Detecção de Ataques DDoS utilizando Machine Learning

Uma comparação entre modelos LTSM, KNN, NC, SVM, Random Forest, XGBoost

Helano Fontenele, Matheus Rocha, Vitor Rosa

Departamento de Teleinformática

Universidade Federal do Ceará

22 de novembro de 2023

UNIVERSIDADE FEDERAL DO CEARÁ

Sumário

UNIVERSIDADE FEDERAL DO CEARÁ

O que é DDoS?

DDoS significa "Distributed Denial of Service"e é um tipo de ataque cibernético no qual vários computadores são usados para sobrecarregar um sistema, rede ou serviço, tornando-o inacessível. O objetivo é negar o serviço a usuários legítimos, enviando uma quantidade massiva de tráfego, geralmente usando botnets. As motivações podem incluir política, rivalidade comercial ou simplesmente causar danos. Empresas usam estratégias como firewalls e sistemas de detecção para se proteger contra esses ataques.

UNIVERSIDADE FEDERAL DO CEARÁ

Tipos de DDoS

- UDP Flood: Sobrecarrega a largura de banda com pacotes UDP.
- HTTP Flood: Esgota os recursos do servidor web com muitas solicitações HTTP.
- TCP SYN/ACK Flood: Satura os recursos do servidor ao explorar o processo de estabelecimento de conexão TCP.
- DNS Amplification: Amplifica ataques usando servidores DNS abertos.
- NTP Amplification: Amplifica ataques explorando servidores NTP.
- Slowloris: Esgota recursos mantendo conexões abertas lentamente.
- ICMP Flood: Sobrecarrega a largura de banda com solicitações ICMP.

UNIVERSIDADE FEDERAL DO CEARÁ

Sobre os Datasets explorados

UNIVERSIDADE FEDERAL DO CEARÁ

CAIDA "DDoS Attack 2007"Dataset

O conjunto de dados CAIDA "DDos Attack 2017"contém registros de tráfego de um ataque DDoS que ocorreu em 4 de agosto de 2007, com duração aproximada de uma hora. O conjunto possui cerca de 21 GB de dados separados em registros de 5 minutos cada. Pertence ao San Diego Supercomputer Center,localizado UC San Diego campus de La Jolla, CA.

ISCXIDS2012 Dataset

O conjunto de dados ISCX-IDS-2012 abrange sete dias de atividade de rede, detalhando eventos diários, suas datas e tamanhos.

- Dia, Data, Descrição, Tamanho (GB)
- Sexta-feira, 11/6/2010, Atividade Normal. Sem atividade maliciosa, 16.1
- Sábado, 12/6/2010, Atividade Normal. Sem atividade maliciosa, 4.22
- Terça-feira, 15/6/2010, Ataque Distribuído de Negação de Serviço usando uma Botnet IRC, 23.4

Pertence ao Canadian Institute for Cybersecurity - University of New Brunswick

Dados prórprios

Para fins de demonstração da utilização de um modelo de machine learning para detecção de ataques DDoS, capturamos dados da nossa própria rede em dois momentos: tráfego normal e sob ataque DDoS.

Sobre o LTSM

As Long Short-Term Memory (LSTM) networks são uma classe de redes neurais recorrentes (RNNs) projetadas para lidar com o problema de vanishing gradient ao adicionar uma memória de curto prazo para a rede. Sendo muito utilizadas para problemas que envolvem dados baseados em séries temporais.

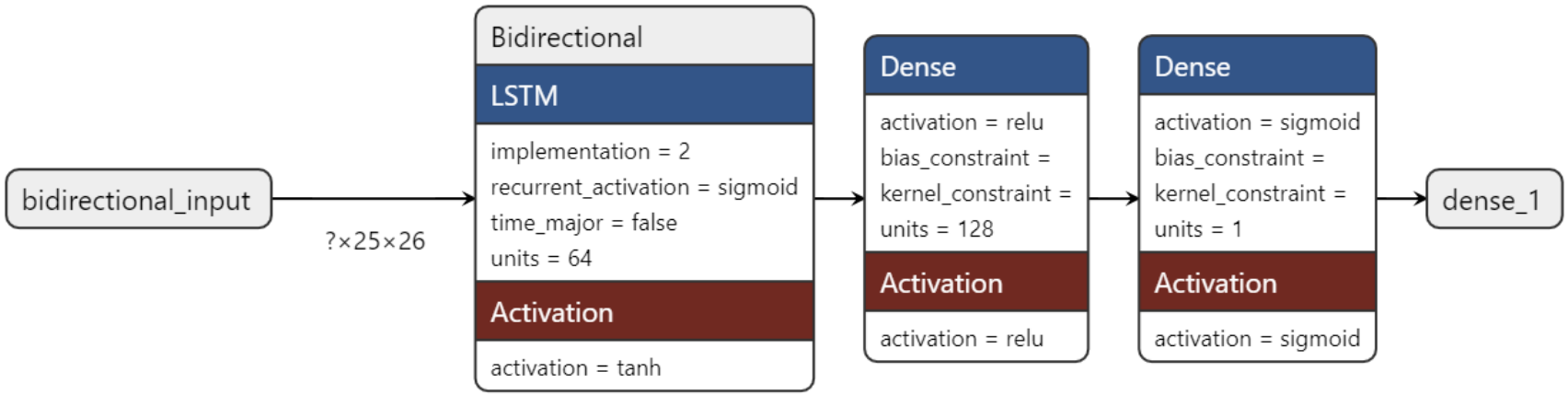
UNIVERSIDADE FEDERAL DO CEARÁ

Sobre os outros modelos utilizados

- KNN: (K-Nearest Neighbors). Classificador simples baseado em distância.
- NC: (Nearest Centroid). Classificador simples baseado em distância a protótipos/representantes de cada classe.
- SVM: (Suport Vectors Machine). Hiperplano ótimo para separação entre as classes.
- Random Forest: Comitê/Ensemble de árvores de decisão.
- XGBoost: Algoritmo baseado em árvores com aumento de gradiente.

Arquitetura da Rede LTSM

Figura: Arquitetura da Rede LTSM



Fonte: Feito pelos autores

Pré-processamento dos Dados

- Remoção do IP emissor e IP destino dos pacotes, evitando enviesamento do modelo.
- Agrupando pacotes em séries temporais com janelamento a cada 25 pacotes.
- 40% dos dados para treino, 60% dos dados para teste.

