Classificação de Risco de Vulnerabilidades de Segurança via Processos Gaussianos e Aprendizado Ativo

Davyson S. Ribeiro, Rafael Lemos, Francisco R. P. da Ponte, César Lincoln C. Mattos, Emanuel B. Rodrigues

UNIVERSIDADE FEDERAL DO CEARÁ

Agenda

- Motivação
- Contribuição
- Conceitos
- Conjunto de Dados
- Estrutura do Modelo
- Estratégias de Seleção de Amostras
- Cenários Experimentais e Resultados
- Conclusões

Motivação

- Importância da Gestão de Vulnerabilidades
- Desafios na Rotulagem de Vulnerabilidades

Contribuição

- Investigar a viabilidade de utilizar **Procesos Gaussianos** combinados com **Aprendizado Ativo** para classificar vulnerabilidades de segurança quanto ao risco de exploração.

Contribuição

Tabela 1: Comparação entre os trabalhos relacionados e o presente artigo.

	Detecção de Vulnerabilidades	Classificação de Risco	Classificadores de Aprendizado de Máquina	Medição de Incerteza - Aprendizado Ativo
Kashyap et al. [2022]	✓		GPR	
Sun et al. [2023]	✓		Bert, Bert-AL Bert-SSL, ASSBert	Entropy
Kure et al. [2022]		✓	KNN, NN, DT, RF LR, NBM, NB	
Elbaz et al. [2021]		✓	CRFs	Least Confident
Ponte et al. [2023a]		✓	RF, GB, RL SVC, MLP	Entropy
Este Trabalho		✓	GP	Entropy, Least Confident, BSB, GPLCB, Random

Conceitos

- Processos Gaussianos:
  - Métodos de aprendizado supervisionado não paramétricos que modelam distribuições probabilísticas sobre funções

Conceitos

- Aprendizado Ativo:
  - Técnica que permite ao modelo selecionar as amostras mais informativas para rotular, minimizando a quantidade de dados rotulados necessários

Conjunto de Dados

- CVEJoin[1]
  - Conjunto com mais de 200 mil amostras
  - 208 Amostras rotuladas por especialistas
  - 29 Atributos
  - 4 Classes de Riscos de Vulnerabilidades
    (Baixo, Moderado, Importante e Crítico)

[1] https://github.com/rodrigoparente/cvejoin-security-dataset

Estrutura do Aprendizado Ativo com Processos Gaussianos

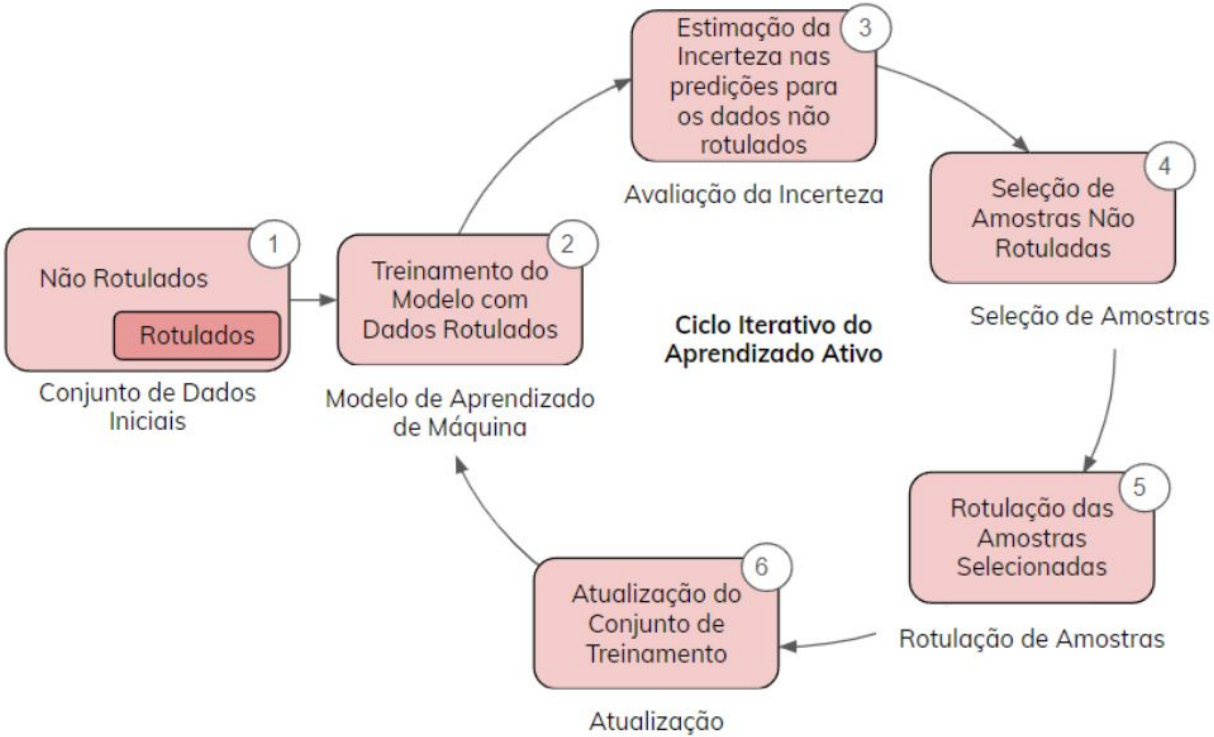


Figura 1: Ciclo iterativo do Aprendizado Ativo.

