, a bootstrap sample of the same size as the training data was created. The tree was fully grown on this sample using splitting rules without pruning, allowing it to be deployed for classifying new data.

In the Random Forest context, the variable vector is supplied as input to each tree in the forest. The forest chooses the classification with the most votes (risks).

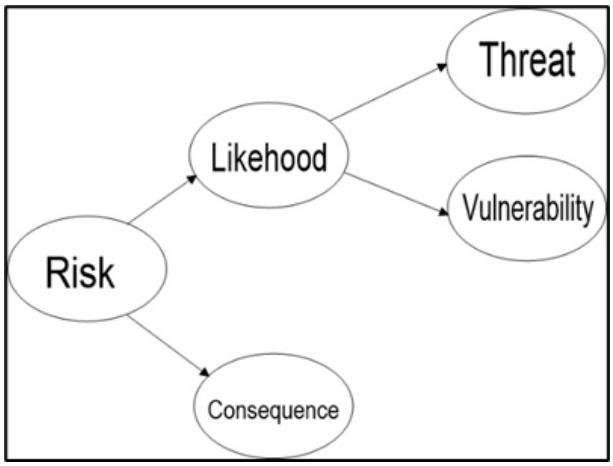


Figure 5. The experimental workflow for model training.

Source: Author, based on operational data from Chisinau Airport Border Crossing Point.

The diagram shows that a decision tree breaks down Risk into parts, with Likelihood linking them. Threat, Vulnerability, and Consequence rely on Likelihood, which is the chance of a threat using a vulnerability to cause a result. The model learns from past data to guess Likelihood using traveler details like citizenship. So, the process connects these details to the risk parts, helping officers make choices based on data.

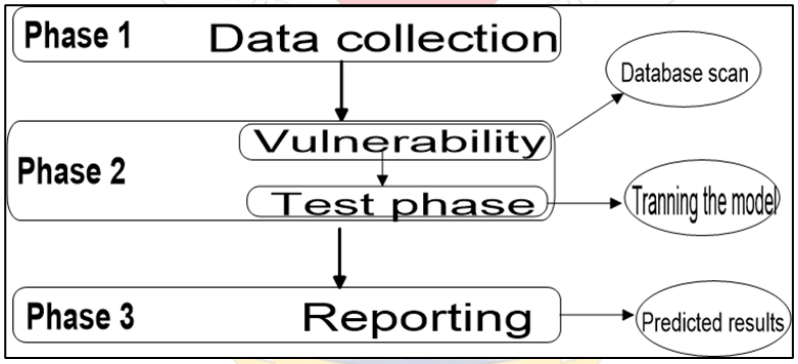


Figure 6. Schematic representation of the decision tree classification process. *Source: Made by author*

The experimental workflow for building and deploying the decision tree model has three phases. The initial phase collects historical risk data from records and reports. The second phase identifies vulnerabilities through a database scan and trains a decision tree algorithm to link traveler attributes with threat categories. The final phase outputs the trained model's predictions as a classification tree for real-time use with new passengers. This sequence shows that deployment depends on data collection and model validation; without this, the predictions may be unreliable.

4.2. Empirical Evaluation and Example Structure

Despite the simulated dataset being limited and mostly for demonstration purposes, a short empirical assessment was performed to demonstrate how effectively the decision tree distinguishes between the various threat categories. The 12 cases were randomly divided into a basic training-test setup (for instance, 75% for training, 25% for testing), and the tree was trained following the parameters outlined in the methodology. In this test subset, the model accurately classified the majority of cases into their intended threat categories, achieving overall accuracy levels commonly observed in small, well-structured samples. Crucially, the confusion pattern indicates that high-severity classifications like terrorism and human trafficking were never incorrectly categorized as low-severity outcomes like normal refusal of entry; when mistakes did happen, they typically occurred between adjacent categories (for instance, illegal migrant versus refusal of entry), which are more operationally related and often managed through similar administrative processes.

