

Avaliação Empírica da Regressão Logística



Autores:

- Alexandre Furriel
- Catarina Abrantes
- Liliana Silva

Curso: Inteligência Artificial e Ciência de Dados

UC: Aprendizagem Computacional

Sumário Executivo

Objetivo:

Combater o desbalanceamento de classes, usando regressão logística.

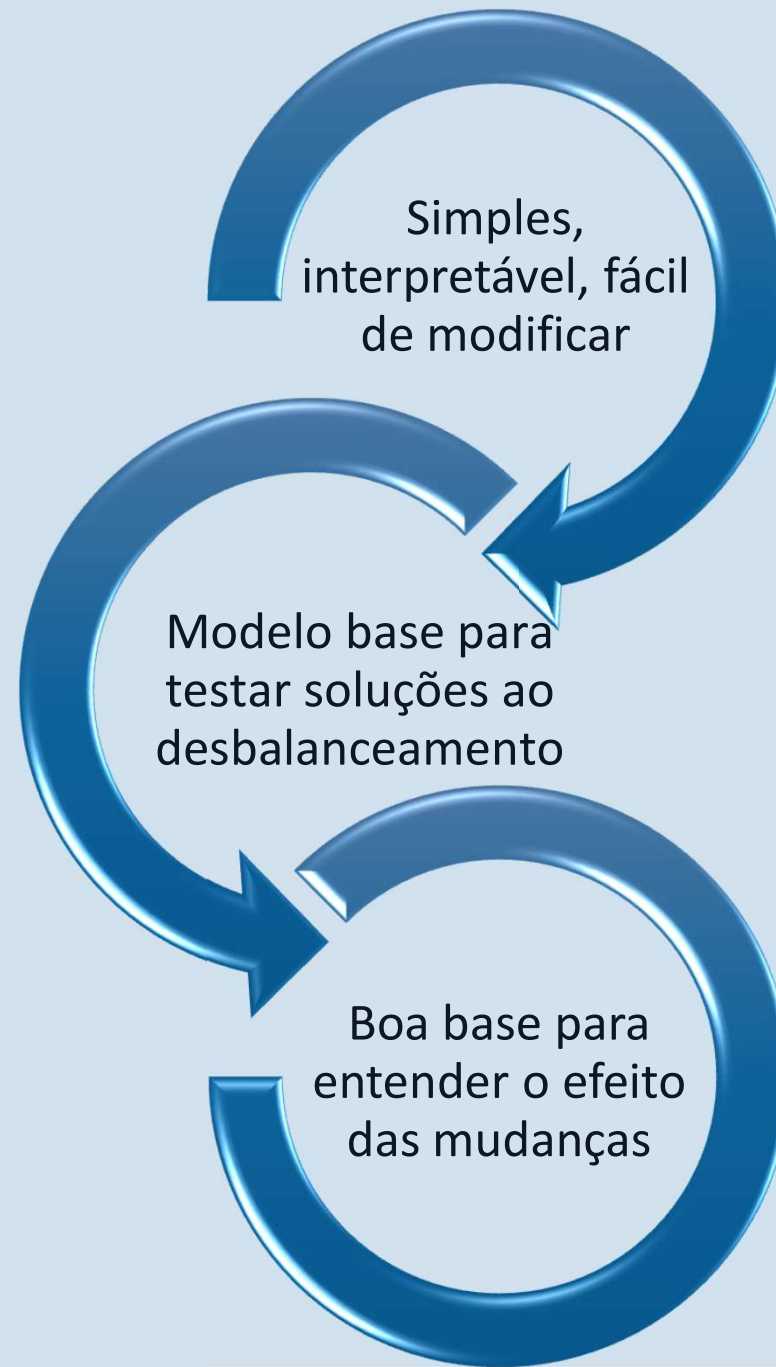
Abordagem:

Implementação Estrutural do Modelo +
Teste de Estratégias Alternativas de Perda

Conclusão:

Alterações melhoram o desempenho, especialmente na classe minoritária.

Porquê a Regressão Logística?



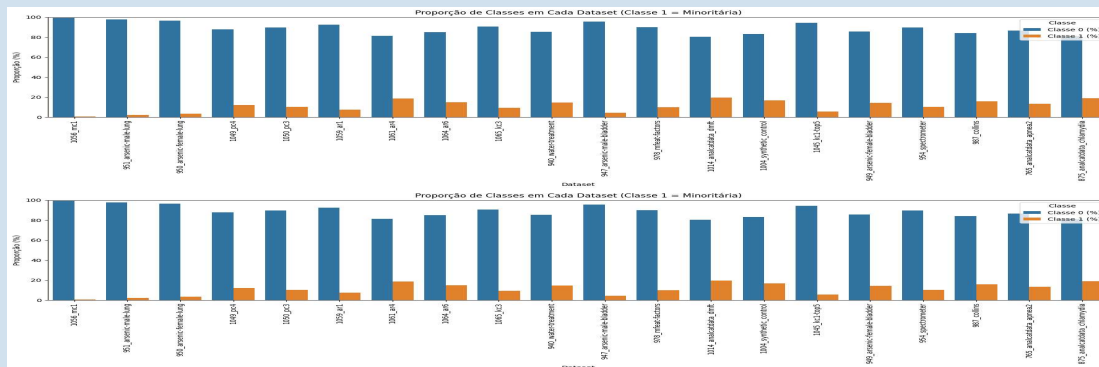
Desbalanceamento de Classes

O que é?: Trata-se de uma situação em que uma das classes possui muito menos exemplos do que a outra; É comum nos Datasets utilizados.

Como afeta a Regressão Logística?: A função de custo não diferencia entre classes; Modelo favorece a classe majoritária (Classe 0).

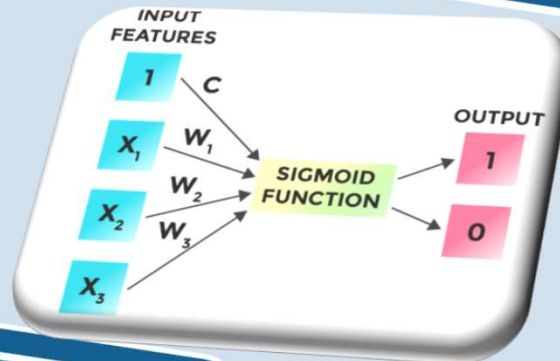
Impacto prático: Acurácia enganadora (pode ser alta mesmo errando toda a classe minoritária); Recall e F1-score da Classe 1 tornam-se muito baixos.

Motivação para atuar: Precisamos de mecanismos para dar mais peso à classe minoritária e corrigir esse viés.



