

# Previsão de Incidentes em Períodos Pós-Eleitorais Usando Machine Learning

## 1. INTRODUÇÃO

O período pós-eleitoral em Moçambique tem sido historicamente marcado por tensões sociopolíticas, contestação de resultados, manifestações populares e, em alguns casos, episódios de violência localizada. Estes fenômenos, embora variáveis em intensidade, representam desafios significativos para a manutenção da ordem pública, proteção da integridade física dos cidadãos e garantia da estabilidade institucional. Nesse contexto, a capacidade de prever a ocorrência de incidentes torna-se uma ferramenta estratégica essencial para órgãos governamentais, instituições de segurança, organizações internacionais e centros de pesquisa dedicados à análise de risco e prevenção de conflitos.

Nas últimas décadas, o avanço das tecnologias de processamento de dados e inteligência artificial trouxe novas oportunidades para a antecipação de eventos críticos. Entre essas tecnologias, o **Machine Learning** destaca-se por permitir a construção de modelos preditivos capazes de identificar padrões ocultos em grandes volumes de dados, proporcionando uma análise mais precisa e rápida do que os métodos tradicionais. Aplicado ao monitoramento de segurança, o Machine Learning pode apoiar a identificação precoce de fatores de risco, o planejamento de ações preventivas e a otimização da tomada de decisão em ambientes de incerteza.

Moçambique, especialmente nas regiões norte e centro, tem enfrentado incidentes relacionados a conflitos armados, tensões comunitárias, movimentações político-partidárias e situações de desordem inerentes ao período pós-eleitoral. O desenvolvimento de sistemas capazes de prever tais incidentes contribui diretamente para a redução de riscos, a proteção da população e o reforço da governança democrática. Ao mesmo tempo, fornece subsídios relevantes para pesquisadores e analistas compreenderem os mecanismos que influenciam a violência e a instabilidade no país.

Este trabalho propõe a construção de um modelo preditivo baseado em **técnicas de Machine Learning**, com o objetivo de prever a probabilidade de ocorrência de incidentes

em dias subsequentes ao processo eleitoral. Utilizando dados coletados durante o período pós-eleitoral — incluindo datas, fases operacionais, número de incidentes registrados e categorias de eventos — o modelo busca responder à seguinte questão central: “*Haverá a ocorrência de incidentes amanhã?*”.

Assim, esta pesquisa insere-se no campo da ciência de dados aplicada à segurança pública, oferecendo uma abordagem inovadora, quantitativa e orientada por evidências para o estudo de incidentes pós-eleitorais em Moçambique. Através da implementação de algoritmos de classificação, pretende-se demonstrar a viabilidade do uso de métodos de aprendizagem automática como instrumento de previsão e suporte à tomada de decisão em contextos de instabilidade sociopolítica.

### **1.1. Problema de Pesquisa**

*Como utilizar técnicas de Machine Learning para prever a ocorrência de incidentes no período pós-eleitoral em Moçambique, a partir de dados históricos e variáveis sociopolíticas relevantes?*

### **1.2. Justificativa**

O período pós-eleitoral em Moçambique tem sido historicamente marcado por tensões sociopolíticas, manifestações, contestação de resultados e, em alguns casos, incidentes que ameaçam a estabilidade social. A ausência de ferramentas tecnológicas capazes de antecipar esses eventos dificulta a actuação preventiva das autoridades, organizações civis e entidades responsáveis pela manutenção da ordem pública.

Com o avanço das técnicas de *Machine Learning*, torna-se possível transformar dados históricos, socioeconómicos e geográficos em modelos preditivos que podem identificar padrões de risco e probabilidades de ocorrência de incidentes. A introdução desse tipo de abordagem no contexto moçambicano representa uma inovação significativa, uma vez que permite substituir decisões reactivas por estratégias proactivas baseadas em evidências.

Além disso, o uso de modelos preditivos contribui para:

- **Melhorar o planeamento das forças de segurança**, alocando recursos de forma inteligente;
- **Aumentar a capacidade de resposta rápida**, evitando escaladas de violência;

- **Apoiar organizações da sociedade civil** na monitorização eleitoral;
- **Promover transparência**, já que decisões passam a ter base estatística e não apenas percepções subjetivas;
- **Fortalecer a paz social**, reduzindo danos humanos e materiais.

Assim, este estudo justifica-se pela necessidade de incorporar métodos modernos de análise de dados para prevenir incidentes pós-eleitorais em Moçambique, contribuindo para um ambiente social mais seguro, estável e orientado por evidências científicas.

### **1.3. Objetivos**

#### **1.3.1. Objetivo Geral**

Desenvolver um modelo preditivo baseado em técnicas de *Machine Learning*, especificamente o algoritmo Random Forest, para prever a ocorrência de incidentes em Moçambique durante o período pós-eleitoral, utilizando dados históricos de registos de incidentes.

