Opinnäytetyö AMK

Tieto- ja viestintätekniikka

2023

Santeri Hautala

Tekoälyn rooli tietoturvassa

TURKU AMK
TURKU UNIVERSITY OF
APPLIED SCIENCES

Opinnäytetyö AMK | Tiivistelmä

Turun ammattikorkeakoulu

Tieto- ja viestintätekniikka

2023 | 42

Tekijä Santeri Hautala

Tekoälyn rooli tietoturvassa

Tekoäly on yksinkertaisesti tietokoneen tai ohjelman kyky tehdä itsenäisiä älykkäitä toimintoja. Tämän opinnäytetyön tavoite oli selvittää, mikä on tekoälyn rooli ja merkitys tietoturvan näkökulmasta. Selvityksessä perehdyttiin mitä on tekoäly, miten sitä hyödynnetään tietoturvassa nyt ja mahdollisesti tulevaisuudessa. Mikä on tietoturvan merkitys nykypäivänä ja mitkä ovat yleisimmät tietoturvauhat. Mitä etuja tekoäly tuo tietoturvaan. Mihin tekoälyä on väärinkäytetty ja miten rikolliset ovat käyttäneet tekoälyä kyberhyökkäyksissä.

Opinnäytetyön raportointimuodoksi valittiin tutkimuksellinen opinnäytetyö ja lähteinä käytettiin lukuisia tiedeartikkeleita ja aiheeseen liittyvää kirjallisuutta. Lähteiden painoarvoina pidettiin niiden luotettavuutta, ajankohtaisuutta sekä aihetta ja sen sisältöä. Koska tekoäly ja tietoturva ovat laajoja käsitteitä, työ rajattiin ajankohtaisimpiin ja tekijän oman mielenkiinnon mukaan eri aihealueisiin.

Tuloksista selvisi, että tekoäly tuo paljon haasteita ja riskejä tietoturvaan. Tekoälyä voidaan myös hyödyntää parantamaan tietoturvaa. Tekoälyn yleisimmät sovellukset tietoturvassa ovat IDS-järjestelmissä joissa käytetään koneoppimis- ja keinotekoisia neuroverkkoja. Tekoälyä on käytetty myös rikollisten toimesta kyberhyökkäyksissä. Se tuo myös riskejä moneen sovellukseen, mikä mahdollistaa sen väärinkäytön esim. GPT-3-kielimallissa.

Asiasanat:

haittaohjelmat, koneoppiminen, kyberturvallisuus, kyberrikollisuus, neuroverkot, syväoppiminen, tekoäly, tietoturva, verkkohyökkäykset

Bachelor’s Thesis | Abstract

Turku University of Applied Sciences

Information and Communcation Technology

2023 | 42

Author Santeri Hautala

Role of AI in cybersecurity

Artificial intelligence (AI) is simply the ability of a computer or program to perform independent intelligent actions. The objective of this thesis was to find out about the role and importance of artificial intelligence from the point of view of cyber security. The thesis also aimed to find out how it is used in cyber security now and possibly in the future. What the importance of cyber security is today and what are the most common cyber security threats. What advantages does AI bring to cyber security. What artificial intelligence has been misused for and how criminals have used artificial intelligence in cyber-attacks.

To achieve the objective of this thesis numerous scientific articles and related literature were used as sources. The criteria for selecting the sources were considered their reliability, topicality, and the topic and its content. As AI and cybersecurity are broad concepts, the work was divided into the most current and different subject areas according to the author's own interest to limit the research work to keep it within a reasonable length.

The results showed that artificial intelligence brings a lot of challenges and risks to cyber security. Artificial intelligence can also be used to improve cyber security. The most common applications of artificial intelligence in cyber security are in Intrusion Detection Systems using machine learning and artificial neural networks. Artificial intelligence has also been used by criminals in cyber-attacks and it also brings risks to many applications, which enables its misuse in e.g., when using the GPT-3 language model.

Keywords:

artificial intelligence, cybersecurity, cyber-crime, cyber-attacks, data security, deep learning, machine learning, malware, neural networks

Sisältö

Kuvat

Käytetyt lyhenteet

AE	autoenkooderi (eng, AutoEncoder)
AI	tekoäly (eng. Artificial Intelligence)
AIDS	poikkeamien havaitsemiseen perustuva tunkeilijan havaitsemisjärjestelmä (eng. Anomaly detection-based Intrusion Detection System)
ANN	keinotekoiset neuroverkot (eng. Artificial Neural Networks)
BGP	internetin reititysprotokolla (eng. Border Gateway Protocol)
CNN	konvoluutioneuroverkko (eng. Convolutional neural network)
DBN	syvä uskomusverkko (eng. Deep Belief Network)
DL	syväoppiminen (eng. Deep Learning)
DT	päätöspuu (eng. Decision Tree)
ENISA	Euroopan unionin verkko- ja tietoturvavirasto (eng. European Union Agency for Cybersecurity)
FAR	väärien hälytysten määrä (eng. False alarm rate)
GAN	generatiivinen kilpaileva verkosto (eng. Generative adversarial network)
HIDS	isäntäpohjainen tunkeilijan havaitsemisjärjestelmä (eng. host-based Intrusion Detection System)
IDS	tunkeilijan havaitsemisjärjestelmä (eng. Intrusion Detection System)

KNN k:n lähimmän naapurin menetelmä (eng. K-Nearest Neighbor)

ML koneoppiminen (eng. Machine Learning)

NIDS verkkopohjainen tunkeilijan havaitsemisjärjestelmä (network-based Intrusion Detecion System)

NIST Yhdysvaltain standardisointi- ja teknologiainstituutti (eng. National Institute of Standards and Technology)

NLP neurologinen ohjelmointi (eng. Neuro-Linguistic Programming)

SIDS allekirjoitukseen perustuva tunkeilijan havaitsemisjärjestelmä (eng. Signature-based Intrusion Detection System)

SVM tukivektorikone (eng. Support Vector Machine)

1 Johdanto

Tekoäly on nykypäivänä lähes välttämätön osa jokaisen ihmisen arkea ja yritysten tietoturvaa. Jatkuvasti kehittyvät tietoturvauhat, trendit sekä kasvavat tietoverkot tekevät toimivan puolustuksen tekemisestä erittäin haastavaa. Suuren kasvun seurauksena tietoturvatiimit joutuvat käsittelemään valtavan määrän dataa puolustautuakseen mahdollisilta uhilta. Tietoverkkojen suojaus on myös vaikeutunut huomattavasti hyökkääjien kehittyessä ja laajentaessa hyökkäystekniikoitaan. Tekoälyn avulla voidaan käsitellä valtava määrä dataa lyhyessä ajassa ja paikantamaan erilaisia tietoturvauhkia, joita ihmisen on mahdoton tehdä lyhyessä ajassa. Tekoälyn tuomat mahdollisuudet ja uhat tietoturvaan sekä kyberrikollisuuteen ovat olleet jo pitkään puheenaiheena.

Tämän työn tarkoituksena on tutkia

- tekoälyn roolia ja merkitystä tietoturvan näkökulmasta.
- lyhyesti mitä on tekoäly ja sen alalajit.
- tekoälyn sovelluksia tietoturvassa nyt ja tulevaisuudessa.
- tekoälyn tuomat hyödyt ja haasteet tietoturvaan.
- tietoturvan merkitys ja yleisimmät tietoturvauhat.
- tekoälyn väärinkäyttö ja rikollisuus tietoturvassa.

2 Yleiskatsaus tekoälyyn

2.1 Tekoäly lyhyesti

Tekoäly *(*engl. AI, Artificial intelligence) on noussut valtamediaan ja ihmisten tietoisuuteen nykypäivänä mm. elokuvien, tv-sarjojen, pelien ja uutisten kautta. Tekoälyn määritelmä on suhteellisen tuore, sen keksi John McCarthy vuonna 1956 Darthmouthin Konferenssissa, jossa tekoäly terminä määriteltiin simuloimaan ihmisen älykästä käyttäytymistä mahdollisimman tarkasti. (Huawei Technologies Co., Ltd", 2023, 1.)

Tekoälyä on luokiteltu monella eri tavalla, mutta yleisimmin se jaetaan kahteen luokkaan, heikko tekoäly ja vahva tekoäly.

Vahva tekoäly tarkoittaa koneita, joilla on oma itsetietoisuus, maailmankuva ja arvomaailma. Ne pystyvät ratkomaan ongelmia itsenäisesti. (Huawei Technologies Co., Ltd", 2023, 5.)

