

Koneoppimiseen perustuva Internet-liikenteen luokittelu tietoturvan näkökulmasta

Abdirahman Jibril

Aalto-yliopisto

KANDIDAATINTYÖ 2026

Sähkötekniikan korkeakoulu

Koneoppimiseen perustuva Internet-liikenteen luokittelu tietoturvan näkökulmasta

Abdirahman Jibril

Koulutusohjelma: Teknistieteellinen kandidaattiohjelma

Pääaine: Informaatioteknologia

Korkeakoulu: Sähkötekniikan korkeakoulu

Päiväys: 15. heinäkuuta 2026

Tekijä	Abdirahman Jibril
Työn nimi	Koneoppimiseen perustuva Internet-liikenteen luokittelu tietoturvan näkökulmasta
Päiväys	15. heinäkuuta 2026
Sivumäärä	28
Pääaine	Informaatioteknologia
Koodi	ELEC3015
Valvoja	TkT Pasi Lassila
Työn ohjaaja(t)	FT Pasi Sarolahti

Tiivistelmä

Internet-liikenteen luokittelu on tärkeä osa verkkojen valvontaa ja tietoturvaa, mutta salauksen yleistymisen ja verkkoliikenteen jatkuva muuttuminen vaikeuttavat portteihin ja pakettien sisältöön perustuvaa tunnistamista. Tämän kandidaatintyön tavoitteena oli selvittää, miten koneoppimista voidaan hyödyntää Internet-liikenteen luokittelussa tietoturvan näkökulmasta, miten luokittelua voidaan käyttää tietoturvan tukena ja millaisia rajoitteita menetelmiin liittyy.

Työssä tarkasteltiin aikaisempaa tutkimusta valvotusta ja valvomattomasta oppimisesta sekä syväoppimisesta. Tarkastelu kohdistui erityisesti sovellus- ja protokollaliikenteen luokitteluun, poikkeavan liikenteen havaitsemiseen ja haitallisen liikenteen tunnistamiseen sekä niissä käytettäviin liikennepiirteisiin ja aineistoihin.

Tarkastelun perusteella koneoppiminen mahdollistaa liikenteen luokittelun myös silloin, kun salaus estää pakettien hyötykuorman suoran tarkastelun. Luokittelussa voidaan hyödyntää esimerkiksi pakettien kokoja ja suuntia, ajoitusta, yhteyksien kestoa sekä liikennevirtojen tilastollisia ominaisuuksia. Luokittelutieto voi tukea verkon valvontaa auttamalla tunnistamaan sovelluksia, normaalista poikkeavaa käyttäytymistä ja mahdollisesti haitallista liikennettä. Yhtä yleisesti parasta menetelmää ei kuitenkaan voida nimetä, sillä mallien suorituskyky riippuu käytetystä aineistosta, piirteistä ja luokittelutehtävästä. Syväoppiminen voi vähentää käsin tehtävän piirteiden muodostamisen tarvetta, mutta myös perinteiset koneoppimismenetelmät voivat olla kilpailukykyisiä.

Keskeisiksi käytännön haasteiksi tunnistettiin aineiston laatu ja edustavuus, verkkoliikenteen muuttuminen ajan myötä, uudet ja tarkoituksellisesti muokatut hyökkäykset, salausprotokollien kehittyminen sekä mallien selitettävyys ja laskennalliset vaatimukset. Työn keskeinen johtopäätös on, että koneoppimispohjaista liikenteen luokittelua ei tule arvioida pelkän luokittelutarkkuuden perusteella. Myös yleistettävyys, kestävyys, käsittelynopeus ja ylläpidettävyys vaikuttavat siihen, kuinka luotettavasti menetelmää voidaan hyödyntää todellisissa ja muuttuvissa verkkoympäristöissä.

Avainsanat	koneoppiminen, Internet-liikenne, liikenteen luokittelu, tietoturva, salattu liikenne
Kieli	suomi

Author	Abdirahman Jibril
Title of thesis	Machine Learning-Based Internet Traffic Classification from a Cybersecurity Perspective
Date	July 15th 2026
Pages	28
Major	Information Technology
Code	ELEC3015
Supervisor	D.Sc. (Tech.) Pasi Lassila
Instructor	Ph.D. Pasi Sarolahti

Abstract

Internet traffic classification supports network monitoring and cybersecurity, but increasing encryption and changing network traffic make traditional port- and payload-based identification more difficult. The aim of this bachelor's thesis was to examine how machine learning can be used for Internet traffic classification from a cybersecurity perspective, how classification can support cybersecurity, and what limitations are associated with these methods.

The thesis examined previous research on supervised and unsupervised learning and deep learning. The review focused on application and protocol traffic classification, anomaly detection, malicious traffic identification, and the traffic features and datasets used in these tasks.

