Samen aanjagen van vernieuwing

Technologieverkenning AI voor Cybersecurity dienstverlening van SURFnet

Auteur(s): Duuk Baten

Versie: 1.0

Datum: april 2019

Moreelsepark 48
3511 EP Utrecht

Postbus 19035
3501 DA Utrecht

088 - 787 30 00
info@surf.nl
www.surf.nl

"If you think technology can solve your security problems

then you don't understand the problems and

you don't understand the technology"

Bruce Schneier

Secrets and Lies (2015)

Inhoudsopgave

Introductie

AI is op het moment dé technologische ontwikkeling om in de gaten te houden. Zo noemt Gartner zowel autonomous agents als augmented analytics als top strategische technologie trends voor 2019, en 6 uit de 10 trends worden mede gedreven door ontwikkelingen in AI.[1] Zo lijkt dit ook de trend te zijn binnen security. Bedrijven zoals Darktrace en Cylance bieden beide AI of machine learning gedreven securityoplossingen aan en ook IBM heeft een op Watson draaiende dienst die 'cognitive security' aanbiedt. De belofte lijkt dan ook groot, zo zegt men de 'unknown unknowns' en zero-day infecties automatisch te kunnen detecteren en onschadelijk maken. Deze ontwikkelingen bieden kansen voor het veilig houden onderwijsinstellingsnetwerken. Nu de recente ontwikkelingen in artificial intelligence (AI) ook het cyber security werkveld raken is het van belang om te kijken in welke mate deze beloften relevant zijn voor SURFnet en haar instellingen.

Dat deze beloftes aanslaan blijkt wel uit een recent onderzoek van het Ponemon Institute in opdracht van Hewlett-Packard, waarin naar voren kwam dat 62% van de respondenten denkt dat AI de effectiviteit van hun cybersecurity-team gaat verhogen.[2] Echter, uit datzelfde onderzoek komt ook naar voren dat slechts 25% daadwerkelijk AI-oplossingen toepast. Dat laatste is interessant omdat het suggereert dat de potentie van AI vooral gebaseerd is op de verwachte werking, volgens de geraadpleegde security experts, en niet de daadwerkelijke toepassing. AI is een lastige term die op het moment veel hype veroorzaakt. Zo blijkt wel uit het onderzoek dat een Engelse investeringsmaatschappij heeft gedaan naar het label AI bij start-up bedrijven. Zij zagen dat 1 uit 12 Europese startups in perceptie of marketing gebruik maakt van 'AI', maar dat daarvan uiteindelijk maar 60% ook daadwerkelijk AI-technologieën gebruikte.[3] AI is een hippe technologie en een verstandige zakelijke keuze, want volgens datzelfde onderzoek vingen de 'AI' start-ups 10-15% meer investeringen. Waar security through obscurity niet werkt, doet sales through obscurity dat wel.

Of bedrijven zoals Darktrace en Cylance ook tot deze categorie behoren is lastig in te schatten. Op het eerste gezicht lijken ze wel daadwerkelijk 'AI' toe te passen, maar zonder inzicht te krijgen in hun applicaties is dat lastig te verifiëren. Het is dus niet onomstreden of AI al duidelijk bruikbare security toepassingen heeft. Daarom blijft het verstandig om met een kritische noot naar de ontwikkelingen op AI-gebied te kijken, maar daarbij niet uit het oog te verliezen dat er ook zeker nuttige toepassingen zijn. Dat is dan ook het doel van dit verslag, het inzichtelijk krijgen wat 'AI' precies inhoudt voor cybersecurity maar ook een verkenning te doen van de mogelijkheden en kansen die hier liggen voor de SURFnet dienstverlening.

Hierin is de rol van het team security binnen SURFnet vooral gericht op het ondersteunen van onze instellingen in het verhogen van de veiligheid van hun netwerk. Om dit te doen bieden ze verscheidene diensten aan zoals SURFcert, SURFcertificaten, SURFmailfilter, SURFaudit, normenkaders, Cybersave Yourself, het SURF cyberdreigingsbeeld etc. Daarin heeft SURFnet ook de rol van het vergaren en verspreiden van kennis over nieuwe technologische ontwikkelingen zoals AI.

Daarom wordt in dit rapport de opkomst van AI besproken, wat machine learning is, en hoe dit relevant kan zijn voor SURFnet. Dit document is daarmee een verkenning van de ontwikkelingen in kunstmatige intelligentie als oplossing voor uitdagingen in de cyberweerbaarheid.

---

[1] https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/

[2] https://news.arubanetworks.com/press-release/advancedattacks/global-study-finds-artificial-intelligence-key-cybersecurity-weapon-io

[3] The State of AI: Divergence 2019, MMC Ventures (2019). p. 97

Artificial Intelligence

Kunstmatige intelligentie (AI) is een vakgebied van onderzoek en technologie dat zich bezighoudt met het begrijpen van 'intelligentie' en het creëren van technologieën die intelligent zijn dan wel intelligentie simuleren[4]. Dit is een brede definitie, en in de volksmond wordt AI ook gebruikt om te refereren naar 'slimme' technologieën, die op het moment vaak ontstaan uit toepassingen van machine learning. AI is niet nieuw, al sinds de jaren '50 is dit een levendig onderzoeksgebied en ook machine learning en deep learning bestaan al langere tijd.

In die zin zijn ontwikkelingen in AI zowel wetenschappelijk als technologisch. Een aspect van AI gaat over het vormen van kennis over intelligentie en methodes waarmee kennis over de wereld wordt gevormd. Het technologische aspect zien we in de manier waarop dit in computersystemen wordt ontwikkeld. Vaak wordt hierin een onderscheid gemaakt tussen sterke AI, ook wel artificial general intelligence (AGI) genoemd, en zwakke AI. Ontwikkelingen in sterke AI houden zich bezig met het creëren van een bijna menselijke of soms zelfs bovenmenselijke intelligentie in een door mensen gemaakt object. Zwakke AI is de vaak toepassingsgerichte versie van AI waarin men een meer taakgerichte intelligentie creëert. De huidige ontwikkelingen in AI in de niet academische setting gaan bijna zonder uitzondering over dat laatste: hoe kunnen principes van AI toegepast worden om iets intelligent te 'doen'.

In de praktische zin kunnen we dus over AI denken als systemen die cognitieve taken op zich nemen. Idealiter taken waar mensen niet zo goed in zijn. Op die manier wordt nu na fysieke handeling ook steeds meer denkwerk geautomatiseerd. Dit is dan ook precies waar de aantrekkingskracht ontstaat voor AI in cybersecurity. In een wereld waar de hoeveelheid IT-systemen en afhankelijkheden steeds groter wordt, wordt ook het verdedigen van die systemen steeds complexer. AI zou kunnen helpen door delen van dit werk te automatiseren, meer data te kunnen verwerken, en op een hogere resolutie.

