Machine learning techniques for flow-based intrusion detection systems

...

Axel Faes: Bachelorthesis

Onderwerp

Intrusie detectie

- Classificeren/detectie van onverwacht netwerkgedrag
- Extern en intern

IP Flows

- Geaggregeerd vanuit packet data
- specifiek bedoeld voor high traffic systems
- vereist geen in-depth knowledge van het netwerk
- Hoe kan flow data gebruikt worden in een IDS?
- Welke types anomalie kunnen we automatisch detecteren ?

Onderwerp

machine learning technieken

Algoritmes ‘leren’ zelf zonder regeltjes expliciet te programmeren

7 verschillende algoritmes

In hoeverre zijn machine learning technieken inzetbaar voor anomaly detection?

Kunnen we een IDS maken dat out-of-the-box een aanvaardbare ‘hit rate’ biedt ?

Zijn dergelijke technieken bruikbaar in real-life condities ?

Werking:

Stap 1

Aanleren van het model via subset van een learning data set

Stap 2

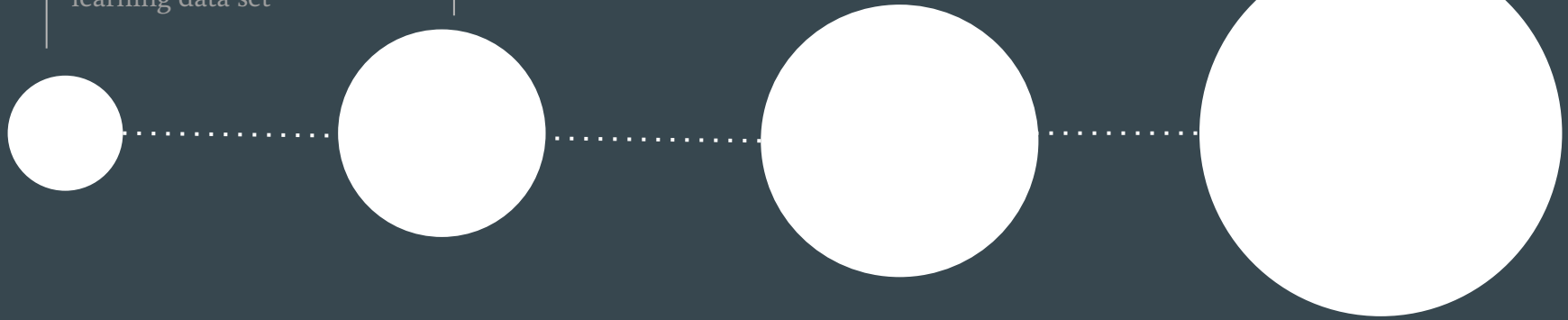
Validatie met gekende test-data uit learning data set

