



Anexo - AvantIntel



Soluções de Vanguarda em Segurança da Informação

Sumário

Conteúdo

Sumário	2
1. Introdução	3
2. Sobre	4
3. Instalação	6
Pré-requisitos	6
Scripts	6
Templates	7
4. Utilização	8
5. Testando	11



Soluções de Vanguarda em Segurança da Informação

1. Introdução

O AvantIntel é o módulo de inteligência artificial para identificação de padrões, anomalias comportamentais de entidades, como por exemplo, usuários, endereços, dispositivos ou qualquer elemento analisado, além de possibilitar a predição de eventos temporais, baseado na análise de dados históricos. Desta forma, é possível gerar alertas ou ou ainda tomar ações corretivas quando for reconhecido um desvio de padrão comportamental de um elemento individual ou de grupos específicos, reconhecido de uma base de milhares de entidades analisadas.



Soluções de Vanguarda em Segurança da Informação

2. Sobre

Com o objetivo de detectar anomalias, o AvantIntel utiliza 3 algoritmos para identificar irregularidades em 5 diferentes conjuntos de parâmetros.

Algoritmos utilizados:

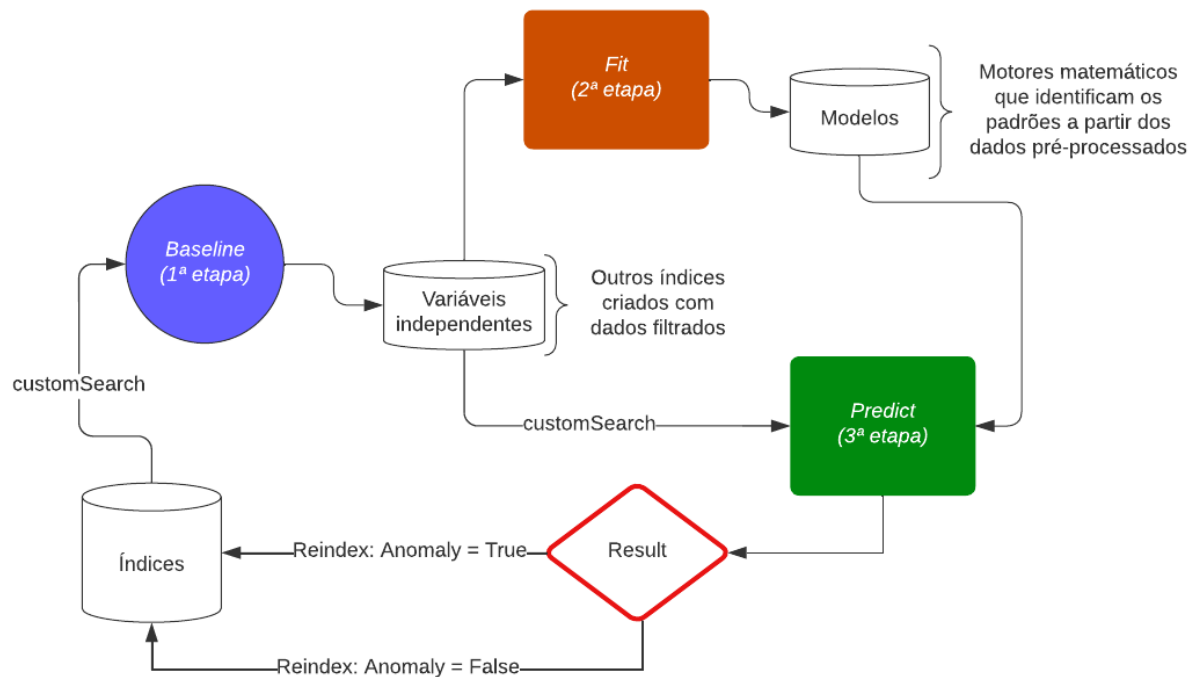
- [KNN](#) (K-ésimo Vizinho mais Próximo)
- [CBLOF](#) (Fator Outlier Local Baseado em Clustering)
- [PCA](#) (Análise de Componentes Principais)

Conjuntos de modelos padrões:

- [ED](#) (Domínio Externo)
- [SLBDW](#) (Logins Suspeitos por Dia da Semana)
- [SLBH](#) (Logins Suspeitos por Hora)
- [SLBIA](#) (Logins Suspeitos por Endereço IP)
- [SLBSN](#) (Logins Suspeitos por Nome de Serviço)
- Outros modelos podem ser criados a partir dos exemplos acima.