Onderdelen van AI

Het mag duidelijk zijn dat AI een breed begrip is. Wat is dan een relevante blik vanuit cybersecurity? Om dit goed in te kunnen schatten is het handig AI te zien als een cybernetisch systeem van verschillende modellen en algoritmen dat probeert waar te nemen, analyseren, en vervolgens te handelen. Zo'n heel systeem gebruikt hiervoor verschillende technieken. Hiervan zijn tegenwoordig de machine learning technieken het bekendste, maar van oudsher werd er veel gebruik gemaakt van symbolische AI. Symbolische AI zijn technieken die redeneren vanuit regels en wetmatigheden op een 'als dit, dan dat' soort wijze. Dit soort systemen vertonen intelligentie door input te verwerken, maar zijn niet in staat om te leren. Machine learning werkt met het detecteren en leren van patronen in data door middel van statistische methoden.

Zowel machine learning als symbolic AI kunnen gebruikt worden in toekomstige security toepassingen. Als we kijken naar systemen die al gebruikt worden zijn expert-systemen in principe symbolisch, waartegen bijvoorbeeld de op Bayes filters gebaseerde spam filters onder machine learning vallen. De uiteindelijke markttoepassingen van AI zijn vaak een combinatie van meerdere technieken in een groter systeem. Het Britse Darktrace, bijvoorbeeld, zegt verschillende supervised, unsupervised, en deep learning technieken toe te passen binnen een Bayesian framework.[5]

---

[4] Of AI echt intelligent kan zijn is een lastige vraag die nog niet beantwoord is, maar dat neemt niet weg dat onderzoekers daarnaar streven. Voor meer naslag rondom dit onderwerp, zie: https://plato.stanford.edu/entries/artificial-intelligence/#WhatExacAI

[5] Machine Learning in the Age of Cyber AI. Darktrace.

Conclusie

Of AI uiteindelijk echt intelligent kan zijn, is een wetenschapsfilosofische vraag. Maar in de context van zwakke AI kan de vraag gesteld worden of AI nuttig is en de juiste antwoorden geeft op de vragen die we er mee proberen te beantwoorden. Uiteindelijk bestaan AI-toepassingen uit een drieslag van data, modellen, en algoritmen; en is het de vraag of die bij elkaar tot een acceptabele graad de waarheid benaderen. In de praktijk zal dat voor een groot deel bestaan uit het zinvol combineren en linken van verschillende algoritmen en systemen. Dat is dan ook de vorm van 'AI' die in dit rapport verder besproken zal worden.

De huidige ontwikkeling die in zwakke AI een centrale rol speelt is machine learning, en daar gaat de volgende sectie dieper op in.

1 Machine Learning

Machine learning (ML) bestaat uit algoritmen die leren op basis van voorgaande data hoe ze om moeten gaan met toekomstige data. ML verwerkt data en leert daarvan, om in de toekomst data weer beter te kunnen verwerken. Hoe dit precies gebeurt, is erg afhankelijk van het type machine learning algoritme, net als wat er precies mee kan worden gedaan. Maar ML-algoritmes zijn vooral bekend van de spam filters, advertentiekeuze algoritmes, en video- of muziekaanbevelingen.

In dit rapport worden vier manieren waarop machine learning kan leren verder uitgelegd. Dit zijn: supervised learning, unsupervised learning, transfer learning, en reinforcement learning. Deze methoden verschillen qua werkwijzen en mogelijke resultaten. Daarna gaan we nog in op een specifiek machine learning algoritme, namelijk deep learning.

1.1 Supervised learning

Supervised learning is op het moment de meest succesvolle vorm van machine learning. Door het algoritme te trainen met datasets met labels, kan het model bepaalde type dingen leren herkennen. Een voorbeeld hiervan is objectherkenning, waarbij het algoritme getraind wordt met een dataset van plaatjes met het label van het object op het plaatje. (Bijvoorbeeld ‘kat’ en ‘hond’, zie kopje Deep Learning).

Een gelabelde dataset ontstaat bijvoorbeeld door reCaptcha's, waar gebruikers om een site te mogen gebruiken de bekende vakjes moeten aanklikken met 'bussen' of 'verkeersborden'. Deze datasets worden vervolgens aan een algoritme gegeven die gaat leren herkennen dat bepaalde eigenschappen van een plaatje betekenen dat het een bepaald label verdient. Dit heet classificatie.

Naast classificatie bestaan er nog twee andere toepassingen van supervised learning: regressie en voorspellen. Regressie benadert de relaties tussen verschillende variabelen en leert zo herkennen welke variabelen in een dataset afhankelijk en onafhankelijk van elkaar zijn. Voorspellingsalgoritmen proberen op een tijdschaal voorspellingen te doen op basis van de trainingsdata die het gegeven is.

Supervised learning heeft op die manier altijd een basiswaarheid waar het vanuit gaat. Dit zit in de labels van de data. Een ding is voor het algoritme namelijk zeker, als het label 'kat' zegt dan is dat een 'kat'. De trainingsset moet dus al alle gegevens bevatten die uiteindelijk in nieuwe data voorspeld moeten worden.

1.2 Unsupervised learning

Unsupervised learning is een zelfstandigere vorm van machine learning, in tegenstelling tot supervised learning gebruiken dit soort algoritmen geen gelabelde datasets om getraind te worden. Wat unsupervised learning algoritmes doen, is zoeken naar structuur in een dataset. Dat gebeurt vaak op twee verschillende manieren, het clusteren van data en het reduceren van dimensies in een dataset.

Het clusteren van data is een handig functie die vaak een gebruikt wordt in systemen die vergelijkbare datapunten proberen te vinden. Een voorbeeld hiervan kunnen de Netflix recommendations zijn waarbij men films krijgt aangeraden die gelijkenden ook gekeken hebben. In een clustering algoritme worden sets van vergelijkbare data gegroepeerd op basis van criteria. Zo kan een clustering algoritme dat een set katten- en hondenplaatjes krijgt, nooit zeggen 'dit is een kat' en 'dit is een hond', maar wel netjes alle plaatjes van katten en honden op twee aparte stapeltjes leggen (indien het algoritme goed is afgesteld).

Een andere functie is het reduceren van dimensies in een dataset. Dit heeft als nut dat de hoeveelheid variabelen die nodig zijn om nuttige informatie te verwerken kleiner wordt, en daarmee de hele dataset. Dit

vereist natuurlijk wel dat de ontwikkelaar van het algoritme een idee heeft welke features belangrijk zijn, anders kan het algoritme ook niet naar de juiste feature reduceren.

Ondanks dat deze algoritmen geen gelabelde data gebruiken, is de dataset waarmee ze getraind of afgesteld worden nog steeds erg belangrijk. Een dataset die niet normaal verdeeld is, met bijvoorbeeld 99 kattenplaatjes en 1 hondenplaatje, is waarschijnlijk niet goed in het herkennen van andere hondenplaatjes als horende bij het cluster van dat ene hondenplaatje.

1.3 Transfer learning en reinforced learning

Twee andere varianten van machine learning worden ook steeds meer gebruikt. Transfer learning en reinforcement learning zijn twee manieren om met een gebrek aan goede trainingsdatasets om te gaan en je algoritme te trainen.