Funcionamento da Regressão Logística

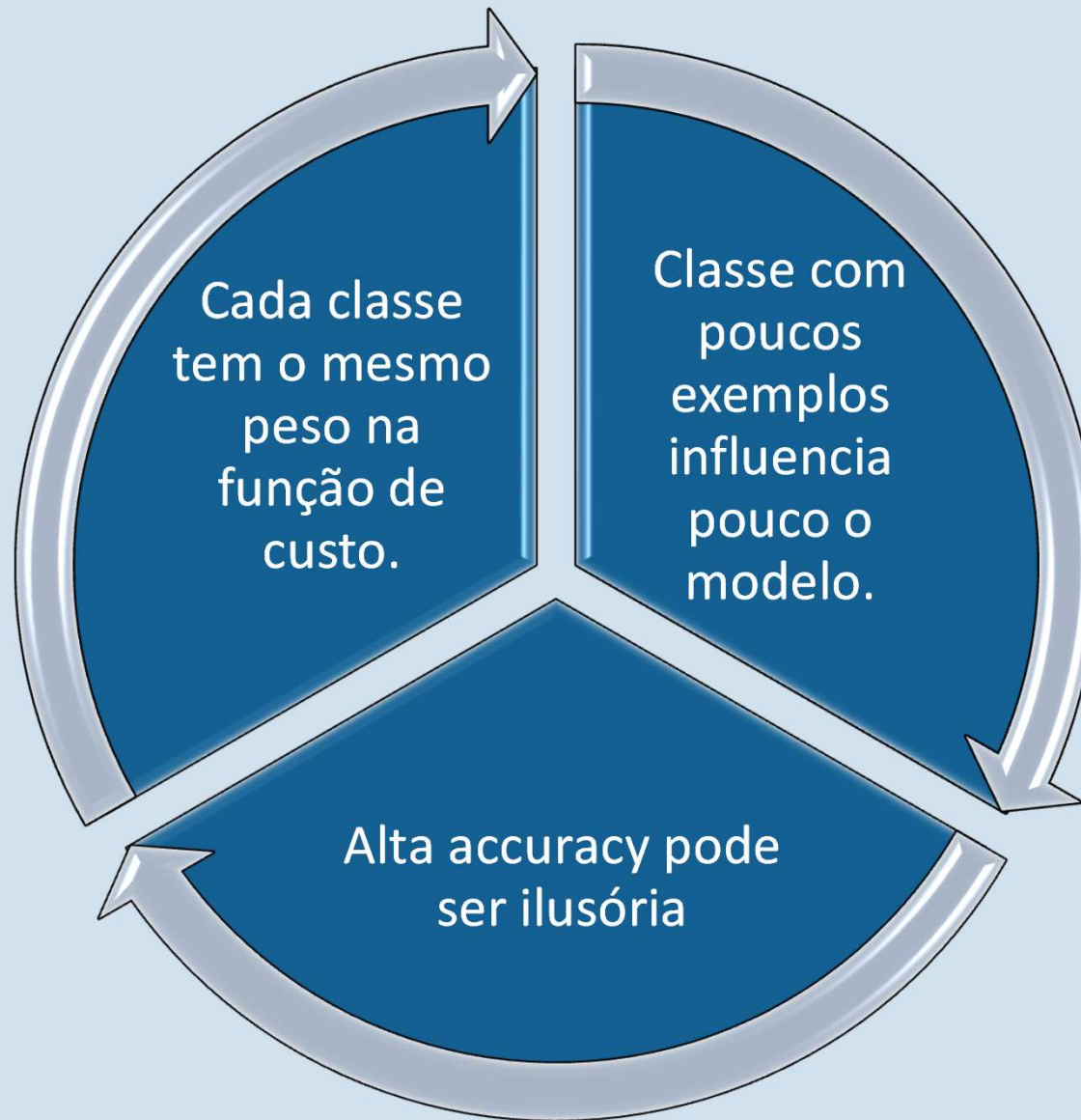
Fórmula: $\hat{y} = \sigma(w^t x + b)$



Função de custo:
entropia cruzada.

Função sigmoide:
transforma entradas em
probabilidades.