Estratégias de Seleção de Amostras no Aprendizado Ativo

- Random
- Least Confident
- BSB (Best and Second Best)
- Entropy
- GPLCB (Gaussian Process Lower Confidence Bound)

Avaliação do Modelo e Cenários Experimentais

Tabela 2: Configurações do Aprendizado Ativo usadas nos experimentos.

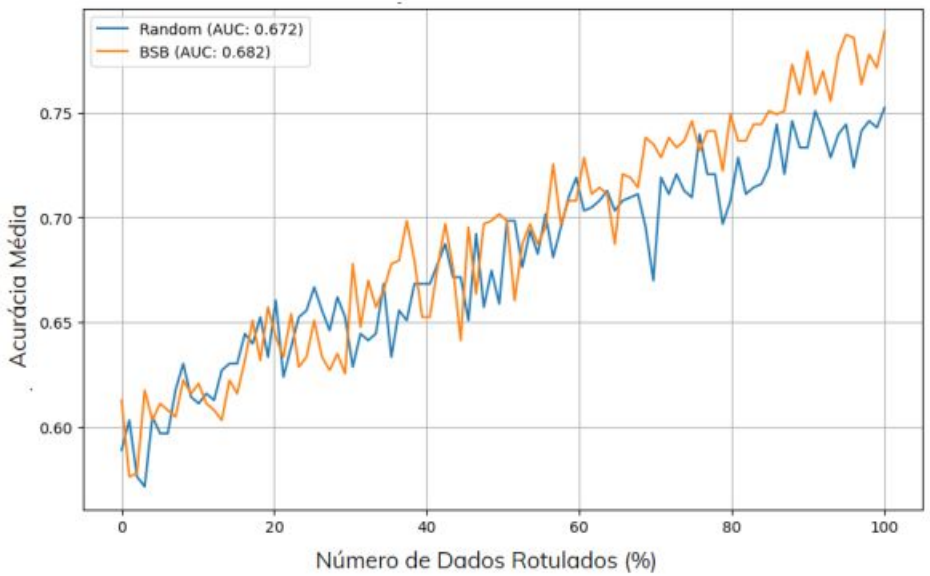
Configurações	**Valor**
Tamanho Inicial	10× Número de classes
Iterações Ativas	Variação de acordo com o cenário
Seleção Ativa por Iteração	Variação de acordo com o cenário
Estratégias de Seleção	Random, Least Confident, Entropy, BSB, GPLCB
Número de repetições independentes	30
Divisão de dados (Treino/Teste)	90%/10%

Cenários Experimentais

Interações com Especialistas VS Quantidade de Vulnerabilidades Rotuladas

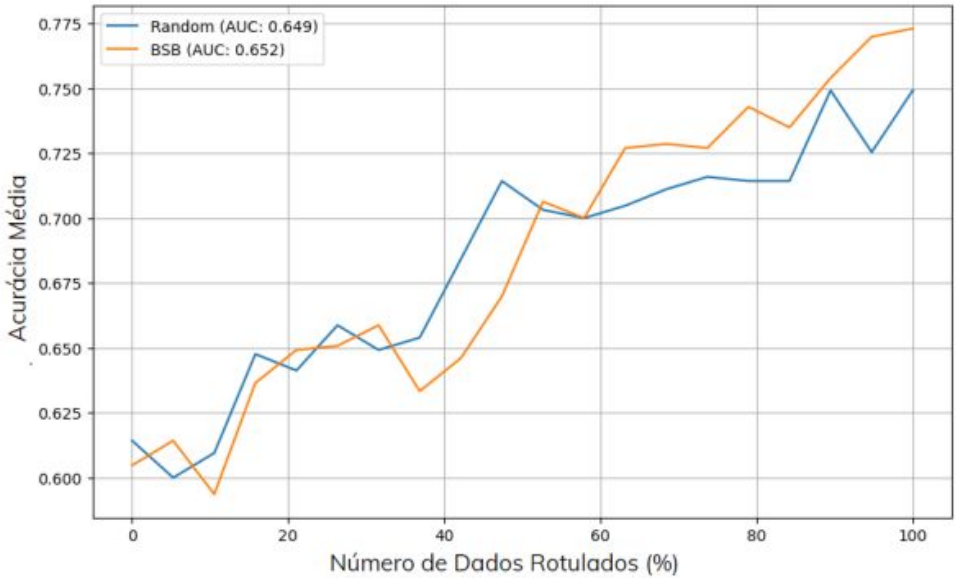
- Cenário I:
  - 100 iterações com 1 dado rotulado por vez
- Cenário II:
  - 20 iterações com 5 dados rotulados por vez
- Cenário III:
  - 10 iterações com 10 dados rotulados por vez

