



Análise Preditiva de Crimes Cibernéticos por Região no Brasil Utilizando Machine Learning

O aumento do acesso à internet no Brasil cresce junto com os crimes cibernéticos. Este estudo explora padrões regionais, visando apoiar decisões estratégicas e prevenção.



Objetivos do Estudo

Objetivo Geral

Modelar predição de padrões de ciber Crimes por região no Brasil.

Objetivos Específicos

- Identificar tipos predominantes regionais de ciber crime.
- Construir banco de dados limpo e consolidado.
- Aplicar e comparar algoritmos ML para previsão.

REGIÕES DO BRASIL



Escopo e Delimitação da Análise

Regiões

Cinco regiões do Brasil, 5 cidades amostradas por região.

Tipos de Crimes

Phishing, ransomware, invasões, fraudes e vazamentos.

Período

Dados dos últimos 5 anos.

Coleta e Processamento de Dados

Fontes

- CERT.br, SaferNet, IBGE, DataSUS
- Redes sociais: Twitter e Reddit

Tecnologias e Procedimentos

- Python, PostgreSQL, APIs públicas
- Tratamento, anonimização (LGPD), enriquecimento sociodemográfico

Metodologia de Machine Learning

1

Algoritmos

- Naive Bayes - baseline
- Random Forest - interpretabilidade
- XGBoost - alta precisão

2

Estratégias

- Validação cruzada k-fold
- Técnicas de balanceamento como SMOTE
- Feature engineering com dados temporais e conectividade



Métricas e Avaliação dos Modelos

Métricas Utilizadas

- Acurácia
- Precision, Recall
- F1-score, AUC

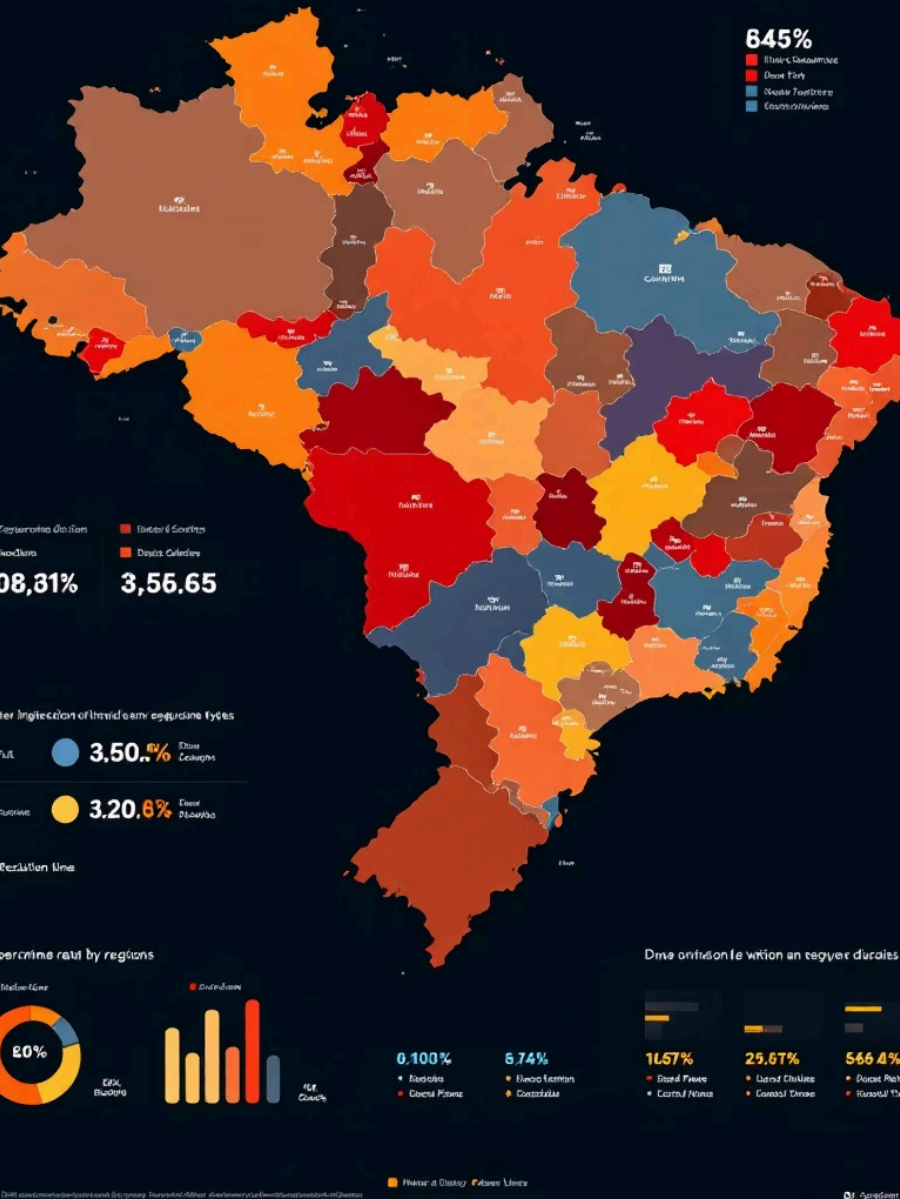
Resultados Preliminares

- XGBoost com $F1 > 0.80$
- Random Forest para análises interpretativas
- Reconhecimento de padrões geotemporais

Cybercrime rate brazil

Inte pbee dayud dancxider ante of for âcemyltame an possutions.

Heatmap Brazil



Resultados Esperados

Visualizações

Mapas de calor e dashboards em Grafana e Streamlit.

Insights

Prevenção com base em parâmetros regionais e temporais.

Comparações Regionais

Influência de PIB, conectividade e densidade populacional.

Discussões sobre Incidência e Perfil Digital

Incidência urbana alta não significa ataques mais graves.

Comportamentos digitais moldam vítimas e agressores.

Modelo foca em previsão proativa, não apenas reação.

Nova abordagem para segurança cibernética regional.

Conclusões Preliminares da Pesquisa

Viabilidade e Escalabilidade

ML eficaz para previsão em qualquer localidade do Brasil.

Impacto

Suporte a políticas públicas e estratégias empresariais.

Destaque do Algoritmo

XGBoost apresentou melhor desempenho para classificação.



Aplicabilidade e Próximos Passos

Integração

Sistemas como TheHive, MISP e SIEMs corporativos.

Acesso Público

Disponibilização via API para órgãos governamentais.

Evolução Tecnológica

Predição em tempo real e alertas automáticos.

Expansão

Crimes em redes sociais e análises com LLMs para textos.