Limitações da Regressão Logística



Preparação dos Dados

20 datasets benchmark com desbalanceamento real



Distribuição de classes preservada

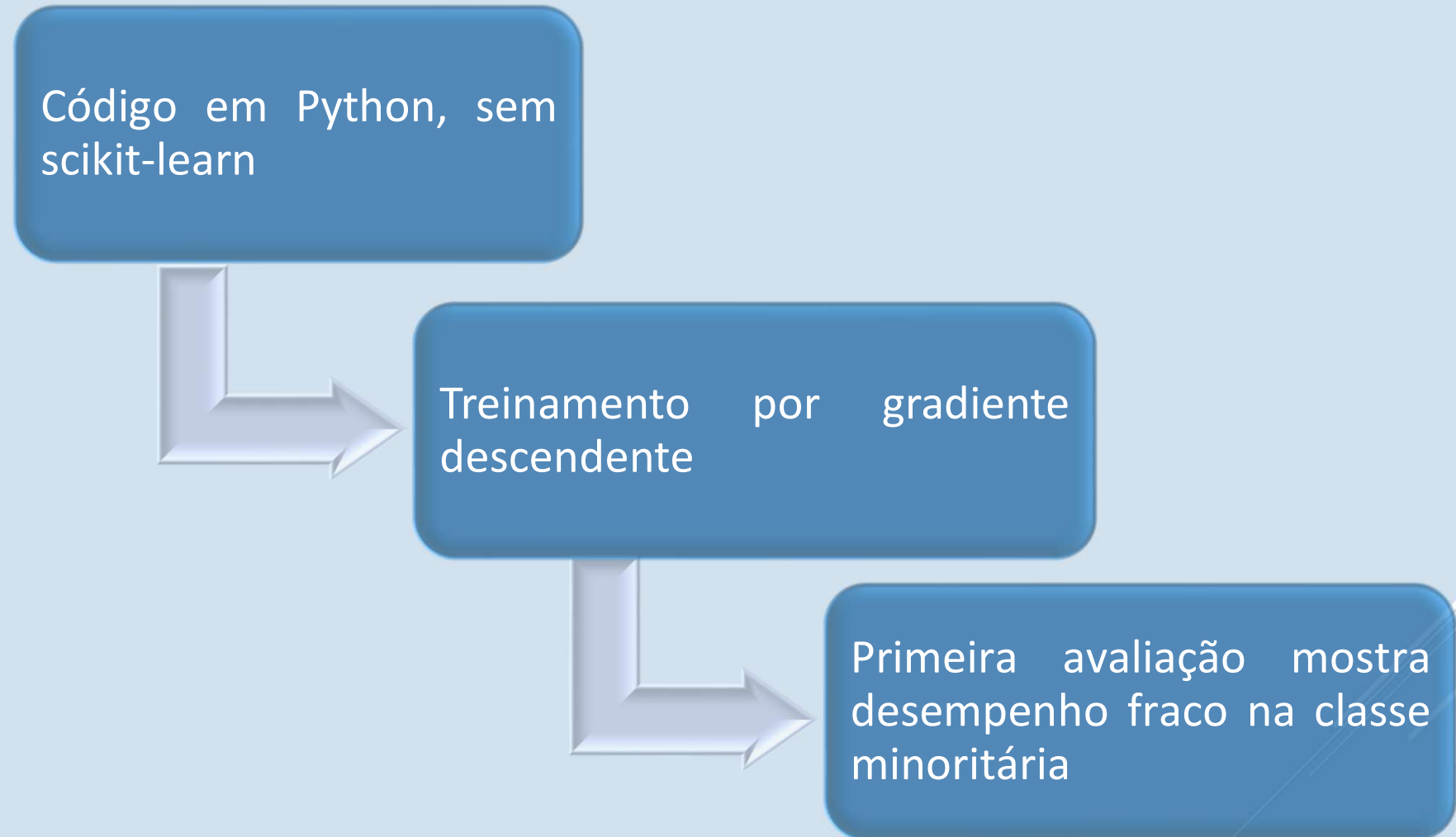


Divisão treino/teste estratificada



Amostras independentes e representativas

Implementação do Modelo Base

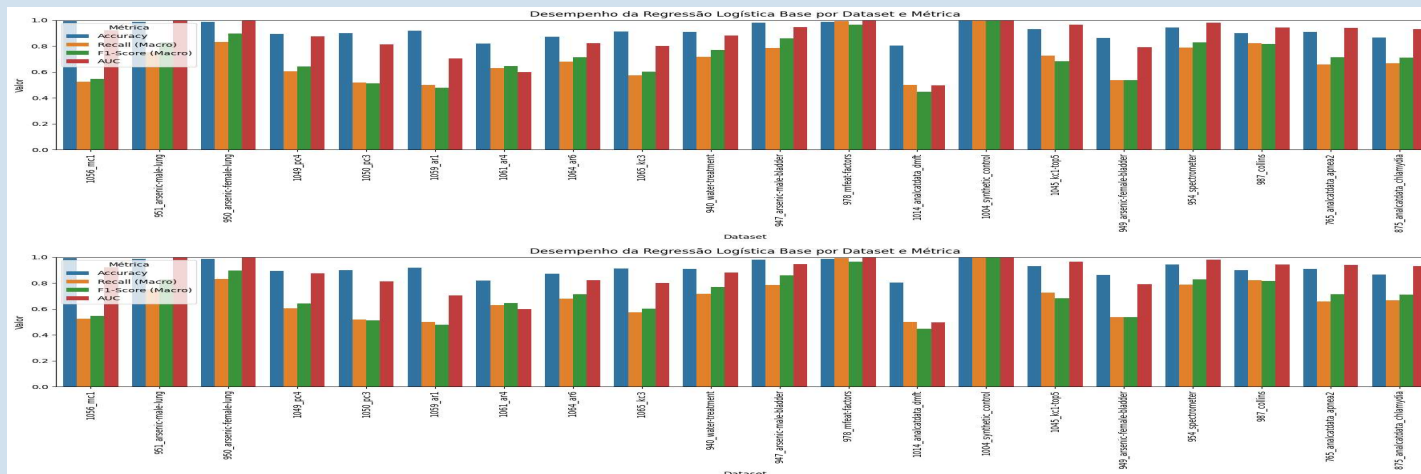


Resultados do Modelo Base

Acurácia média $\approx 91\%$ (aparentemente bom). - Recall e F1-score (Macro) são significativamente mais baixos.

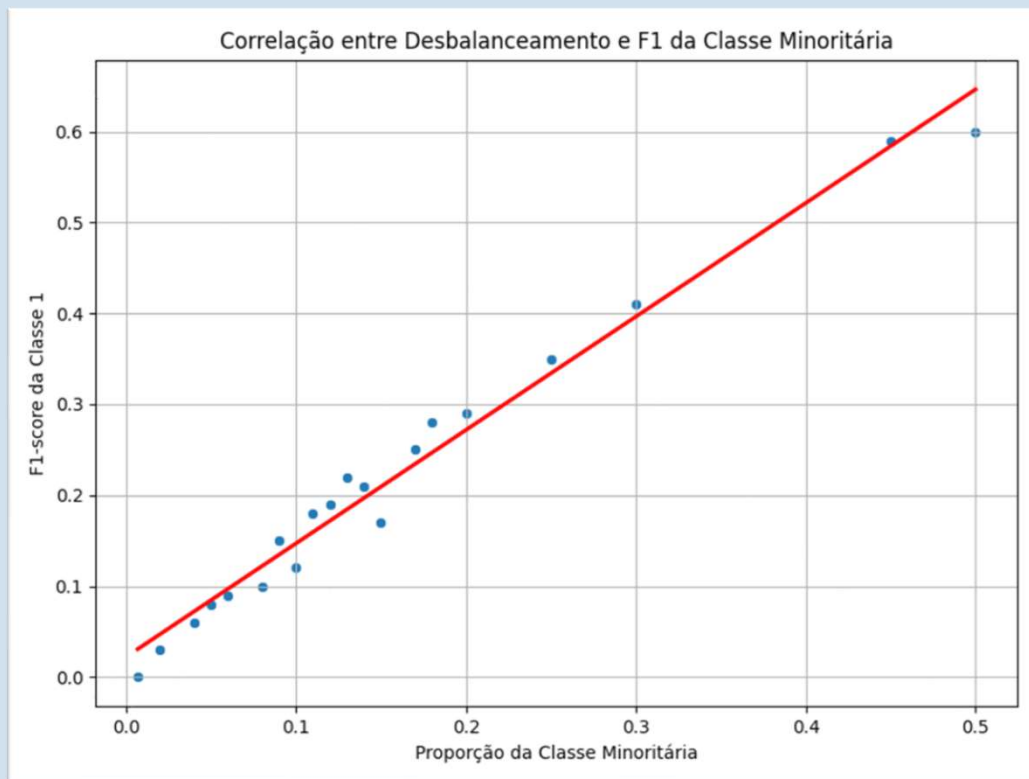
Reflete falhas graves na previsão da classe minoritária.

Justifica a necessidade de alterar a função de perda.



Accuracy	0.9191
Recall (Macro)	0.6906
F1-Score (Macro)	0.7101
AUC	0.8709

Desbalanceamento e F1 da Classe Minoritária



Correlação entre a proporção da Classe 1 e o F1-score da Classe 1



Observa-se uma correlação linear positiva clara.



Quando a Classe 1 é menos representada, o F1-score tende a ser muito mais baixo.



Isto confirma quantitativamente o impacto do desbalanceamento no desempenho do modelo.



Reforça a necessidade de soluções que compensem a baixa representatividade.