Bij transfer learning wordt een model eerst getraind op een andere dataset met vergelijkbare eigenschappen (bijvoorbeeld: foto's van dieren), hierop leert het model. Vervolgens wordt het laatste stuk van het model weer leeggehaald en wordt het op een kleinere set data getraind. Op deze manier kunnen modellen bijvoorbeeld longtumoren leren herkennen. Er is niet een set met longfoto's met tumoren die groot genoeg is om een heel algoritme te trainen; maar door het eerst op bijvoorbeeld dierenfoto's te leren hoe randen en vormen herkend moeten worden, kan daarna de finishing touch gedaan worden met medische foto's. Bij transfer learning worden dus de waarden gekopieerd van het ene model naar een ander model voor een ander doeleinde. Op deze manier bespaar je trainingstijd voor een model. Een voorbeeld hiervan is het detecteren van borstkanker[6].

Bij reinforcement learning wordt het model getraind door middel van feedback op de prestaties. De beslissingen gemaakt door het model hebben een invloed op de beslissingen in de toekomst. Hierbij is ook een beloningsfunctie nodig die het model vertelt hoe 'goed' een bepaalde beslissing is. Deze vorm van ML heeft enorme hoeveelheden data nodig, dus werkt vooral goed voor problemen die gesimuleerd kunnen worden, omdat dan de data zelf-generatief is. Succesvolle toepassingen hier zijn bijvoorbeeld AlphaGo.[7] Deze vorm aanpak is op het moment nog niet goed toe te passen buiten de research setting.

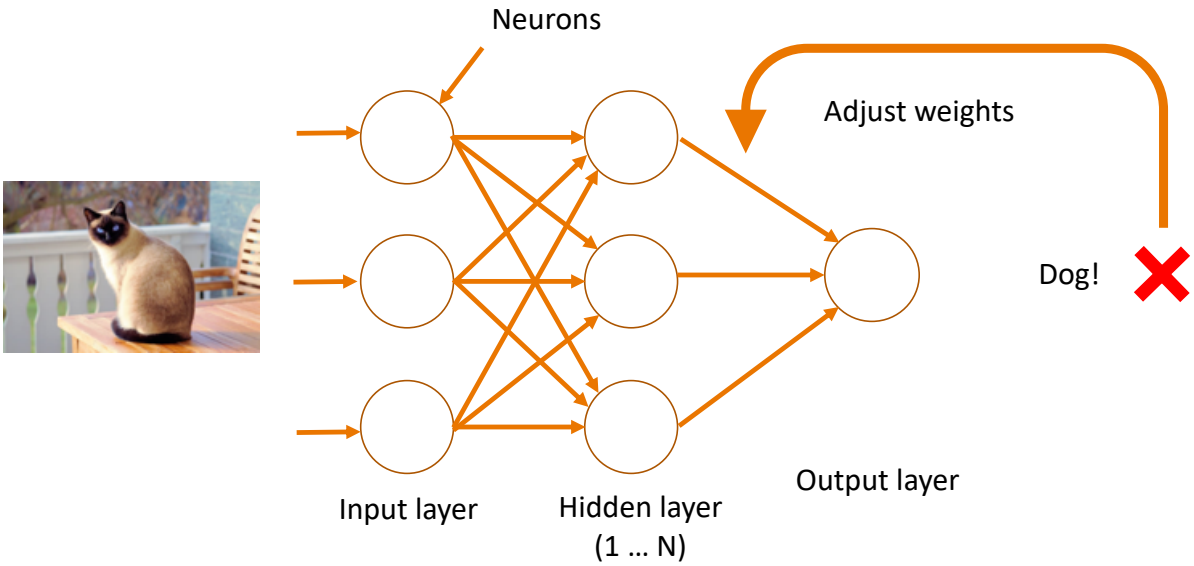
1.4 Een machine learning algoritme uitgelicht: neurale netwerken

Neural Networks (NN) is een techniek waarmee machine learning gedaan kan worden. Neurale netwerken zijn geïnspireerd op het menselijke brein. Zoals het brein bestaat uit verschillende neuronen bestaat een NN uit allemaal verschillende nodes die met elkaar verbonden zijn en 'signalen' aan elkaar doorgeven. NNs zijn in de afgelopen jaren enorm in populariteit gestegen door de grotere beschikbaarheid van rekenkracht, de toepassing van NNs die uit tien tot hondertallen lagen bestaan is men ook wel Deep Learning (DL) gaan noemen. DL is dus een machine learning algoritme en kan zowel worden toegepast voor supervised of unsupervised learning.

In het geval van onderstaand figuur, probeert een supervised learning algoritme katten en honden te herkennen. Hierbij past het zelf, volgens een algoritme, de waarden aan van de verschillende nodes in het netwerk. Om zo vervolgens tot de juiste conclusie te komen, passend bij het label van de gelabelde dataset.

[6] https://www.computer.org/csdl/proceedings/aipr/2017/1235/00/08457948-abs.html

[7] https://en.wikipedia.org/wiki/AlphaGo



Figuur 1: Deep Learning neuraal netwerk. Credits: Caspar van Leeuwen (SURFsara).

Zo kan een DL algoritme zelf bepalen of een verwachting accuraat is of niet. Zo ontdekt DL zelf de parameters die nodig zijn voor het model om tot conclusies te komen. Dit betekent dat DL bijvoorbeeld ook parameters kan ontdekken die de ingenieur die het ontwikkeld heeft zelf niet kent. Dit laatst wordt soms de black box van AI genoemd. Succesvolle DL toepassingen zijn bijvoorbeeld de Netflix recommendations en gezichtsherkenning.

1.5 Verschillende algoritmen

Onderstaande tabel is een verzameling van verschillende soorten ML-algoritmen overeenkomstig met de machine learning toepassingen waarvoor ze gebruikt kunnen worden. Een uitgebreide lijst van algoritmes kan gevonden worden op de site van het R-Project.[8]

[8] https://cran.r-project.org/web/views/MachineLearning.html

	Supervised learning			**Unsupervised learning**			**Reinforce-ment learning**
	Classification	Regression	Forecasting	Clustering	Dimension reduction	Feature extraction	
Naïve Bayes classifier	1						
Linear regression		1					
Support Vector Machine	1						
K-means clustering				1			
ANN	1	1					1
Logistic regression	1						
Decision trees	1	1					
Random forest	1	1					
Nearest neighbours							
Gradient boosting tree	1	1					
DB scan				1			
Gaussian mixture model				1			
k-modes				1			
Principle component analysis					1		
Singular value decomposition					1		
Latent Dirichlet Analysis					1		
Hierarchical clustering				1			
Deep neural networks	1	1				1	
Long Short-Term Memory	1		1				

Tabel 1: Een overzicht van verschillende algoritmen gematcht met de toepassing die ze hebben binnen machine learning.

