

EXTRAÇÃO DE INFORMAÇÕES

E

ANÁLISE DE SENTIMENTOS

Luciano Ignaczak

Sócio-Fundador

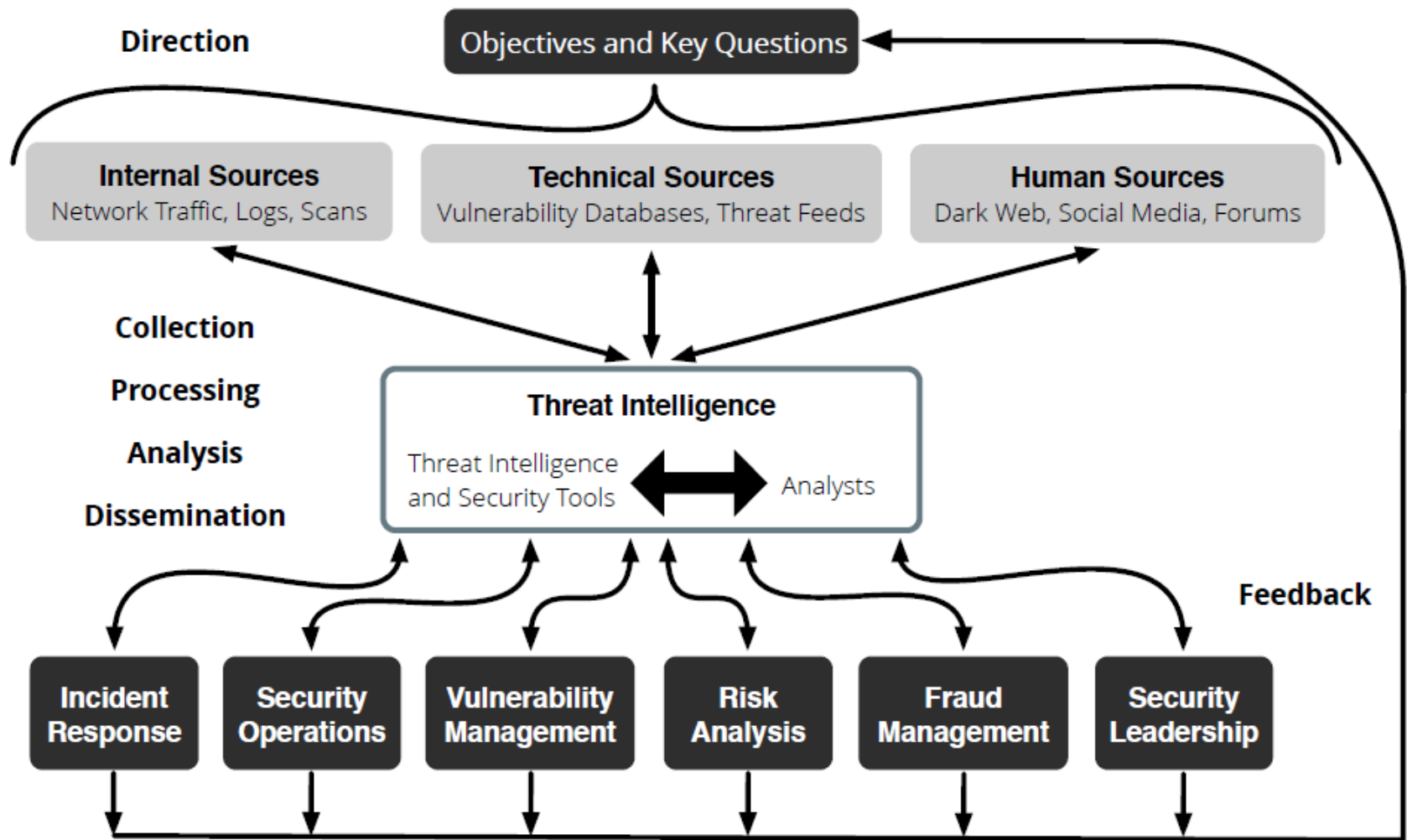
ignaczak@icybersec.com

Márcio Garcia Martins

Sócio-Fundador

marciogm@icybersec.com

Cyber Threat Intelligence (CTI)



Fonte: Recorded Future. **Threat Intelligence Handbook**. 2018.