Stap 3

Testing met real-world gelabelde data

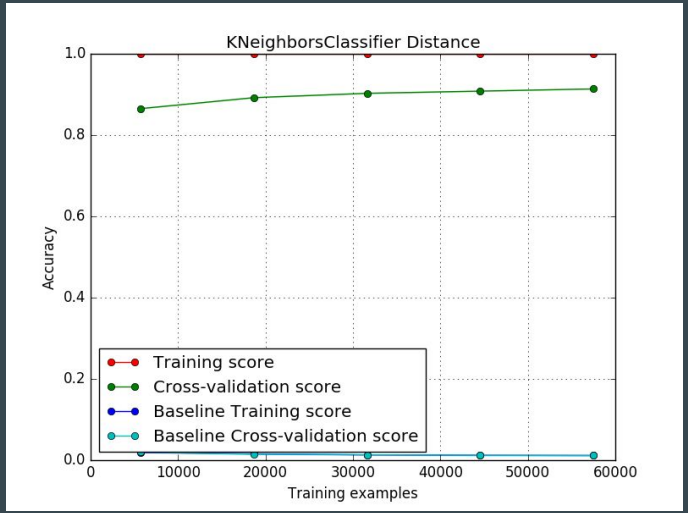
Stap 4

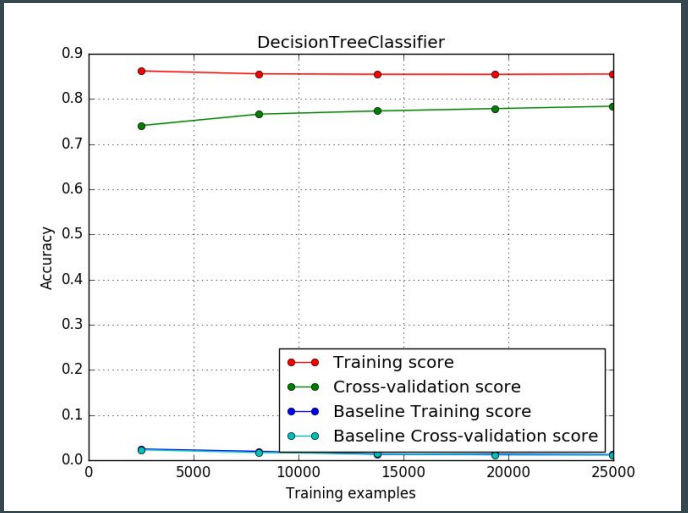
Validatie met ongelabelde real-world data