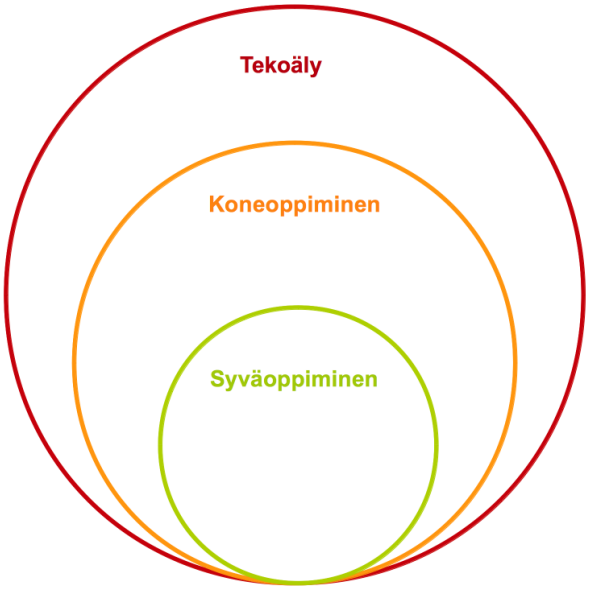
Heikko tekoäly puolestaan tarkoittaa koneita, jotka eivät kykene itsenäiseen ongelmanratkaisuun tai älykkääseen ajatteluun. Heikolla tekoälyllä ei ole itsetietosuutta, kuten vahvalla tekoälyllä. Nykypäivän käytännön sovellukset perustuvan heikkoon tekoälyyn. (Huawei Technologies Co., Ltd", 2023, 5.)

Tekoäly koostuu monesta eri osaamisalueesta, joihin lukeutuu mm. matematiikka, koneoppiminen, logiikka-ajattelu ja tiedonkäsittely.

Tunkeilijan havaitsemisjärjestelmät (eng, Intrusion Detection Systems, IDS) ovat tämän hetken ajankohtaisimpia tietoturvan sovelluksia rinnastettuna tekoälyllä.

Koneoppimiseen perustuvat algoritmit pystyvät käsittelemään suuren määrän dataa lyhyessä ajassa, jonka ansiosta niitä käytetään tunkeilijan havaitsemisjärjestelmissä.

Koneoppiminen ja syväoppiminen ovat tekoälyn peruspilareita (kuva 1). Koneoppimisen ja tekoälyn termien käyttö on vuosien aikana risteytynyt ja ne määritellään nykypäivänä samaksi asiaksi. (Zeadally ym. 2020.)



Kuva 1. Koneoppiminen ja syväoppiminen ovat tekoälyn alakenttiä (Lehto ym. 2019).

2.2 Tekoälyn käsitteitä

2.2.1 Koneoppiminen

Koneoppiminen (eng. Machine learning, ML) on tekoälyn osa-alue, joka perustuu algoritmeihin. Koneoppiminen koulutetaan syöttämällä dataa, jonka ansiosta se pystyy kehittymään. Koneoppivat järjestelmät pystyvät ennustamaan ja reagoimaan tapahtumiin itsenäisesti käyttäen apunaan siihen syötettyä dataa. (Zeadally ym. 2020.)

Koneoppimien luokitellaan yleisesti kahteen eri luokkaan niiden ominaisuuksien perusteella: ohjattu oppiminen (eng. supervised learning) ja ohjaamaton oppiminen (eng. unsupervised learning). (Zeadally ym. 2020.)

Ohjattu oppiminen

Ohjatussa oppimisessa data koulutetaan yleensä manuaalisesti ihmisten toimesta syöttämällä koulutettu data algoritmiin. Algoritmista tehdään matemaattinen malli, joka automatisoi datanäytteiden luokittelun. Ennen datan syöttämistä se koulutetaan havaitsemalla siinä olevia luokkia ja malleja. Koulutettu data merkataan joko haitalliseksi tai lailliseksi riippuen sen luokituksesta. (Zeadally ym. 2020.)

Ohjaamaton oppiminen

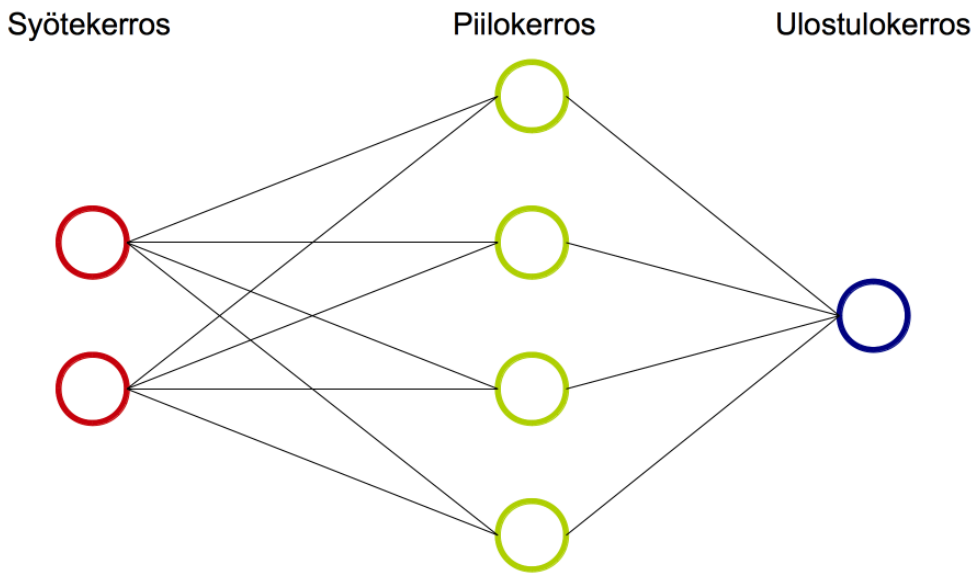
Ohjaamattomassa oppimisessa algoritmit määrittelevät syötetyn datan yhtenäisyyden ja hajanaisuuden, jonka perusteella tehdään luokat. Luokat määritellään syötetyn datan laadun perusteella luokkien sisällä ja niiden modulaarisuuksien välillä. (Zeadally ym. 2020.)

2.2.2 Syväoppiminen ja keinotekoiset neuroverkot

Syväoppiminen (eng. Deep learning, DL) kuuluu koneoppimisen algoritmeihin ja sen toiminta perustuu monikerroksisista kytköksistä dataan ja jopa jäsentimettömästä sekä luokittelemattomasta datasta.

Keinotekoisia neuroverkkoja (eng. Artificial neural networks, ANN) käytetään esimerkiksi puheentunnistuksessa ja luokittelussa. (Romney ym. 2019.) Sen toiminta mukailee lineaarisen regression yhtälöä, jonka avulla syötetyt tiedot painotetaan ja iteroidaan saadakseen virhevälin tarpeeksi pieneksi sekä mahdollisimman lähelle haluttua tavoitearvoa. Kun virhearvo on tarpeeksi pieni, algoritmi tuottaa matemaattisen yhtälön, joka antaa syötetylle datalle

informatiivisen arvon kuten luokan. Keinotekoisten neuroverkkojen tarkoituksena on simuloida ihmisten aivojen toimintaa käyttämällä matemaattisia malleja ja algoritmeja. Neuroverkot koostuvat ulostulo- ja syötekerroksesta jonka välissä voi olla yksi tai useampi piilokerros (kuva 2). (Zeadally ym. 2020.)



Kuva 2. Neuroverkko (Lehto ym. 2019).

2.3 Yleiset käyttökohteet

Tekoälyä käytetään melkein jokaisella toimialalla. Esimerkiksi pankit, vähittäiskaupat, terveydenhuolto, tuotantotalous käyttävät tekoälyä hyödykseen. Tekoälyn arjen sovelluksia ovat mm. hakukoneet, digitaaliset avustajat, nettiostokset, mainonta, navigointi, autot ja sosiaalisen median alustat.

3 Yleiskatsaus tietoturvauhkiin ja haasteisiin

Verkkoon kytkettyjen laitteiden ja sovellusten määrä lisääntyy päivittäin, korostaen tietoturvan merkitystä nykypäivänä. Lisääntyvät tietouhat tuovat haasteita laitevalmistajille, sovelluskehittäjille sekä tietoturva-alan ammattilaisille. Pienentääkseen tietoturvauhkien riskiä, käyttäjiä tulee valistaa turvallisesta nettikäyttäytymisestä samalla sovelluskehittäjien sekä laitevalmistajien pitää olla tietoisia haavoittuvuuksista ja päivittää järjestelmät mahdollisimman nopeasti. (Ozkaya 2019, 18.)