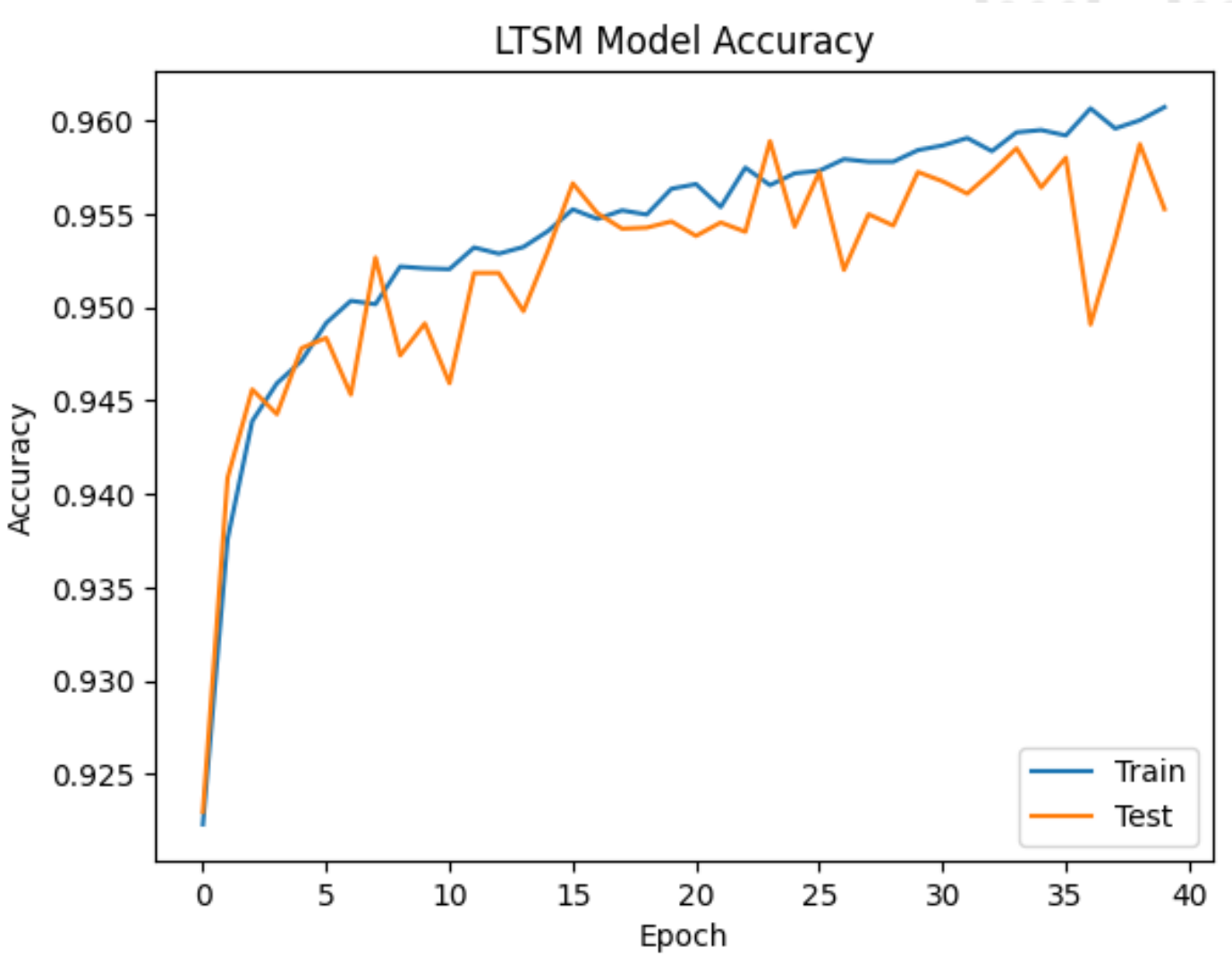
UNIVERSIDADE FEDERAL DO CEARÁ

Resultados obtidos com a LTSM

UNIVERSIDADE FEDERAL DO CEARÁ

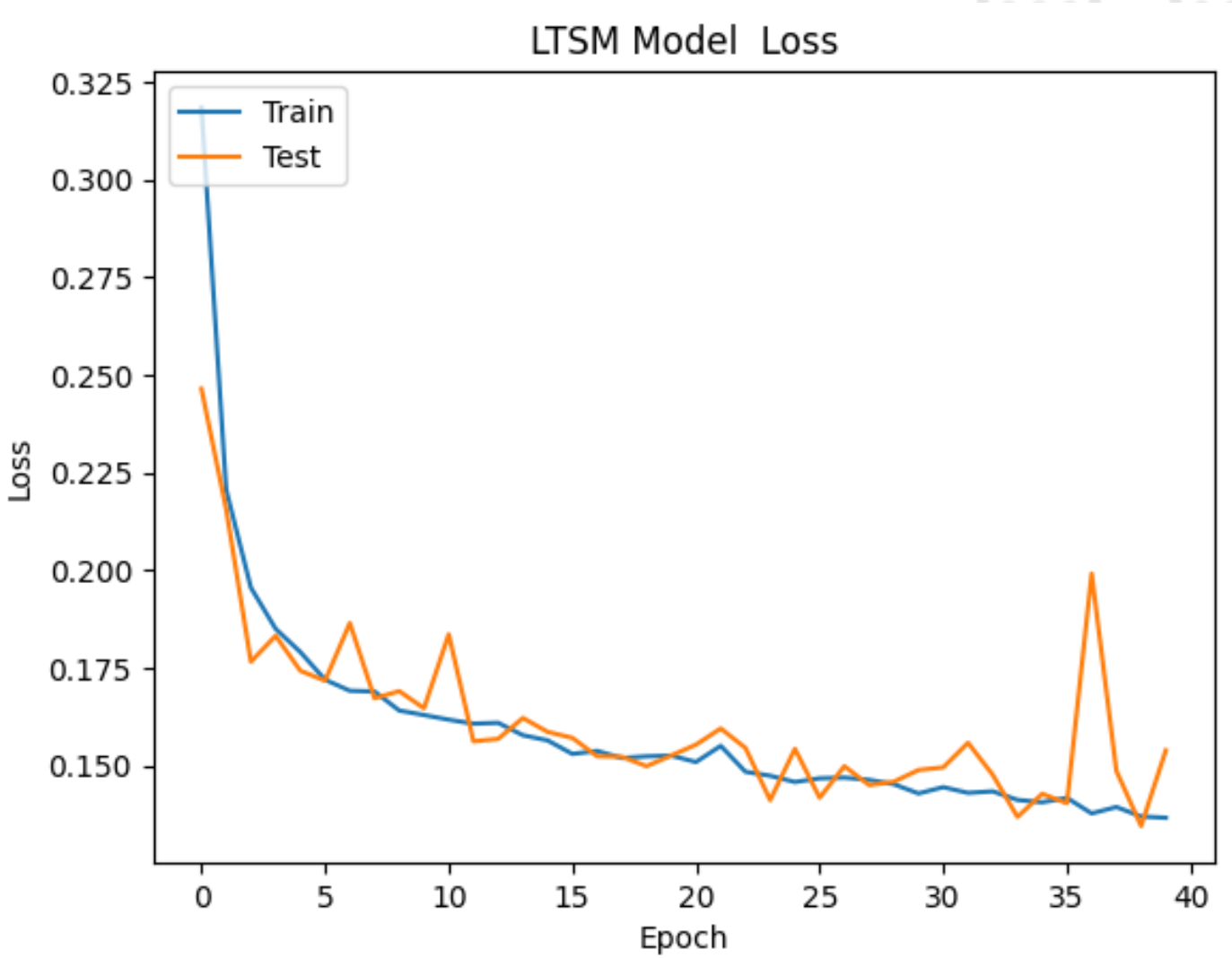
Acurácia

Figura: Acurácia durante as 40 épocas de treinamento



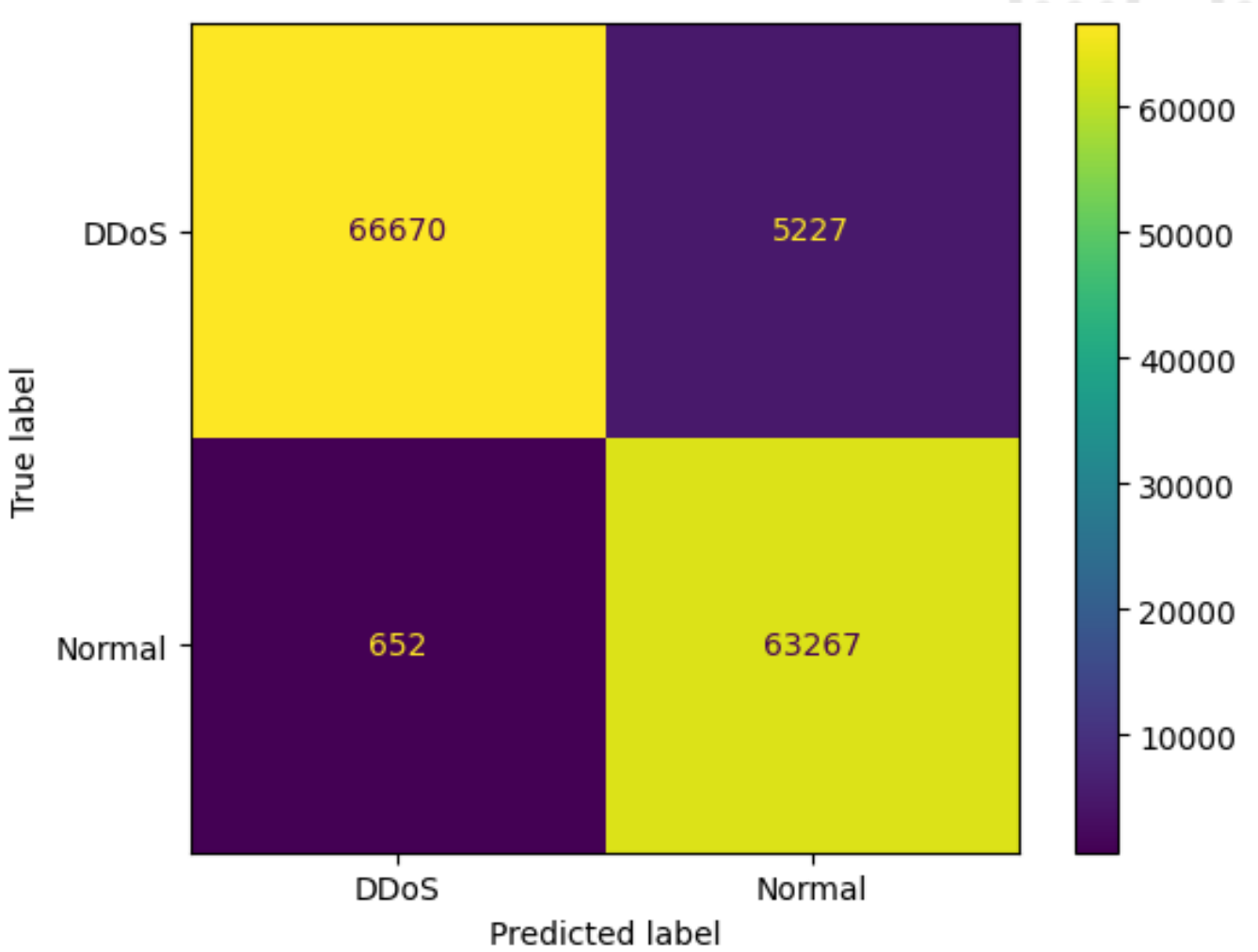
Loss

Figura: Loss durante as 40 épocas de treinamento



Matriz de Confusão

Figura: Matriz de Confusão para a LTSM



Fonte: Feito pelos autores

Índices

- Acurácia: 96%
- Precisão: 99% para DDoS; 92% para Fluxo Normal
- F1-Score: 0.96
- Recall: 0.93 para DDoS; 0.99 para Fluxo Normal

