

SLIDE 01 · CAPA DO RELATÓRIO

Classificação de notas bancárias com Machine Learning como base para discutir IA com rigor.

Este deck retoma o relatório acadêmico original, reorganiza os achados em linguagem visual e transforma os resultados em uma peça pronta para circulação profissional.

UNIFOR · Engenharia da Computação

Mateus Gomes Macário

Líder Projetos · Abril de 2026



O relatório foi reorganizado em sete blocos fáceis de navegar.

01

Problema e objetivo do estudo

02

Dataset, features e protocolo de validação

03

Resultados de acurácia, pr

04

Latência e custo de inferência

05

Recomendação por cenário de uso

06

Ponte entre pesquisa, meta

Detectar notas falsas é um problema operacional. Validar IA antes da produção também.

Contexto do estudo

- notas falsas geram prejuízo direto e exigem inspeção confiável
- o dataset usa quatro features extraídas por Transformada Wavelet
- o objetivo foi comparar precisão e eficiência entre famílias clássicas

Objetivo principal

Comparar cinco classificadores implementados do : oferece melhor aderência para ambientes controlac discussão de governança de IA.

O dataset já nasce limpo,
balanceado e adequado
para classificação
supervisionada.



Features utilizadas

Variância

Assimetria

Curtose

Entropia

FONTE

UCI Machine Learning Repository · Banknote Authentication

CLASSES

55,5% autên

VALIDAÇÃO

10-fold cross validation com random_state=42

LEITURA

base peque
separação c

Cada feature foi resumida como faixa observada, média e dispersão.

Variância

DP 2,84

-7,04

média 0,43

6,82

Assimetria

-13,77

média 1,92

Curtose

DP 4,31

-5,29

média 1,40

17,93

Entropia

-8,55

média -1,19

Foram comparadas duas famílias com cinco configurações distintas.

K-NN

K-NN Euclidiana

cálculo de distância para toda a base

- treino virtualmente instantâneo
- máxima precisão em ambiente controlado

K-NN

K-NN Manhattan

mais robusto a outliers que a Euclidiana

- treino instantâneo
- equilíbrio intermediário

K-NN

K-NN Chebyshev

distância máxima por eixo

- treino instantâneo
- laboratórios e validação c

NAIVE BAYES

Naive Bayes Univariado

muito rápida, mas com hipótese fraca

- modelo probabilístico simples
- baseline probabilístico

NAIVE BAYES

Naive Bayes Multivariado

rápida e adequada para escala

- estimativa de média e covariância
- produção em larga escala

O protocolo foi simples, robusto e fácil de explicar para qualquer banca.

10 folds

Fold 01

Fold 02

Fold 03

Fold 04

Fold 05

Fold 06

Fold 07

Fold 08

Fold 09

Fold 10

Cada modelo foi avaliado em todas as partições, reduzindo o risco de resultado acidental e melhorando a leitura estatística da comparação.

Métricas monitoradas

Acurácia

Precisão

F1-score

Tempo de treino

Tempo

O relatório não olha apenas para acerto. Ele também considera o classificador em uma rotina real.

K-NN dominou a precisão; Naive Bayes Multivariado tornou a escala viável.