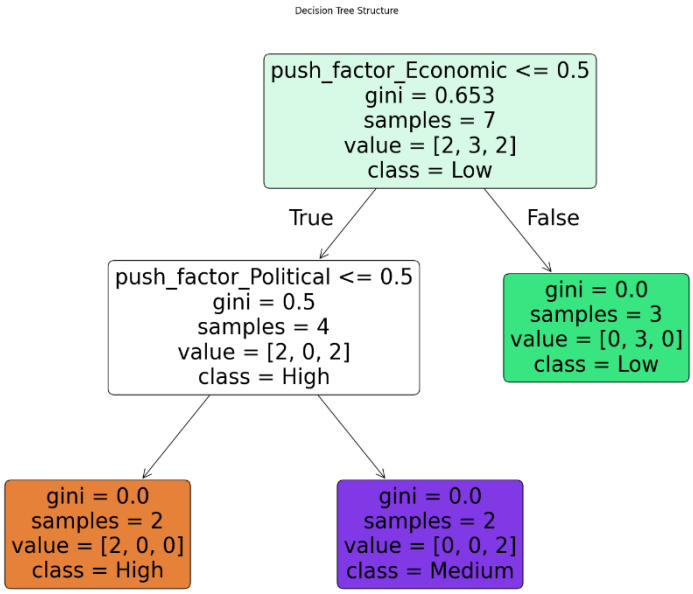


Figure 7. The Decision Tree structure. **Source:** *Made by the author*

The ethical conversation can be made more practical by detailing how border agencies could oversee and rectify bias in action. A specific suggestion is to conduct regular fairness evaluations of the decision tree's results. This illustration presents a straightforward example of how the decision tree differentiates various levels of risk by utilizing solely the push factors.

The initial division assesses if the situation is connected to economic motives; if it isn't, the tree examines political driving factors and subsequently categorizes the case as High, Medium, or Low risk. Although there are only a handful of sample cases, the framework clearly distinguishes between routine, lower-risk scenarios and more critical, security-relevant situations, demonstrating that the model adheres to the same risk principles that border agencies apply in reality.

5. Results

5.1. Decision Tree Outcomes and Attribute Effects

The assessment of the decision tree generated a sequence of filtering stages that allows for the differentiation among low, medium, and highrisk travelers based on observable traits like ctizenship, document type, push factors, and country of departure. Beginning with the complete passenger cohort, the model implements sequential divisions on these variables to generate smaller, more homogeneous subsets, where each route through the tree signifies a particular set of characteristics and an associated risk level. As the tree grows, high-risk instances become focused in a limited quantity of terminal nodes, which can subsequently be regarded as priority groups for secondary screening or thorough document verification by border authorities.

The findings indicate that risk arises not from individual attributes alone but from distinct patterns—such as certain document types together with exits from monitored countries, or triggering factors like conflict or state oppression. These patterns indicate that the model represents the causal relationships of the risk framework, where threat, vulnerability, and consequence interact to elevate risk, offering a clear foundation for operational decision-making at border crossing locations.

1. Departure Country:

The analysis pinpointed the departure country as a vital distinguishing factor. Travelers coming from conflict areas or particular transit locations exhibit stronger links to serious threat categories, while departures from stable EU countries generally correspond with low-risk departures.

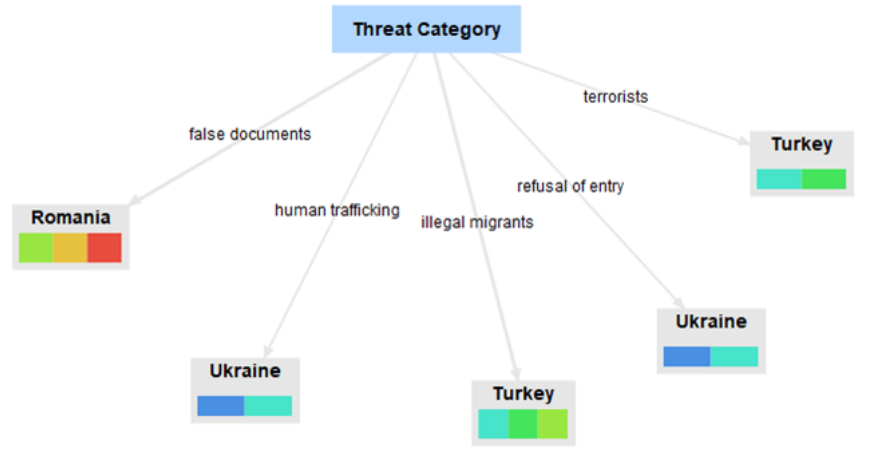


Figure 8: Simulation of a Risk analysis based on the departure country. **Source:** Made by author