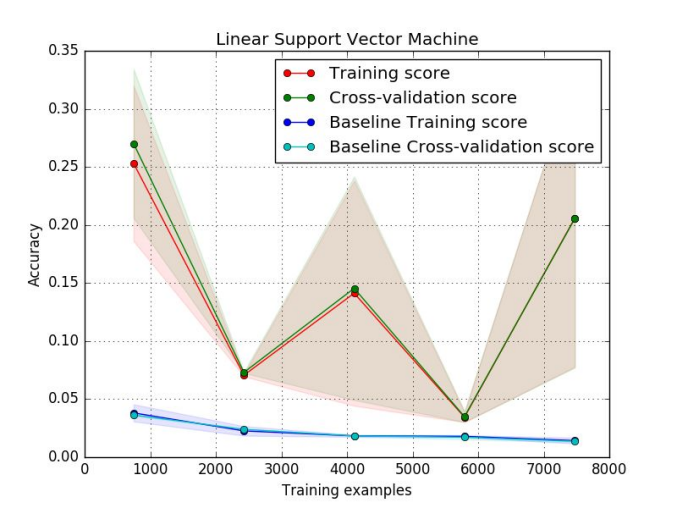
Demo

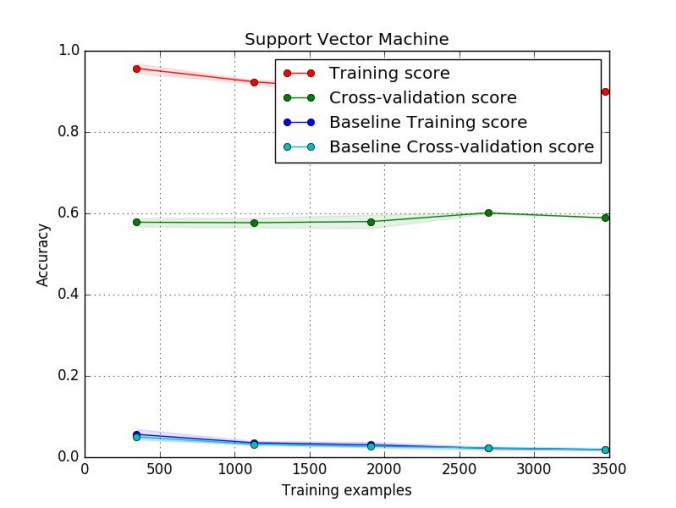
Stap 1: Learning curves



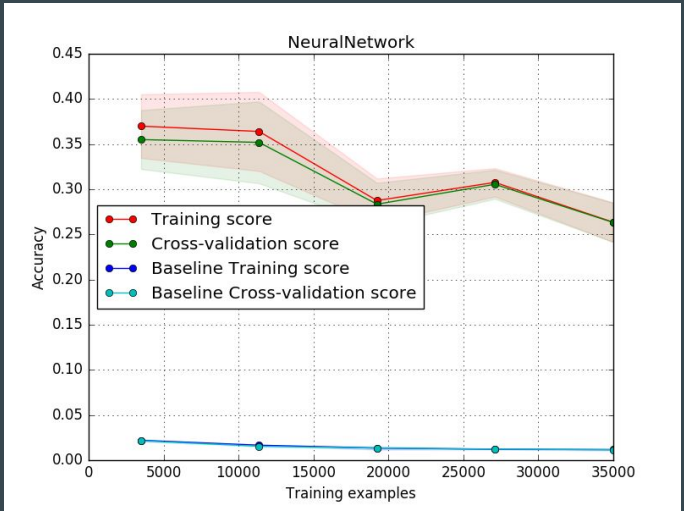


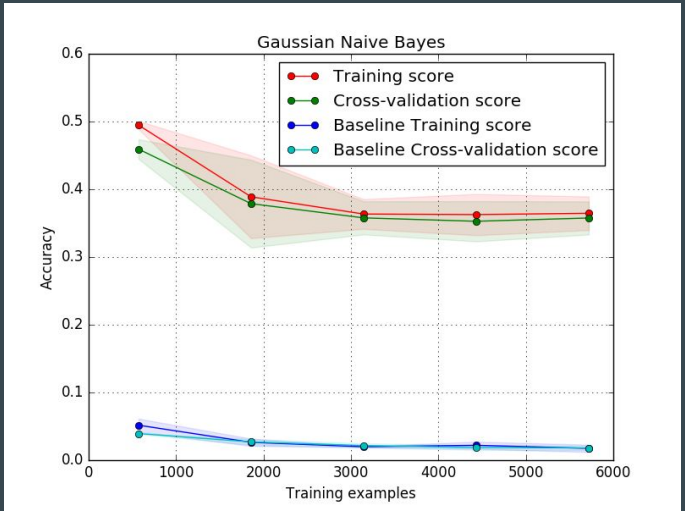
Stap 1: Learning curves



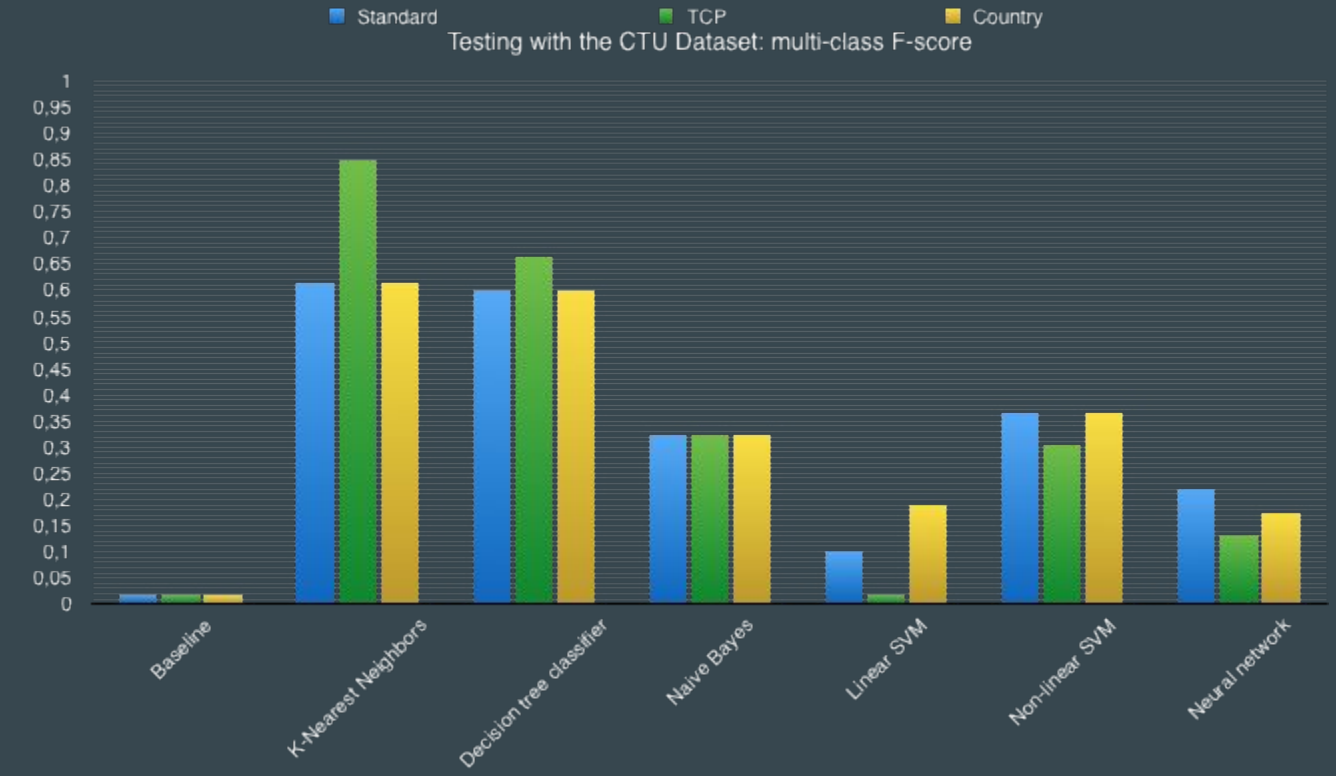


Stap 1: Learning curves

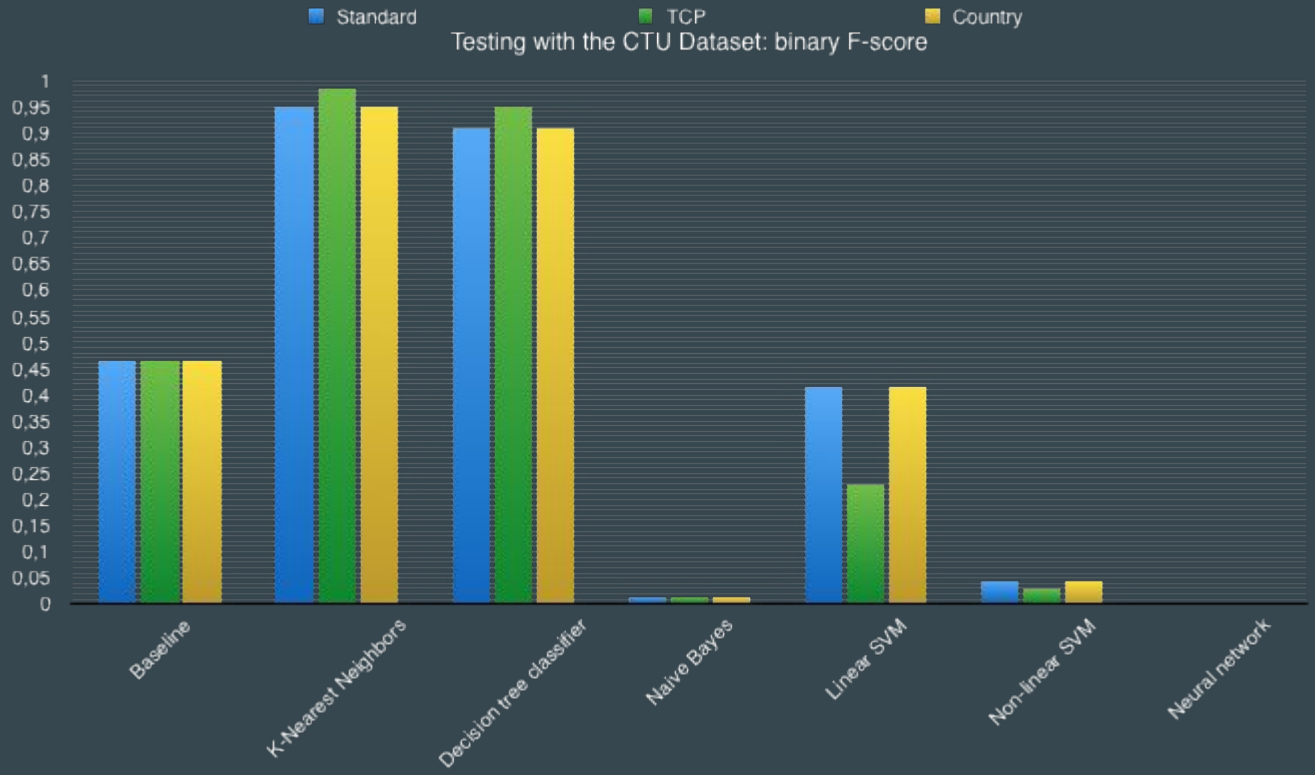




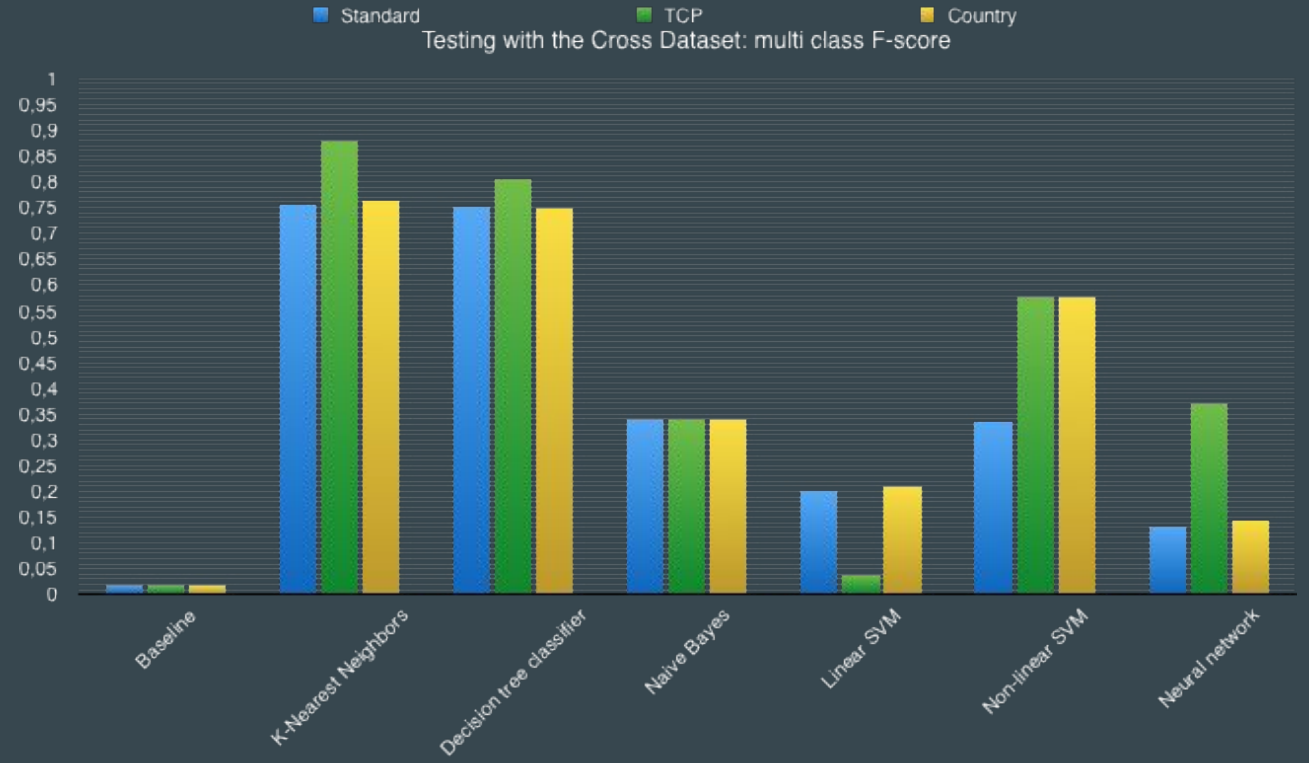
Stap 2: CTU dataset



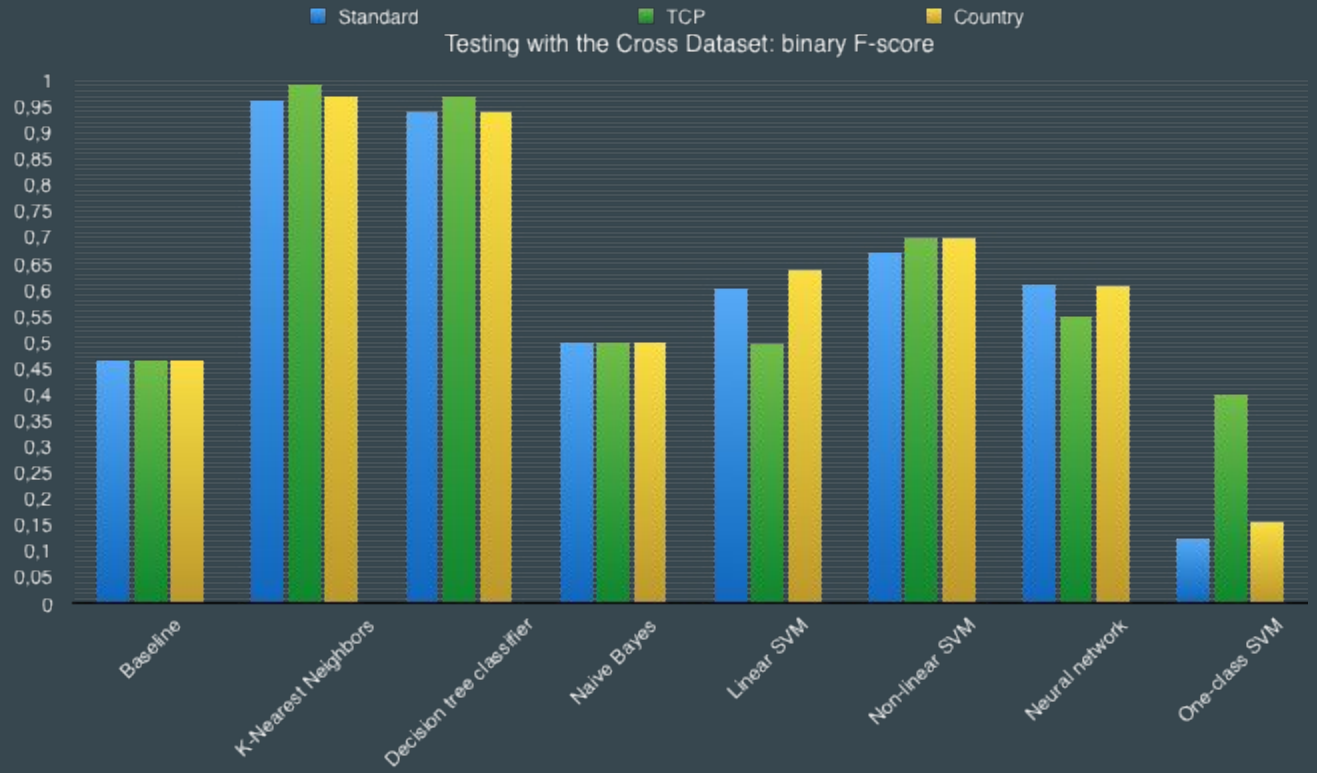
Stap 2: CTU dataset



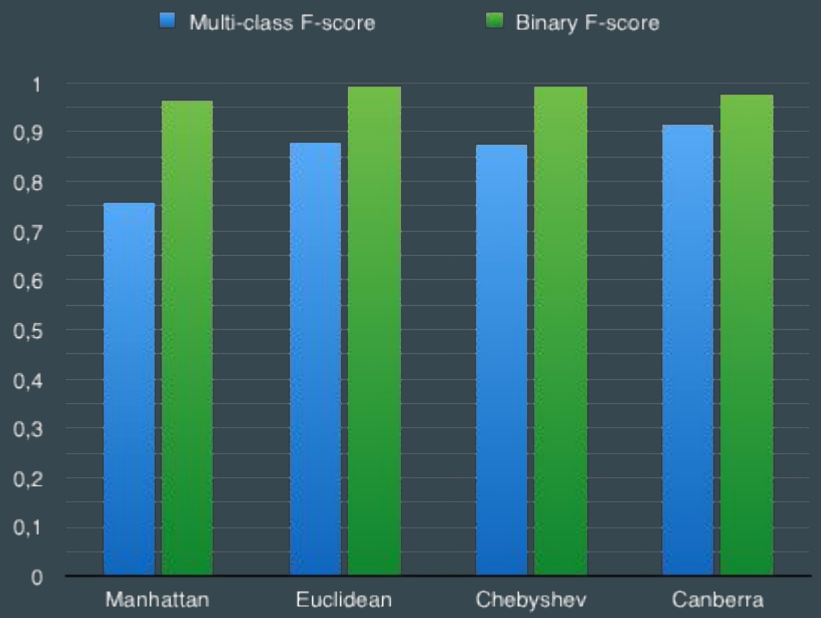
Stap 3: Cross dataset

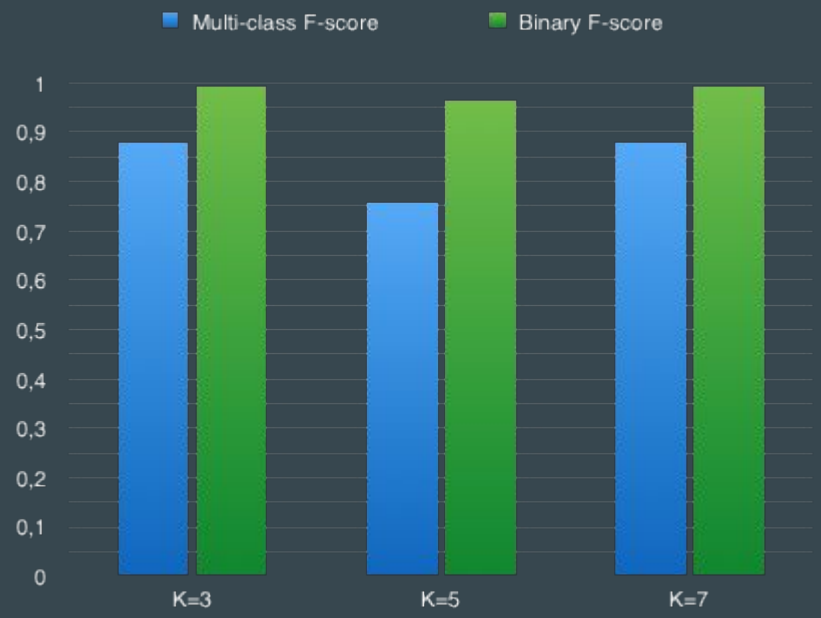