A detecção de anomalias identifica itens ou eventos que são significativamente diferentes da maioria dos dados. Portanto, com a finalidade de separar as irregularidades, primeiramente os dados são coletados e separados em 5 conjuntos de parâmetros por meio de uma [baseline](#), consequentemente modelos são criados a partir de um [fit](#) e por fim a detecção de anomalias é efetivada pelo [predict](#).

As etapas de funcionamento podem ser visualizadas pelo seguinte diagrama:



3. Instalação

O procedimento de instalação consiste no envio dos arquivos do AvantIntel para o AvantData e na verificação dos pré-requisitos.

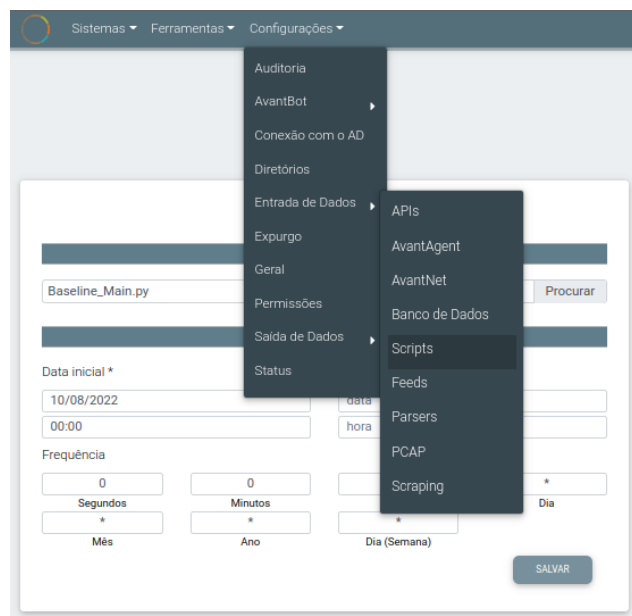
Pré-requisitos

- Ter o serviço do RA2D ativado
 - Para verificar: `avantsystem$ systemctl -l status ra2d`
- Ter as entradas relacionadas ao modelo já definidas e recebendo, com dados suficientes no índice para formar uma amostra que possibilite a criação da baseline e o reconhecimento de padrões;

Scripts

Os scripts disponibilizados na área de scripts (*Baseline_Main.py*, *Fit_Main.py* e *Predict_Main.py*) devem ser enviados para o ambiente de execução de scripts do AvantData.

Para enviar, vá em **Configurações > Entrada de Dados > Scripts**. Carregue cada arquivo de uma vez no campo **Entrada** e ajuste a **Frequência** de execução conforme o desejado.



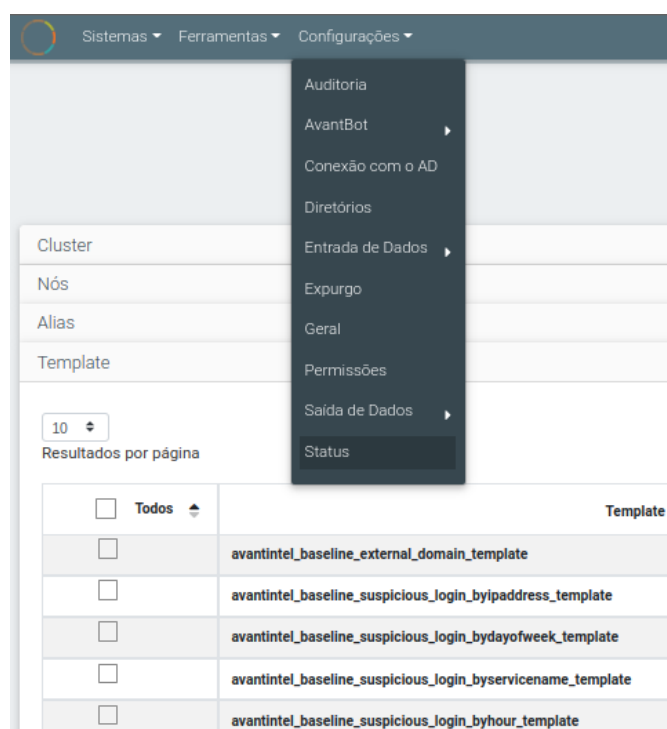
Como a ordem de execução dos scripts deve ser Baseline_Main.py > Fit_Main.py > Predict_Main.py, recomenda-se ajustar a **Frequência** seguindo a ordem correta. Pode-se colocar, por exemplo, o baseline para ser executado em todo minuto 0, o fit em todo minuto 5 e o predict em todo minuto 10. Favor atentar ao volume de dados analisados e ao sizing do AvantData para a correta definição do tempo de execução. Pode ser necessária a execução de configurações de escalonamento do sistema para atender as demandas dos modelos.