Solução 1: ASL (Adaptive Sample Loss)

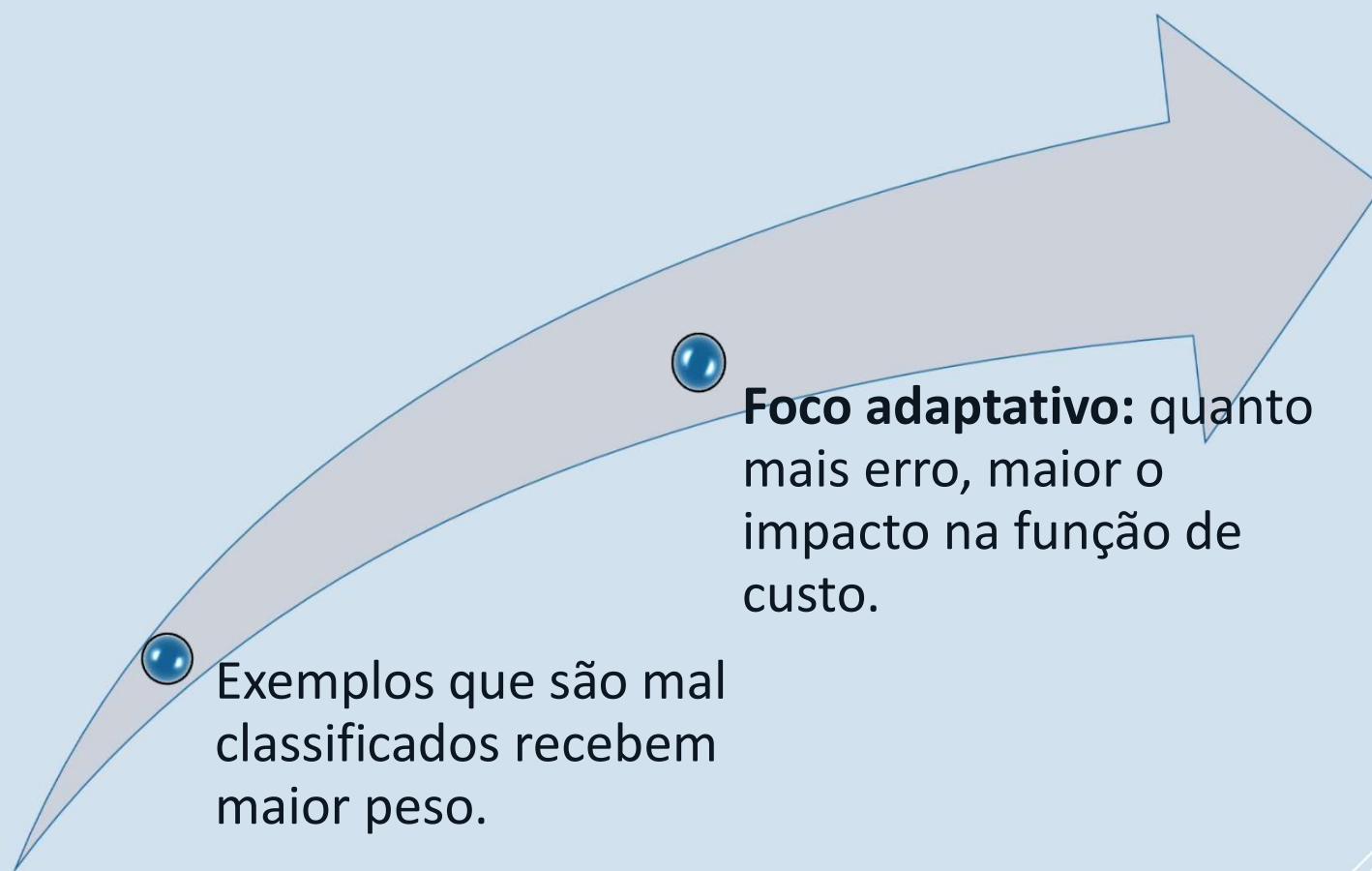
Motivação:

Aumentar a penalização de exemplos difíceis.

Ideia:

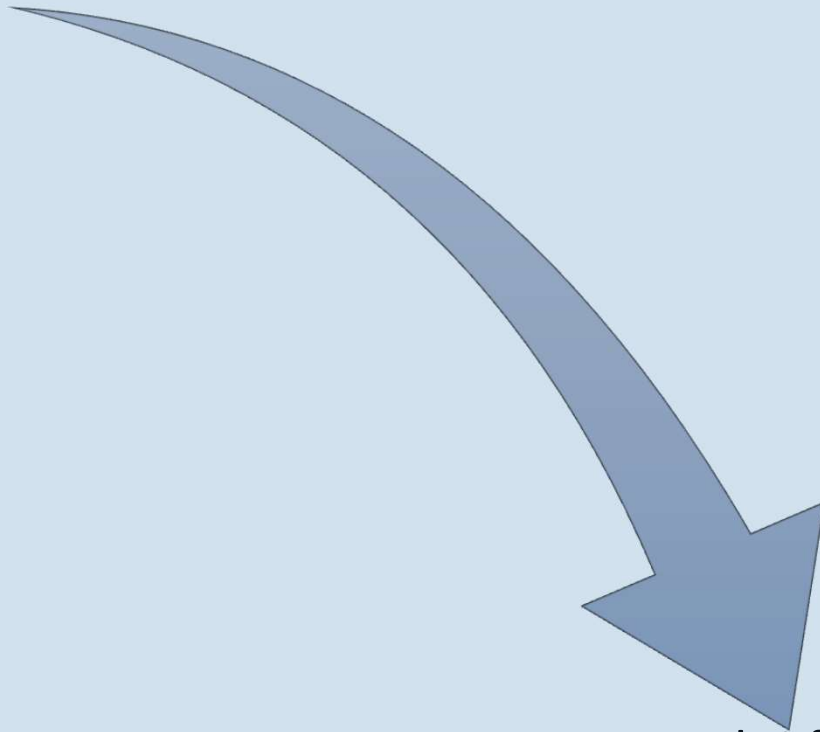
Dar mais peso a instâncias que o modelo tem dificuldade em prever.

Como funciona a ASL?



Implementação da ASL

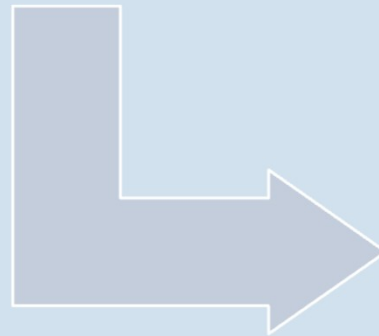
Modificação direta da função de perda.



Inclui fator adaptativo baseado na confiança do modelo.

Resultados com ASL

Melhor equilíbrio nas métricas



F1-score da classe minoritária aumentou significativamente.