Extração de Informações (IE)

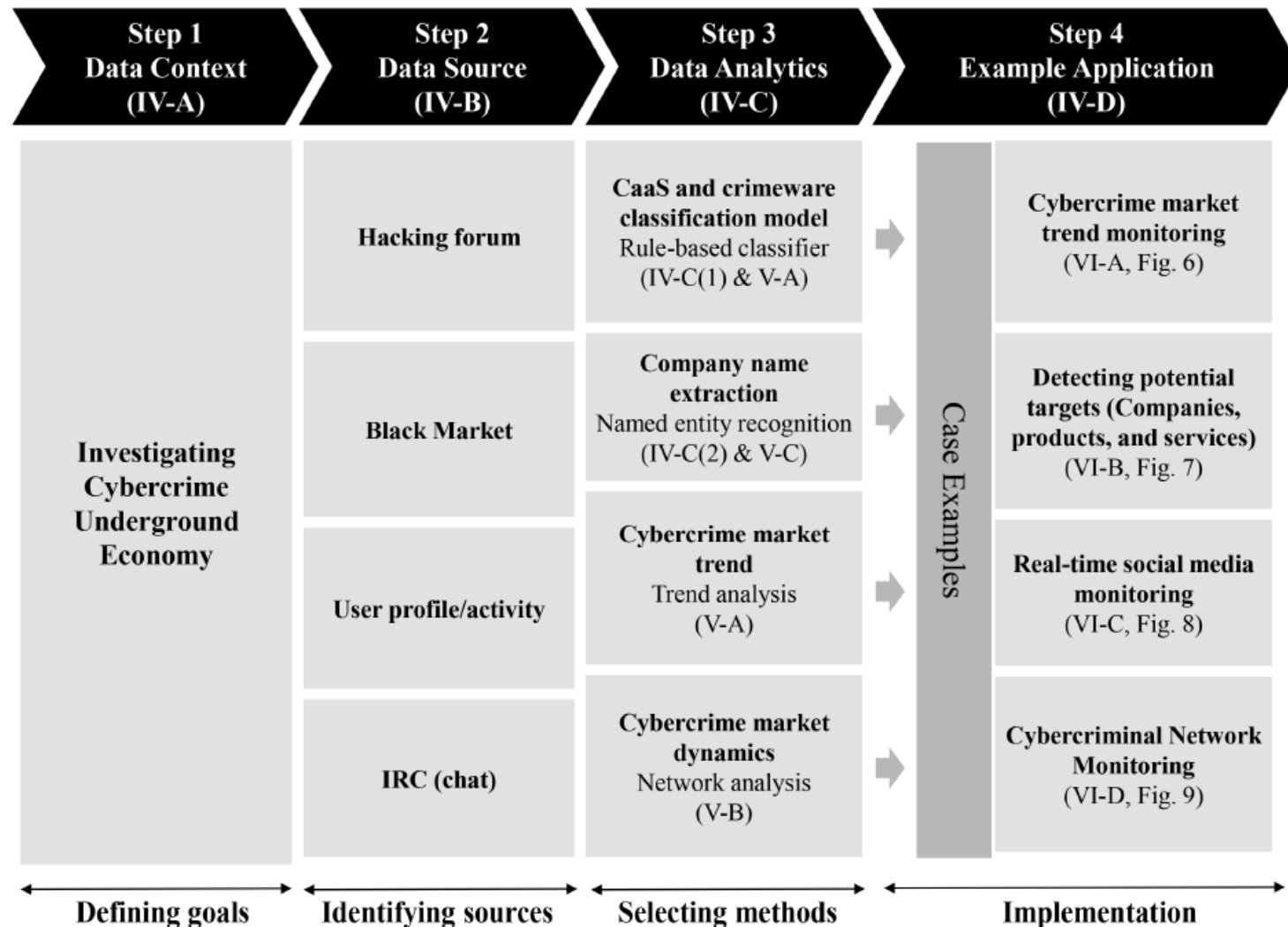
- O objetivo da extração de informações é descobrir informação estruturada em uma coleção de documentos não estruturados;
 - Através da extração de informações é possível realizar:
 - Named Entity Recognition – NER – tem como objetivo a identificação de entidades em um texto e sua classificação em tipos previamente definidos;
 - Extração de relações – consiste na detecção e caracterização de relações semânticas entre entidades no texto.
-