K-NN Euclidiana

K-NN

ACURÁCIA

100,00%

PRECISÃO

100,00%

F1

100,00%

K-NN Manhattan

K-NN

ACURÁCIA

99,93%

PRECISÃO

99,86%

F1

99,93%

K-NN Chebyshev

K-NN

ACURÁCIA

100,00%

PRECISÃO

100,00%

F1

100,00%

Naive Bayes Univariado

Naive Bayes

ACURÁCIA

83,97%

PRECISÃO

84,10%

F1

81,42%

Naive Bayes Multivariado

ACURÁCIA

PRECISÃO

F1

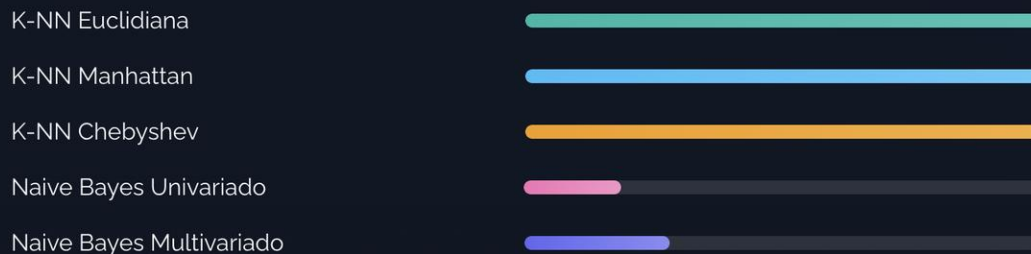
Naive Bayes reduz drasticamente o custo de inferência no tempo de teste.

GANHO DE ORDEM DE GRANDEZA

~140X

O Naive Bayes Multivariado entrega uma latência muito menor sem derrubar a taxa de acerto a um nível impraticável.

Tempo de teste por classificador



A escolha certa depende do volume, do risco e do custo que a operação tolera.

MÁXIMA PRECISÃO

K-NN Euclidiana ou Chebyshev

Para validações altamente controladas e baixa tolerância a erro, a família K-NN se destaca com classificação perfeita no experimento.

PRODUÇÃO EM LARGA ESCALA

Naive Bayes Multivariado

Quando o problema envolve grande volume, latência e custo operacional, o equilíbrio entre 98,54% de acurácia e tempo de resposta muito menor muda a decisão.

ESTRATÉGIA HÍBRIDA

Naive Bayes para triagem limitrofes

A combinação usa velocidade e reserva precisão máxima p

A melhora do Naive Bayes Multivariado acontece porque as features carregam relação entre si.

Matriz qualitativa de correlação

	Var.	Ass.	Cur.	Ent.
Var.	1,0	alta	moderada	baixa
Ass.	alta	1,0	alta	inversa
Cur.	moderada	alta	1,0	moderada
Ent.	baixa	inversa	moderada	1,0

Leitura rápida

- o K-NN se beneficia da separabilidade clara entre
- o Naive Bayes Univariado perde força ao ignorar
- o Naive Bayes Multivariado se aproxima do topo a

Cada família ganha em um eixo e paga preço em outro.

K-Vizinhos

VANTAGENS

- simples de explicar
- sem hipótese sobre distribuição dos dados
- desempenho excelente em dados bem separados

LIMITAÇÕES

- não escala bem em memória e latência
- sensível a features irrelevantes
- depende de normalização e métrica correta

Naive Bayes

VANTAGENS

- rápido para treinar e testar
- baixo custo computacional
- alto valor em produção massiva

LIMITAÇÕES

- hipótese de independência
- falha com dados desbalanceados
- requer normalização
- não funciona bem com dados contínuos

O relatório recomenda classificador por situação, e não por torcida tecnológica.

K-NN Euclidiana ou Chebyshev

Quando a prioridade é acertar tudo.

- baixa tolerância a erro
- dataset pequeno ou médio
- validação laboratorial e auditoria final

Exemplo: decisão crítica em laboratório forense.

K-NN Manhattan

Quando existe espaço para um compromisso fino.

- dados com ruído
- outliers relevantes
- pipeline de médio porte

Exemplo: triagem com busca de near-perfect accuracy.

Naive Bayes Univariado

Quando a simplicidade é o foco, mesmo com perda de qualidade.

- baseline rápido
- hardware limitado
- teste exploratório

Exemplo: protótipo inicial em ambiente restrito.

Naive E

Quando p
na conta.

- alto vol
- feature
- escala

Exemplo: :
SLA defini

A pesquisa mostra que precisão e eficiência precisam ser lidas juntas.

Dados e features definem o teto

A Transformada Wavelet foi decisiva para separar padrões entre notas autênticas e falsas. Em IA moderna, a mesma lógica vale para contexto, estrutura e qualidade de entrada.

Correlações importam

A diferença entre Naive Bayes Univariado e Multivariado mostra que ignorar relações entre variáveis custa desempenho real.