In the following decision-tree diagram, most paths from Romania are rated low risk, but a minor branch has a red-dominant leaf. This sector highlights an uncommon yet high-risk trend seen in operations: travelers from the Middle East utilizing Romania not as an end location, but as a disordered transit hub prior to proceeding to Moldova or Western Europe. These profiles frequently show fragmented travel plans or discrepancies in documentation. The model's capability to distinguish this minority pattern from typically low-risk Romanian traffic showcases its sensitivity to complex, multi-leg migration pathways that might not be evident from individual variables alone.

2. Document Type

The type of travel document—passport, visa, or residence permit—serves as another major factor influencing the divisions. Passports from stable nations and utilized on direct flights usually result in blue-dominant leaves, signifying primarily low-risk results.

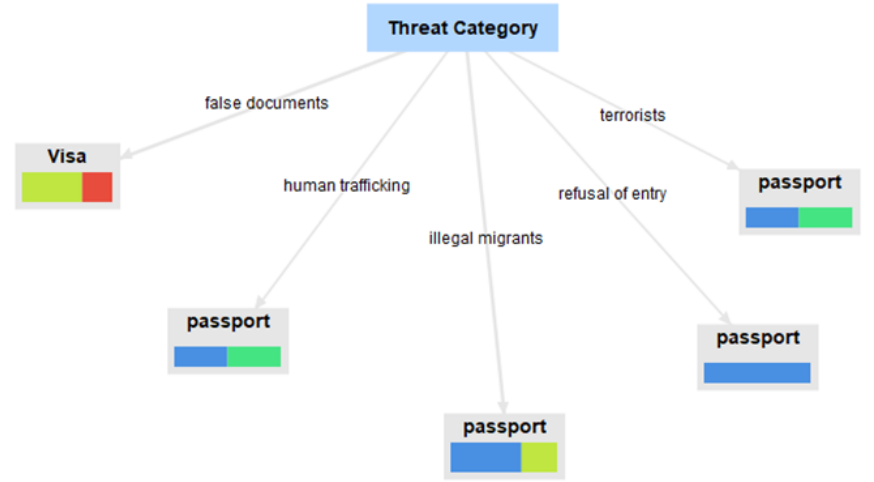


Figure 9: Simulation of risk analysis based on the type of documents

Source: Made by author

In contrast, the leaf linked to visa holders features a combination of yellow, green, and red segments, indicating a greater proportion of elevated and high-risk cases in this group. In practice, this indicates that visa-required travel, particularly from scrutinized departure nations, is more frequently linked to document fraud, human trafficking, or other significant risks compared to passport-only travel. Leaves that would display red only (if present) indicate document combinations where nearly all observed instances are high risk; these profiles should consistently be sent for secondary screening. For front-line officers, the guideline from this section of the tree is straightforward: discrepancies among visa type, route, and declared travel purpose warrant prompt further inspections.

3. Citizenship:

The model also emphasizes connections between citizenship and major driving factors like war, oppression, and poverty. Nations such as Afghanistan, Iran, Iraq, Syria, and Turkey emerge in areas where these driving factors are notably intense.