In 1998, Larry Page and Sergey Brin founded Google Inc.

*FounderOf(Larry Page, Google Inc.),
FounderOf(Sergey Brin, Google Inc.),
FoundedIn(Google Inc., 1998).*

Fonte: AGGARWAL, CHARU C. **Mining Text Data**. 2012.

Aplicação de IE em CTI

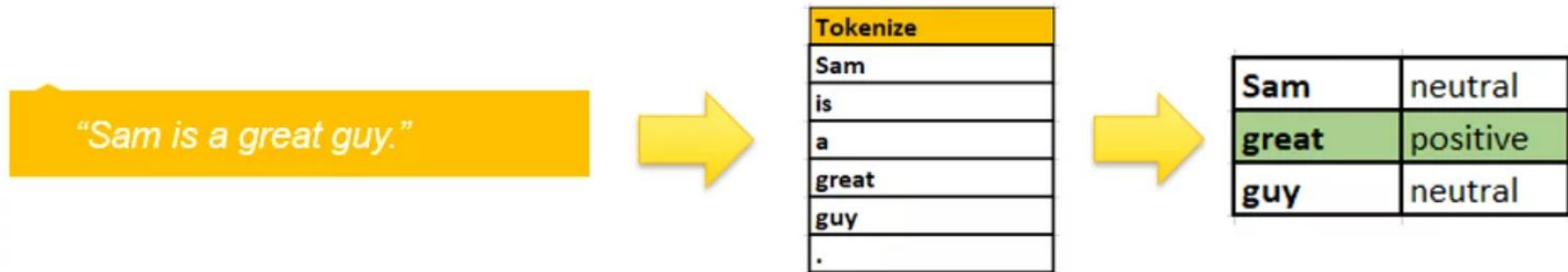


Fonte: AN, JUNGKOOK and KIM, HEE-WOONG. A Data Analytics Approach to the Cybercrime Underground Economy. 2018.

Análise de Sentimentos (SA)

- O objetivo da análise de sentimentos é explicar se um texto contém emoções positivas ou negativas;
- Análise de sentimentos pode identificar as emoções de um documento ou de partes do documento, por exemplo, de cada sentença que forma um texto;
- Duas abordagens podem ser aplicadas para realizar análise de sentimentos:
 - Dicionário (léxico): atribui polaridade para as palavras de um dicionário previamente criada.
 - Aprendizado de máquina: baseada no treinamento de um classificador usando um dataset com sentimentos anotados.

