Campos do dataset SVM:

TARGET

target

DIVERSOS

os\_timestamp, node\_name, iteration

CPU

mean\_os\_cpu\_ctx\_switches, mean\_os\_cpu\_guest, mean\_os\_cpu\_guest\_nice,mean\_os\_cpu\_idle, mean\_os\_cpu\_interrupts, mean\_os\_cpu\_iowait, mean\_os\_cpu\_irq, mean\_os\_cpu\_nice, mean\_os\_cpu\_soft\_interrupts, mean\_os\_cpu\_softirq, mean\_os\_cpu\_steal, mean\_os\_cpu\_syscalls, mean\_os\_cpu\_system, mean\_os\_cpu\_user,

DISCO

mean\_os\_disk\_discard\_io, mean\_os\_disk\_discard\_merges, mean\_os\_disk\_discard\_sectors, mean\_os\_disk\_discard\_ticks, mean\_os\_disk\_in\_flight, mean\_os\_disk\_io\_ticks, mean\_os\_disk\_read\_io, mean\_os\_disk\_read\_merge, mean\_os\_disk\_read\_sectors, mean\_os\_disk\_read\_ticks, mean\_os\_disk\_time\_in\_queue, mean\_os\_disk\_write\_io, mean\_os\_disk\_write\_merge, mean\_os\_disk\_write\_sectors, mean\_os\_disk\_write\_ticks,

MEMÓRIA

mean\_os\_mem\_nr\_active\_anon, mean\_os\_mem\_nr\_active\_file, mean\_os\_mem\_nr\_inactive\_anon, mean\_os\_mem\_nr\_inactive\_file, mean\_os\_mem\_nr\_mapped, mean\_os\_mem\_pgfault, mean\_os\_mem\_pgfree, mean\_os\_mem\_pgmajfault, mean\_os\_mem\_pgpgin, mean\_os\_mem\_pgpgout, mean\_os\_mem\_pgreuse,

REDE

mean\_os\_net\_bytes\_recv, mean\_os\_net\_bytes\_sent, mean\_os\_net\_dropin, mean\_os\_net\_dropout, mean\_os\_net\_errin, mean\_os\_net\_errout, mean\_os\_net\_num\_connections, mean\_os\_net\_packets\_recv, mean\_os\_net\_packets\_sent,

PROCESSO

mean\_process\_cpu\_children\_system, mean\_process\_cpu\_children\_user, mean\_process\_cpu\_iowait, mean\_process\_cpu\_system, mean\_process\_cpu\_user, mean\_process\_disk\_read\_bytes, mean\_process\_disk\_read\_chars, mean\_process\_disk\_read\_count, mean\_process\_disk\_write\_bytes, mean\_process\_disk\_write\_chars, mean\_process\_disk\_write\_count, mean\_process\_mem\_data, mean\_process\_mem\_dt, mean\_process\_mem\_lib, mean\_process\_mem\_resident, mean\_process\_mem\_shared, mean\_process\_mem\_size, mean\_process\_mem\_text, mean\_process\_net\_rx\_bytes, mean\_process\_net\_rx\_compressed, mean\_process\_net\_rx\_drop, mean\_process\_net\_rx\_errs, mean\_process\_net\_rx\_fifo, mean\_process\_net\_rx\_frame, mean\_process\_net\_rx\_multicast, mean\_process\_net\_rx\_packets, mean\_process\_net\_tx\_bytes, mean\_process\_net\_tx\_carrier, mean\_process\_net\_tx\_colls, mean\_process\_net\_tx\_compressed, mean\_process\_net\_tx\_drop, mean\_process\_net\_tx\_errs, mean\_process\_net\_tx\_fifo, mean\_process\_net\_tx\_packets,