Velocidade também

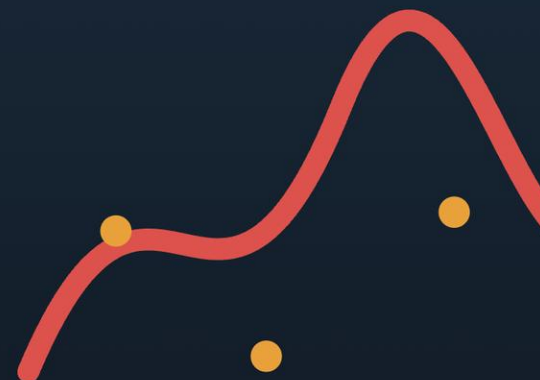
100% de acurácia impressio
transacional mudam a esco

O mesmo raciocínio estatístico do estudo serve para avaliar prompts e agentes.

Underfitting



Overfitting



Toda decisão de modelo é uma negociação entre erro sistemático e sensibilidade ao ruído.

ALTO VIÉS

subajuste

equilíbrio

sol

ALTA VARIÂNCIA

LEITURA PARA ML

features e regularidade do dataset mudam o ponto ótimo

O dataset de notas falsas é pequeno, limpo e com boa engenharia de features.

LEITURA PARA PROMPTS

exemplos demais também podem viciar o comportamento

Meta prompt bom não é o mais longo. É o que generaliza melhor para cenários novos.

LEITURA PARA PRODUTO

medir evita confundi

Equipe madura documenta produção.

Criar meta prompts sem medir é repetir o erro de um modelo sem validação.

O caminho frágil

- prompt ajustado só por impressão subjetiva
- testes curtos e não versionados
- ausência de dataset de avaliação
- aprovação por entusiasmo, não por evidência

O caminho de engenharia

- métricas definidas antes da iteração
- casos comuns, edge cases e casos adversariais
- registro das alterações e reexecução completa
- observabilidade, segurança e documentação des

O mesmo deck vira um framework simples para avaliar IA antes da homologação.

FRAMEWORK DEVSECOPS PARA IA RESPONSÁVEL



01

Definir métricas

acurácia factual, formato, segurança, conformidade e rastreabilidade.

02

Criar da

casos cor adversari

03

Avaliar sistematicamente

execução repetível, scoring claro e comparação entre versões

04

Iterar c

ajuste cir decisão e

Meta prompts e fine-tuning não competem no vazio. Eles respondem a maturidades diferentes.

Meta prompts

SETUP

baixo a médio

ITERAÇÃO

rápida

GOVERNANÇA

precisa de dataset e critérios

QUANDO USAR

times que estão estruturando a camada de IA

Fine-tuning

SETUP

alto

ITERAÇÃO

mais

GOVERNANÇA

dataset versionado e operação madura

QUANDO USAR

com
esca

Os princípios da pesquisa continuam válidos quando o assunto vira engenharia de IA.

Dados e features importam

Entradas boas continuam sendo o ativo mais subestimado de qualquer pipeline inteligente.

Medir é essencial

Sem benchmark, qualquer melhoria vira narrativa e não evidência.

Overfitting não desapareceu

Ele só mudou de roupa: pode morar no modelo, no dataset e no prompt.

Produção

Rastreamento adversário.

Pesquisa aplicada vira ativo público quando ganha método, narrativa e distribuição.

Referências centrais

- UCI Machine Learning Repository · Banknote Authentication Data Set
- UNIFOR · Disciplina de Inteligência Artificial Computacional
- Machine Learning Mastery · Bias-Variance Trade-Off
- GitLab 2024 Developer Survey · AI, segurança e produtividade
- Black Duck 2024 DevSecOps Report

URLs de distribuição

- <https://www.liderprojetos.com/provadeconceito/pesquisa/ia>
- <https://www.liderprojetos.com/provadeconceito/pesquisa/ia/devsecops/apresentacao>
- <https://www.liderprojetos.com/provadeconceito/pesquisa/ia>
- <https://www.liderprojetos.com/provadeconceito/pesquisa/ia>

Esta versão existe para PDF e distribuição. A apresentação edi separadamente para leitura navegável.