2 Machine learning en data

Uiteindelijk maakt een machine learning algoritme een model van de werkelijkheid. Zoals met elk model is dit een incomplete benadering. Dit model wordt, in tegenstelling tot andere methodes, niet direct door een ontwerper gemaakt maar ontstaat uit de data die het algoritme in gaat. Via de keuze van de data en het afstellen van het algoritme beïnvloedt een ingenieur dus indirect het ML-model. Voor supervised learning gebeurt dat explicieter, omdat deze algoritmen getraind worden op een gelabelde dataset. Bij unsupervised learning is dat lastiger, omdat het algoritme getraind wordt op data waarvan de labels niet bekend zijn. Dit betekent echter niet dat de ontwerpers van de ML-algoritmen zomaar elk soort data kunnen gebruiken. De distributie van een dataset beïnvloedt bijvoorbeeld voor een groot deel het resultaat van het algoritme. Het begrijpen en juist verwerken van de data is om die reden cruciaal.

Het verzamelen van data is daarom een belangrijk maar uitdagend onderdeel van werkzaamheden in machine learning. Data is het nieuwe olie, het is nodig om deze algoritmen goed te laten functioneren. Daarom is het belangrijk een goede data infrastructuur op te zetten[9], hieronder valt het verzamelen van data maar ook het begrijpen van data punten en die kunnen vertalen naar doelen of oplossingen. Een van de redenen dat AI voor cybersecurity in academische literatuur niet zo succesvol lijkt te zijn ten opzichte van claims van marktpartijen, zou kunnen zijn dat marktpartijen agressiever en flexibeler zijn in het verzamelen van data om hun algoritmes op te trainen en evalueren. Waardoor marktpartijen sneller en met betere data hun algoritmes kunnen ontwikkelen.

Om goed inzicht te krijgen in de problemen die ML kan oplossen is expertise nodig over het probleemdomein. Domeinkennis is de soort expertise die noodzakelijk is om de juiste data op de juiste manier door machine learning te laten verwerken. De grootste uitdaging ligt daarbij in wat men 'feature engineering' noemt, het achterhalen en ontwerpen van de features die belangrijk zijn voor een algoritme. Het gaat daarbij om het uitzoeken op welke manier data gepresenteerd moet worden aan een algoritme om de juiste oplossing voor het probleem te krijgen. Je moet bijvoorbeeld je input verwerken in cijfers waarmee het algoritme kan rekenen. Een algoritme kan niet direct een plaatje ‘bekijken’, dus daarom verwerk je pixels en kan je bijvoorbeeld de ‘feature’ ontwikkelen die kijkt naar hoe helder een pixel is ten opzichte van omringende pixels om te contrast te ‘zien’. In het geval van netwerkanalyse betekent dit bijvoorbeeld het begrijpen dat IP-adressen geen integers zijn waarmee gerekend kan worden, maar wel identifiers en daarmee een persoonsgegeven.

Goed gedefinieerde features zijn cruciaal voor het algoritme om te begrijpen wat belangrijk is. Zonder features is er geen begrip voor wat data betekent. Uiteindelijk zijn er mensen nodig die betekenis geven aan de informatie, en is er dus kennis nodig van het domein waarin een algoritme zich begeeft. Alleen die domeinkennis kan vertellen welke labels of clusters van een algoritme daadwerkelijk relevant zijn.

[9] https://www.gartner.com/smarterwithgartner/how-to-get-artificial-intelligence-right/

2.1 Gelabelde data

De op dit moment meest betrouwbare algoritmes zijn vaak afhankelijk van gelabelde data. Dit brengt voor security een aantal uitdagingen met zich mee. Er is namelijk een huidig gebrek aan goede beschikbare (en gelabelde) netwerkdata[10,11], daarnaast is het ook twijfelachtig of die data gemakkelijk te krijgen valt.

Het blijkt namelijk erg lastig om goede (netwerk)data te verzamelen. Dit levert een aantal specifieke problemen op voor supervised learning. De afhankelijkheid van het algoritme van goede data is erg groot, als er slechte data in gaat komt er ook slechte data uit. Het hebben van een trainingsdataset is nodig voor het trainen van het model. Dit betekent een dataset waarin een goede distributie is verzameld van de verschillende datapunten die van belang zijn voor je algoritme. Die datapunten moeten vervolgens allemaal gelabeld zijn, zodat het algoritme later dit label kan voorspellen. Bijvoorbeeld 'deze combinatie van IP-range, tijd tussen pakketverzendingen, en grote pakketjes' krijgt als label "c&c[12] netwerkverkeer". Vervolgens kan het netwerkverkeer in de gaten gehouden worden door een algoritme dat een bepaalde handeling uitvoert als het weer eenzelfde combinatie van datapunten tegenkomt dat als label 'c&c netwerkverkeer' krijgt.

Als een supervised learning algoritme op deze manier data nodig heeft, wordt duidelijk dat de kwaliteit van de data erg belangrijk is. Als er geen dataset is die een combinatie aan netwerkdatapunten kan koppelen aan een voor security relevant label, dan gaat dat algoritme dat later ook niet kunnen doen. Elke infectie die dan ook gemist is in de trainingsdata zal ook gemist worden door het uiteindelijke model als het gebruikt wordt. Daarnaast moet een dergelijk classificatie-algoritme, nadat het getraind is, nog steeds up-to-date gehouden worden door hertrainingen.

Deze uitdagingen zijn ten dele nog steeds aanwezig bij het gebruik van unsupervised learning. Waar deze algoritmen niet getraind worden op een trainingsdataset, dient hun succes nog steeds geëvalueerd te worden. Dit betekent dat er dus uiteindelijk nog steeds kennis moet ontstaan over de juistheid van, bijvoorbeeld, de ontstane clusters ‘malware’ en ‘geen malware’.

2.2 Output verificatie en interpretatie

Een ander belangrijk onderdeel van kennis opbouw via ML is het juist verifiëren en interpreteren van de resultaten. Hoe kom je tot een algoritme dat kloppende resultaten levert en ben je in staat die resultaten ook te controleren.