Stap 3: Cross dataset



Stap 3: K-nearest neighbors





Stap 4: real-world testing

	K-Nearest Neighbours	Decision tree classifier
F-score	0,7633	0,0155

Vragen?

Doelstelling ID systemen

Classificeren/detectie van onverwacht netwerkgedrag

Extern:

- port scans
- ssh connection attempts
- side-effect verkeer (ICMP, IRC)
- high volume DDoS
- low volume DDoS

Intern:

- botnets (communication with master/assist with DDoS attacks)
- worms (bij binnendringen/uitbreken van systeem)

Typische werking van bestaande ID systemen

Veel voorkomende technieken:

- Signature-based detecties (op basis van rule matching op inhoud van flow/packet data)
- Anomaly detection

Algemene doelstelling : patroonherkenning

- Binaire classificatie: malicious vs non-malicious
- Classificatie voor specifieke type van malicious behaviour

Specifieke eigenschappen v.h. te ontwikkelen systeem

Detectie louter gebaseerd op flow data:

- i.t.t packet- en log-based ID systemen
- specifiek bedoeld voor high traffic systems
- vereist geen in-depth knowledge van het netwerk

Gebruik van machine learning technieken

- Kosten-efficiënt inzetten in bestaande netwerken
- Algoritmes ‘leren’ zelf zonder regeltjes expliciet te programmeren

De onderzoeksvragen van de bachelorproef

In hoeverre zijn machine learning technieken inzetbaar voor anomaly detection (welke technieken werken goed/niet goed) ?

Hoe kan flow data gebruikt worden in deze technieken?

Kunnen we een IDS maken dat out-of-the-box een aanvaardbare ‘hit rate’ biedt ?

Welke types anomalie kunnen we automatisch detecteren ?

Is (automatische) klassificatie van de anomalie mogelijk ?

Zijn dergelijke technieken bruikbaar in real-life condities ?

Training data

Geannoteerde data sets zijn specifiek bedoeld om het algoritme een model aan te leren (stap 1 en 2)

- worden typisch opgedeeld in disjuncte subsets
- 1 subset specifiek om een model aan te leren
- 1 subset om dit model te kunnen valideren

Momenteel gebruikte datasets

CTU-13 dataset (stappen 1, 2 en 3):

- Bevat botnet, normaal en background traffic
- Zeer gedetailleerd geclassificeerd

Tracelabel dataset van UTwente (stappen 1, 2 en 3):

- Bevat traffic geclassificeerd als malicious door honeypot
- Bevat ftp, http, ssh, icmp, irc verkeer

EDM dataverkeer (stap 4):

- Unlabeled data
- Manuele verificatie van classificatie