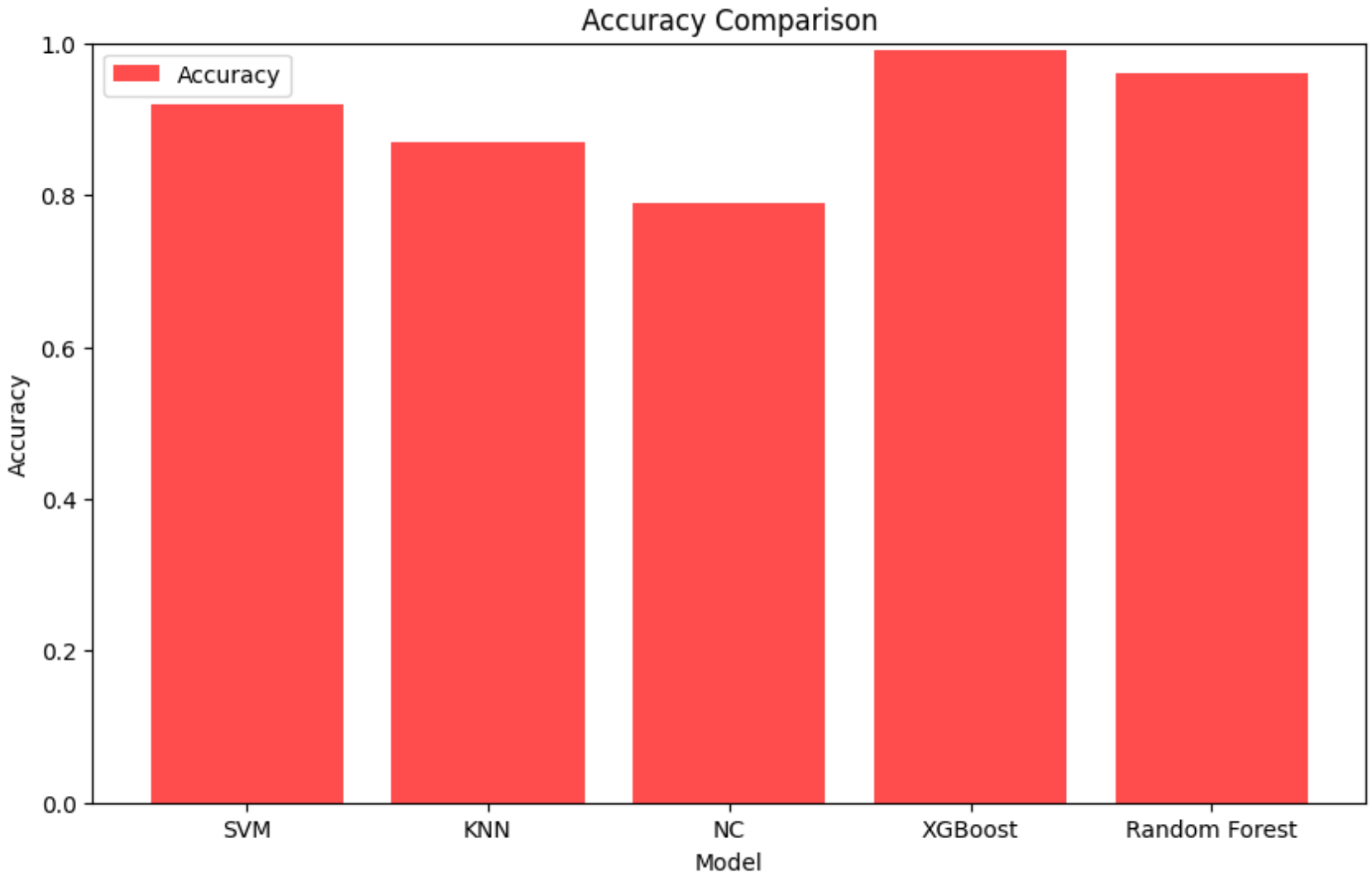
UNIVERSIDADE FEDERAL DO CEARÁ

Resultados obtidos com outros classificadores

UNIVERSIDADE FEDERAL DO CEARÁ

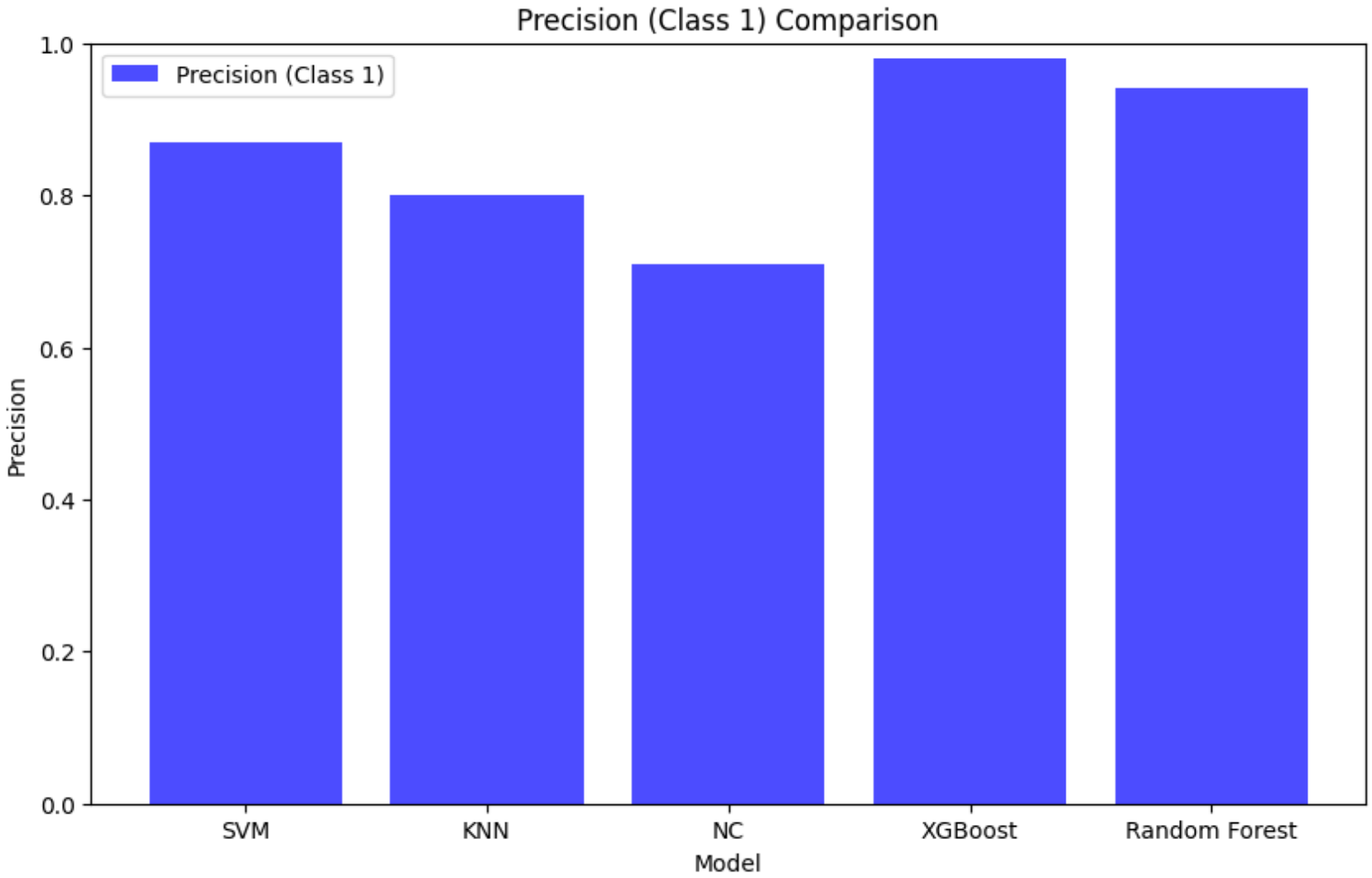
Acurácia