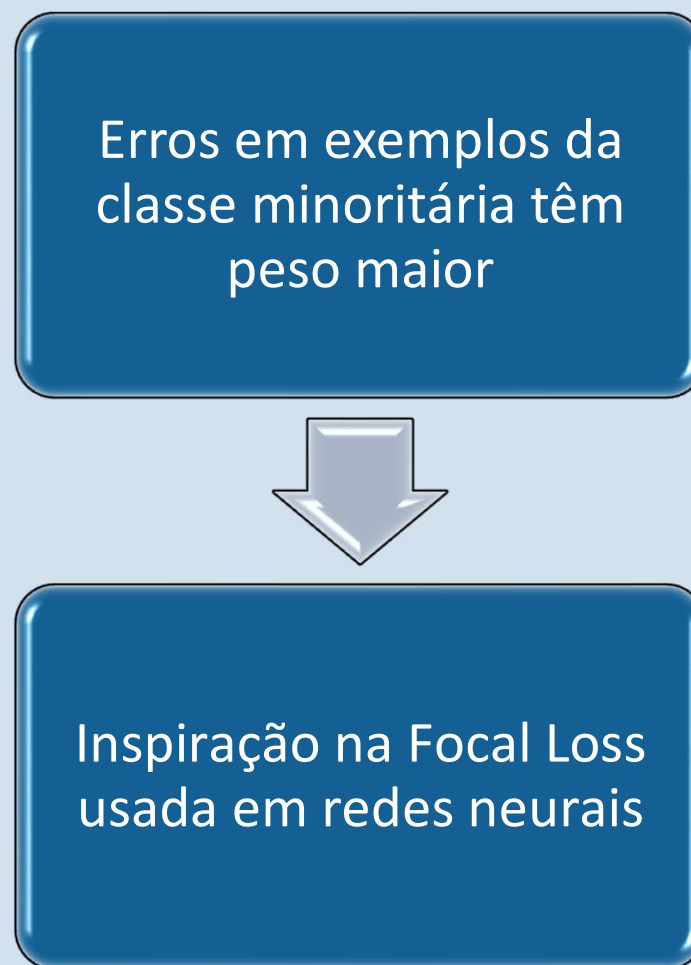
Dataset	Modelo	Accuracy	Recall (Macro)	F1-Score (Macro)	AUC
dataset_1056_mc1.csv	Original	0.993310	0.525000	0.545940	0.923245
dataset_1056_mc1.csv	AFL	0.992958	0.524823	0.543687	0.929184
dataset_1056_mc1.csv	ASL	0.970070	0.811170	0.609486	0.920585

Solução 2: AFL (Asymmetric Focal Loss)

Motivação: Corrigir a assimetria entre classes.

Ideia: Penalizar mais os erros na classe minoritária.

Como funciona a AFL?



Implementação da AFL

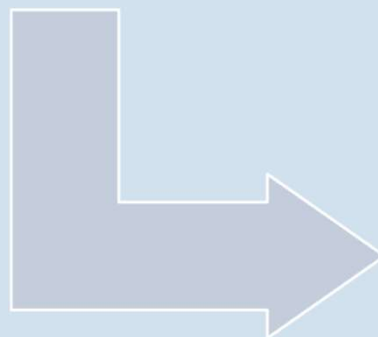
Função de perda
assimétrica com foco em
exemplos da classe de
interesse.



Fácil integração com o
modelo original.

Resultados com AFL

AFL apresentou melhorias a nível do recall e F1-score



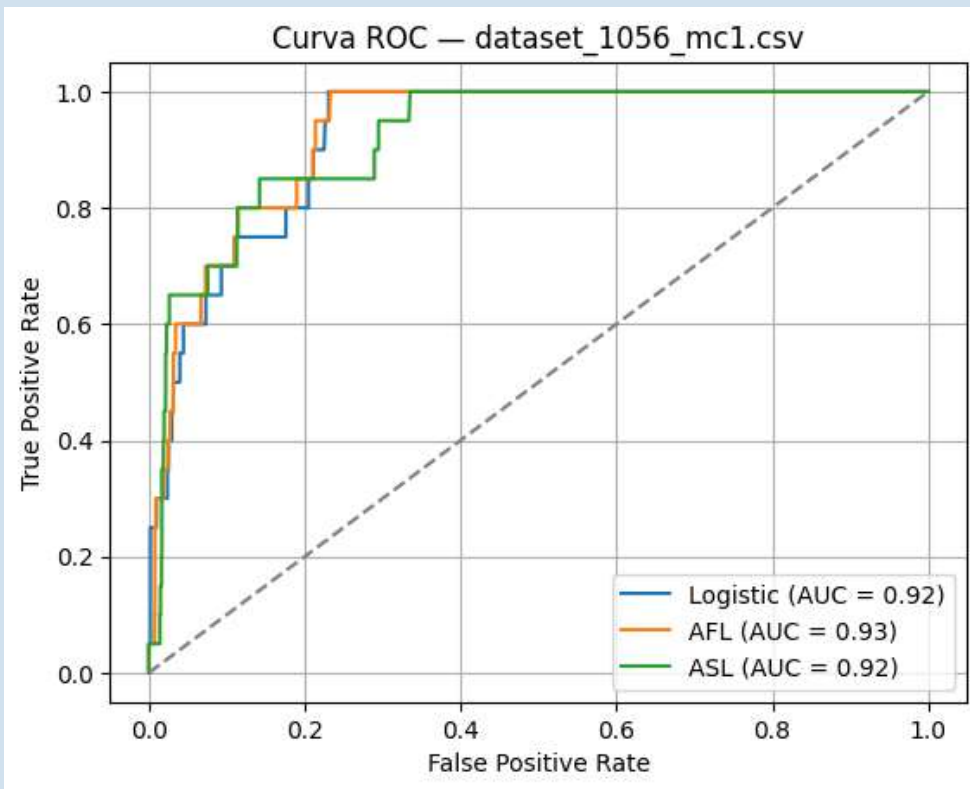
Modelo tornou-se mais sensível aos exemplos minoritários.

Dataset	Modelo	Accuracy	Recall (Macro)	F1-Score (Macro)	AUC
dataset_1056_mc1.csv	Original	0.993310	0.525000	0.545940	0.923245
dataset_1056_mc1.csv	AFL	0.992958	0.524823	0.543687	0.929184

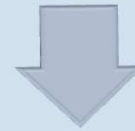
Análise Comparativa

Característica	Regressão Logística (Original)	AFL (Asymmetric Focal Loss)	ASL (Adaptive Sample Loss)
Objetivo	Maximizar a verosimilhança (log loss)	Focar mais em amostras difíceis e minoritárias	Ajustar dinamicamente a penalização conforme o tipo de amostra
Sensibilidade ao Desbalanceamento	Alta (pouco robusta)	Média-Alta (melhora recall da classe minoritária)	Alta (melhora recall e precisão com penalização adaptativa)
Mecanismo	Penaliza erros igualmente	Penaliza mais os falsos negativos da classe minoritária	Modula o gradiente conforme a confiança e classe
Complexidade	Baixa	Média	Alta
Evolução	—	Evolução da focal loss para dados desbalanceados	Evolução ainda mais adaptativa da focal loss (como AFL++)
Melhor desempenho	Classe majoritária	Classe minoritária (com risco de overfitting leve)	Classe minoritária com controlo mais refinado

Curvas ROC – Avaliação da Separação entre Classes



A curva ROC mostra a capacidade do modelo em separar as classes.



ASL e AFL apresentam maiores áreas sob a curva (AUC).