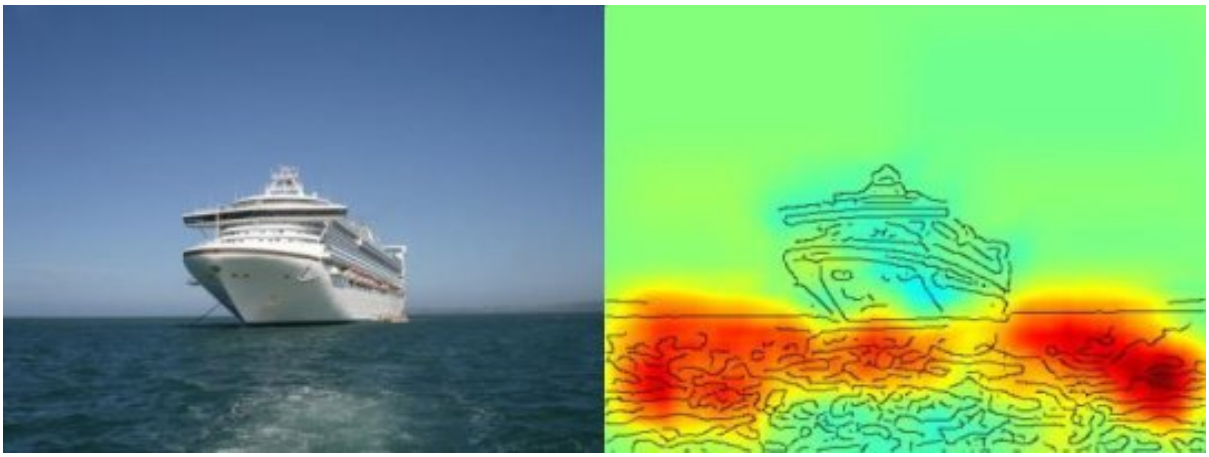
Modellen hebben bijvoorbeeld de neiging om de trainingsdata te ‘overfitten’, waarbij het model extra aspecten in de data herkent die helpen het juiste label te voorspellen die alleen in de trainingsdataset voorkomen. Dit is vaak lastig te controleren. Het is daarom ook belangrijk om niet alleen de fit van een algoritme kritisch te bekijken, maar ook waarom het past. In het onderstaande plaatje wordt bijvoorbeeld aangegeven in een heatmap op welke pixels het algoritme zich baseert in het doen van de voorspelling dat er een schip op het plaatje staat. Dit plaatje suggereert dat het model de beslissing of er wel of geen schip op het plaatje staat gebaseerd heeft op de pixels die de zee representeren en niet het schip zelf. Dit kan een signaal zijn dat het algoritme niet schepen herkent, maar een andere feature van de dataset die ook correleert met de verkregen labels van de trainingsdata-set.

---

[10] Apruzzese, Giovanni, et al. "On the effectiveness of machine and deep learning for cyber security." 2018 10th International Conference on Cyber Conflict (CyCon). IEEE, 2018. https://ieeexplore.ieee.org/abstract/document/8405026

[11] https://raffy.ch/blog/2018/08/07/ai-ml-in-cybersecurity-why-algorithms-are-dangerous/comment-page-1/#comment-258516 Slide 16/41

[12] Command and Control: het beheren van meerdere geïnfecteerde systemen.



Figuur 2: Heatmap van gebruikte pixels in een bootherkennings-machine learning model[13].

Dat algoritmes soms rare sprongen maken komt duidelijk terug uit recente casussen. Zo werd de zelflerende chatbot Tay van Microsoft op Twitter binnen 24 uur racistisch en nazistisch gedrag aangeleerd[14]. Dit soort manipulatief beïnvloeden van de dataset die een algoritme verwerkt heet *data poisoning*, omdat je via de input data van een algoritme probeert het normgedrag van het algoritme te veranderen. Bijvoorbeeld door veel legitieme mails te versturen met spam-achtige karakteristieken, of juist spam mails met legitieme karakteristieken, met als doel zo het algoritme te beïnvloeden. Adverserial machine learning is een andere manier om een algoritme te beïnvloeden. In adverserial machine learning kan bijvoorbeeld een kleine laag ruis op een plaatje ervoor zorgen dat het, voor mensen, overduidelijke plaatje van een panda als een mensaap herkend wordt.[15] Als op dit soort manieren beveiligingsalgoritmen gemanipuleerd kunnen worden dan brengt dat grote risico's met zich mee voor de cyberweerbaarheid.

2.3 Conclusie

Machine learning maakt modellen die de wereld benaderen. Dit kan op verschillende manieren gebeuren met verschillende soorten datasets. Voor machine learning is het daarom cruciaal om na te denken over:

- Het probleem dat je wilt oplossen

- Welke oplossing daarbij past

- Welke data daarvoor nodig is

Om dit goed te doen is het van belang dat AI-expertise en domein-expertise samen komen. Dat is noodzakelijk om tot de juiste vraag te komen en de juiste features te identificeren die deze vraag kunnen beantwoorden.

[13] https://www.sciencedaily.com/releases/2019/03/190312103643.htm

[14] https://qz.com/646825/microsofts-ai-millennial-chatbot-became-a-racist-jerk-after-less-than-a-day-on-twitter/

[15] https://www.volkskrant.nl/wetenschap/overal-hangen-beveiligingscamera-s-hoe-betrouwbaar-zijn-de-interpretaties-die-computers-maken-van-de-beelden-~bef1fd8b/

3 De toekomst van AI voor cybersecurity

AI zal zich in de toekomst steeds verder ontwikkelen. Als de beloftes waar worden gemaakt, zal het inderdaad een general purpose technologie worden die alle delen van de maatschappij zal doordringen. Ook voor het security werkveld zal het dan waarschijnlijk zijn dat het een grotere rol gaat spelen. Ontwikkelingen in data-analyse en automatisering zijn namelijk inherent aan het cybersecurity werkveld. Hoe dat zal gebeuren en welke kansen dat biedt voor SURFnet wordt in dit gedeelte verder besproken.

3.1 Toekomstige ontwikkelingen

Belangrijk in die toekomstige ontwikkelingen is tijdig inzien in welke richting ze zich ontwikkelen. Garry Kasparov heeft daar een specifieke visie op, hij beschrijft een toekomst van wat hij augmented intelligence noemt. Wat hij daar mee bedoelt is een toekomst waarin mensen en machine intelligence elkaar gaan aanvullen. Op een manier gelijkend aan hoe schakers nu gebruik maken van schaakcomputers kunnen mensen dan gebruik maken van de functionaliteiten die AI te bieden heeft. En de mensen zelf vullen dan als unieke factor hun creativiteit toe. Op die manier kunnen ze samen verdere stappen maken dan afzonderlijk.

In het cybersecurity werkveld kunnen we ons dit als volgt voorstellen: cybersecurity experts gebruiken hun creativiteit om problemen te analyseren en kunnen daarbij gebruik maken van machine learning en AI om juist die aspecten op te pakken waar mensen slecht in zijn zoals repetitieve taken, gedetailleerde taken of continue opletten gedurende langere tijd. Dit is niet een toekomstbeeld waarin security experts verdwijnen, maar wel een grotere reikwijdte gaan krijgen door aangevulde cognitieve capaciteiten.

3.2 De drijfveren achter AI innovatie

Op het moment is veel van de ontwikkeling in AI te vinden in machine learning, maar niet omdat de technologie nieuw is. De statistische methodes achter bijvoorbeeld neurale netwerken bestaan al langere tijd. De grootste redenen dat we nu op dit moment grote sprongen in AI ervaren, hebben niet alleen te maken met de ontwikkelingen in AI maar vooral ook met de grote beschikbaarheid van data die in het laatste decennium is verzameld evenals de enorm gegroeide rekenkracht van ook voornamelijk grafische processing units (GPUs). Als de rekenkracht van computers dan ook niet blijft groeien zullen de resultaten van dit soort technologieën daar onder te lijden hebben.[16] Dit is een interessante kans voor AI-ontwikkeling in Europa, waar men op het moment achterloopt op Amerika en China qua hoeveelheid dataverzameling en rekenkracht voor AI en machine learning. Volgens de Nederlandse AI-alliantie ALLAI zou dit dan ook een kans bieden voor Europese AI-initiatieven, omdat dit een reset knop kan zijn die ontwikkelingen in AI meer afhankelijk maakt van de creativiteit van ontwikkelaars dan grote hoeveelheden data.[17]

3.3 Toepassingen voor SURFnet

Ontwikkelingen in AI en machine learning zullen waarschijnlijk in de toekomst ook een rol gaan spelen in security voor SURFnet en haar instellingen. Voor SURFnet is het van belang om inzicht te krijgen in welke toepassingen voor haar rol relevant zijn, en waar kansen liggen voor waardeverhogende dienstverlening in de toekomst.