Figura: Acurácia dos outros modelos



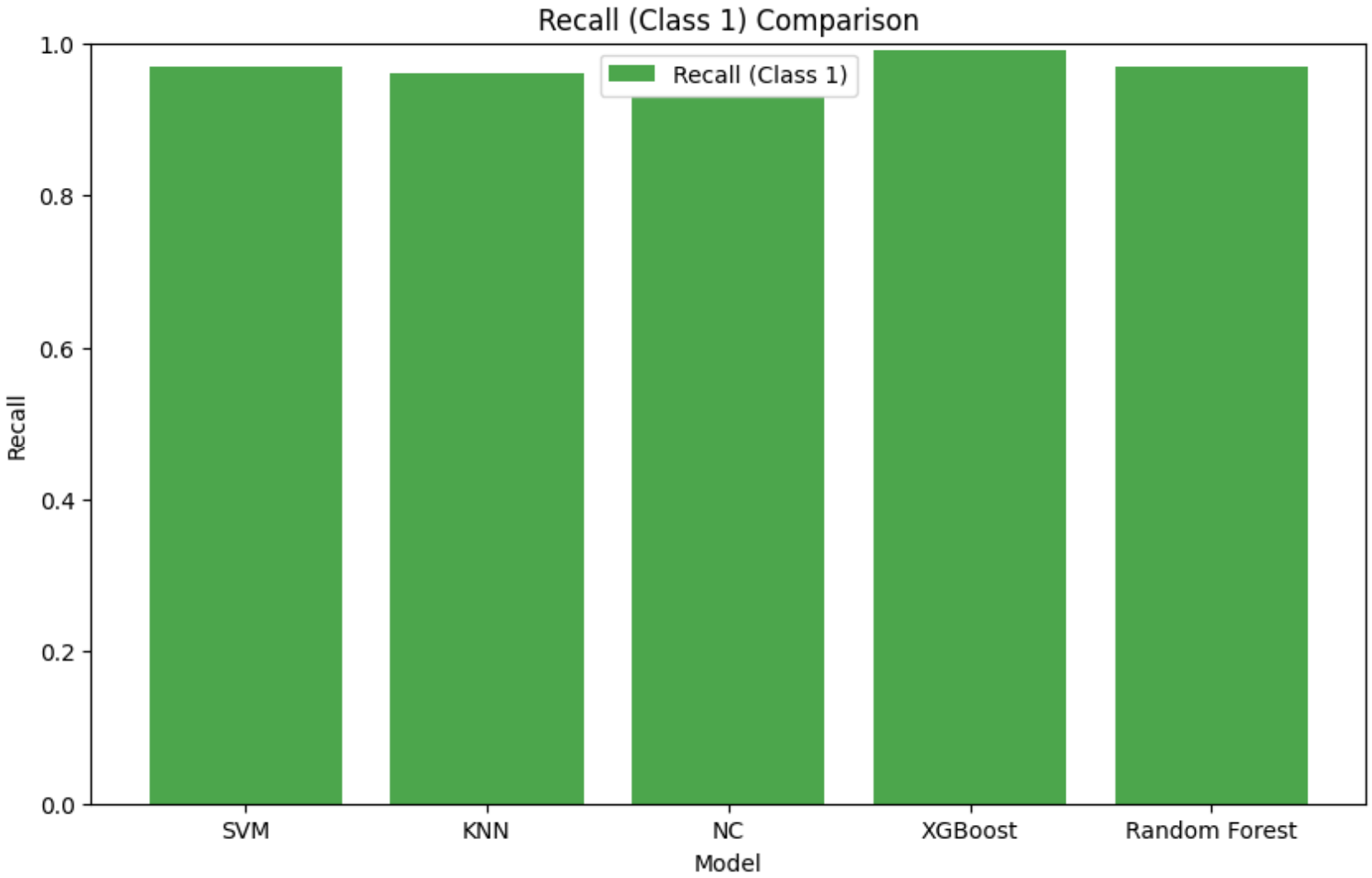
Precisão

Figura: Precisão dos outros modelos



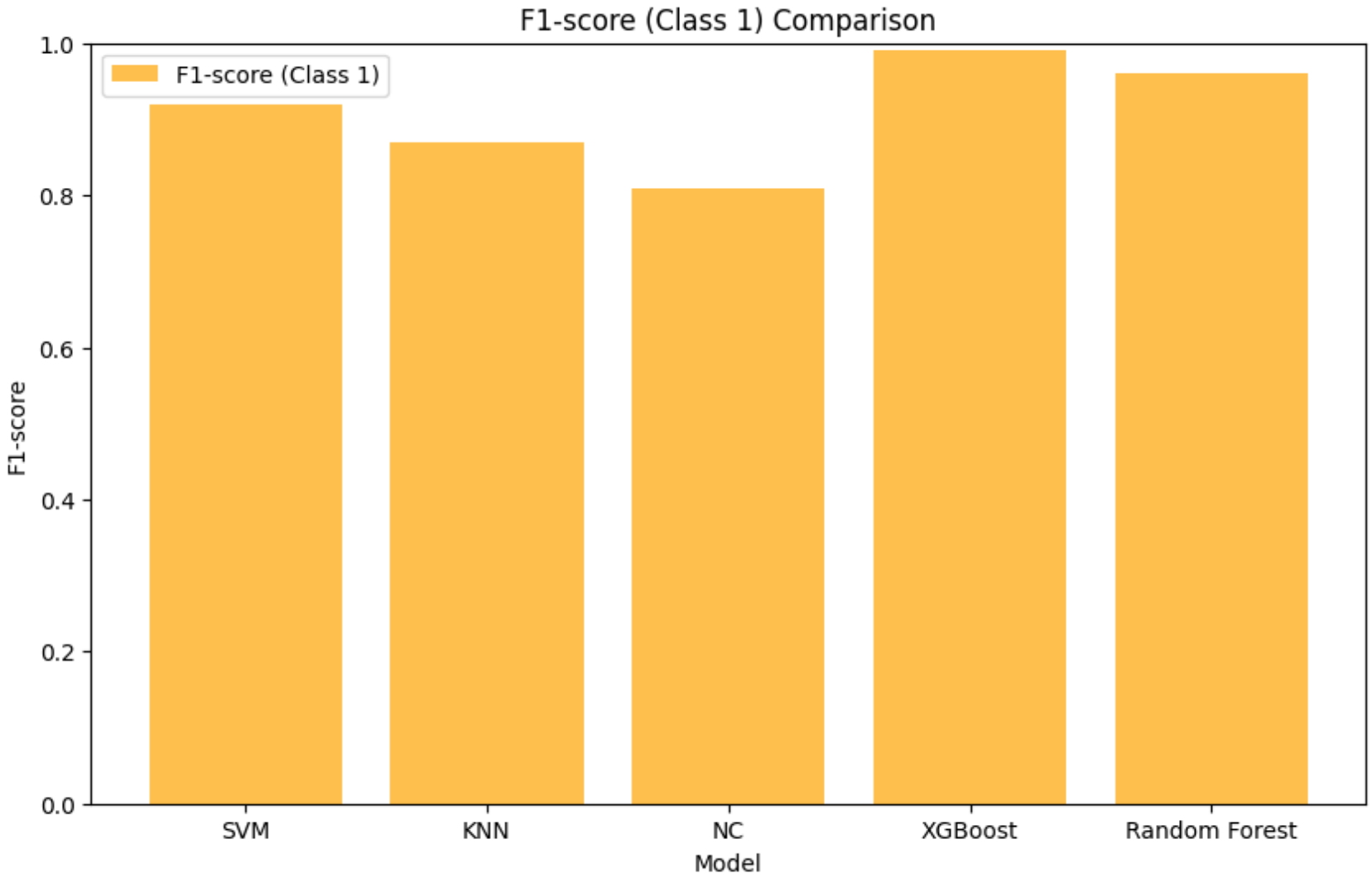
Recall

Figura: Precisão dos outros modelos