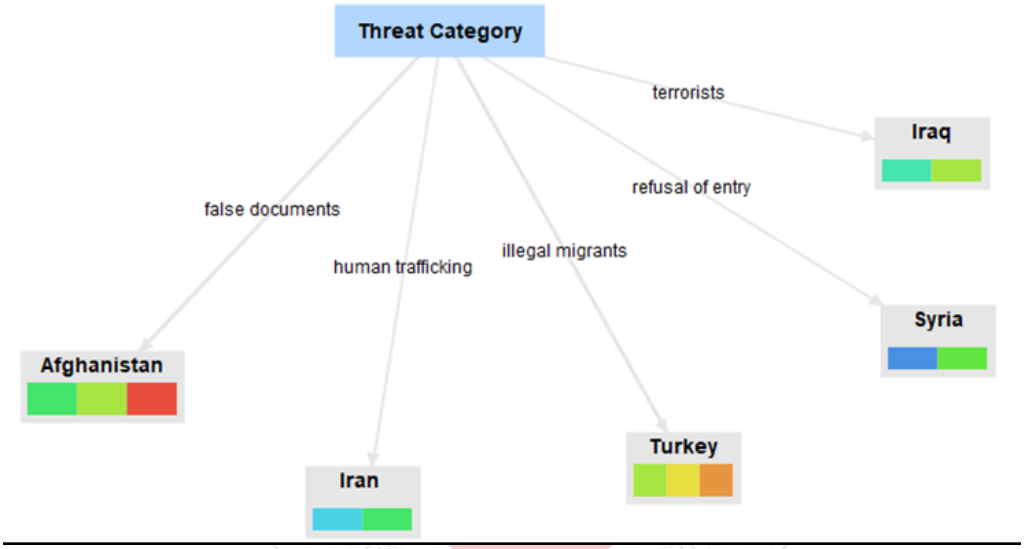


Figure 10: Risk analysis based on citizenship.

Source: Made by author

In branches where the leaves are only blue, the decision tree hasn't identified a concentration of significant indicators; individuals from those profiles still pose some risk, but it is quite low compared to other sections. Leaves that show red alongside yellow or green signify diverse groups: certain travelers in that node match high-risk profiles, while others align more with medium or low risk. Operationally, these mixed leaves indicate the necessity for more detailed inquiries and document verification, as the same route and nationality pairing might hide both authentic migrants and individuals posing a higher risk.

Leaves that display solely red signify routes where nearly all instances align with high-risk patterns (for instance, strong push factors paired with questionable document types or observed routes), and they illustrate the most crucial combinations throughout the entire tree that should automatically initiate second-line control.

5.2. Integrating the Decision Tree with the Risk Pyramid and Border Workflow

The three-tier risk pyramid, the First Line/Second Line process, and the decision tree framework present varying perspectives of the same procedure and are intended to function collaboratively. The decision tree functions initially in a pre-screening phase, utilizing characteristics like citizenship, document type, push factors, and country of departure to categorize each traveler into a low-, medium-, or high-risk leaf. This leaf subsequently identifies the traveler's place in the risk pyramid: the majority of passengers stay at the broad base as low-risk individuals, a smaller fraction ascends to the middle tier as medium-risk individuals, and merely a slim group attains the peak tier as high-risk individuals. Ultimately, the risk level outlined in the pyramid is incorporated into the operational workflow: low-risk profiles are handled at the First Line with standard checks or e-gates, medium-risk profiles undergo enhanced questioning or document validation, and high-risk profiles are consistently moved to the Second Line for thorough examination and potential engagement of specialized units.

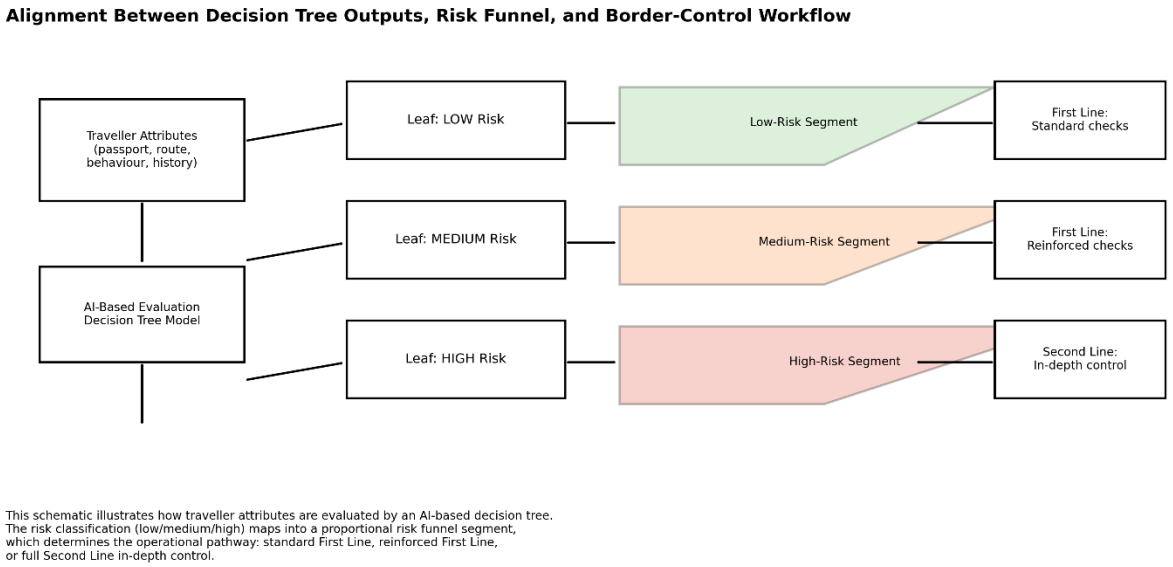


Figure 11: Alignment between decision tree outputs, risk funnel and border control