3.1 Hyökkäyspinta-ala

Tietoturvaan liittyen hyökkäyspinta-alan (eng, attack surface) käsite kuvaa kaikkia mahdollisia haavoittuvuuksia, jotka voivat mahdollistaa luvattoman pääsyn verkkoon, tietokantoihin tai järjestelmiin. Nämä haavoittuvuudet tunnetaan myös nimellä hyökkäysvektorit (eng, attack vector), ja ne voivat sisältyä eri komponentteihin, ohjelmistoihin, tietoverkkoihin ja loppukäyttäjiin. Mitä enemmän hyökkäysvektoreita on, sitä suurempi on hyökkäyspinta-ala ja sitä suurempi on riski joutua kyberhyökkäyksen kohteeksi. Riskiä voidaan pienentää pienentämällä hyökkäyspinta-alan kokoa rajoittamalla hyökkäysvektorien määrää. (Ozkaya 2019, 10.)

3.2 Uhkamaisema

Uhkamaisema (eng, threat landscape) on tietoturvassa usein käytetty termi, johon kuuluu ajankohtaisimmat tietouhat, tietoturvatrendit sekä tiedot mahdollisista hyökkääjistä. Tietoturvayritykset, kuten Kyberturvallisuuskeskus, ENISA (Euroopan Union verkko- ja tietoturvavirasto) ja NIST (Yhdysvaltain standardisointi- ja teknologiainstituutti) julkaisevat kuukausittaisia ja viikoittaisia uhkamaisemaraportteja. Tietoturva-alan osaajien tulee olla tietoisia uusimmista

uhkamaisesta jatkuvasti muuttuvien työkalujen, hyökkäysten ja haavoittuvuuksien takia pysyäkseen ajan tasalla. (Ozkaya 2019, 11.) Seuraavaksi käymme ENISAN vuoden 2022 uhkamaisemaraportista läpi viimevuoden yleisimpiä kyberuhkia.

3.2.1 Haittaohjelma

Haittaohjelmat (eng. malware) ovat aina olleet tietouhkien kärkipäässä vuodesta toiseen. Niiden toiminta perustuu järjestelmien ja sovelluksien luvattomaan tiedonpääsyyn, häiritsemiseen, varastamiseen tai vahingoittamiseen. Haittaohjelmia on erityyppisiä joihin lukeutuu esimerkiksi, virukset, troijalaiset, vakoiluohjelmat sekä kiristyshaittaohjelmat. Nykypäivän kehittyneimmät haittaohjelmat pystyvät toimiaan huomaamattomasti järjestelmissä jopa vuosia. Internetistä ladatut haitalliset sovellukset ja liitteet, jotka hyödyntävät järjestelmien heikkoutta ja tietoturvaa ovat yleisimpiä keinoja haittaohjelmien pääsyn kohteisiin. (ENISA 2022.)

3.2.2 Käyttäjän manipulointi, kalastelu

Kalasteluhyökkäykset (eng. social engineering, phishing) pyrkivät huijaamaan käyttäjiä esittämällä luotettavaa tahoa, kuten sosiaalisen median palvelujen henkilökuntaa tai pankkia. Kalasteluhyökkäysten tarkoitus on saada käyttäjä luovuttamaan arkaluontoista sisältöä hyökkääjälle, kuten pankkikortin numeron tai käyttäjätilin salasanan. Sähköpostiviestit ovat yleisimpiä keinoja kalasteluhyökkäysten levittämiseen niiden helpon jakelun ja aidolta näyttävien viestien takia. (ENISA 2022.)

3.2.3 Palvelunestohyökkäys

Palvelunestohyökkäykset (eng. Denial of Service, Dos, DDoS) kohdistuvat verkkoinfrastruktuurin komponentteihin ja palveluihin tarkoituksenaan ylikuormittaa verkkolaitteet ja estää käyttäjien pääsy palveluihin. (ENISA 2022.)

3.2.4 Internetin palvelunestouhat

Internetin palvelunestouhat sisältävät joukon eri uhkia, joiden tarkoituksena on estää käyttäjien pääsyy palveluihin. Näihin uhkiin lukeutuu esimerkiksi palvelunestohyökkäykset ja BGP (Border Gateway Protocol) kaappaukset. (ENISA 2022.)

3.2.5 Tietouhat

Tietouhkien (eng. threats against data) tarkoituksena on saada luvaton pääsy tietolähteisiin. Tietouhat voivat manipuloida ja häiritä järjestelmien käytöstä sekä paljastaa arkaluontoisia tietolähteitä. Tietouhat voidaan myös luokitella tietomurroiksi ja tietovuodoiksi. Tiedon tahaton jakelu johtuen esimerkiksi inhimillisestä virheestä lasketaan tietovuodoksi. Pääsyy arkaluontoisiin tietoihin ja sen tahalliseen levittämiseen kutsutaan tietomurroksi. Tietouhan sattuessa ne ovat yleensä myös monen muun hyökkäyksen alkukohta, kuten kiristyshaittaohjelmat sekä palvelunestohyökkäykset. (ENISA 2022.)

3.2.6 Harhaanjohtavan ja väärän tiedon levitys

Väärän ja harhaan johtavan tiedon levitys (eng. Disinformation – misinformation) on yleistynyt internetin saatavuuden ja sosiaalisen median käytön kasvun seurauksena. Nopea sisällön luonti ja jakaminen saavuttaen suuren määrän ihmisiä lyhyessä ajassa on tehnyt siitä yhden merkittävimmistä uhista vuosien aikana. Houkuttelevan kuuloiset uutiset ja linkit saavat ihmiset lukemaan ja jakamaan sisältöä ottamatta huomioon sen alkuperää ja luotettavuutta. (ENISA 2022.)

3.2.7 Kiristyshaittaohjelmat

Kiristyshaittaohjelmat (eng. Ransomware) ovat vuodesta toiseen yleisimpiä hyökkäystyyppejä ja niiden toiminta perustuu kohteen tai sen omaisuuden kaappaamiseen ja vapauttamiseen lunnaita vastaan. (ENISA 2022.)

3.2.8 Toimitusketjuhyökkäykset

Toimitusketjuhyökkäysten (eng, Supply-chain attacks) tarkoituksena on häiritä organisaatioiden ja toimittajien välistä toimintaa kohdentamalla hyökkäykset toimittajaan ja toimittajan organisaatioon sekä omaisuuteen. Toimitusketjuhyökkäykset muodostuvat yleensä vähintään 2 eri hyökkäyksestä, joista ensimmäinen kohdistuu toimittajaan, jonka avulla hyökkääjät pääsevät käsiksi toimittajan kohdeorganisaatioihin ja omaisuuteen. (ENISA 2022.)

4 Tekoälyn tarve tietoturvassa

Tietoturvan tarve on nykypäivä suurempi kuin koskaan. Päivittäin lisääntyvät verkkoon kytketyt laitteet, sovellukset ja tietouhat korostavan tietoturvan merkitystä ja tarvetta. Verkkoon kytkettyjen laitteiden ja sovellusten määrän kasvun seurauksena verkkoliikenne on kasvanut niin suureksi, että nykypäivän tietoturvaratkaisut eivät kykene käsittelemään kaikkea verkkoliikennettä nopeasti, jonka seurauksena tietoturvajärjestelmät toimivat puolustaen hyökkäyksiä niiden sattuessa. (Zeadally ym. 2020.)

Jotta tietoturva voi kehittyä, tutkijat ovat koittaneet rinnastaa tekoälyä ja koneoppimista tietoturvaratkaisuihin. Yksi suurimmista syistä tekoälyn rinnastettavaksi tietoturvaan on verkossa kasvava tiedon määrä, joka vaatii valtavasti resursseja ja aikaa analysoida ja tunnistaa kuviot, epämääräisyydet tai tunkeilijat. Koneoppimiseen perustuvien järjestelmien avulla pystytään käsittelemään suuren määrän dataa kerralla parantaen reaktioaikaa ja ennakoimaan mahdollisia hyökkäyksiä. Tekoälyn tuoman kehityksen myötä myös kyberrikolliset ovat koittaneet etsiä keinoja, miten tekoälyä voidaan hyödyntää tekemällä uusista hyökkäyksistä tehokkaampia ja vaikeasti havaittavia. (Zeadally ym. 2020.)