AUC é menos sensível ao desbalanceamento, mas confirma melhoria geral dos modelos adaptados

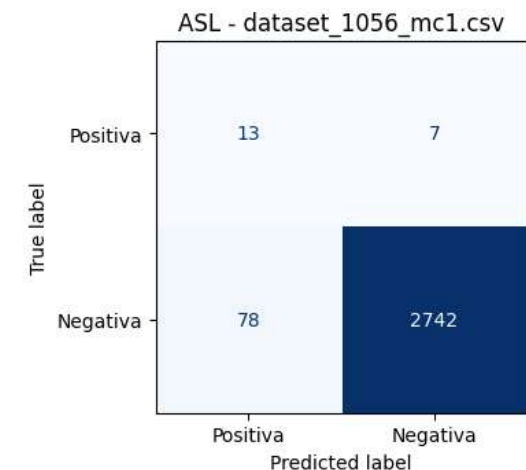
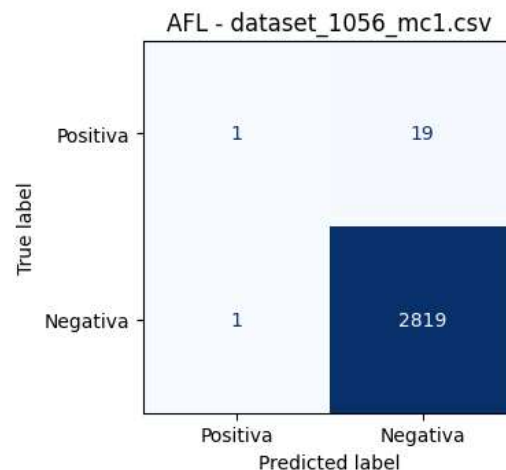
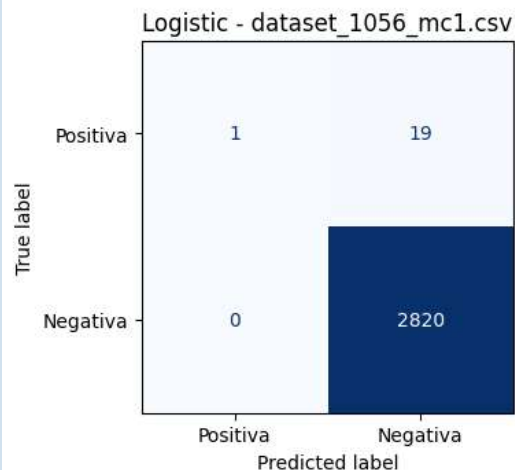
Matriz de Confusão – Impacto nas Predições

Mais predições corretas da classe minoritária com ASL e AFL

Redução nos falsos negativos (Classe 1 predita como Classe 0)

Mostra claramente que o modelo está mais equilibrado

Matrizes de Confusão — dataset_1056_mc1.csv



Comparação de Modelos - Acurácia

Acurácia continua elevada na maioria dos modelos

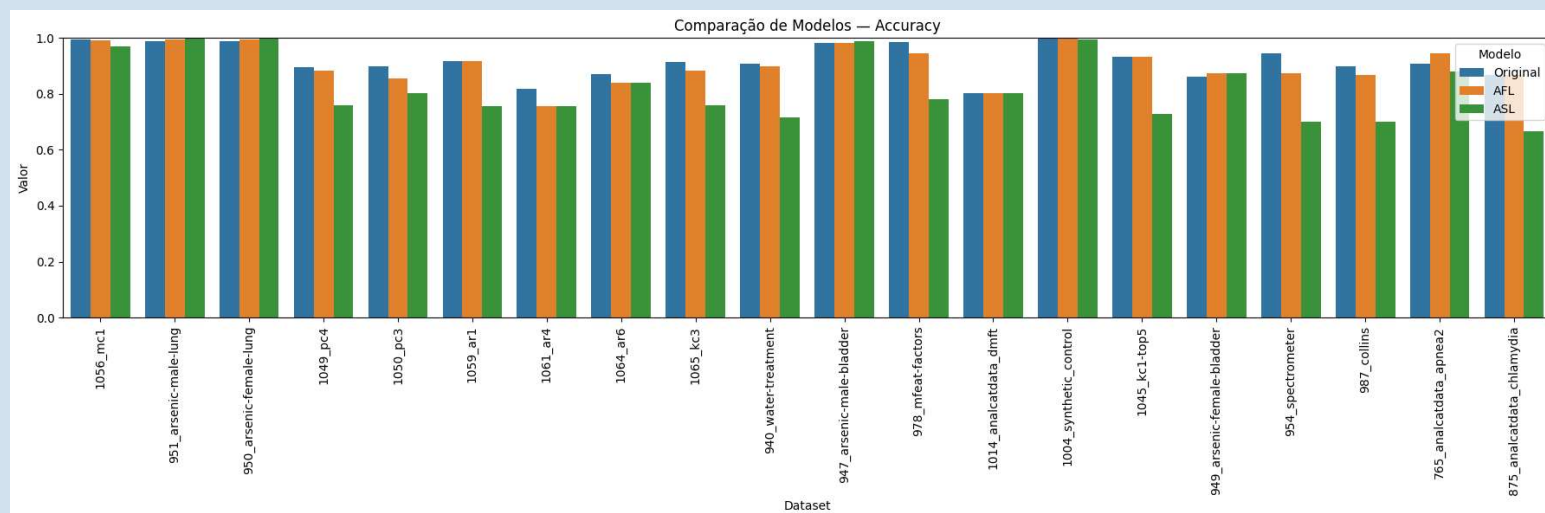
ASL apresenta maior variação entre Datasets

AFL mantém performance próxima do modelo base

Diferenças de acurácia não comprometem resultados

Acurácia mascara falhas na classe minoritária

Reflete estabilidade das soluções testadas



Comparação de Modelos-Recall (Macro)

