Comprehensive Report: Predicting Medal Performance for Great Britain at the 2025 World Athletics Championships in Tokyo

Introduction

This report outlines a structured analytical approach used to predict the potential medals Great Britain's track and field team could win at the 2025 Outdoor World Athletics Championships in Tokyo. The purpose of this analysis is to guide strategic decisions concerning training priorities, resource allocation, and performance strategy.

Objective

The primary objective is to predict:
- How many medals Great Britain is likely to win.
- The specific events where these medals might be secured.

Data Sources

The analysis utilized the following datasets:
- **GBR Competition Results (2019-2024)**: Historical performance data for British athletes.
- **Global Competition Results**: International athlete performance data for benchmarking.
- **GBR Athlete List & Global Athlete List**: Detailed athlete information.

Methodology

Data Understanding and Preparation

The datasets underwent a systematic preparation process:
1. **Data Loading & Exploration**: Initial analysis to identify the data structure and key variables.
2. **Data Integration**:
   - Combined athlete details with competition results.
   - Consolidated GBR and global datasets, introducing a 'Flag' to distinguish between national and global data.

Data Cleaning

The cleaning process included:
- **Filtering**: Focused on relevant track and field events, retaining only valid performance records.
- **Column Dropping**: Removed irrelevant columns such as "ValidPerformance," "CompetitionName," and "EventWind."
- **Standardisation & Formatting**: Ensured consistent data formats (dates, numerical values, categorical labels).

- **Duplicate & Missing Values Treatment**: Addressed special characters, duplicated rows, and filled missing values strategically to maintain dataset integrity.

Data Analysis & Visualization

Cleaned data facilitated meaningful insights:
- Evaluated historical GBR medal trends (2019-2024).
- Identified events where GBR historically excelled or lagged behind global competitors.
- Visualized performance gaps, team composition trends, and medal predictions using heatmaps and bar charts.

Threshold-Based Statistical Approach:

Due to inconsistencies and unreliability in the dataset's available features, applying traditional machine learning algorithms was impractical. To address this challenge, a threshold-based statistical approach was developed. This method involved establishing performance benchmarks using historical global top-three performances by event type and gender. British athletes meeting or surpassing these thresholds were identified as medal contenders. This innovative statistical method provided clear, interpretable predictions despite data limitations and ensured the robustness of the analysis.

Performance Analysis & Medal Prediction

Performance Benchmarking

A comprehensive benchmarking exercise compared GBR athletes against global top-three performances:
- **Event Classification**: Categorized events into 'time-based' and 'distance-based' for targeted analysis.
- **Global Top-3 Performers Identification**: Isolated the best three performers per event, providing a robust performance benchmark.

Medal Prediction Model

- **Threshold Setting**: Defined medal-contending thresholds based on the global top-three performances per event type and gender.
- **Identifying Medal Contenders**: GBR athletes whose performances matched or exceeded these thresholds were highlighted as potential medalists.

Key Predictions & Insights

Analysis predicted:
- GBR potentially winning **8 medals** across **6 events** at Tokyo 2025.
- Highlighted events include the 10000m, 5000m, 1500m, 800m, 400m, and Pole Vault.
- Notable trends identified include a consistent dominance of female GBR athletes in historical medal achievements.

Limitations

The prediction analysis acknowledges several limitations:
- Dependent exclusively on historical data provided.
- Fixed medal threshold criteria; predictions do not reflect a performance range.
- Does not incorporate emerging global or GBR talent, current athlete fitness, or injury statuses.
- Excludes real-world dynamics like recent performance improvements or declines and athlete retirements.

Strategic Questions for Further Analysis

To further refine the analytical model, the following questions were identified for UK Athletics Coaches and the Performance Leadership Team:
1. What additional athlete-specific or environmental factors significantly impact performance?
2. How is current athlete fitness and form being evaluated and tracked leading up to major competitions?
3. What specific interventions or support mechanisms are available for athletes who are borderline medal contenders or are currently underperforming?

Recommendations

- Integrate real-time athlete monitoring data (health, injury, recent form).
- Conduct regular reviews of emerging talent to update prediction models dynamically.
- Collaborate closely with coaching and performance teams to align data-driven insights with practical training and strategic decisions.

Conclusion

This detailed analytical process provides a robust foundation for predicting potential GBR medal performances at the 2025 World Athletics Championships. Strategic alignment of data insights with domain expertise will ensure optimal resource allocation and targeted athlete development, enhancing GBR’s competitive advantage on the global stage.