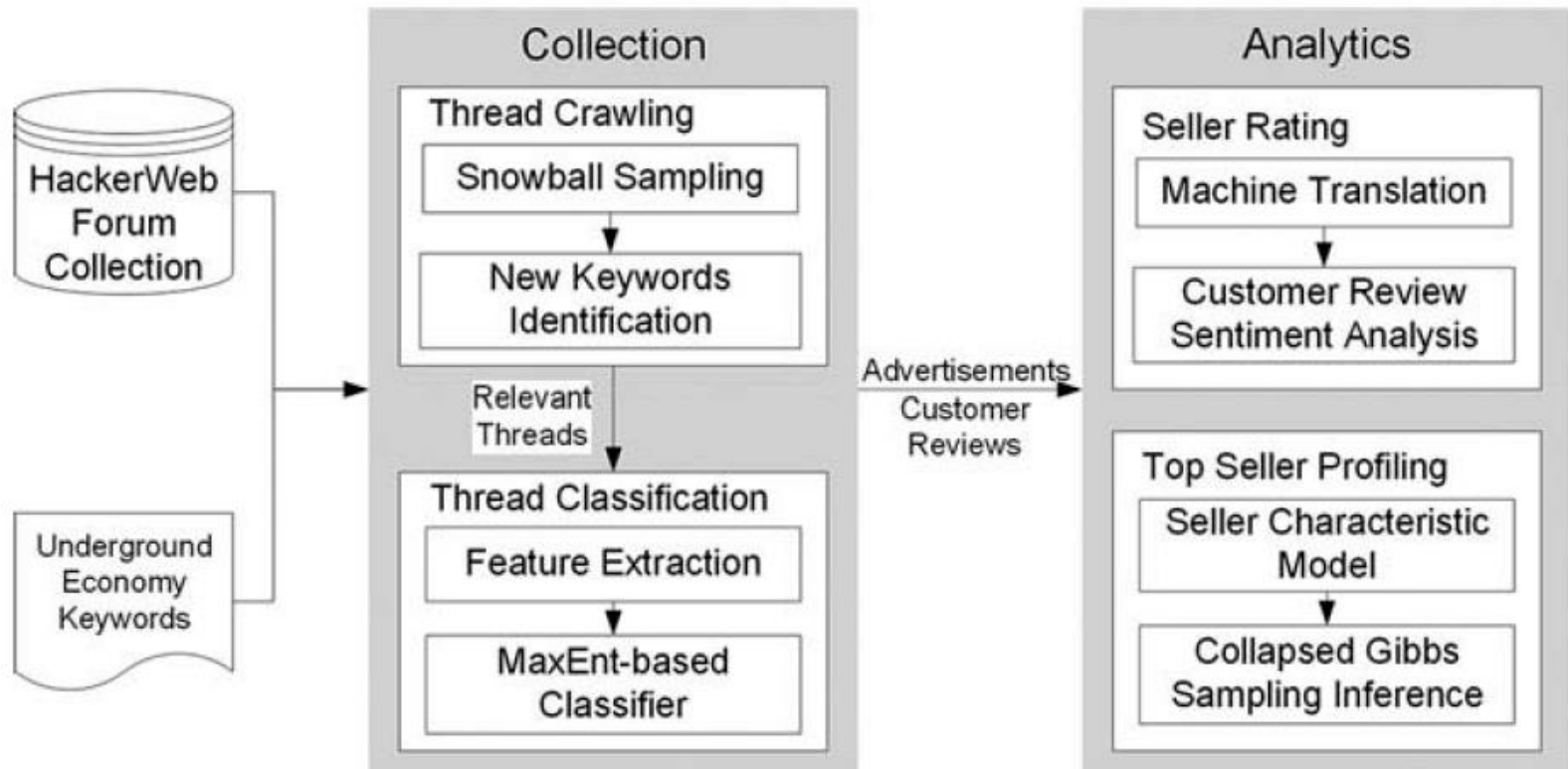
Análise de Sentimentos



Word	Sentiment	Valence/Intensity
okay	Positive	0.9
good	Positive	1.9
great	Positive	3.1
horrible	Negative	-2.5
:(Negative	-2.2
Sucks (or the slang: sux)	Negative	-1.5