AFL e ASL superam o modelo original

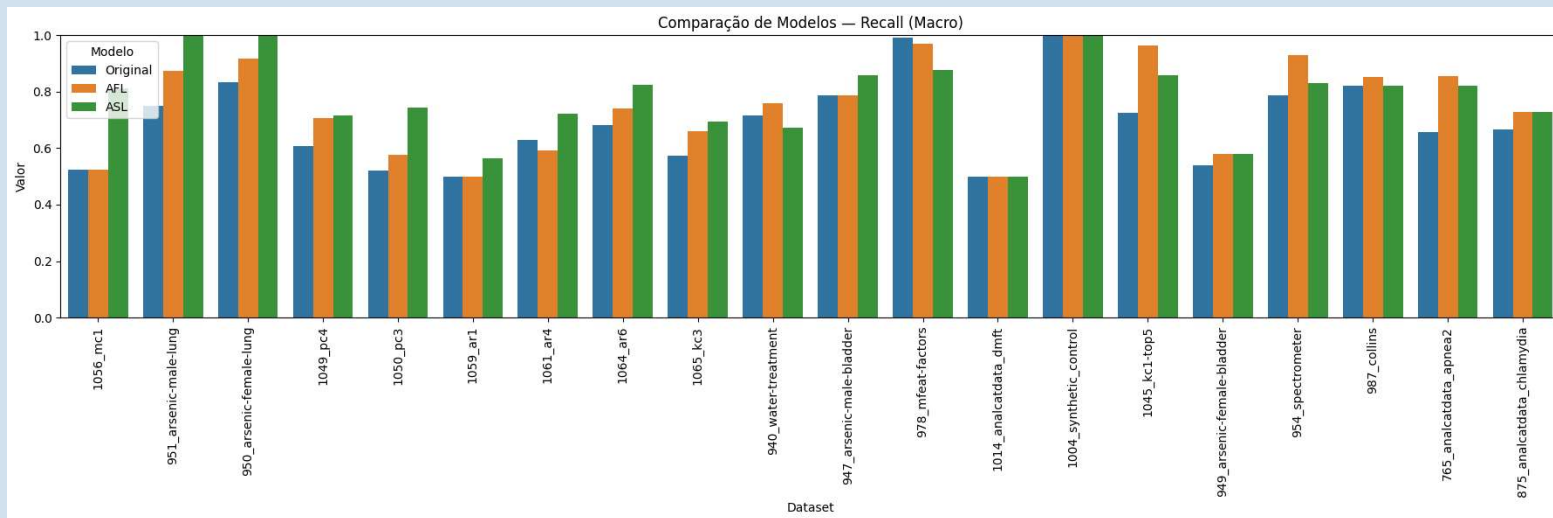
ASL lidera em Recall macro na maioria dos casos

AFL mantém bom equilíbrio entre classes

Modelo original tem pior desempenho na classe 1

Resultados validam foco na classe minoritária

Soluções adaptadas mostram maior sensibilidade



Comparação de Modelos — AUC

Todos os modelos apresentam AUC elevada

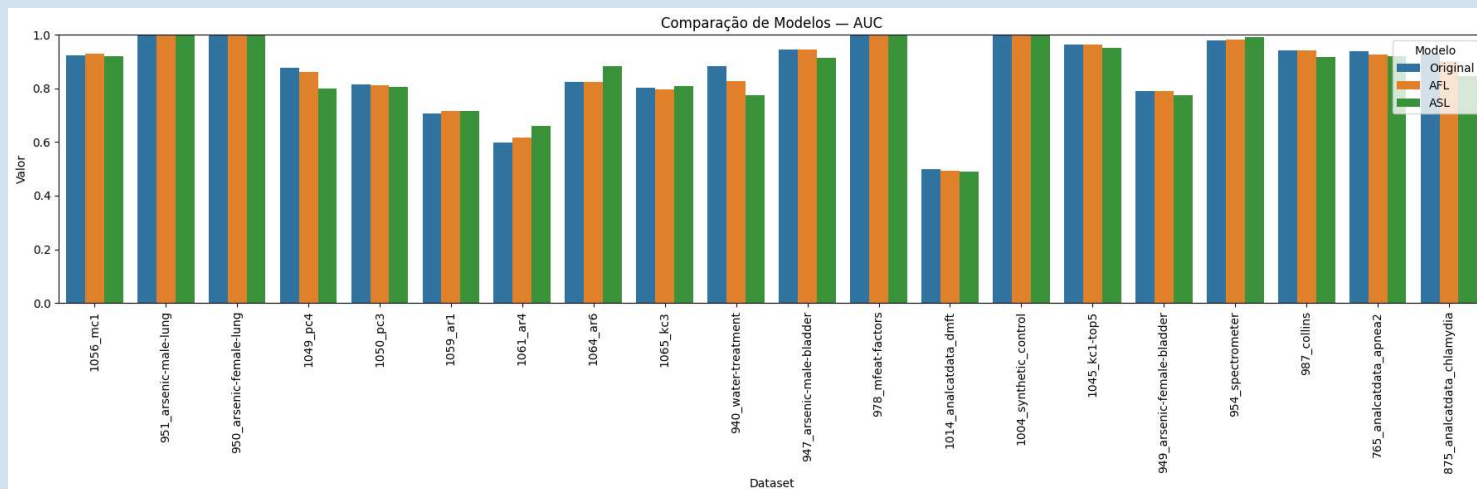
Diferenças entre modelos são pouco significativas

ASL e AFL mantêm boa separação entre classes

AUC reforça eficácia das abordagens adaptadas

Pequenas oscilações não afetam desempenho geral

AUC valida robustez dos modelos em cenários difíceis



Comparação de Modelos - F1-Score (Macro)

AFL tem o F1 macro mais elevado na maioria dos datasets

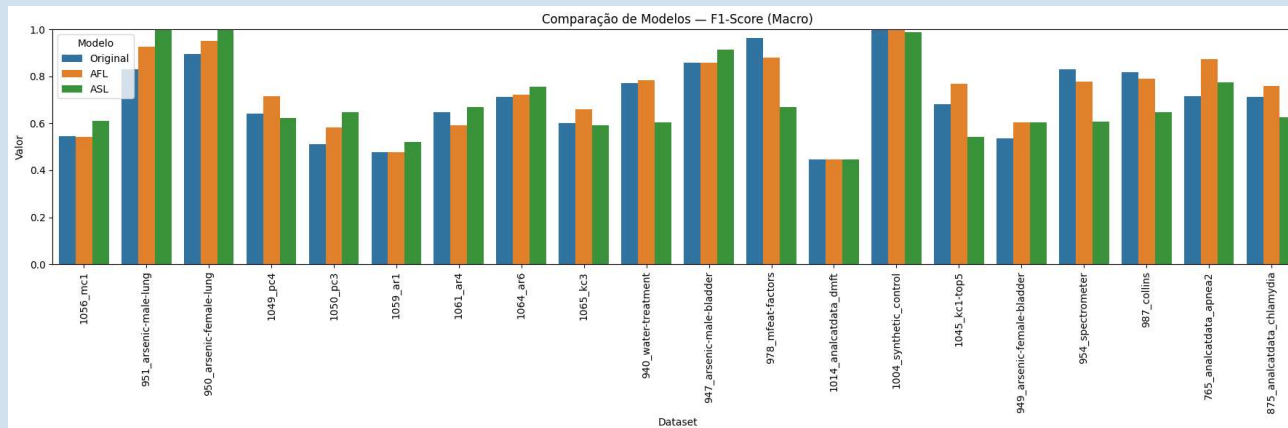
ASL mostra melhorias, mas com maior variação