The findings show that machine learning enables traffic classification even when encryption prevents direct inspection of packet payloads. Classification can use packet sizes and directions, timing, connection duration, and statistical flow features. This information can support network monitoring by helping identify applications, abnormal behavior, and potentially malicious traffic. No single method is universally best, because performance depends on the dataset, selected features, and classification task. Deep learning can reduce the need for manually designed features, but traditional machine learning methods can also remain competitive.

The main practical challenges identified were data quality and representativeness, changes in network traffic over time, new and intentionally modified attacks, evolving encryption protocols, model explainability, and computational requirements. The main conclusion is that machine-learning-based traffic classification should not be evaluated only by classification accuracy. Generalizability, robustness, processing speed, and maintainability also affect how reliably the methods can be used in changing network environments.

Keywords	machine learning, Internet traffic, traffic classification, cybersecurity, encrypted traffic
Language	Finnish

Sisällys

Esipuhe	iv
Käytetyt lyhenteet	v
1 Johdanto	1
2 Internet-liikenteen luokittelu tietoturvassa	3
2.1 Perinteiset luokittelumenetelmät	3
2.2 Salatun liikenteen tuomat haasteet	4
3 Koneoppiminen verkkoliikenteen analyysissä	5
3.1 Valvottu oppiminen	5
3.2 Valvottoman oppiminen	6
3.3 Syväoppiminen	6
3.4 Piirteiden valinta ja aineistot	7
4 Koneoppimis pohjaiset käyttökohteet tietoturvassa	9
4.1 Sovellus- ja protokollaliikenteen luokittelu	9
4.2 Poikkeavan liikenteen havaitseminen	10
4.3 Haitallisen liikenteen tunnistaminen	12
5 Menetelmien rajoitteet ja käytännön haasteet	16
5.1 Aineiston laatu ja yleistettävyyys	16
5.2 Muuttuvat hyökkäystavat ja salattu liikenne	17
5.3 Selitettävyyys, reaaliaikaisuus ja käyttöönotto	19
6 Yhteenveto ja johtopäätökset	23
6.1 Vastaukset tutkimuskysymyksiin	23
6.2 Keskeiset havainnot	23
6.3 Jatkotutkimusmahdollisuudet	24
Lähteet	25

Esipuhe

Haluan kiittää työni ohjaajaa FT Pasi Sarolahtea ja valvojaa TkT Pasi Lassilaa heidän ohjauksestaan ja tuestaan kandidaatintyön aikana. Erityiskiitos Saira Järviselle hänen tekstipajastaan, dioistaan ja kannustavasta palautteestaan, joista on ollut suuri apu kirjoittamisessa.

Otaniemi, 15. heinäkuuta 2026

Abdirahman Jibril

Tekoälytyökalujen käyttö

ChatGPT:tä on käytetty aiheen rajaamisen, työn rakenteen suunnittelun ja lähdeaineiston jäsentämisen tukena. Claudea on käytetty termien kääntämiseen, kielenhuoltoon ja LaTeX-muotoiluun sekä Grammarlyä kieliopin tarkistamiseen. Näitä tekoälytyökalujen käyttötapoja lukuun ottamatta tämän työn teksti, analyysi, havainnot ja johtopäätökset ovat tekijän itsensä tuottamia.

Käytetyt lyhenteet

CFS	correlation-based feature selection; korrelaatioon perustuva piirteiden valinta
CNN	convolutional neural network; konvoluutioneuroverkko
DDoS	distributed denial-of-service; hajautettu palvelunestohyökkäys
DNN	deep neural network; syvä neuroverkko
DPI	deep packet inspection; pakettien syvätarkastus
ECH	Encrypted ClientHello; TLS-kättelyn ClientHello-tietoja salaava mekanismi
F1	F1 score; precision- ja recall-arvojen harmoninen keskiarvo
GRU	gated recurrent unit; portitettu toistuva yksikkö
IDS	intrusion detection system; tunkeutumisen havaitsemisjärjestelmä
IoT	Internet of Things; esineiden internet
IP	Internet Protocol; Internet-protokolla
k-NN	k-nearest neighbors; k:n lähimmän naapurin menetelmä
LSTM	long short-term memory; pitkän ja lyhyen aikavälin muistiverkko
ML	machine learning; koneoppiminen
PCAP	packet capture; verkkopakettien kaappaustiedosto
QUIC	QUIC; UDP-pohjainen salattu kuljetusprotokolla
RNN	recurrent neural network; toistuva neuroverkko
SNI	Server Name Indication; palvelimen verkkotunnuksen ilmaiseva TLS-laajennus
SSL	Secure Sockets Layer; salausprotokolla
SSH	Secure Shell; suojattu etäyhteysprotokolla
SVM	support vector machine; tukivektori-kone
TCP	Transmission Control Protocol; kuljetuskerroksen yhteydellinen protokolla
TLS	Transport Layer Security; tietoliikenteen salausprotokolla
UDP	User Datagram Protocol; kuljetuskerroksen yhteydetön protokolla