---

[16] Er zijn zorgen dat de Wet van Moore (een tweejaarlijkse verdubbeling van het aantal transistoren per oppervlakte) zal afremmen als we dichter bij de minimum dimensies komen van silicium bewerking.

[17] http://allai.nl/there-is-no-ai-race/ // https://www.technologyreview.com/s/612768/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/

Use case 1: detectie van afwijkingen

Het detecteren van afwijkingen in een patroon, ook wel 'anomaly detection' genoemd, is een van de meest genoemde AI-toepassingen voor security op het moment van schrijven. Partijen zoals Darktrace en Cylance beloven door afwijkingen in internetgedrag te herkennen van alles te kunnen waarnemen en voorkomen; van data exfiltratie tot zero-day exploits. Het idee hierachter is dat een onbekende bedreiging van het netwerk ('unknown unknown') nog steeds ander gedrag vertoont dan het standaard gedrag op het netwerk. Door het 'normale' gedrag als een norm te definiëren kan door middel van machine learning algoritmen gekeken worden naar afwijkend gedrag, waarna dat gedrag kan worden ingeperkt. Op deze manier kan het netwerk beschermd worden door te kijken naar afwijkend gedrag in plaats van rules en signatures.

In zijn belofte is deze use case erg interessant. Dit zou namelijk betekenen dat exploits, waarvan geen rule of signature bekend is, alsnog gedetecteerd kunnen worden aan de hand van hun gedrag. Hiervoor zijn wel een aantal lastige afhankelijkheden:

- Definiëren van afwijkend of normaal gedrag
- Verzamelen van genoeg netwerkdata om het algoritme te trainen
- Omgaan met false positives/false negatives

In de praktijk blijkt dit niet gemakkelijk. Voor de afdeling netwerkdiensten is een testopdracht uitgezet bij het bedrijf BIT-students die afwijkingen probeerde te ontdekken in de data van Splunk[18] en InfluxDB[19]. Waar dit tot zekere hoogte lukte (een betrouwbaarheid van 78,4% op de testdata van Splunk) liepen zij tegen een boel problemen aan zoals het gemis van gelabelde data en onduidelijkheid wanneer iets een afwijking is. Het is ook maar de vraag of voor deze toepassing machine learning modellen de juiste aanpak zijn, uit een literatuurreview van data-analyse methodes voor voorspellingen in het time-sequence domein bleken machine learning modellen vaak slechter te scoren dan traditionele statistische methodes[20].

Use case 2: AIRT incident tracking

AIRT is de applicatie die gebruikt wordt door SURFcert om incidenten te registreren. Dit soort melding systemen zijn belangrijk om overzicht te krijgen van wat er in het netwerk speelt en dat verder te kunnen onderzoeken. In de werkwijzen van AIRT of de applicatie die AIRT gaat vervangen kan ruimte zijn voor een aantal ML-toepassingen.

Zo zou een toepassing van een clustering algoritme een rol kunnen spelen in het categoriseren van soortgelijke meldingen. Hieraan zouden verschillende prioriteiten of automatische handelingen gekoppeld kunnen worden. Dit is een lastig proces dat niet altijd even eenvoudig verloopt (zie BIT students met Splunk meldingen) vooral ook omdat de data in de meldingen zinnig verwerkt moet worden.

[18] Chronologische database met netwerkmeldingen.

[19] Een database die het netwerkverkeer over tijd bijhoudt.

[20] Makridakis S, Spiliotis E, Assimakopoulos V (2018) Statistical and Machine Learning forecasting methods: Concerns and ways forward. PLoS ONE 13(3): e0194889. https://doi.org/10.1371/journal.pone.0194889

3.4 Evaluatie van toepassingen

Het SURFnet Strategisch Marketingplan voor Security en Privacy dienstverlening geeft inzicht in de huidige sterktes en gaten in de SURFnet dienstverlening. Dit wordt gedaan aan de hand van 7 verschillende aspecten van security:

Policy: in beleid of procedures vastleggen hoe goede security en privacy wordt ingevuld en bereikt.

Prediction: de capaciteit om toekomstige aanvallen te voorspellen.

Prevention: de capaciteit om aanvallen te voorkomen of tegen aanvallen beschermd te zijn.

Detection: de capaciteit om aanvallen te detecteren (zowel lopende als afgelopen aanvallen).

Mitigation: de capaciteit om aanvallen af te weren wanneer ze gaande zijn en uitbreiding van de aanval, schade en vervolgschade te beperken.

Recovery: de capaciteit om te herstellen van opgelopen schade door aanvallen.

Awareness: bewustzijn bij gebruikers en welke risico's ze lopen.

Om inzicht te krijgen welke mogelijkheden er allemaal zijn met AI voor het dienstenportfolio van SURFnet is een brainstorm georganiseerd. De ideeën uit deze brainstorm worden samen met suggesties uit deskresearch in de volgende tabel behandeld om een eerste verkenning te doen van de mogelijkheden.

Idee	Impact	Techniek	Benodigdheden	Status
Automatische tips bij Benchmark Assessment	Policy, prevention	Automatisering, geen AI? IFTTT[21]	Tips	Nu
Gepersonaliseerde phishing mail	Awareness, Prediction	Bayes filtering?	Data over gebruiker, mandaat	Nu
UBA context gebaseerde 2e factor authenticatie	Prevention	Patroonherkenning of classificatie	Software op client	1-3
Identificeren 'onveilig' gedrag van gebruikers	Awareness, prevention	Classificatie	Definitie 'onveilig', data individuele gebruikers	1-3
Foutcorrectie ('to' -> 'bcc')	Prevention, awareness	Automatisering	Verwerking in mailclient of serverside	Nu
Herkenning afwijkend gedrag	Detection	Patroonherkenning over tijd (regression)		1-3
OSINT analyse	Prediction, prevention			Nu