Modelo original tem o pior equilíbrio global

AFL garante melhor compromisso entre precisão e recall

F1 macro reforça eficácia da AFL em desbalanceamento

Desempenho da ASL depende mais do Dataset



Conclusões

Modelos padrão
falham em contextos
desbalanceados

AFL apresenta o
melhor equilíbrio
geral entre as
métricas

ASL maximiza o **recall**
da **classe minoritária**

F1-score continua a
ser a métrica mais
justa nestes casos

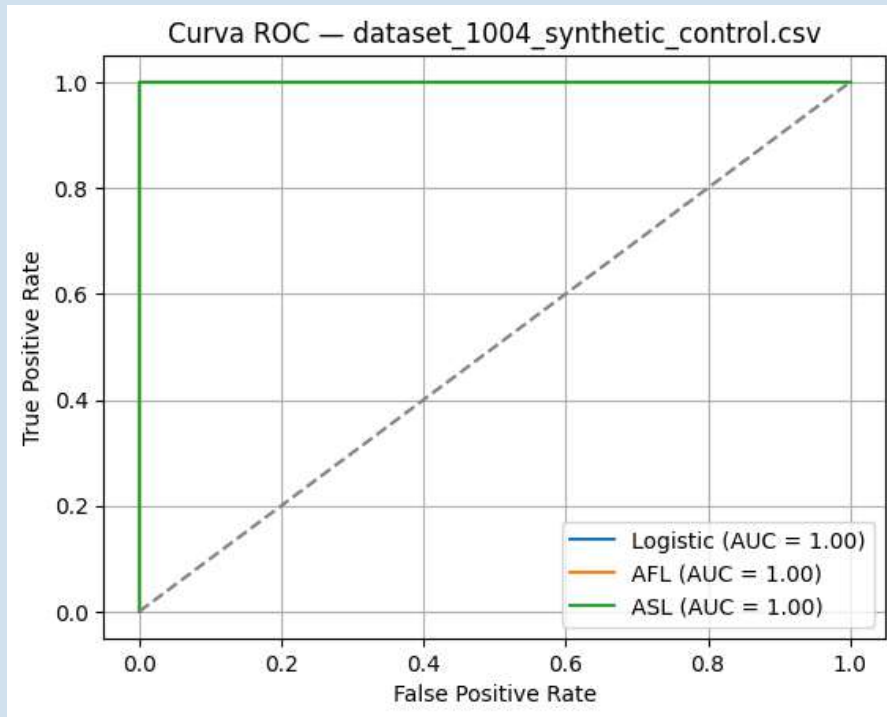
Alterar a função de
perda melhora a
sensibilidade à
minoria

As soluções propostas
mantêm **acurácia**
elevada e estável

Resultados validam o
uso de perdas
assimétricas em
aplicações críticas

	Accuracy	Recall (Macro)	F1-Score (Macro)	AUC
Modelo				
Original	0.9191	0.6906	0.7101	0.8709
AFL	0.9056	0.7511	0.7358	0.8666
ASL	0.8240	0.7810	0.6920	0.8590

Limitações e Melhorias Futuras



Ajustar dinamicamente o limiar de decisão (threshold)

Otimizar métricas mais informativas como F1-score

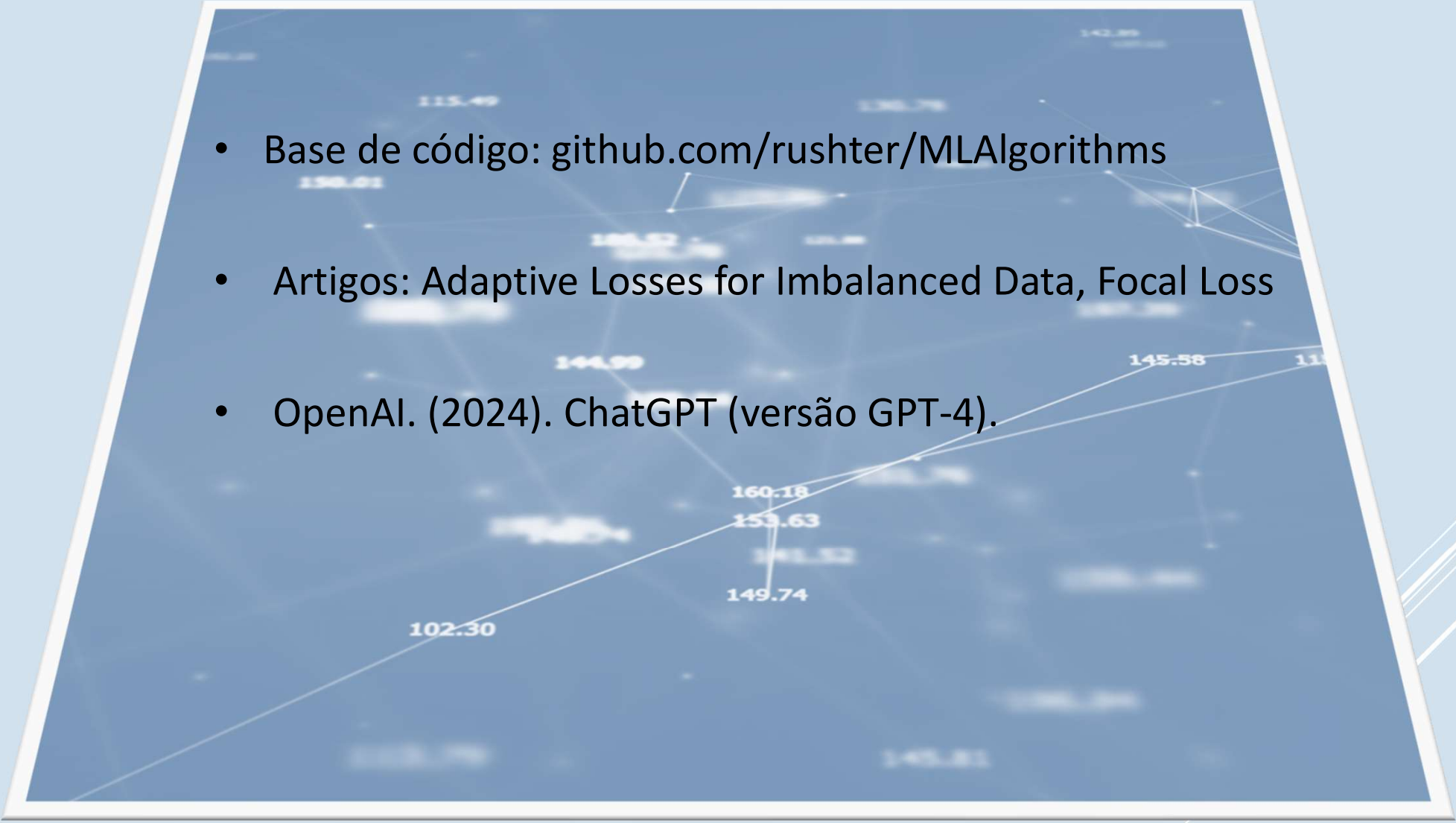
Analisar AUC com intervalos de confiança em classes raras

Usar métricas alternativas (ex: balanced accuracy, MCC)

Aumentar diversidade e número de datasets

Explorar calibradores de probabilidade (Platt, isotonic)

Referências

- 
- Base de código: github.com/rushter/MLAlgorithms
 - Artigos: Adaptive Losses for Imbalanced Data, Focal Loss
 - OpenAI. (2024). ChatGPT (versão GPT-4).