4.1 Kehitys

Tekoälyn käyttö oli pääosin valtion puolustusjärjestelmissä ennen 90-lukua. Yhdysvaltain puolustusministeriö on jo 1970-luvulta lähtien sijoittanut tekoälyllä toimiviin automaattisiin asejärjestelmiin, kuten lennokkeihin. Tekoälyn automatisoinnin, nopean päätöksenteon ja laskentatehon ansiosta sen katsottiin parantavan kansallista turvallisuutta. Esimerkkinä Yhdysvaltain ilmavoimien tiedustelukone U-2 toimii käyttäen tekoälyyn perustuvaa ARTU-algoritmia jonka avulla se hallitsee sen sensoreita ja navigointia. (Michael & Wingfield 2021, 90–91.)

Tekoälyä on käytetty myös internet bottien muodossa 90-luvulta lähtien. Bottien tarkoitus oli etsiä internetistä haluttua sisältöä ja suojata Wikipedia-artikkeleita manipuloinnilta. Internetin moninpeleissa sekä kalastelu- ja spam-viesteissä on myös käytetty tekoälyllä toimivia botteja antaakseen pelaajille kilpailullisia etuja ja levittääkseen haittaohjelmia. (Zeadally ym. 2020.) Tekoälyn yleistyessä myös yksityiset henkilöt voivat hyödyntää tekoälyllä toimivia tietoturvasovelluksia, kuten antivirus ohjelmia. (Michael & Wingfield 2021, 90–91).

4.2 Tietoturvan haasteet tekoälylle

Tietoturvan tarve nykypäivänä on korostunut kasvavien tietoturvauhkien ja verkkoon kytkettyjen laitteiden sekä sovellusten myötä. Tietoturva-alalla on huutava pula ammattilaisista, joka johtuu osittain korkean osaamistason vaatimisesta ja aiemmasta työkokemuksesta alalla. Tämä on johtanut tietoturva-alan henkilöstön vaihtuvuuteen työpaikkojen välillä. Kehittyvien tekoälypohjaisten työkalujen, kuten ChatGPT:n käyttö, joka pohjautuu GPT-3 kielimalliin on mahdollistanut uusia hyökkäystapoja kyberrikollisille ja lisännyt amatöörihakkerien määrää. ChatGPT:n avulla voidaan luoda lyhyessä ajassa muun muassa erittäin aidolta vaikuttavia kalastelusähköposteja ja haittaohjelmia. Ratkaistakseen tietoturvaan liittyvät ongelmat ja uhat, tekoälyn mahdollisuuksia tulee tutkia ja kehittää parantamaan tietoturvaa samalla vapauttaen tietoturva-alan ammattilaisten resursseja keskittymään tärkeimpiin uhkiin ja toimenpiteisiin. Tekoälyn tarkoitus on aluksi toimia tietoturva-alan osaajien kanssa parantaakseen puolustusta, mutta loppuvaiheessa sen päätavoitteena on täysin automatisoitu uhkien tunnistus ja reagointi tietoturvaan liittyviin ilmoituksiin. (Bresniker ym. 2019, 45–46.)

Koneoppimiseen ja syväoppimiseen perustuvia tietoturvaratkaisuja tulee kehittää automatisoidakseen tietoturvaan liittyviä tehtäviä vapauttaakseen tietoturva-alan henkilöstön resursseja. Jotta tekoäly tietoturvassa voi kehittyä, sen tulee analysoida talteen otettua tietoa ja opetella uusia hyökkäys- ja puolustustaktiikoita. Tietoturvan automaation avulla voidaan vastata tapahtumiin reaaliajassa ja laajentaa yritysten tietoturvaa entisestään. (Bresniker ym. 2019, 45–46.)

Tekoälyn ja koneoppimisen soveltaminen tietoturvaan luo haasteita sen monimutkaisuuden takia. Tekoälyllä vahvistettujen tietoturvaratkaisujen luotettavuutta ja kestävyyttä tulee testata riittävästi ennen käyttöönottoa. (Tan & Karri, 2020). Huolimattomasti koulutettu tekoäly voi ohittaa tunnistamaan joitain hyökkäyksiä ja uhkia. (Chan, 2019.) Hyökkäyspinta-alan laajentuessa erilaisia

menetelmiä tulee kehittää tekoälyn avulla jolla voidaan ennaltaehkäistä ja torjua kyberhyökkäyksiä. (Bresniker ym. 2019, 45–46.)

4.3 Vaikutus ja hyödyt tietoturvaan

Tekoälyn ansiosta käsitellä suuren määrän dataa lyhyessä ajassa verrattuna perinteisiin tietoturvaratkaisuihin se tuo mukanaan lukuisia etuja tietoturvaan niin yksityisille henkilöille kuin myös yrityksille. Tekoälyllä toimivat antivirus ohjelmat parantavat loppukäyttäjien tietoturvaa. Koneoppimiseen perustuvat tunkeilijan havaitsemisjärjestelmät parantavat yritysten tietoturvaa nopeuttamalla verkkoliikenteen analysointia ja pienentämää myös uusien uhkien reaktioaikaa. Yritysten tehtäviä voidaan automatisoida tekoälyn avulla ja se voi myös vapauttaa tietoturvahenkilöstä keskittymään kriittisimpiin uhkiin. Tehtävien automatisoinnin avulla yritysten tietoturvan tehokkuus ja koko voi kasvaa tekoälyn pystyessä tunnistamaan laajemmin eri hyökkäystekijöitä.

Tekoälyn soveltaminen organisaatioiden tietoturvaan tuo mukanaan ainakin seuraavia etuja (Murugesan, 6. 2022):

- Nopeuttaa tapausten vasteaikaa ja lisää turvallisuutta tekoälyn tuoman laskentatehon ja kyvyn ansiosta havaita uudet hyökkäykset.

- Tekoälyn kyky kvantifioida riskit nopeasti, joka nopeuttaa analyytikoiden päätöksentekoa tietopohjaisilla lieventämisoperaatioilla johtaa nopeutettuun havaitsemis- ja vastesykliaikaan.

- Lisääntyvät kustannussäästöt organisaatioissa haitallisten hyökkäysten, ehkäisyn, lieventämisen ja reaaliaikaisen suojauksen ansiosta.

- Virheiden määrän pienenee manuaalisissa ja osittain manuaalisissa prosesseissa.

- Vapauttaa tietoturvahenkilöstä keskittymään tärkeimpiin uhkiin nopeuttamalla esimerkiksi väärien positiivisten ilmoitusten läpikäyntiä.

- Toistuvat Manuaaliset ja kuluttavat työtehtävät voidaan automatisoida vapauttaen tietoturvahenkilöstä keskittymään kriittisimpiin uhkiin parantaen myös työilmapiiriä sekä tyytyväisyyttä.

- Organisaatioiden tietoturvakäytännön taso nousee kehittyneemmän suojauksen anisosta parantaen samalla brändi-imagoa ja asiakastyytyväisyyttä.

- Pienentää riskien mahdollisuutta ja samalla parantaa tietoturvan puolustusta.

5 Tunkeilijan havaitsemisjärjestelmät

Tunkeilijan havaitsemisjärjestelmät (eng. Intrusion Detection Systems, IDS) tarkkailevat verkkoliikennettä epänormaalilta käytökseltä. Sen toiminta perustuu verkkoliikenteen seuraamiseen ja poikkeamien ilmoittamiseen verkon ylläpitäjälle. IDS:n avulla voidaan vahvistaa tietoverkkoja asentamalla se strategisesti järkeviin paikkoihin verkkotopologiassa. IDS ei itsestään estä kyberhyökkäyksiä, mutta oikealla sijoittamisella ja verkkoliikenteen seuraamisella se voi parantaa koko verkon tietoturvaa yhdistettynä muihin verkossa oleviin tietoturvaratkaisuihin. (Bhatia ym. 2020.)

Ensimmäiset IDS sovellukset tulivat 80-luvulla, jonka jälkeen on kehitetty lukuisia erilaisia IDS järjestelmiä parantaakseen verkon tietoturvaa. IDS yleisimpiä heikkouksia ovat olleet sen tehottomuus havaita väärien positiivisten hälytyksiä ja nollapäivä hyökkäyksiä (zero-day attacks). Suuren verkkoliikenteen määrä ja kehittyvät tietoturvauhat ovat luoneet haasteita IDS sovelluksille. Tekoälyyn perustuvia IDS sovelluksia, jotka käyttävät koneoppimista ja syviä neuroverkkoja on ehdotettu tutkijoiden toimesta viimeisen vuosikymmenen aikana parantaakseen IDS:n tehokkuutta. (Ahmad ym. 2021.)