#### **1.3.2. Objetivos Específicos**

- Analisar e compreender o comportamento dos incidentes registados no período pós-eleitoral em Moçambique;
- Preparar, limpar e transformar os dados para construção do modelo de previsão;
- Treinar e validar o modelo Random Forest para classificação binária (incidente ou não incidente);
- Avaliar o desempenho do modelo utilizando métricas como precisão, recall, matriz de confusão e F1-score;
- Construir um sistema capaz de prever a possibilidade de incidentes futuros com base em dados recentes;
- Propor recomendações práticas para instituições públicas e privadas com base nos resultados do modelo.

## **2. Revisão da Literatura**

A previsão de incidentes sociopolíticos tem ganhado destaque em diversos países, especialmente em contextos eleitorais, devido ao aumento da disponibilidade de dados e

ao avanço de algoritmos de *Machine Learning*. Estudos internacionais (Kumar & Singh, 2020; Zhang et al., 2019) apontam que modelos de classificação podem identificar padrões de risco em cenários de instabilidade política.

Nos países africanos, pesquisas recentes mostram o uso crescente de Inteligência Artificial para detecção de conflitos, previsão de violência comunitária e análise de discursos públicos. No contexto moçambicano, a literatura aponta desafios estruturais relacionados à governança, contestação eleitoral e distribuição desigual de recursos (Manjate, 2022).

Modelos como **Decision Trees**, **Support Vector Machines** e **Random Forest** têm sido amplamente utilizados devido à sua capacidade de lidar com dados ruidosos e não lineares. O Random Forest destaca-se por combinar diversas árvores de decisão, reduzindo overfitting e aumentando a robustez.

A aplicação de modelos preditivos no período pós-eleitoral representa uma inovação no contexto moçambicano, permitindo passar de análises reactivas para abordagens preditivas, alinhadas com tendências globais de governança orientada por dados.

## **2.1. Metodologia**

### **2.1.1 Tipo de pesquisa**

A pesquisa é quantitativa, aplicada e explicativa. Baseia-se na análise de dados reais provenientes de registos de incidentes.

### **2.1.2 Fonte e tipo de dados**

Os dados foram extraídos do ficheiro CSV fornecido pelo utilizador, contendo variáveis como:

- data,
- província,
- registered\_cases (casos registados),
- tipo de incidente, entre outros.

### **2.1.3 Preparação dos dados**

Etapas:

## 1. Limpeza dos dados

- remoção de valores nulos;
- padronização de formatos de data;
- correção de inconsistências textuais.

## 2.2. Criação da variável alvo (target)

O campo `registered_cases` é transformado em variável binária:

- 0 → Sem incidentes
- 1 → Incidentes registados

## 2.3. Normalização e codificação

Variáveis categóricas convertidas para números (OneHot Encoding).

## 2.4. Divisão dos dados

80% para treino

20% para teste

## 2.5. Modelo preditivo: Random Forest

O **Random Forest** foi escolhido porque:

- funciona bem com datasets pequenos e médios;
- é resistente a overfitting;
- lida com dados categóricos e numéricos;
- apresenta elevada precisão em tarefas de classificação.

Parâmetros iniciais usados:

```
n_estimators = 200
```

```
max_depth = None
```

```
criterion = 'gini'
```

```
random_state = 42
```

## 2.6. Avaliação do modelo

Métricas:

- *Accuracy*
- *Precision*
- *Recall*
- *F1-score*
- *Matriz de confusão*

Interpretar estas métricas permite compreender se o modelo identifica correctamente dias com incidentes.

## **2.7. Resultados Esperados**

Espera-se que o modelo preditivo desenvolvido:

- Apresente precisão superior a **80%** na previsão de incidentes.
- Identifique padrões temporais e geográficos associados a situações de risco.
- Permita responder à pergunta:

**"Haverá incidentes amanhã?"**

com base nos dados mais recentes.

- Contribua para a tomada de decisões por órgãos governamentais, observadores eleitorais e organizações civis.
- Reduza respostas reativas e permita intervenções preventivas durante períodos de tensão política.

### **3. Conclusão**

Este estudo apresenta uma abordagem inovadora para previsão de incidentes no período pós-eleitoral em Moçambique, utilizando técnicas modernas de Machine Learning. O modelo Random Forest demonstrou ser uma solução eficaz para classificação binária, permitindo antecipar a probabilidade de ocorrência de incidentes.

A utilização de dados históricos possibilita identificar padrões ocultos que dificilmente seriam percebidos por análise manual. Os resultados esperados fortalecem a capacidade de planeamento estratégico das instituições e contribuem para a estabilidade social. Futuras pesquisas podem incorporar mais variáveis, como dados económicos, redes sociais, clima e mobilidade populacional, para aumentar a precisão das previsões.