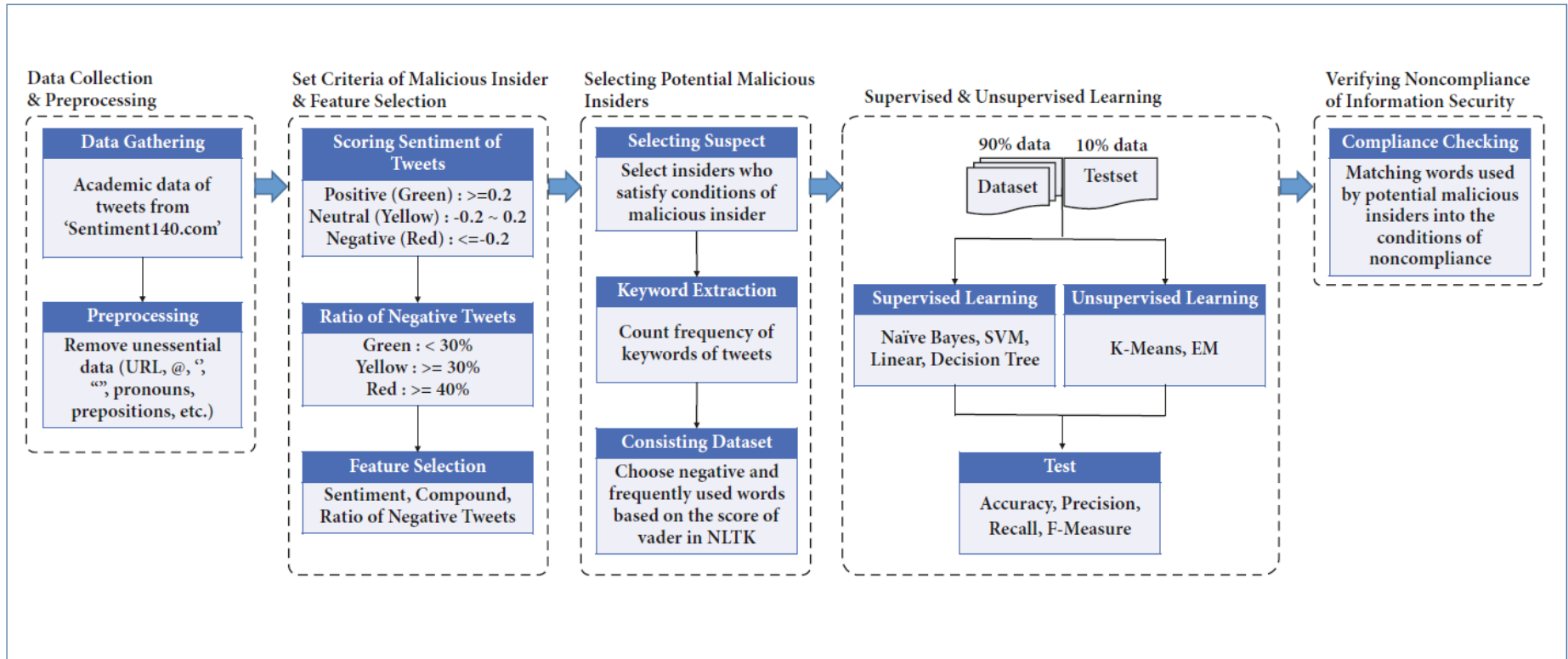
Fonte: ALPHABOLD. **Sentiment Analysis**: the Lexicon based Approach. 2019.