Koneoppimiseen perustuva IDS riippuu suuresti ominaisuussuunnittelusta saadakseen hyödyllistä tietoa verkkoliikenteestä. Syviin neuroverkkoihin perustuva IDS on puolestaan hyvä oppimaan automattisesti monimutkaisia ominaisuuksia raakadatasta syvän rakenteensa ansiosta. Koneoppimiseen liittyvät metodit verkkopohjaisille tunkeilijan havaitsemisjärjestelmille on vielä alkuvaiheissa ja siihen liittyvää tutkimusta on vielä valtavasti jäljellä, jotta voidaan tehokkaasti havaita tunkeilijat verkossa. (Ahmad ym. 2021.)

5.1 IDS luokittelu

IDS luokitellaan yleisesti sen käytön ja havaitsemismetodien perusteella. (Ahmad ym. 2021.)

5.1.1 Käyttöönottomenetelmään perustuva IDS

IDS luokitellaan edelleen kahteen eri luokkaan riippuen sen sijainnista ja käyttötarkoituksista verkkotopologiassa. Nämä luokat ovat verkkopohjainen (eng. Network-based IDS, NIDS) ja isäntäpohjainen IDS (eng. Host-based IDS, HIDS). NIDS toimii verkossa ja sen tarkoituksena on seurata verkkoliikennettä sekä suojata verkkoa ja siinä olevia laitteita tunkeutumiselta. HIDS toimii puolestaan yhdellä laitteella valvoen kaikkia sen toimintoja mahdollisilta tietoturvarikkomuksilta ja epäilyttävältä toiminnalta. (Ahmad ym. 2021.)

5.1.2 Tunnistusmenetelmään perustuva IDS

Tunkeilijoiden havaitsemisjärjestelmät, jotka perustuvat tunnistuspohjaisiin menetelmiin voidaan jakaa kahteen eri luokkaan, allekirjoitukseen perustuva IDS (eng. Signature-based IDS, SIDS) ja poikkeamien havaitsemiseen perustuva IDS (eng. Anomaly detection-based IDS, AIDS). (Ahmad ym. 2021.)

SIDS toiminta perustuu suuren ylläpidettävään tietokantaan, joka sisältää allekirjoitettuja datamalleja. Datamallit mukailevat hyökkäysmalleja, joiden avulla pystytään torjumaan jo tiedossa olevia hyökkäyksiä. Uusien hyökkäyksien sattuessa tietokantaan lisätään uusi allekirjoitettu datamalli perustuen tapahtuneeseen hyökkäykseen. Tietokannan kasvaessa hyökkäysten tunnistaminen parantuu, joka tekee siitä erittään tehokkaan menetelmän. Haittapuolena tietokannan koko kasvaa ja sen ylläpitäminen vie enemmän resursseja. Sen kyky havaita uusia hyökkäyksiä on myös huono, koska

tietokannasta puuttuu uusien hyökkäysten allekirjoitetut datamallit. (Ahmad ym. 2021.)

AIDS toimii puolestaan tunnistamalla epänormaalia toimintaa vertaamalla sitä valmiiksi tehtyihin profiileihin, joissa määritellään mikä toiminta on normaalia ja epänormaalia. Sen avulla pystytään havaitsemaan uusia tuntemattomia hyökkäyksiä tehokkaasti, koska ne poikkeavat jo valmiiksi määritetyistä profiileista. Tehokkaan tunnistamisen seurauksena AIDS:in haittapuolena on suuri väärien hälytysten määrä (eng false alarm rate, FAR), koska normaalin ja epänormaalin profiilin välistä rajaa on vaikea löytää tunkeutumisen havaitsemiseksi. (Ahmad ym. 2021.)

5.2 Tekoälyn sovellukset verkkopohjaisille tunkeilijan havaitsemisjärjestelmille

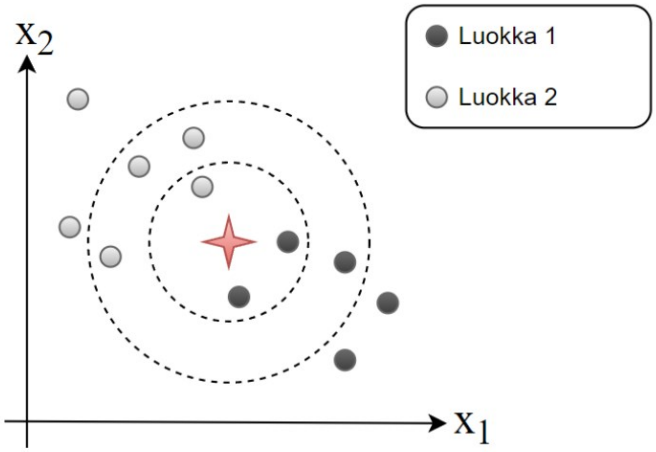
Tässä osioissa käydään läpi yleisimpiä tekoälyyn pohjautuvia menetelmiä ja algoritmeja, joita käytetään verkkopohjaisessa tunkeilijan havaitsemisjärjestelmien suunnittelussa.

5.2.1 Päätöspuu

Päätospuu (eng. Decision tree, DT) kuuluu valvotun koneoppimisen algoritmeihin. Päätöspuun nimi tulee sen puunmuotoisesta rakenteesta ja se koostuu oksista sekä lehtisolmuista. Oksat edustavat jotain sääntöä tai päätöstä, kun taas jokainen lehtisolmu edustaa mahdollista lopputulosta tai luokittelua. Päätöspuun algoritmi karsii puun rakenteesta kaikki turhat oksat ja valitsee lehtisolmun, jonka avulla se saavuttaa parhaan mahdollisen lopputuloksen syötetyn datan perusteella. Yleisimpiä Päätöspuun algoritmeja ovat CART ID3 (Ahmad ym. 2021.)

5.2.2 K:n lähimmän naapurin menetelmä

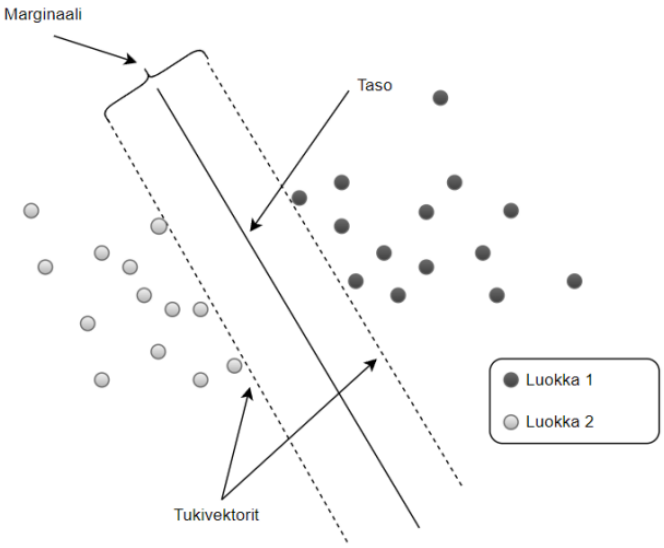
K:n lähimmän naapurin menetelmä (eng. K-Nearest Neighbors , KNN) on ohjattuun koneoppimiseen kuuluva algoritmi. KNN käyttää datapisteiden välistä etäisyyttä tehdäkseen datan liittyviä ennustuksia tai luokittelua. (Ahmad ym. 2021.) Kuvassa kolme esimerkki KNN luokittelusta.



Kuva 3. KNN datan luokittelu (Lähde 2022).

5.2.3 Tukivektorikone

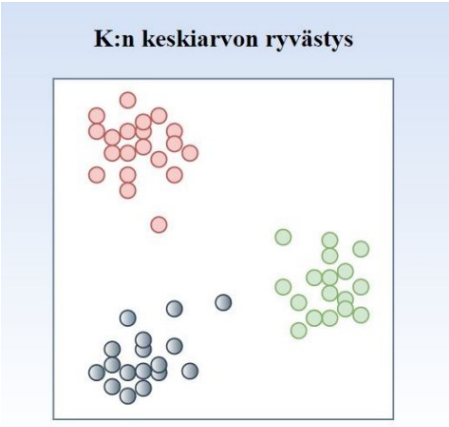
Tukivektorikone (eng. Support vector machine, SVM) on ohjattu koneoppimisen algoritmi, jota käytetään syötetyn datan luokitteluun ratkomalla lineaarisia ja epälineaarisia ongelmia. Tukivektorien avulla syötetty data erotetaan ja jaetaan kahteen eri luokkaan. Luokat on eroteltu tasolla, joka on valittu pitääkseen luokkien välisen marginaalin mahdollisimman pienenä (kuva 4). Verkkopohjaisissa tunkeilijan havaitsemisjärjestelmissä tukivektorikonetta voidaan käyttää tehostamaan ja parantamaan järjestelmän tarkkuutta. (Ahmad ym. 2021.)