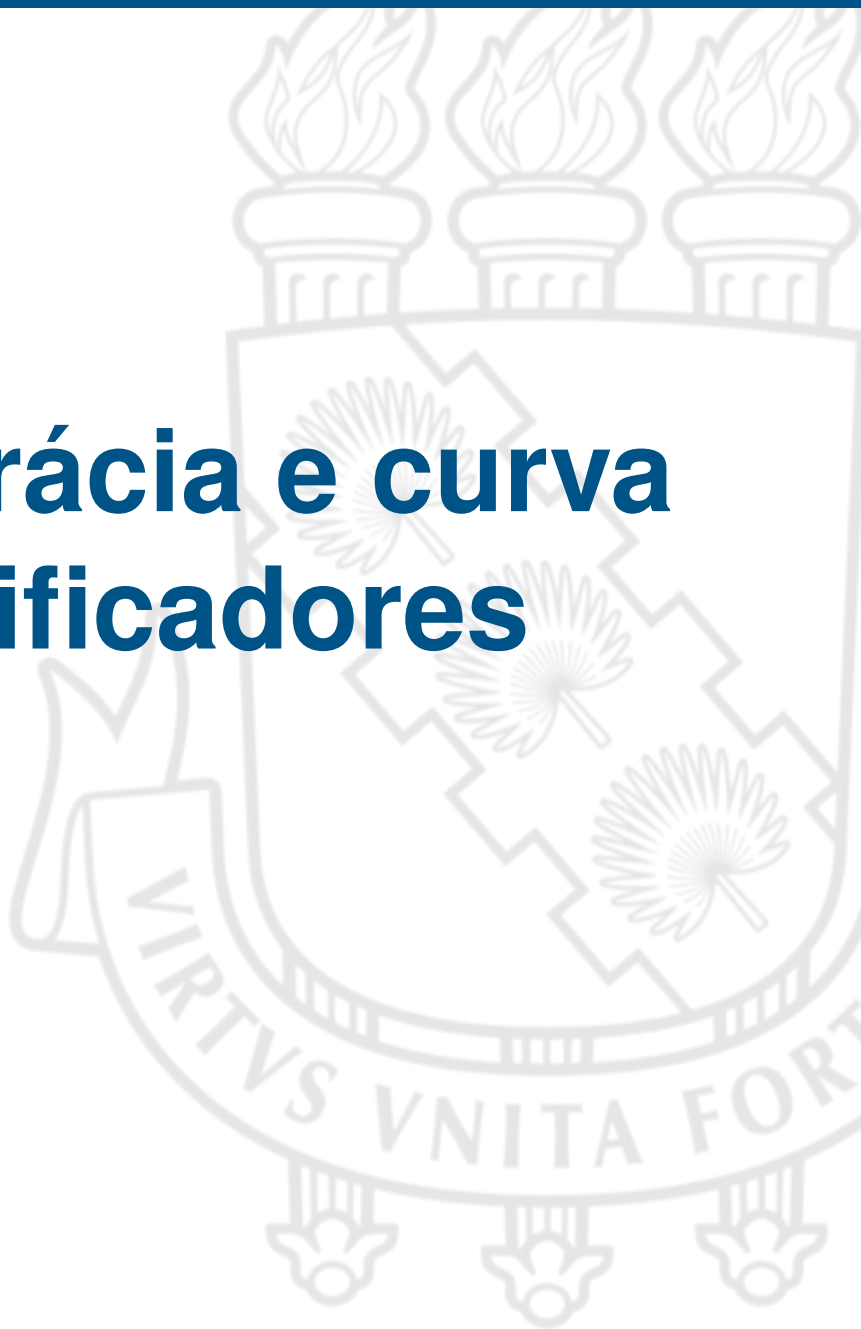


F1-Score

Figura: Precisão dos outros modelos



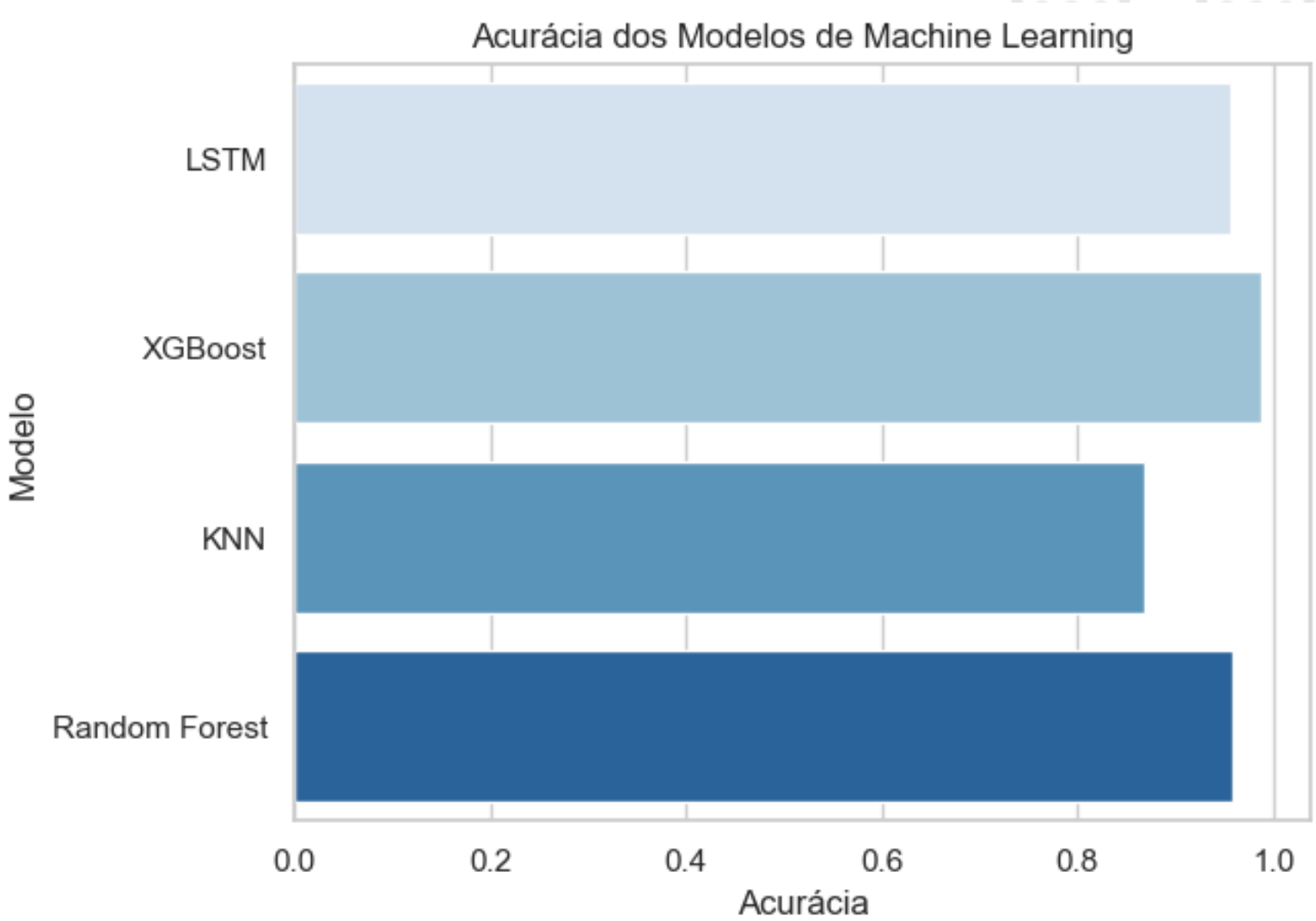
Comparando acurácia e curva ROC dos classificadores