CONTAINER

mean\_container\_cpu\_nr\_periods, mean\_container\_cpu\_nr\_throttled, mean\_container\_cpu\_system, mean\_container\_cpu\_throttled\_time, mean\_container\_cpu\_user, mean\_container\_disk\_8:0\_async, mean\_container\_disk\_8:0\_discard, mean\_container\_disk\_8:0\_read, mean\_container\_disk\_8:0\_sync, mean\_container\_disk\_8:0\_total, mean\_container\_disk\_8:0\_write, mean\_container\_mem\_active\_anon, mean\_container\_mem\_active\_file, mean\_container\_mem\_cache, mean\_container\_mem\_inactive\_anon, mean\_container\_mem\_inactive\_file, mean\_container\_mem\_mapped\_file, mean\_container\_mem\_pgfault, mean\_container\_mem\_pgmajfault, mean\_container\_mem\_pgpgin, mean\_container\_mem\_pgpgout, mean\_container\_mem\_rss, mean\_container\_mem\_unevictable, mean\_container\_net\_rx\_bytes, mean\_container\_net\_rx\_compressed, mean\_container\_net\_rx\_drop, mean\_container\_net\_rx\_errs, mean\_container\_net\_rx\_fifo, mean\_container\_net\_rx\_frame, mean\_container\_net\_rx\_multicast, mean\_container\_net\_rx\_packets, mean\_container\_net\_tx\_bytes, mean\_container\_net\_tx\_carrier, mean\_container\_net\_tx\_colls, mean\_container\_net\_tx\_compressed, mean\_container\_net\_tx\_drop, mean\_container\_net\_tx\_errs, mean\_container\_net\_tx\_fifo, mean\_container\_net\_tx\_packets,

https://claude.ai/chat/13c5ae50-7314-4387-9ebf-0c04bbecd4da

**Informação Mútua vs Ganho de Informação:** São conceitos relacionados! Informação mútua mede quanto "saber o valor de X reduz a incerteza sobre Y". É uma medida mais robusta que lida bem com features contínuas.

ANOTAÇÕES DE REUNIÃO COM OS PROFESSORES

Treinamento(40), Teste(30) e validação(30)

De maneira aleatória considerando a distribuição das classes

Treinando vários modelos: SVM, Decision Tree, Knn,

Medição: matriz de confusão, curva ROC, F1

Calcular o ganho de informação das features.

O modelo será gerado a partir dos dados de validação.

Utilizar a validação no grid\_search para definir os melhores parâmetros do modelo.

Preparação dos dados

Separação em treino, teste e validação

Fazer o ganho de informação no modelo de treino

Fazer gridsearch com os dados de validação (que modelo vou treinar)

SVM, Random Forest, KNN, Xboost,

Treinar os modelos com treino e teste

Fazer a matriz de confusão e demais métricas

https://claude.ai/chat/90e0995d-91a8-452f-a3c4-ee70ee2dce92

**Papel de Cada Conjunto**

| **Conjunto** | **Uso** | **Análise de Features** |
| --- | --- | --- |
| **Treino** | Treinar modelo e selecionar features | ✅ SIM |
| **Validação** | Ajustar hiperparâmetros | ❌ NÃO |
| **Teste** | Avaliação final | ❌ NÃO |

**Diferenças Importantes**

| **Aspecto** | **Ganho de Informação** | **Informação Mútua** |
| --- | --- | --- |
| **Simetria** | ❌ Não: IG(Y,X) vs IG(X,Y) | ✅ Sim: MI(X,Y) = MI(Y,X) |
| **Foco** | Redução de entropia do **target** | Dependência **mútua** |
| **Uso comum** | Árvores de decisão | Seleção de features |
| **Interpretação** | "X ajuda a prever Y" | "X e Y se relacionam" |
| **Normalização** | Geralmente não | Pode ser normalizada (0-1) |

Treinamento

Texto

O conteúdo gerado por IA pode estar incorreto.

Texto

O conteúdo gerado por IA pode estar incorreto.

Tabela

O conteúdo gerado por IA pode estar incorreto.

Texto, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Gráfico, Gráfico de linhas

O conteúdo gerado por IA pode estar incorreto.

Interface gráfica do usuário, Texto, Aplicativo, Email

O conteúdo gerado por IA pode estar incorreto.

Gráfico, Gráfico de linhas

O conteúdo gerado por IA pode estar incorreto.

Texto

O conteúdo gerado por IA pode estar incorreto.

Uma imagem contendo Forma

O conteúdo gerado por IA pode estar incorreto.

Interface gráfica do usuário, Texto, Aplicativo

O conteúdo gerado por IA pode estar incorreto.