[21] If This Then That

Binary code analysis	Detection			3-5, 5+
Big data verwerking van Known attacks	Prevention	Clustering, dimension reduction	Data van known attacks	Nu
ML, geanonimiseerde (DNS) data de-anonimiseren (onderzoek)	Prevention	?		Nu
Simuleren van 'fout' gedrag op netwerk	Awareness, prevention		Verwerken van netwerkdata, definieren 'fout' gedrag	1-3
End-point data-transfer detectie	Detection	Patroonherkenning over tijd (regression)		3-5
Lange termijn infectie detectie	Detection	Patroonherkenning over tijd (regression)	Lange termijn opslag van data, netwerksensoren	1-3
Firewall loganalyse	Detection		Firewall-as-a-Service project af, of logtoegang van instellingsfirewall	1-3
IDS/IDP-loganalyse	Detection	Regression, classification	Installatie van IDS/IDP, gelabelde data	1-3
Forensics	Recovery	Clustering	ML-expertise bij SURFcert	Nu
Intelligent Blue-team agent	Detection, mitigation	?		5+
Automated Cyber Intelligence	Prevention, Policy, Prediction	Cognitive computing, taal analyse	Inkoop e.g. Watson of zelf ontwikkelen	1-3 of 5+
Intelligent Agent klantsupport	Mitigation, Recovery	Taalanalyse, chatbot		1-3
Bug detectie in SURFnet software[22]	Prevention	?		3-5
Monitoring meldingen verwerken	Detection	Classification/clustering	Meenemen in AIRT vervangingen	Nu

Definities van tijdsperiodes:

Nu: Als er nu middelen beschikbaar komen om hieraan te werken, zou SURFnet hier direct mee aan de slag kunnen gaan.

1-3: Het meeste werk hierin is beperkt tot de academische omgeving, voordat dit toepasbaar is moet er nog voorbereidend werk of onderzoek gedaan worden.

[22] https://www.schneier.com/blog/archives/2019/01/machine_learnin.html

3-5: Heeft nog veel voeten in de aarde. Afhankelijk van onderzoek en/of grote veranderingen in relatie SURFnet met instellingen.

5+: Dit zou, in de toekomst, misschien mogelijk kunnen worden.

Toekomstige SURFnet ontwikkelingen

Ook binnen SURFnet vinden een aantal ontwikkelingen plaats die relevant kunnen zijn voor de toepassing van AI in cybersecurity.

Strategisch niveau

On Campus impuls: SURF heeft in de Meerjarenagenda 2019-2022 de ambitie uitgesproken om de SURF-dienstverlening een "integraal onderdeel te laten zijn van de campus ICT-infrastructuur"[23]. Hierbij horen pilotprojecten waarbij SURFnet dieper het campusnetwerk van instellingen in gaat met apparatuur en diensten (o.a. NFV, Smart Campus IoT). Dit betekent een groeiend netwerk en meer mogelijkheden tot data verzameling en mogelijk ingrijpen binnen het instellingsnetwerk.

Technisch niveau

SURFnet 8 migratie: Met de migratie naar SN8 verandert het SURFnet netwerk ingrijpend. Dit kan gevolgen hebben voor de securitydienstverlening, maar ook kansen bieden voor ML of AI-toepassingen. Op dit moment verwerkt SURFnet Netflow data van de twee SN7 corerouters. Met de migratie naar een router-based in plaats van switch-based netwerk, zal de hoeveelheid routers van twee core-routers omhooggaan naar zo'n 300 kleinere routers. Dit zal een grote verhoging in de hoeveelheid Netflow data om te verwerken tot gevolg hebben.

Vervanging AIRT: AIRT wordt door SURFcert gebruikt als een monitoring en ticketing systeem voor haar verantwoordelijkheden. Op dit moment wordt er gekeken naar de vervanging van AIRT. Hier zouden mooie kansen kunnen liggen voor het toepassen van ML-modellen om bijvoorbeeld meldingen en tickets te sorteren of zelfs automatisch te verwerken.

Netwerksensoren: In verschillende projecten wordt er gekeken naar de mogelijkheid om sensoren binnen het netwerk te plaatsen. Zoals een intrusion detection system (IDS) met een port-duplicator bij de ingang van een instellingsnetwerk te hangen die via MISP[24], of door SURFnet geplaatste honeypots. Dit zijn beide zowel technische implementaties van On Campus, als nieuwe bronnen van gegevens. De soort real-time dataverwerking die hier plaatsvindt, biedt kansen voor een machine learning toepassing.

[23] Doelen 2JP Impuls On Campus 2019-2020. Roel Rexwinkel, Walter van Dijk, en Jan Bakker. 22-01-2019. P.1. SURF-2-5154

[24] Malware information sharing platform

3.5 Kansen voor SURFnet AI-dienstverlening

Op basis van de voorgaand besproken bevindingen, het Strategisch Marketingplan, eigen observaties en gesprekken met SURFnet medewerkers kan er een SWOT-analyse gemaakt worden welke een basis kan vormen voor latere aanbevelingen.

SWOT

Sterktes	Zwaktes	Kansen	Bedreigingen
Betrouwbaarheid: SURFnet is een betrouwbare partij voor instellingen, met o.a. privacy hoog in het vaandel.	**Kennis**: Op het moment is de hoeveelheid AI (ontwikkel) kennis beperkt binnen SURFnet.	**Outsourcing bij instellingen**: instellingen zijn steeds meer genegen om ICT-dienstverlening te outsourcen naar de cloud of SURF. Dit geeft ruimte voor nieuwe SURF proposities.	**Concurrentie uit de markt**: De markt concurrentie loopt voor op SURFnet als het gaat om AI/ML-toepassingen voor security. Daarnaast lijkt de trend te ontwikkelen richting totaaloplossingen.
Breed personeelsbestand: SURFnet heeft medewerkers die in staat zijn zowel de business als de techniek kant van ontwikkeling te begrijpen. Deze overbruggende rol is cruciaal om noodzaak te vertalen in een technische oplossing.	**Slagkracht**: Een combinatie van weinig vrije uren voor nieuwe ontwikkeling en beperkte gebruiksklare kennis, zorgt ervoor dat SURFnet niet snel aan de slag kan met AI.	**Specifieke doelgroep**: AI/ML heeft veel data nodig en is domein-specifiek. SURFnet bedient een specifieke doelgroep, waardoor ze een maatoplossing kan aanbieden.	**Tekort aan AI en securitypersoneel**: Er lijkt een groeiend tekort aan werknemers te zijn met vaardigheden in zowel AI als security. Dit kan lastig blijken voor toekomstige projecten als SURFnet niet voldoende capaciteit heeft.
Data: SURFnet heeft toegang tot veel data die ook nog bijzonder doelgroep specifiek is. Dat is een unieke positie waar gebruik van gemaakt kan worden.		**Samenwerking**: De gecombineerde expertise van SURFnet T&S en NWS met de AI-expertise bij SURFsara en onze instellingen biedt kansen in samenwerking.	

4 Conclusies

AI is een opkomende technologie. Ook voor het cybersecurity vakgebied zal dit veranderingen brengen. Dit rapport heeft onderzocht hoe AI en ML zich verhoudt tot het security werkveld en waar de aandachtspunten moeten liggen voor SURFnet in de toekomst. AI is een erg breed onderwerp dat zowel wetenschap als techniek omvat. AI als technologie in cybersecurity doet een aantal mooie beloften, zoals het automatiseren van je beveiligingswerk, het detecteren van afwijkingen in netwerkverkeer, en het detecteren en reageren op zero-day infecties. Deze beloften zijn echter niet zomaar gemakkelijk waar te maken.