Resultados Cenário I - Acurácia Média e AUC em função do número de dados rotulados



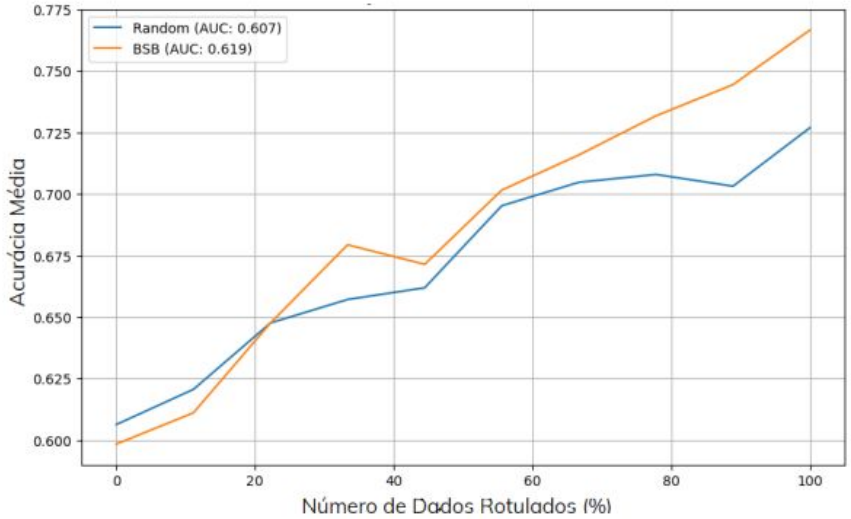
Estratégia	Acurácia $\mu \pm \sigma$	Precisão $\mu \pm \sigma$	Recall $\mu \pm \sigma$	F1-score $\mu \pm \sigma$
BSB	0.78 ± 0.05	0.83 ± 0.06	0.78 ± 0.05	0.78 ± 0.06
Entropy	0.77 ± 0.05	0.86 ± 0.06	0.77 ± 0.05	0.77 ± 0.06
GPLCB	0.76 ± 0.05	0.80 ± 0.02	0.76 ± 0.06	0.76 ± 0.07
Least Confident	0.78 ± 0.04	0.82 ± 0.06	0.77 ± 0.04	0.78 ± 0.05
Random	0.75 ± 0.05	0.78 ± 0.06	0.75 ± 0.05	0.74 ± 0.06

Resultados Cenário II - Acurácia Média e AUC em função do número de dados rotulados