Kuva 4. Kuvassa datan luokittelu käyttäen tukivektoreita (Lähde 2022).

5.2.4 K:keskiarvon ryvästys

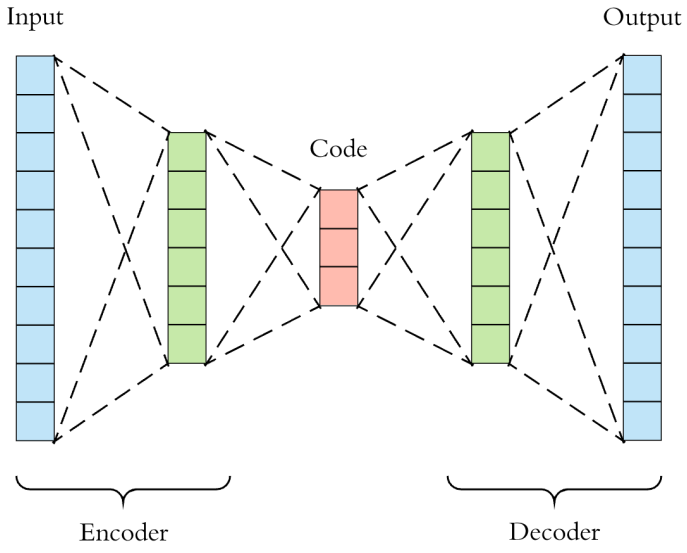
K:n kesiarvon ryvästys (eng. K-means clustering) on yksi koneoppimisen algoritmeista, joka kuuluu ohjaamattoman oppimisen perheeseen. Sen toiminta perustuu poimimalla syötetystä datasta samankaltaiset tiedot jakamalla ne samankaltaisiin ryhmiin, jotka sijaitsevat klusterien sisällä. Klusterien sisällä olevien ryhmien määrä kertoo k:n arvon. Esimerkkinä (kuva 5), jossa k:n arvo on kolme. Algoritmin tarkoituksena on pienentää datapisteiden ja niiden keskipisteiden välisten etäisyyksien summaa klusterin sisällä. (Ahmad ym. 2021.)



Kuva 5. Klusteri, jonka sisällä on kolme ryhmää (Merimaa 2021).

5.2.5 Autoenkooderi

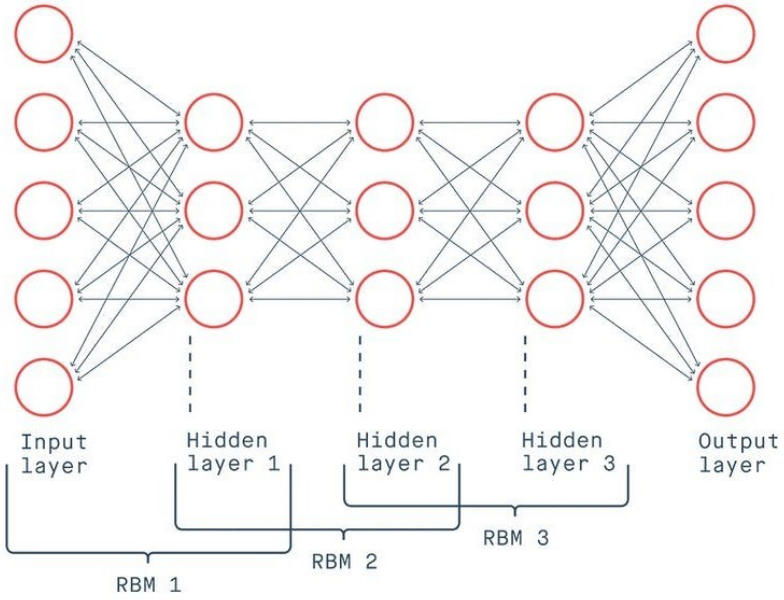
Autoenkooderi (eng, AutoEncoder, AE) kuuluu ohjaamattoman neuroverkkojen joukkoon. Sen toiminta perustuu syötetyn tiedon pakkaamiseen saavuttaakseen mielekkään lopputuloksen ja lopuksi purkamaan sen mahdollisimman lähelle alkuperäistä muotoa (kuva 6). (Ahmad ym. 2021.)



Kuva 6. Autoenkooderi (Dertat 2017).

5.2.6 Syvä uskomusvekko

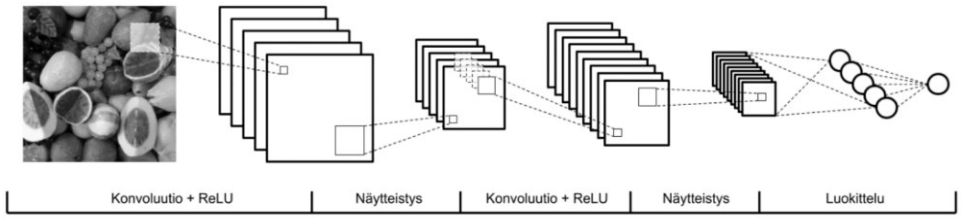
Syvä uskomusverkkojen tehtävä (eng. Deep belief network, DBN) tunkeilijan havaitsemisjärjestelmissä on luokittelu ja ominaisuuksien poimiminen. DBN koostuu luokittelukerroksesta ja useista päällekkäin pinotuista ohjaamattomista verkoista, kuten RBM (Restricted Boltzmann Machines) kerroksista, sekä Autoenkoodereista (kuva 7). (Ahmad ym. 2021.)



Kuva 7. Syvä uskomusverkko ja useat RBM kerrokset (Kalita 2022).

5.2.7 Konvoluutioneuroverkko

Konvoluutioneuroverkko (eng. Convolutional neural network, CNN) on keinotekoisiin neuroverkkoihin perustuva malli, jota käytetään pääosin kuvien tunnistamiseen ja tietokonenäköön. Tunkeilijanhavaitsemisjärjestelmissä sitä käytetään ohjattuun ominaisuuksien poimimiseen ja luokitteluun. (Ahmad ym. 2021.) Konvoluutioneuroverkko koostuu monesta eri kerroksista, joihin lukeutuu näytteistys-, konvoluutio, aktivointi sekä luokittelukerrokset (kuva 8). (Paananen 2018, 12)



Kuva 8. Konvoluutioneuroverkon rakenne (Paananen 2018).

6 Tekoälyn väärinkäyttö ja rikollisuus tietoturvassa

Tekoälyn vaikutus yhteiskuntaan on ollut jo pitkään eri akateemisten tutkimusten ja poliittisten raporttien aiheena. Kehitykset koneoppimisessa ja syvissä neuroverkoissa ovat tuoneet positiivia kehityksiä esimerkiksi lääketieteen alalla käyttäen syviin neuroverkkoihin perustuvia työkaluja syövän tunnistamiseen. Tekoälyn kehitys tuo myös uusia uhkia tietoturvaan mahdollistaen entistä monimutkaisempia ja vaikeasti havaittavia kyberhyökkäyksiä. Hyökkäyspinta-alan laajentuessa kyberrikolliset koittavat rinnastaa nykyisiä hyökkäysmenetelmiä tekoälyn avulla. Sosiaalisen median alustat ja palvelut käyttävät myös tekoälyä hyödykseen algoritmien avulla niiden sisällön tuotannossa. Tekoälyä käytetään myös maan puolustukseen kehittääkseen automaattisia asejärjestelmiä, kuten lennokkeja. Tämä kehitys on mahdollistanut sen väärinkäytön ja rikollisen toiminnan, joita tässä kappaleessa käsitellään. (Blauth ym. 2022.)

6.1 Eheyshyökkäykset

Koneoppimisen yleistyessä viimevuosina hyökkääjät ovat alkaneet manipuloimaan koneoppimisen malleja ja taustalla olevaa dataa tekemällä koneoppimismallit alttiille eheyshyökkäyksille (eng. Intregrity attacks). Eheyshyökkäyksien tarkoituksena on sisällyttää väärää tietoa järjestelmään korruptoidakseen datan ja heikentääkseen sen luotettavuutta. Tekoälyn mallin yksi haavoittuvuuksista on vastakkaisista esimerkeistä (*eng. adversarial examples*), jotka ovat haitallisia syötteitä suunniteltu huijaamaan koneoppimisen malleja, mikä voi aiheuttaa datan väärinluokittelua. Joissain tapauksissa häiriöt voivat olla liian pieniä ihmisten huomattavaksi, mutta ne voivat silti aiheuttaa tekoälyn tekemään virheitä. Yksi esimerkki koneoppimisen eheyshyökkäyksistä on myrkytyshyökkäys (*eng. poisoning attack*), jossa hyökkääjä koittaa vaikuttaa datan kouluttamiseen muokatakseen ennustettavan mallin tuloksia syöttämällä

korruptoituneita pisteitä koulutusprosessissa. Toisin sanoen myrkyllisiä malleja voidaan syöttää harjoitusdataan manipuloidakseen luokittelua, joka voi johtaa ei toivottuihin lopputuloksiin. (Blauth ym. 2022.)