Om meer inzicht te krijgen in wat AI nu precies inhoudt is er dieper ingegaan op machine learning. ML is een techniek van AI en gaat over het verwerken van data door een zelflerend model. Data speelt een centrale rol in machine learning. Gestructureerde dataverzameling en datamanagement zijn daarom cruciaal. De manier waarop een data set samengesteld wordt en/of gelabeld, is voor een groot deel bepalend voor de resultaten van een machine learning model. Voor machine learning toepassingen is het daarom belangrijk om te specificeren:

- Wat het probleem is dat opgelost moet worden
- Welke oplossing daarbij passend is
- Welke data nodig is om die oplossing te bieden

Voor cybersecurity zijn hier nog een aantal uitdagingen. Zo lijkt goede kwaliteit netwerkverkeer data op dit moment niet (publiek) beschikbaar[25]. Daarnaast is het grootschalig verzamelen van data een ingewikkeld privacyvraagstuk, zeker als op basis van (verwerkte) data ook beslissingen worden genomen over het gebruik van het netwerk. Mocht SURFnet verder gaan met AI-toepassingen dan betekent dit een noodzaak tot reflectie op de eigenschappen van de data en de impact op privacy van gebruikers.

Om de mogelijkheden van AI en ML zo goed mogelijk op te pakken is een combinatie van domeinkennis evenals AI-kennis cruciaal. Door het duidelijk stellen van problemen kan er gezocht worden naar kansen voor ontwikkeling. Er zijn een boel ideeën voor AI-innovatie voor security die nu of binnen 3 jaar mogelijk kunnen zijn, het is de vraag of deze interessant genoeg zijn om verder uit te werken. Eigenlijk moeten we kijken naar welke data we waar verwerken, hoe we dat doen, welke informatie daaruit voortkomt, en welk probleem we daarmee (kunnen) oplossen.

Veranderingen binnen SURFnet bieden ook kansen voor keuzes in het toepassen van AI binnen het security werkveld. Strategische ambities zoals On Campus kunnen zorgen voor zowel een diepere inbedding binnen het instellingsnetwerk als eenvoudigere toegang tot AI-expertise. Op een technisch vlak bieden de migratie naar SN8, de vervanging van AIRT, en het plaatsen van netwerksensoren mogelijkheden voor AI-ontwikkeling.

De trend bij instellingen om ICT-dienstverlening te outsourcen biedt samen met onze unieke verbindende positie en een specifieke doelgroep mogelijkheden om een op instellingen toegespitste dienst te ontwikkelen. Het is daarbij zaak om te zorgen voor kennisopbouw rondom AI binnen SURFnet en het creëren van meer slagkracht voor nieuwe (AI) dienstontwikkeling. Het succes van opkomende security marktpartijen in het aanbieden van een totaaloplossing[26] kan daarin een bedreiging vormen voor de SURFnet securitydienstverlening. Er moet dan ook een beslissing worden gemaakt of SURFnet een security totaaloplossing wil gaan aanbieden of niet. De markttrend lijkt hier heen te gaan, en als SURFnet het niet aanbiedt dan gaan instellingen wellicht zoeken naar andere oplossingen. In het Haalbaarheidsonderzoek SOC

[25] Apruzzese, Giovanni, et al. "On the effectiveness of machine and deep learning for cyber security." 2018 10th International Conference on Cyber Conflict (CyCon). IEEE, 2018. https://ieeexplore.ieee.org/abstract/document/8405026

[26] Die al dan wel of niet op AI-toepassingen gebaseerd is.

wordt geconcludeerd dat een SOC-oplossing te veel in het interne netwerk van een instelling zou zitten, maar een gehoste SIEM een mogelijkheid is om de gewenste functionaliteit aan te bieden. Mochten hierin verdere stappen ondernomen worden, dan kan ML daar ook een rol in spelen.

4.1 Aanbevelingen

Op basis van de bevindingen uit dit document zijn de volgende aanbevelingen geformuleerd in twee secties, een innovatief perspectief en een praktisch perspectief.

Innovatief perspectief

Als de inschatting is dat instellingen in de toekomst ambiëren naar een totaaloplossing voor security, dan kan het een unieke propositie zijn voor SURFnet om te kijken of er een completer cybersecuritypakket aangeboden kan worden met daarin ML- en AI-technologieën. Zoals benadrukt in dit rapport is data namelijk cruciaal voor het goed werken van deze technologieën, en SURFnet zou een vertrouwde partner kunnen zijn die de data van meerdere instellingen gebruikt om een dergelijk systeem te trainen. Hierdoor zal het eindproduct beter passen bij de specifieke use case die hoger onderwijs IT levert en beschermen wij instellingen en hun data tegen het in zee gaan met grotere commerciële partijen.

Praktisch perspectief

Op basis van al het voorgaande zijn de volgende aanbevelingen geformuleerd:

- AI-kennisopbouw binnen SURFnet
    - Bij medewerkers door middel van trainingen en ruimte om te proberen
    - Door een flexibele schil met ML/AI developers die gebruikt kunnen worden voor het uitwerken van PoCs en pilots
    - Centraliseer AI-expertise om het zo toegankelijk te maken voor heel SURFnet
    - Organiseer AI- en/of ML-trainingen voor medewerkers en instellingsdeelnemers in SURFacademy

- Datamanagement prioriteren
    - Inventariseer databronnen en hun mogelijke betekenis
    - Investeer in een goede datamanagementstructuur SURFnet breed
    - Formuleer een data-strategie: welke data heeft SURFnet en welke data kan het krijgen
    - Werk verder aan het delen van data met onderzoekers, en maak gebruik van de resultaten

- AI-dienstontwikkeling voor cybersecurity
    - Noem niet alles AI, maar specificeer de toepassing
    - Zoek de samenwerking met SURFsara
    - Zorg voor een duidelijke probleemstelling
    - Zorg voor intensief betrokken domein-experts

- SURFnet kan een rol als bruggenbouwer nemen en zoeken naar de vertaalslag van problemen naar mogelijke oplossingen met de middelen die beschikbaar zijn.
    - De rol als PM zal nog meer over de *inhoud* van een product of dienst gaan, voor het ontwerp is AI/ML-expertise nodig

- Werk vanuit de unieke positie die SURFnet heeft

    - Onderzoek of SURFnet een rol kan spelen in het beschikbaar stellen van een kwalitatieve publieke dataset met netwerkverkeer ter vervanging van of aanvulling op de veel gebruikte DARPA 1998 en 1999 datasets[27]

- Experimenteer
    - Intern. Bijvoorbeeld in een onderzoekend project met SURFsara, T&S, en NWD
    - Met inkoop. Start bijvoorbeeld eens een trial van Darktrace op het Onweer netwerk

[27] A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. Anna L. Buczak and Erhan Guven. *IEEE Communications Surveys & Tutorials, Vol. 18, No. 2, Second Quarter 2016* (p. 1157)