Capa

Índice

[1. Introdução 3](#_Toc211534254)

[2. Dados em análise 3](#_Toc211534255)

[3. KDD Process – Step 1 3](#_Toc211534256)

# Introdução

TBC

# Dados em análise

O dataset em análise pode ser encontrado no [Kaggle](https://www.kaggle.com/datasets/teamincribo/cyber-security-attacks/data?select=cybersecurity_attacks.csv) e contém 40 mil entradas com 25 colunas correspondentes a métricas numéricas e simbólicas relacionadas com tráfego de rede.

|  |  |  |  |
| --- | --- | --- | --- |
| Métrica | Descrição | Exemplo | Numérica / Simbólica |
| Timestamp |  | 2023-05-30 06:33:58 | Numérica |
| Source IP Address |  | 103.216.15.12 | Simbólica |
| Destination IP Address |  | 103.216.15.12 | Simbólica |
| Source Port |  | 31225 | Numérica |
| Destination Port |  | 17616 | Numérica |
| Protocol |  | ICMP | Simbólica |
| # Packet length |  | 503 | Numérica |
| Packet type |  | Data | Simbólica |
| Traffic type |  | HTTP | Simbólica |
| Payload Data |  | Text…… | Simbólica |
| Action Taken |  | Logged | Simbólica |
| Severity Level |  | Low | Simbólica |
| User Information |  | Sumar Rana | Simbólica |
| Device Information |  | Mozilla/5.0 (compatible; MSIE 8.0; Windows NT | Simbólica |
| Network Segment |  | Segment A | Simbólica |
| Geo-Location data |  | Jamshedpur, Sikkim | Simbólica |
| Proxy Information |  | 150.9.97.135 | Simbólica |
| Firewall Logs |  | Log Data | Simbólica |
| IDS/IPS alerts |  | NaN / Alert data | Simbólica |
| Log Source |  | Server | Simbólica |

# KDD Process – Step 1