UNIVERSIDADE FEDERAL DO CEARÁ

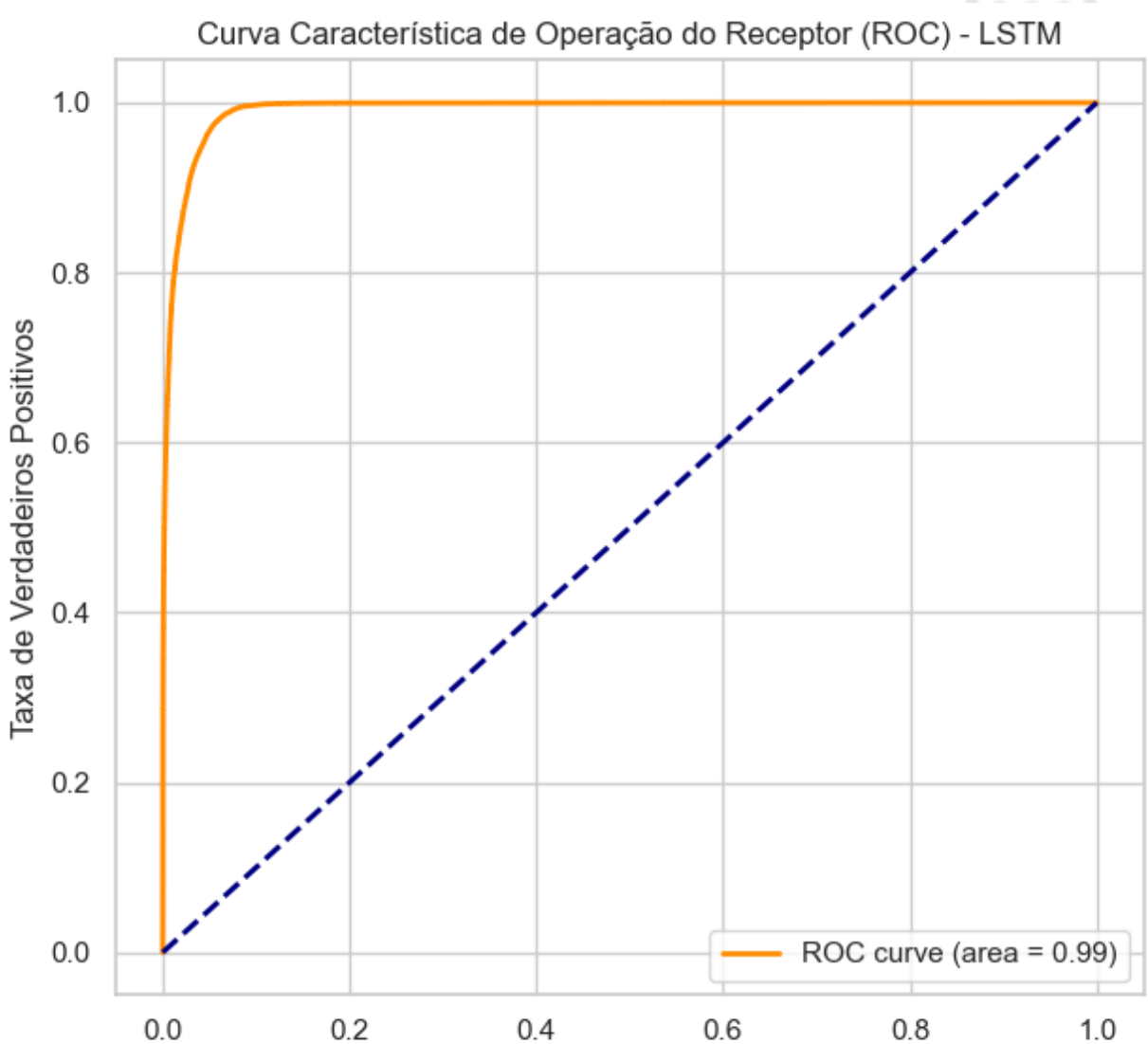
Acurácia

Figura: Comparação da acurácia



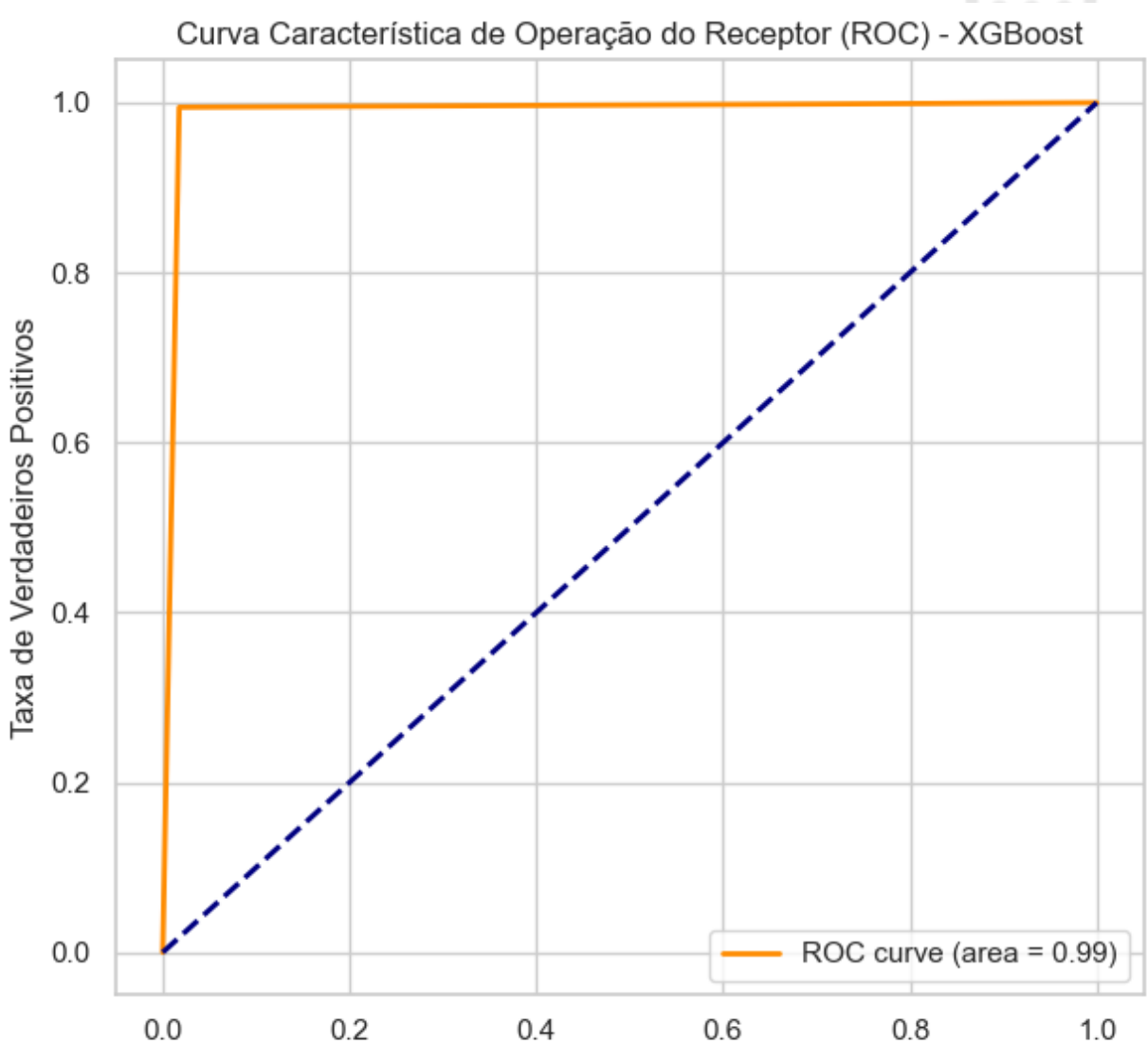
Curva ROC

Figura: Curva ROC LTSM



Curva ROC

Figura: Curva ROC XGBoost