6.2 Tekoälyn käytön seurauksesta johtuvat tapaukset

Tekoälyn käyttämät mallit voivat antaa eriäviä lopputuloksia niistä mitä kehittäjä on odottanut useasta eri syystä. Neuroverkkoihin perustuvat mallit voivat tahattomasti paljastaa ja muistaa yksityiskohtia, joka voi johtaa ongelmiin varsinkin, kun koulutuksessa käytetty data on arkaluontoista tai yksityistä. Käytännössä koulutusprosessin aikana kyseiset mallit saattavat muistaa yksityiskohtia, jotka eivät ole olennaisia sen tehtävänantoon liittyen. Tekniikat, joilla voidaan varmistaa datan yksityisyys ja estääkseen haitallisia seuraamuksia tahattomasta datan muistamisesta ja sen paljastamisesta ovat välttämättömiä. (Blauth ym. 2022.)

6.3 Jäsenyydenhäirintähyökkäys

Jäsenyydenhäirintähyökkäykset perustuvat koneoppimiseen käytettyjen mallien uudelleenrakentamiseen ja paljastamiseen. Jäsenyydenhäirintähyökkäyksiä voidaan käyttää myös GAN (Generative adversarial networks) malleja vastaan. GAN on yksi syväoppimisen mallin luokista, jota käytetään monissa eri sovelluksissa, kuten väärennettyjen verkkosivujen osoitteiden luomisessa. (Blauth ym. 2022.)

6.4 Väärän tiedon levitys ja uutisointi

Väärän tiedon levitys on yleistynyt vauhdilla sosiaalisen median ansiosta sekä ihmisten tavasta kuluttaa ja jakaa sisältöä tavoittaen miljoonia ihmisiä lyhyessä ajassa. Palvelut, kuten Facebook näyttävät uutisia käyttäjilleen, jotka perustuvat algoritmeihin. Algoritmi seuraa käyttäjän avaamia linkkejä ja katsomia sivustoja näyttäen samantyyppistä uutta sisältöä myös jatkossa. Tämän tyyppisten algoritmien vaarana on kiihdyttää väärän tiedon levitystä. Väärän tiedon levityksellä voi olla myös isoja vaikutuksia valtion tasolla vaikuttaen demokratiaan. Tekoälyllä toimivien autoregressiivisten kielimallien kuten GPT-3 työkalujen avulla voidaan luoda uskottavalta vaikuttavaa tekstiä lyhyessä ajassa. GPT-3 toiminta perustuu syväoppimiseen. Kielimallien kehittyessä ja yleistyessä väärän ja oikean tiedon tunnistaminen on vaikeutunut. Esimerkkinä nettisivusto ”NotRealNews.net”, joka on tehty käyttäen pelkästään tekoälyn avulla. (Blauth ym. 2022.)

6.5 Syväväärennökset

Syväväärennökset (eng. Deepfakes) ovat tekoälyn kehityksen mukana yleistyneet vauhdilla. Syväväärennösten tarkoituksena on huijata katsojaa esittäytymällä toisena henkilönä muuttaen käyttäjän ulkonäköä ja ääntä. Yleisimmät syväväärennöksen kohteet ovat suositut henkilöt, kuten poliitikot ja julkkikset. Se mahdollistaa myös monen erityyppisen haitallisen käytön, kuten väärän tiedon levityksen, propagandan ja kiusaamisen. (Blauth ym. 2022.)

6.6 Haittaohjelmat

Haittaohjelmat (eng. malware) ovat olleet jo pitkään mukana kyberhyökkäyksissä jo 1970-luvulta lähtien. Haittaohjelmien käytön nousu ja tekoälyn kehitys ovat nykypäivän tietoturvan yksi suurimmista huolista. Satoja tuhansia uusia saastuttavia sovelluksia ja haittaohjelmia rekisteröidään päivittäin. Uusien suojauskeinojen käyttöönotto vaikeutunut tekoälyn tuoman kehityksen myötä tuoden monimutkaisuutta järjestelmiin ja kyberuhkiin. Tekoäly rinnastettuna haittaohjelmiin tekevät niistä vaikeasti havaittavia ja entistä tehokkaampia, jonka seurauksena se on noussut suureksi huolenaiheeksi tietoturvassa. Nykypäivän haittaohjelmat, jotka toimivat tekoälyllä eivät ole kuitenkaan kovin kehittyneitä ja niitä tutkitaan akateemisissa tutkimustöissä ja käytännön konsepteissa. IBM:n vuonna 2018 esittelemä DeepLocker järjestelmä rinnastaa haittaohjelman tekoälyllä ja samalla parantaa sen välttelymahdollisuuksia. DeepLocker toimii myös tutkimalla tekoälyjärjestelmien selitettävyyttä käyttääkseen sitä hyödykseen. Käyttämällä syviä neuroverkkoja se pystyy valitsemaan kohteensa automaattisesti ja salaamaan tarkoituksensa, kunnes se on päässyt haluttuun päämäärään. Tämän tyyppisten tekoälypohjaisten haittaohjelmien yksi suurimmista riskeistä on niiden mahdollisuus saastuttaa monia järjestelmiä ilman niiden huomaamista. (Blauth ym. 2022.)

6.7 Automaattiset asejärjestelmät

Tekoälyn tuoman nopean päätöksenteon ja laskentatehon ansiosta sitä on koitettu hyödyntää armeijan asejärjestelmissä jo 1950-luvulta lähtien. Tekoälyllä toimivat automaattiset asejärjestelmät (eng, autonomous weapons systems, AWS) mahdollistavat esimerkiksi lentokoneiden ja laivojen itsenäisen ohjauksen, sekä automaattisen hyökkäyksen ja kohteiden valitsemisen, mutta ne tuovat mukanaan myös riskejä. Tekoälyllä toimivat asejärjestelmät oppivan siihen syötetyllä datalla tunnistamaan ja tekemään päätöksiä itsenäisesti. Rikolliset voivat manipuloida ja myrkyttää dataa, jota käytetään automaattisten

asejärjestelmien kouluttamiseen. Myrkytetyllä datalla toimiva lennokki voi vaihtaa alkuperäistä kohdetta aiheuttaen valtavia tuhoja. Vaikka automaattisia asejärjestelmiä kehitellään jatkuvasti, ne ovat herättäneet paljon keskusteluita ja huolia niiden luotettavuuden, tarkkuuden ja mahdollisten seurausten takia joutuessaan rikollisten käsiin. (Blauth ym. 2022.)

7 Loppuyhteenveto

Opinnäytetyön tarkoituksena oli selvittää tekoälyn roolia ja vaikutuksia tietoturvaan nyt ja tulevaisuudessa. Tutkimuksessa ilmeni tekoälyn olevan välttämätön osa tietoturvaa nykyisin ja tulevaisuudessa, koska verkossa liikkuva tiedon määrä ja uusien kyberuhkien lisääntyminen vaatii jatkuvaa seurantaa tietoturva-alan osaajilta joista on kasvava pula. Tekoäly pystyy tekemään nopeita laskelmoitua päätöksiä sekä parantamaan tietoturvaa vapauttaen tietoturva-alan ammattilaisten resursseja, jolloin he voivat keskittyä tärkeimpiin tietoturvaan liittyviin toimenpiteisiin. Tekoälyn tuoman laskentatehon ja automatisoinnin avulla uhkien havainnointi paranee ja reaktioaika voi pienentyä sadoista tunneista sekunteihin.

Tekoäly on tuonut myös paljon haasteita ja riskejä tietoturvaan sen lisäämän monimutkaisuuden takia järjestelmiin sekä sovelluksiin. Koneoppimiseen käytetyn datan koulutus ja validointi voi johtaa virheisiin ja ohittamaan joitain hyökkäyksiä, koska yritykset haluavat saada tuotteensa mahdollisimman nopeasti markkinoille. Nykypäivänä tekoälyn ajankohtaisimmat sovellukset tietoturvassa ovat yleisesti IDS-järjestelmissä, joissa käytetään koneoppimista ja keinotekoisia neuroverkkoja. Tekoälyä on myös väärinkäytetty eri sosiaalisen median alustoilla, ja kyberrikolliset ovat hyödyntäneet tekoälyä mm. kalastelusähköposteilla.