Aplicação de SA em CTI



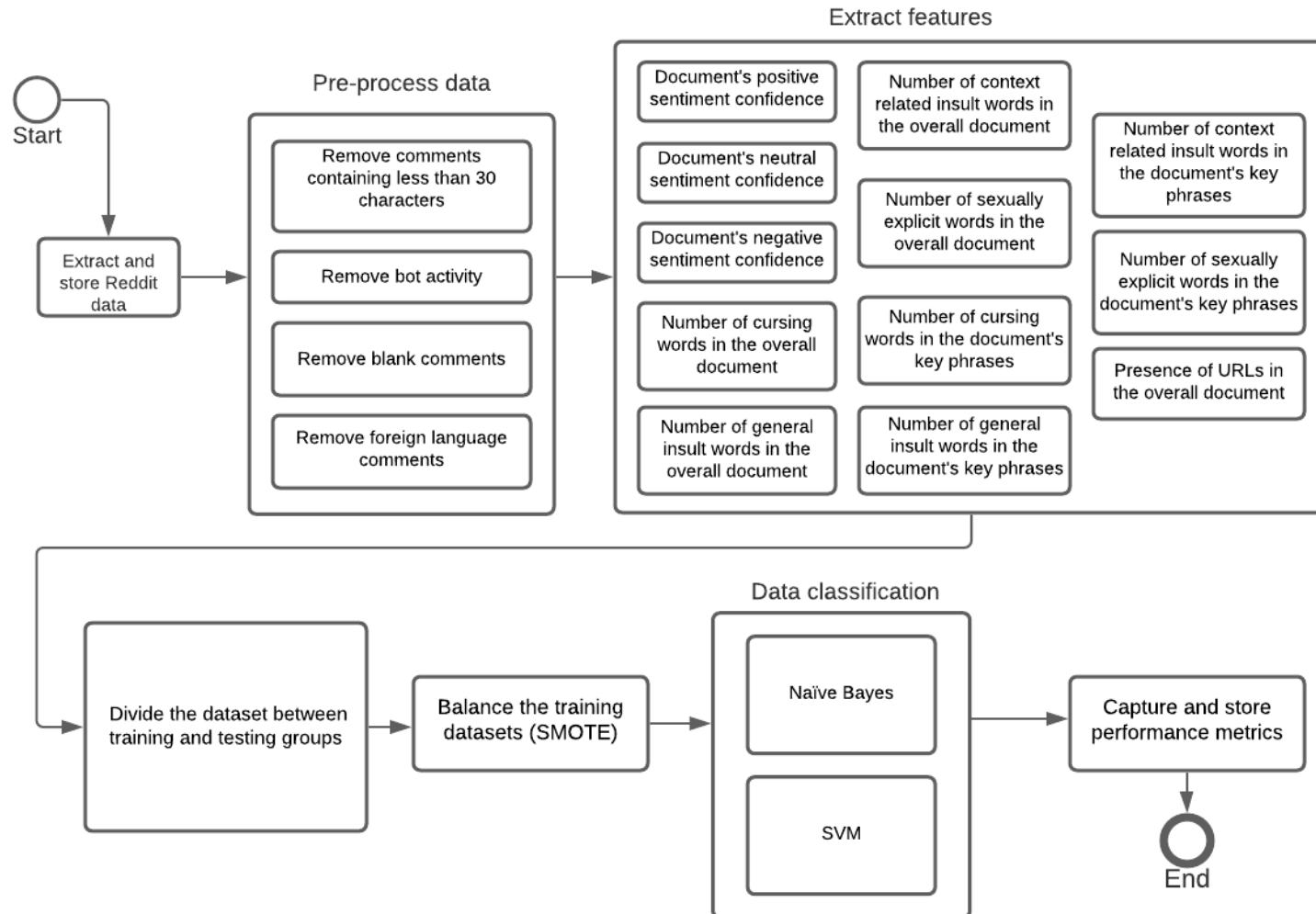
Fonte: LI, WEIFENG et al. **Identifying and Profiling Key Sellers in Cyber Carding Community**: AZSecure Text Mining System. 2016

Aplicação de SA em Segurança Cibernética



Fonte: PARK, WON et al. **Detecting Potential Insider Threat**: Analyzing Insiders' Sentiment Exposed in Social Media. 2018

Aplicação de SA e IE



Fonte: EBERHARDT, CAROLINA and IGNACZAK, LUCIANO. **A text mining model proposal for the identification of cyberbullying messages in the Brazilian Portuguese language.** 2021