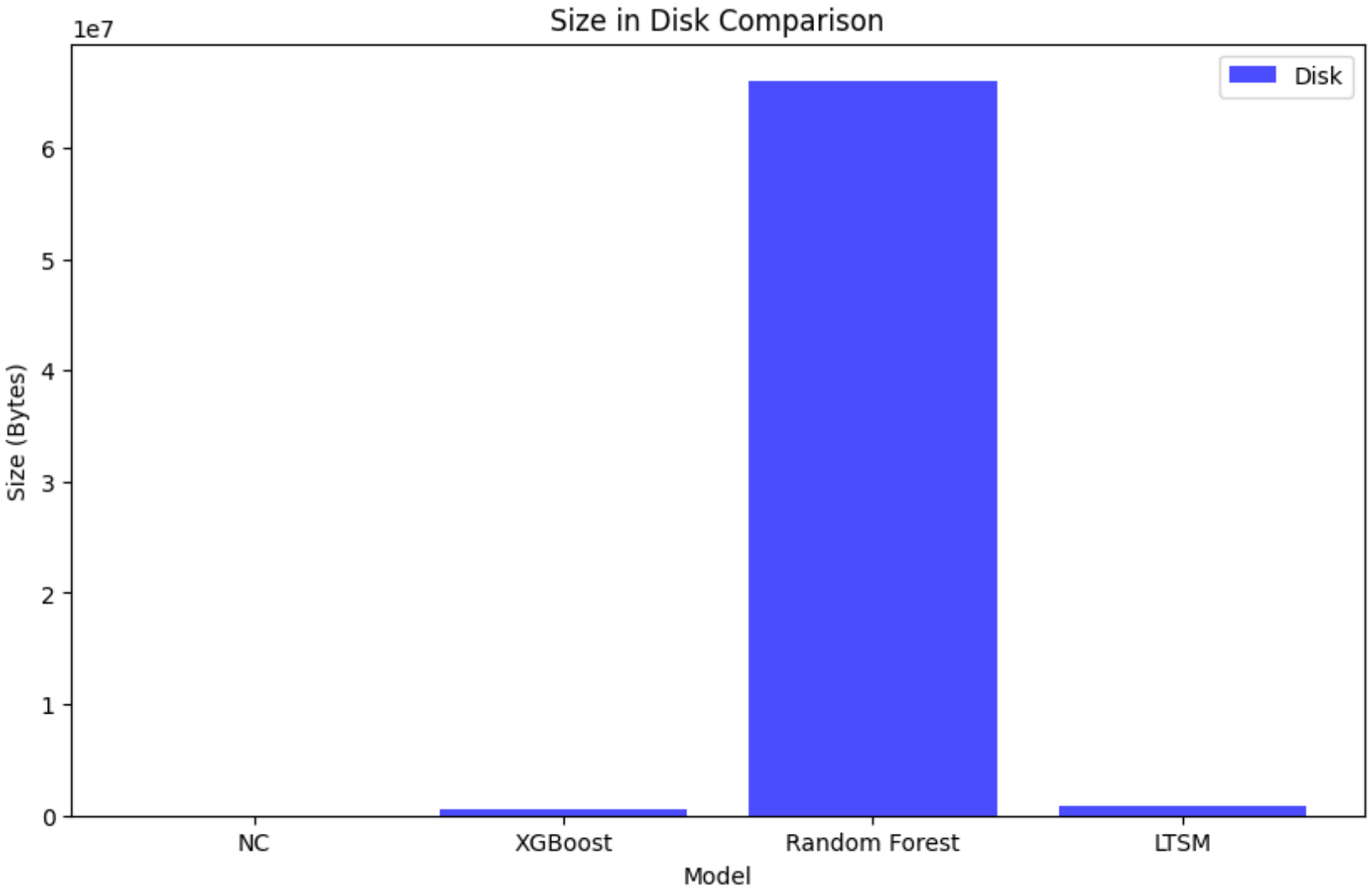


Comparando memória, disco, tempo de cpu dos classificadores

UNIVERSIDADE FEDERAL DO CEARÁ

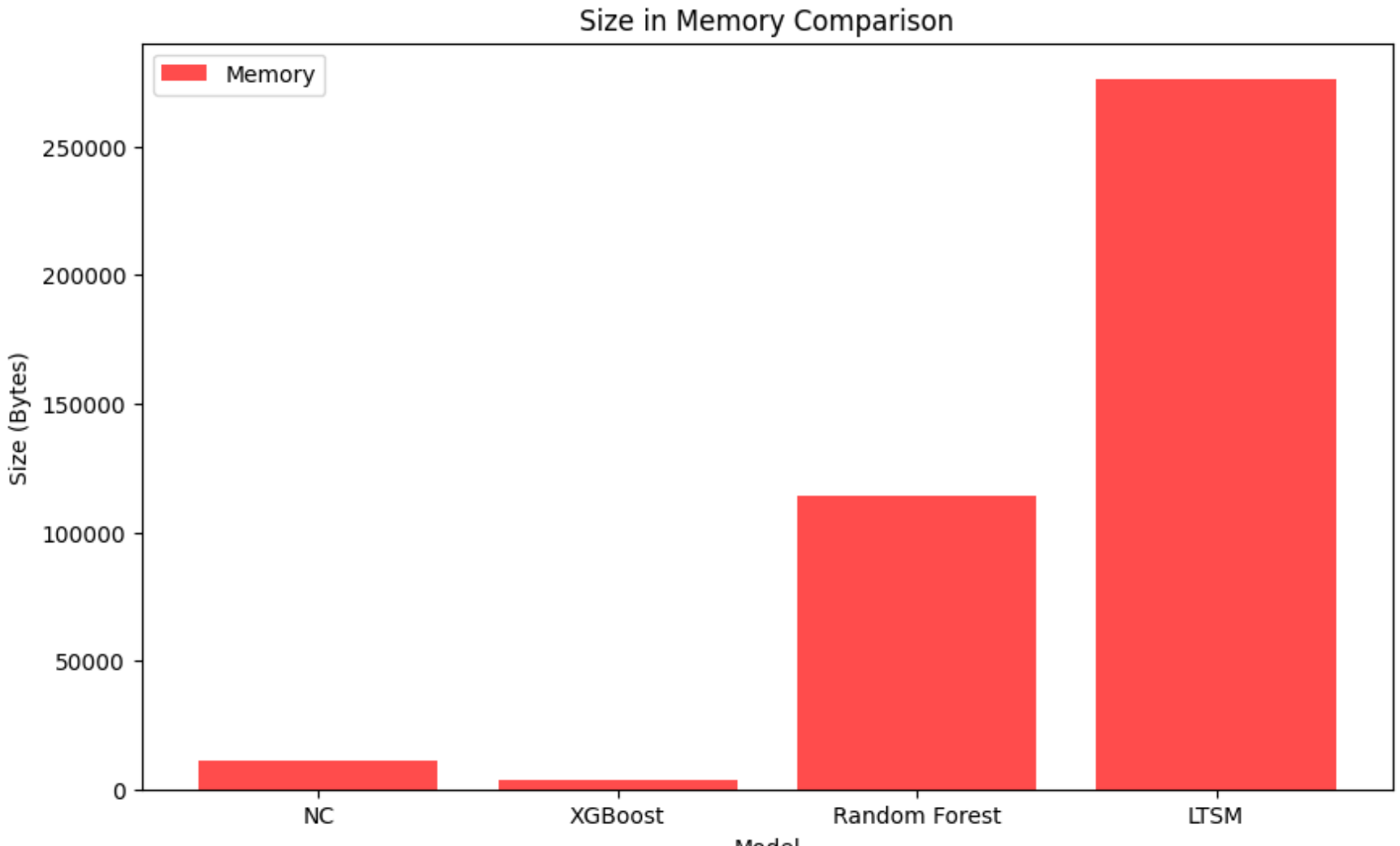
Espaço em Disco

Figura: Comparação do espaço em disco (Excluindo o KNN por motivos óbvios)