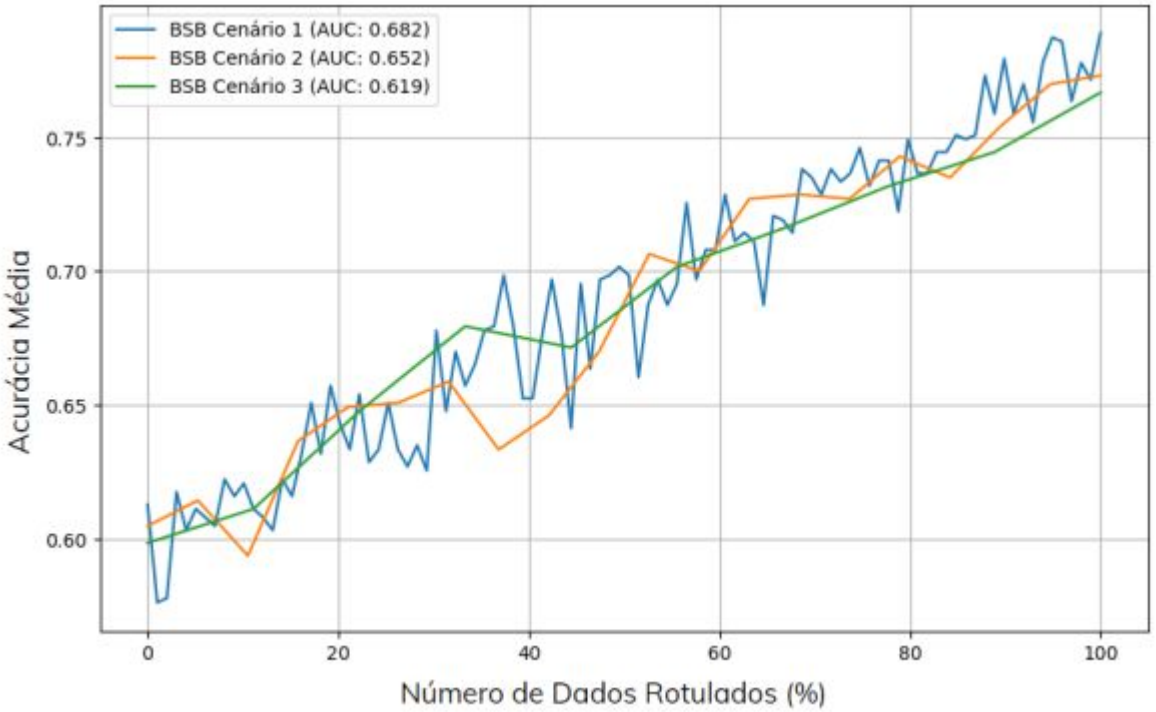
Estratégia	Acurácia $\mu \pm \sigma$	Precisão $\mu \pm \sigma$	Recall $\mu \pm \sigma$	F1-score $\mu \pm \sigma$
BSB	0.77 ± 0.05	0.73 ± 0.06	0.77 ± 0.05	0.76 ± 0.06
Entropy	0.76 ± 0.05	0.72 ± 0.06	0.76 ± 0.05	0.75 ± 0.06
GPLCB	0.75 ± 0.05	0.74 ± 0.02	0.75 ± 0.06	0.75 ± 0.07
Least Confident	0.75 ± 0.04	0.72 ± 0.06	0.75 ± 0.04	0.75 ± 0.05
Random	0.74 ± 0.05	0.72 ± 0.06	0.74 ± 0.05	0.74 ± 0.06

Resultados Cenário III - Acurácia Média e AUC em função do número de dados rotulados



Estratégia	Acurácia $\mu \pm \sigma$	Precisão $\mu \pm \sigma$	Recall $\mu \pm \sigma$	F1-score $\mu \pm \sigma$
BSB	0.76 ± 0.05	0.74 ± 0.07	0.76 ± 0.05	0.76 ± 0.06
Entropy	0.74 ± 0.05	0.74 ± 0.07	0.74 ± 0.05	0.74 ± 0.06
GPLCB	0.74 ± 0.05	0.76 ± 0.07	0.74 ± 0.05	0.73 ± 0.06
Least Confident	0.76 ± 0.05	0.74 ± 0.07	0.76 ± 0.05	0.75 ± 0.04
Random	0.72 ± 0.04	0.71 ± 0.07	0.72 ± 0.04	0.72 ± 0.04

Resultados Final (BsB)- Acurácia Média e AUC em função do número de dados rotulados



Considerações finais

- Otimização do processo de classificação de vulnerabilidades de segurança
- Importância da Incerteza nas previsões é crucial para guiar o processo de rotulagem
- As estratégias BSB e Entropia provaram obter melhores resultados

Trabalhos futuros

- Estender a metodologia para outros domínios de cibersegurança

- Explorar a combinação de redes neurais combinadas com aprendizado ativo para lidar com cenários ainda mais complexos

Obrigado!

Davyson S. Ribeiro
Rafael Lemos,
Francisco R. P. da Ponte,
César Lincoln C. Mattos,
Emanuel B. Rodrigues.

davysonribeiro@alu.ufc.br
emanuel@dc.ufc.br

Equações de Estratégias de Seleção

- Random: Seleciona de Forma Aleatória
- Least Confident: É dado pelo complemento da maior probabilidade média entre as classes

$$\mathrm{lc}(\boldsymbol{x}_*) = 1 - P(y_* = c_* \mid \boldsymbol{x}_*).$$

- BSB: Considera a diferença entre as duas maiores probabilidades preditas para cada amostra.

$$\Delta(\boldsymbol{x}_*) = P(y = c_1 \mid \boldsymbol{x}_*) - P(y = c_2 \mid \boldsymbol{x}_*).$$

Equações de Estratégias de Seleção

- **Entropia:** Quantifica a incerteza associada a uma distribuição de probabilidade

$$H(\boldsymbol{x}_*) = -\sum_c P(y = c \mid \boldsymbol{x}_*) \log P(y = c \mid \boldsymbol{x}_*).$$

- **GPLCB:** Estima a incerteza ao considerar a média das probabilidades preditivas e o desvio padrão associado.

$$\text{GPLCB}(\boldsymbol{x}_*) = 1 - (P(y = c_* \mid \boldsymbol{x}_*) - \beta\sigma_c(\boldsymbol{x}_*)),$$