Työn rajauksena oli tekoäly ja sen alamuodot. Miten tekoälyä sovelletaan tietoturvaan, mitkä ovat tekoälyn tuomat haasteet ja uhat tietoturvassa sekä mitä etuja se tuo tietoturvaan sovellettuna.

Tekoälyyn ja tietoturvaan liittyviä luotettavia lähteitä löytyi paljon aiheen ollessa laaja. Suurin osa lähteistä oli tieteellisiä artikkeleita jotka keskittyivät yksittäisiin osa-alueisiin, kuten koneoppimiseen ja keinotekoisiin neuroverkkoihin ja niiden käyttökohteisiin tietoturvassa.

Työssä saavutettuja tavoitteita olivat tekoälyn ja tietoturvan merkitys ja siihen liittyvien haasteiden ja riskien ymmärtäminen sekä käytännön soveltuvuudet. Opin tarkastelemaan myös paljon erilaisia tieteellisiä artikkeleita ja poimimaan niistä keskeisimmät asiat omaan työhön liittyen.

Tutkimuksen tuloksia voi käyttää suuntaa antavana pohjana yksityisille henkilöille, tai yrityksille kun tekoälyä ja tietoturvaa sovelletaan esimerkiksi sovelluskehityksessä, tietosuojan parantamisessa ja tarkastelussa tai lähtöpohjana valmiiksi kootuista lähteistä ja tutkimustyöstä jatkaakseen aiheen tarkastelua syvällisemmin. Hyviä jatkotutkimisaiheita ovat esimerkiksi tekoälypohjaiset virtuaaliavustajat ja chatbotit.

Lähteet

Ahmad, Z.; Shahid Khan, A.; Wai Shiang, C.; Abdullah, J.; Ahmad, F. 2021. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Trans Emerging Tel Tech*. Vol. 32. Viitattu 19.4.2023. https://doi.org/10.1002/ett.4150

Alshahrani, A & Clark, A. 2022. A Transfer Learning Approach to Discover IDS Configurations Using Deep Neural Networks. 2022 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), Dalian, China, 2022. Vaatii käyttäjätunnuksen. Viitattu 28.3.2023. https://ieeexplore.ieee.org/document/9926695

Bhatia, V.; Choudhary, S.; Ramkumar, K, R. 2020. A Comparative Study on Various Intrusion Detection Techniques Using Machine Learning and Neural Network. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India. Vaatii käyttäjätunnuksen. Viitattu 18.4.2023. https://ieeexplore.ieee.org/document/9198008

Blauth, F. T.; Gstrein, O. J.; Zwitter, A. 2022. Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. IEEE Access. vol. 10, 77110 – 77122. Viitattu 29.3.2023. https://ieeexplore.ieee.org/document/9831441

Bresniker, K.; Gavrilovska, A.; Holt, J.; Milojicic, D.; Tran, T. 2019. Grand Challenge: Applying Artificial Intelligence and Machine Learning to Cybersecurity. Computer. vol. 52, No 12, 45-52. Viitattu 20.3.2023. https://ieeexplore.ieee.org/document/8909930

Chan, L. 2019. Survey of AI in Cybersecurity for Information Technology Management. IEEE Technology & Engineering Management Conference (TEMSCON). Atlanta, GA, USA, 1-8. Vaatii käyttäjätunnuksen. Viitattu 20.3.2023. https://ieeexplore.ieee.org/document/8813605

Dertat, A. 2017. Applied Deep Learning – Part 3: Autoencoders. Towardsdatascience-sivusto. Viitattu 1.5.2023. https://towardsdatascience.com/applied-deep-learning-part-3-autoencoders-1c083af4d798

ENISA threat landscape 2022. Uhkamaisemaraportti. www.enisa.europa.eu-sivusto. Viitattu 21.3.2023. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022

Huawei Technologies Co., Ltd". 2023. Artificial Intelligence Technology. E-Kirja Doabooks-kirjapalvelussa. Ensimmäinen painos. Singapore:Springer Nature. Vaatii kirjautumisen palveluun. Viitattu 9.3.2023. https://directory.doabooks.org/handle/20.500.12854/93980

Kalita, D. 2022. An Overview of Deep Belief Network (DBN) in Deep Learning. Analytics Vidhya-sivusto. Viitattu 1.5.2023. https://www.analyticsvidhya.com/blog/2022/03/an-overview-of-deep-belief-network-dbn-in-deep-learning/

Lehto, M.; Neittaanmäki, P.; Niinimäki, E.; Nyrhinen, R.; Ojalainen, A.; Pölönen, I.; Rautiainen, I.; Ruohonen, T.; Tuominen, H.; Vähäkainu.; Äyrämö, S.; Ogbechie, A.; Savonen, M. 2019. Tekoälyn perusteita ja sovelluksia. Viitattu 30.3.2023. https://tim.jyu.fi/view/kurssit/tie/tiep1000/tekoalyn-sovellukset/kirja#DKUvbnUuGytQ

Lähde, S. 2022. Koneoppiminen teollisessa internetissä ja sen soveltaminen sellu- ja paperiteollisuuteen. Tuotantotalouden kanditaatintyö. Lappeenranta: Lappeenrannan-Lahden teknillinen yliopisto LUT. Viitattu 1.5.2023. https://lutpub.lut.fi/bitstream/handle/10024/164111/Kandidaatintyo%CC%88_Susanna_La%CC%88hde.pdf?sequence=1&isAllowed=y

Merimaa, S. 2021. Koneoppiminen pilvipalveluna. Opinnäytetyö. Sähkö- ja automaatiotekniikan tutkinto-ohjelma. Satakunta: Satakunnan ammattikorkeakoulu. Viitattu 1.5.2023.

https://www.theseus.fi/bitstream/handle/10024/493223/Merimaa_Sesilia.pdf?sequence=2&isAllowed=y

Michael, J. B & Wingfield, T. C. 2021. Defensive AI: The Future Is Yesterday. Computer. vol. 54, No 9, 90-96. Vaatii käyttäjätunnuksen. Viitattu 20.3.2023. https://ieeexplore.ieee.org/document/9524660

Murugesan, S. 2022. The AI-Cybersecurity Nexus: The Good and the Evil. IT Professional. Vol. 24, No 5, 4-8. Viitattu 21.3.2023. https://ieeexplore.ieee.org/document/9967400

Ozkaya, E. 2019. Cybersecurity: The Beginner's Guide: A comprehensive Guide to Getting Started in Cybersecurity. E-kirja ProQuest Ebook Central kirjapalvelussa. Packt Publishing, Limited: Birmingham. Vaatii kirjautumisen palveluun. Viitattu 9.3.2023. https://ebookcentral.proquest.com/lib/turkuamk-ebooks/detail.action?docID=5781046#

Paananen, V. 2018. Neuroverkkojen FPGA-toteutus. Kandidaatintyö. Elektroniikan ja tietoliikennetekniikan tutkinto-ohjelma. Oulu: Oulun yliopisto. Viitattu 29.4.2023. http://jultika.oulu.fi/files/nbnfioulu-201805312377.pdf

Parisi, A. 2019. Hands-On Artificial Intelligence for Cybersecurity: Implement Smart AI Systems for Preventing Cyber Attacks and Detecting Threats and Network Anomalies. E-kirja ProQuest Ebook Central kirjapalvelussa. Packt Publishing, Limited: Birmingham. Vaatii kirjautumisen palveluun. Viitattu 9.3.2023. https://ebookcentral.proquest.com/lib/turkuamk-ebooks/detail.action?docID=5847212.

Romney, G. W.; Guymon, J.; Romney, M. D.; Carlson, D. A. 2019. Curriculum for Hands-on Artificial Intelligence Cybersecurity. 18th International Conference on Information Technology Based Higher Education and Training (ITHET), Magdeburg, Germany. Vaatii käyttäjätunnuksen. Viitattu 20.3.2023. https://ieeexplore.ieee.org/document/8937373

Tan, B & Karri, R. 2020. Challenges and New Directions for AI and Hardware Security. IEEE 63rd International Midwest Symposium on Circuits and Systems

(MWSCAS), Springfield, MA, USA, 2020, 277-280. Vaatii käyttäjätunnuksen. Viitattu 20.3.2023. https://ieeexplore.ieee.org/document/9184612

Zeadally, S.; Adi, E.; Baig, Z.; Khan, I. A. 2020. Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity. IEEE Access. Vol. 8. Viitattu 20.3.2023. https://ieeexplore.ieee.org/document/8963730