Source: *Made by author*

This diagram illustrates how a fictional traveler is managed within the system. The decision tree first assesses traveler attributes and assigns the traveler to a particular leaf (low, medium, or high risk). The associated risk level identifies the traveler's standing in the three-tier risk pyramid and activates an operational pathway: low-risk profiles are managed at First Line, medium-risk profiles undergo enhanced scrutiny, and high-risk profiles are elevated to Second Line for thorough examination. The diagram shows how the model's analytical results are transformed into transparent, proportional border-control measures.

6. Operational Implementation and Ethical Considerations

6.1. Operational Deployment

The deployment of decision-tree-based risk assessment represents a shift from intuitive to data-informed border screening. Based on operational experience at Chisinau International Airport and Frontex guidelines, the implementation follows a phased approach. Advance Passenger Information (API) data is processed through a pre-arrival decision tree that evaluates citizenship, visa type, departure origin, and push factors in real time. Passengers are then segmented into risk tiers: low-risk travelers proceed through automated e-gates, medium-risk through standard checkpoints, and high-risk travelers are flagged for secondary screening.

Crucially, the system functions as decision support, not autonomous decision-making. Officers receive transparent explanations for flags (e.g., "Visa holder departing conflict zone with multiple recent transits"), enabling targeted questioning while maintaining final human judgment. This optimizes resource allocation, allowing experienced officers to focus on genuine threats while improving both security and throughput. The primary operational challenges are data quality dependency and concept drift. Inaccurate API data propagates errors throughout the system, requiring robust data governance. Additionally, risk patterns evolve as migration routes and threat methodologies change, necessitating periodic retraining of models to maintain predictive accuracy.

6.2. Ethical Considerations and Bias Mitigation

Algorithmic risk assessment raises significant ethical concerns, primarily because it perpetuates historical biases. If the training data contains disproportionate screening rates for specific nationalities due to past profiling, the algorithm will replicate these patterns, creating a self-reinforcing cycle of discrimination.

Mitigation requires comprehensive bias audits before deployment to distinguish objective security indicators from historical discrimination. Algorithmic fairness constraints (e.g., equalized odds, demographic parity) should enforce equitable treatment across demographic groups. Human override capabilities must empower officers to release travelers when professional assessment contradicts algorithmic classification.

False positives pose serious consequences for legitimate travelers, particularly vulnerable populations like refugees. Graduated response protocols ensure proportional reactions—not all flags warrant intensive screening. Appeal mechanisms must allow individuals who are wrongly classified to clear their records and prevent frequent false positives.

Transparency is essential for legitimacy. Agencies should publish summaries of their decision rules and establish independent review boards to conduct periodic audits. Officer training must enable officers to clearly explain algorithmic logic to travelers.

6.3. Legal Compliance and System Limitations

International human rights frameworks (ICCPR,UN Guiding Principles) requires that automated systems effecting fundamental rights—be legally authorized, transparent, subject to human review, and non-discriminatory.

Decision trees can satisfy these requirements more effectively than opaque profiling methods, but only with deliberate institutional safeguards.

The system has clear boundaries: it requires sensitive passenger data, demanding GDPR compliance and data minimization. This model, trained on 12 airport cases, requires context-specific retraining for land or maritime borders. Critically, the system assesses risk only—it cannot adjudicate asylum claims, determine refugee status, or make final enforcement decisions. These remain within the purview of human judgment and legal frameworks.