Templates

Os templates para cada item a ser tratado nas anomalias devem ser enviados para o ambiente de templates dentro do AvantData.

Para enviar, vá em **Configurações > Status**. Clique com o botão direito em cima de **Template** e consequentemente em **+Novo Template**. Adicione os templates dentro da pasta [templates](#), de acordo com os modelos que estão sendo aplicados.

Atenção: A aplicação dos templates depende da correta definição dos tipos dos campos, assim como da quantidade de shards primários e possíveis réplicas.







4. Utilização

Após o script de predição ser executado, os índices originais serão atualizados com as informações de detecções das anomalias.

Utilizando como exemplo o índice do AvantAgent, na busca Índice='AvantAgent' CONTÉM='AvantIntel:*' os campos de **Result**, **False Negative** e **False Positive** aparecerão para cada um dos 5 conjunto de parâmetros analisados e para cada um dos 3 algoritmos.

Juntamente com os campos adicionados, um botão **Anomalia** é disponibilizado para mudar manualmente o resultado, seguindo a tabela de cores:

Quantidade de Resultados Positivos	Cor do botão <i>Anomalia</i>
0	 Cinza
1	 Amarelo
2	 Laranja
3 ou mais	 Vermelho

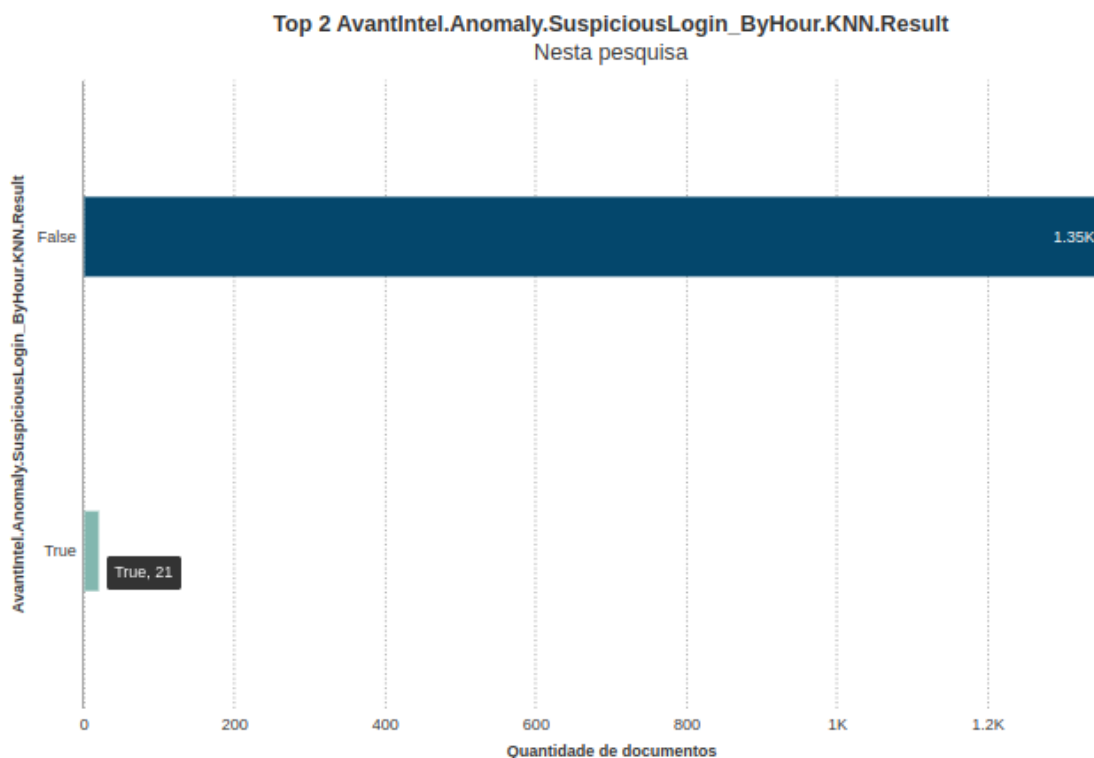
Abaixo um exemplo de como a anomalia é identificada dentro de uma pesquisa rápida:

recordNumber: 52382542 monitorName: AvantAgent_Security idMonitor: 14 agentHostName: AVANTSRV10 TaskDisplayName: Kerberos Service Ticket Operations Task: 14337 Keywords: 0x8020000000000000 Provider Name: Microsoft-Windows-Security-Auditing Correlation ActivityID: Channel: Security Opcode: 0 Execution ProcessID: 816 Guid: {54849625-5478-4994-A5BA-3E3B0328C30D} TimeCreated SystemTime: 2022-07-26T23:43:16.531896000Z EventRecordID: 52382542 Version: 0 ThreadID: 9016 Computer: AVANTSRV10.avantsec.com.br EventID: 4769 Level: 0 Status: 0x0 TicketEncryptionType: 0x12 LogonGuid: {1F838CDF-EBBD-53AF-D891-3C5745873BC3} ServiceName: CLAUDIO\$ TicketOptions: 0x40810000 ServiceSid: S-1-5-21-978060741-3721912390-2439747195-1746 IpAddress: ::ffff:192.168.100.75 IpPort: 55285 TargetUserName: claudio@AVANTSEC.COM.BR TransmittedServices: - TargetDomainName: AVANTSEC.COM.BR Level: 0 LevelDisplayName: Information category: security idAgent: 27 statusCode: System.Diagnostics.Eventing.Reader.EventLogRecord GenerateTime: 26/07/2022 20:44:02 FalsePositive: false FalseNegative: false Result: false FalsePositive: false FalseNegative: false Result: false FalsePositive: false FalseNegative: false Result: false FalsePositive: false FalseNegative: false Result: false FalsePositive: false FalseNegative: false Result: false FalsePositive: false FalseNegative: false Result: false FalsePositive: false FalseNegative: false Result: false FalsePositive: false FalseNegative: false Result: true

_index: avantagent_security-
26.07.2022 _type: avantagent_security _id: 4608830B68B43814EAF8C67DE6F308EB GenerateTime: 26/07/2022 20:44:02 Anomalia

No exemplo acima, o evento foi considerado uma anomalia por um dos algoritmos usados na análise dos dados da amostra. Na expansão do campo result é possível ter uma visão rápida sobre os documentos da amostra, mostrando,

segundo o KNN, neste caso, quantos foram considerados anomalia (True) e quantos foram considerados o comportamento normal (False).



Para redefinir manualmente os campos no botão **Anomalia**, basta selecionar o modelo e o algoritmo, escolhendo se o resultado se enquadra em um **Falso Positivo** ou um **Falso Negativo**. Os resultados mudados servirão como uma classificação para enriquecer os próximos modelos a serem gerados.

Classificação de anomalias

Índice:

palalto-10.08.2022

Tipo:

palalto-10.08.2022

Id:

AYKlpGcDMSXBBExrewkn

Modelo de Anomalia

External_Domain

Tipo de Anomalia

KNN

FALSO NEGATIVO

FALSO POSITIVO

REMOVER CLASSIFICAÇÃO

5. Testando

Para testar o AvantIntel e descobrir erros de funcionamento, pode-se rodar os scripts manualmente. Para isso, deve-se colocar os scripts em uma pasta local e mudar o **base_url** para o endereço de funcionamento do AvantData

```
base_url = 'https://127.0.0.1'
scroll_time = '5m'
time_filter = '3M'
```

As configurações de **scroll_time** e **time_filter** são relativas ao tempo de armazenameno dos dados em scroll e o espaço de tempo de busca nos índices. As unidades de tempo são definidas conforme a tabela:

Unidade	Duração
y	Anos
M	Meses
w	Semanas
h	Horas
H	Horas
m	Minutos
s	Segundos

Baseline_Main.py

Primeiramente o arquivo **Baseline_Main.py** deve ser executado. A sua execução gera os índices que serão utilizados posteriormente para produção dos modelos e predição. Caso a execução seja bem sucedida, novos índices contendo o nome dos templates serão adicionados:

Configurações > Status > Índices.



Soluções de Vanguarda em Segurança da Informação

Fit_Main.py

Sendo Fit_Main.py o segundo arquivo a ser executado, deve gerar uma pasta /ML contendo arquivos com a extensão .pkl. Produz arquivos contendo técnicas de pré-processamento de dados, sendo um *_**ordinalEncoder.pkl**, um *_**standardscaler.pkl** e um modelo *_**estimator.pk** para cada template e para cada algoritmo.

Predict_Main.py

Como último arquivo a ser executado, o [Predict_Main.py](#) deve reindexar os índices originais com os resultados e os falsos positivos contendo o valor de *True* ou *False*.