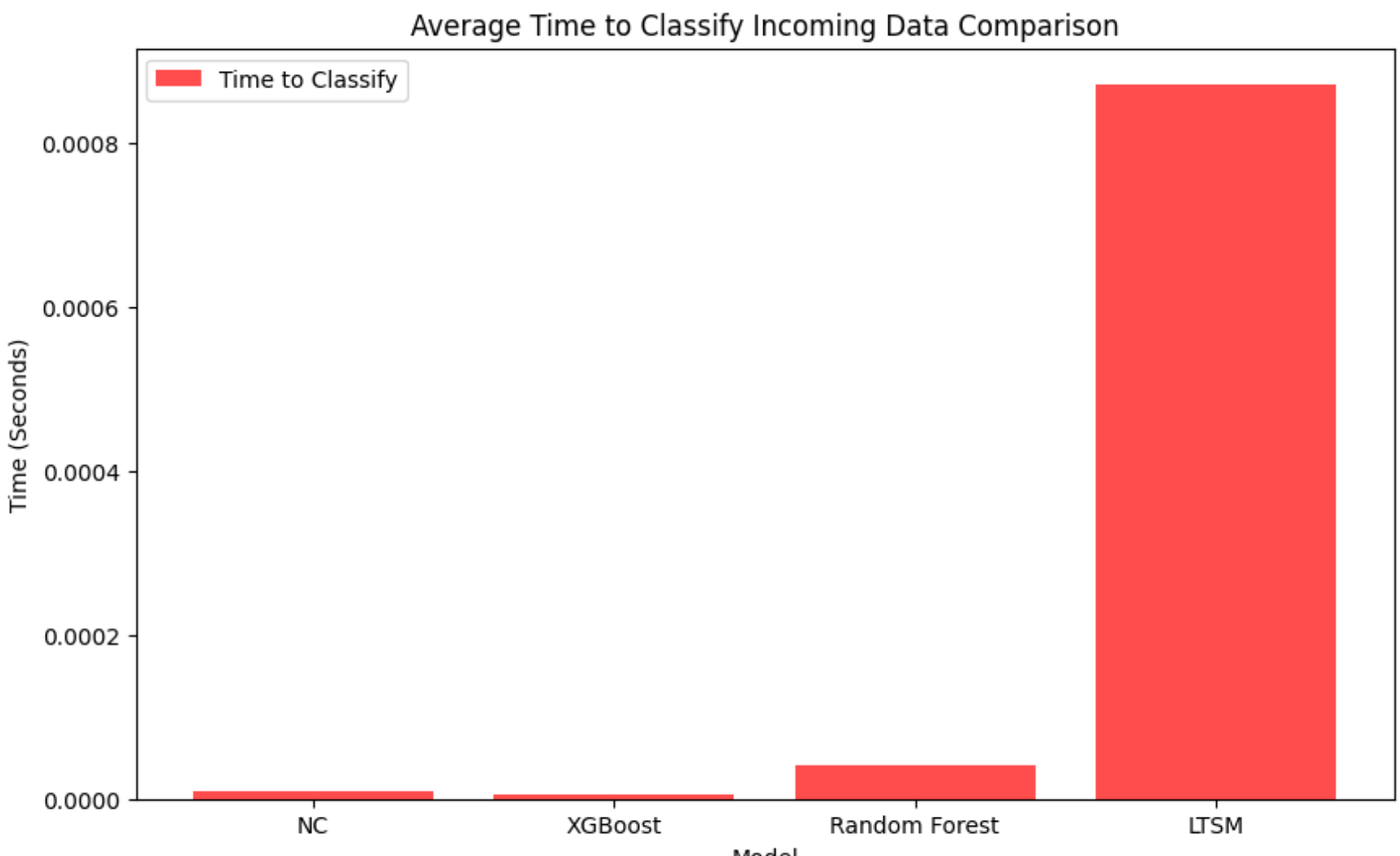
Espaço em Memória

Figura: Comparação do espaço em memória (Excluindo o KNN por motivos óbvios)



Tempo médio para realizar a inferência

Figura: Comparação do tempo médio para realizar uma inferência



Simulação de um ataque DDoS

UNIVERSIDADE FEDERAL DO CEARÁ

Ferramentas Utilizadas

- Python: Linguagem de programação para desenvolvimento dos scripts e modelos.
- PyShark + Wireshark: Para coleta dos dados para geração de um modelo customizado e para detecção em tempo real
- HPing3: Utilizado para simular um DoS com inundação de pacotes TCP.

UNIVERSIDADE FEDERAL DO CEARÁ

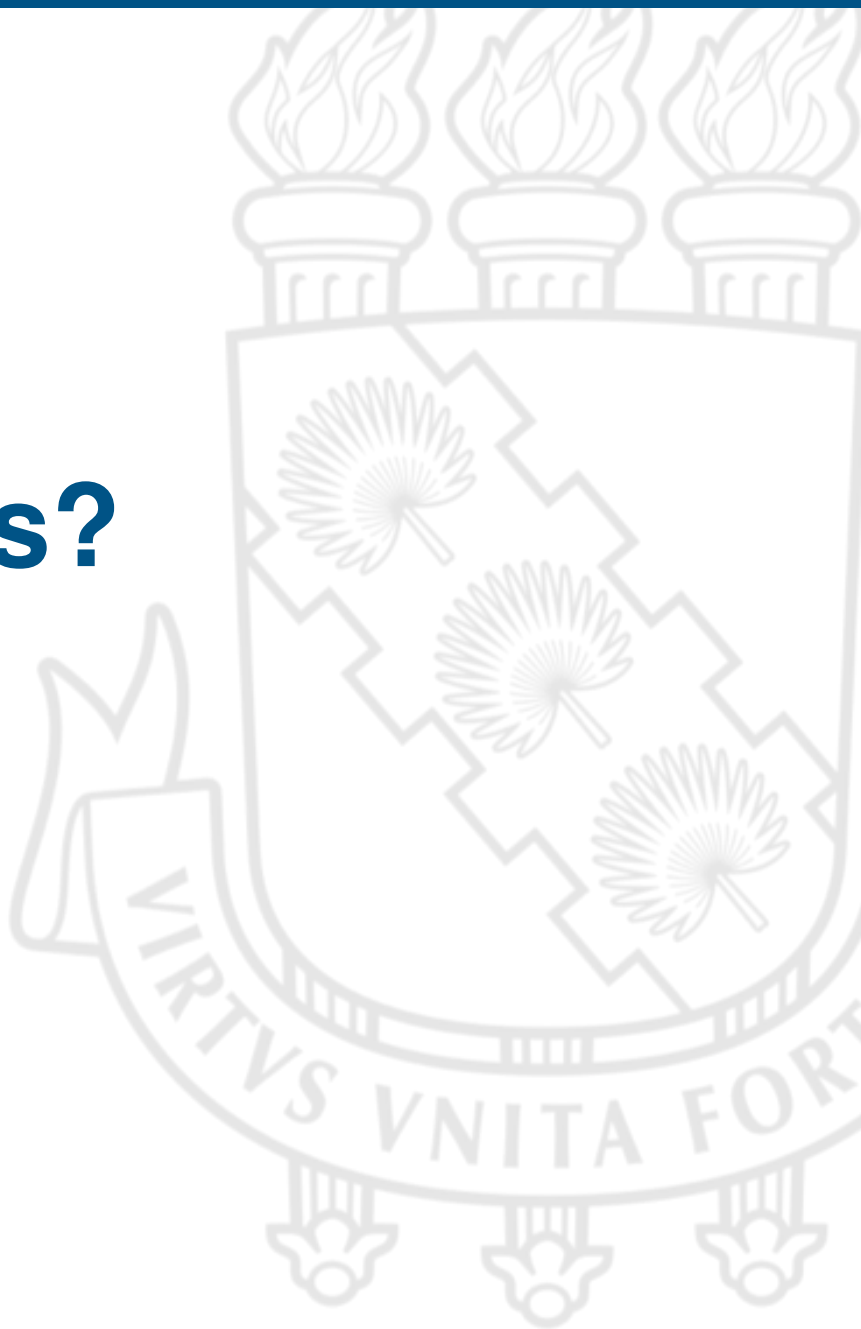
BORA PRO ATAQUE!

Fonte: https://www.peakpx.com/en/hd-wallpaper-desktop-kbwsf

Conclusão

UNIVERSIDADE FEDERAL DO CEARÁ

Dúvidas?



UNIVERSIDADE FEDERAL DO CEARÁ

Obrigado(a) pela Atenção!

UNIVERSIDADE FEDERAL DO CEARÁ