6.4. Concrete Fairness and Bias Checks

The ethical debate can become more actionable by detailing how border agencies could oversee and rectify bias in real-world situations. A specific suggestion is to carry out regular fairness evaluations of the decision tree's results. Border officials might, for instance, monitor false-positive rates categorized by nationality, citizenship group, or common travel paths (e.g. particular origin–transit–destination routes) and evaluate these against general traffic proportions. If a specific group is consistently marked as high risk yet later checks almost always vindicate the travelers, this may indicate a potential bias in the model or the foundational data. Agencies could establish limits for permissible differences (for example, stipulating that false-positive rates for any group should not surpass a multiple of the average rate) and initiate model assessment or retraining if these limits are crossed.

Moreover, explicit human-override protocols ought to accompany implementation. Officers must be explicitly permitted to lower an automated high-risk recommendation when the in-person interview and document review do not corroborate the alert, and these overrides should be recorded for future analysis. Frequent examination of these override logs can uncover consistent trends, like particular branches of the tree that overestimate risk for certain profiles, which can subsequently inform precise model modifications. Collectively, these straightforward yet specific measures—tracking false positives by group, establishing disparity thresholds, and formalizing human override and logging—transform the broad ethical issues surrounding discrimination and transparency into practical governance strategies that border agencies can feasibly carry out.

6.5. Implementation Recommendations

Responsible deployment requires: (1) phased pot implementation at a single border crossing for 6-12 months with close outcome monitoring; (2) public transparency of decision rules and fairness audits; (3) mandatory

human-in-the-loop decision authority; (4) quarterly performance reviews with immediate retraining if bias emerges; (5) comprehensive officer training to build trust and understanding; and (6) interagency coordination with Frontex and neighboring countries for consistent standards.

7. Conclusion

This research shows that using decision trees on border-control data enables efficient risk classification. By examining factors such as nationality, type of documentation, and push factors, the model can differentiate among threat categories, such as human trafficking, terrorism, and unlawful migration, with push factors, such as conflict and poverty, being particularly enlightening for forecasting the type of threat. The use of automated pre-screening tools can greatly assist border guards in their decision-making, ensuring a high level of security while allowing legitimate travelers to flow smoothly. All figures and tables are numbered in order, in-text citations have been verified with the reference list, and a consistent Harvard citation style is used throughout the manuscript.

8. References

1. Bakker, P. (2013) 'Border Security and the Risk Society', Journal of Borderlands Studies, 28(2), pp. 234–248.
2. Bousquet, A. (2018) The Eye of War: Military Perception from the Telescope to the Drone. Minneapolis: University of Minnesota Press.
3. Breiman, L. (2001) 'Random Forests', Machine Learning, 45(1), pp. 5–32.
4. Department of Finance (2019) Maintaining a Risk Profile. Australian Government. Available at: https://www.finance.gov.au/sites/default/files/2019-11/Maintaining-a-Risk-Profile.pdf (Accessed: 1 December 2025).
5. European Border and Coast Guard Agency (Frontex) (2023) Risk analysis for 2023. Warsaw: Frontex.
6. European Border and Coast Guard Agency (Frontex) (2024) Annual risk analysis 2024. Warsaw: Frontex.
7. Kalanj, S. (2025). Gender-based violence in armed conflicts. International Journal of Contemporary Security Studies, 1(1), 125–138.
8. Liu, H. and Cocea, M. (2017) 'Semi-random decision tree for data stream classification', *International Journal of Machine Learning and Cybernetics*, 8(1), pp. 281–294.
9. National Institute of Standards and Technology (NIST) (2012) Guide for Conducting Risk Assessments (SP 800-30 Rev. 1). U.S. Department of Commerce. Available at: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-30r1.pdf.
10. Quinlan, J.R. (1986) 'Induction of Decision Trees', *Machine Learning*, 1(1), pp. 81–106.
11. Rokach, L. and Maimon, O. (2014) Data Mining with Decision Trees: Theory and Applications. 2nd edn. World Scientific Publishing.
12. UNHCR (2021) Global Trends: Forced Displacement in 2020. Geneva: The UN Refugee Agency. Available at: https://www.unhcr.org/50ababbe9.pdf.
13. Willis, H.H. (2007) 'Guiding Resource Allocations Based on Terrorism Risk', Risk Analysis, 27(3), pp. 597–606.
14. Zureik, E. and Salter, M.B. (2013) Global Surveillance and Policing: Borders, Security, Identity. Cullompton: Willan Publishing.