1 Johdanto

Internet-liikenteen määrä on kasvanut nopeasti, ja samalla verkkoon liitettyjen laitteiden, palveluiden ja sovellusten määrä on lisääntynyt. Verkkoliikenne on nykyään keskeinen osa lähes kaikkea digitaalista toimintaa, kuten viestintää, pilvipalveluita, verkkokauppaa, mobiilisovelluksia ja organisaatioiden tietojärjestelmiä. Tämän vuoksi verkkoliikenteen ymmärtäminen ja seuraaminen on tärkeää sekä verkkojen toiminnan että tietoturvan kannalta. Jos liikennettä ei pystytä tunnistamaan tai luokittelemaan riittävän luotettavasti, myös poikkeavan tai haitallisen toiminnan havaitseminen vaikeutuu [1].

Internet-liikenteen luokittelulla tarkoitetaan verkkoliikenteen jakamista eri luokkiin esimerkiksi käytetyn sovelluksen, palvelun, protokollan tai liikenteen yleisen luonteen perusteella. Luokittelun avulla voidaan erottaa toisistaan esimerkiksi tavallinen käyttäjäliikenne, videopalvelut, pilvipalvelut, salattu liikenne ja mahdollisesti haitallinen toiminta. Tietoturvan näkökulmasta luokittelu auttaa muodostamaan paremman kuvan verkossa tapahtuvasta toiminnasta ja tunnistamaan tilanteita, joissa liikenne poikkeaa normaalista [2].

Verkkoliikennettä on perinteisesti tunnistettu esimerkiksi porttinumeroiden, protokollien ja pakettien sisällön perusteella. Nykyisessä verkkoympäristössä nämä menetelmät eivät kuitenkaan aina ole riittävän luotettavia. Sovellukset voivat käyttää muuttuvia tai epätyypillisiä portteja, ja suuri osa liikenteestä on salattua. Tällöin pakettien sisältöön perustuva tunnistaminen vaikeutuu tai ei ole mahdollista lainkaan. Tämän seurauksena liikenteen luokittelussa on alettu hyödyntää enemmän myös pakettien kokoa, ajoitusta, yhteyksien kestoa ja liikennevirtojen käyttäytymistä kuvaavia piirteitä [3, 4].

Koneoppiminen tarjoaa yhden mahdollisen tavan hyödyntää tällaista tietoa. Sen avulla verkkoliikenteestä voidaan etsiä toistuvia piirteitä ja malleja, joiden perusteella liikennettä voidaan luokitella automaattisesti. Koneoppimismenetelmät eivät perustu pelkästään käsin määriteltuihin sääntöihin, vaan ne voivat oppia luokittelua aiemmin kerätyn datan perusteella. Menetelmiä voidaan hyödyntää esimerkiksi sovellusliikenteen tunnistamisessa, poikkeavan liikenteen havaitsemisessa ja haitallisen toiminnan tunnistamisessa [2].

Tämän kandidaatintyön tavoitteena on selvittää, miten koneoppimista voidaan hyödyntää Internet-liikenteen luokittelussa tietoturvan näkökulmasta. Työ toteutetaan kirjallisuuskatsauksena, joten siinä ei rakenneta omaa koneoppimismallia eikä tehdä omaa kokeellista mittausta. Sen sijaan työssä tarkastellaan aikaisempaa tutkimusta ja muodostetaan kokonaiskuva aiheesta käytetyistä menetelmistä, niiden tietoturvaan liittyvistä käyttökohteista sekä keskeisistä haasteista ja rajoitteista. Tarkastelussa painottuu erityisesti se, miten liikenteen luokittelu voi tukea poikkeavan ja haitallisen liikenteen tunnistamista.

Työssä pyritään vastaamaan seuraaviin tutkimuskysymyksiin:

1. Miten koneoppimista käytetään Internet-liikenteen luokittelussa?
2. Miten liikenteen luokittelua voidaan hyödyntää tietoturvan tukena?
3. Mitkä ovat koneoppimiseen perustuvan liikenteen luokittelun keskeiset haasteet ja rajoitteet?

Työ etenee siten, että ensin käsitellään Internet-liikenteen luokittelun taustaa ja sen merkitystä tietoturvassa. Tämän jälkeen tarkastellaan koneoppimisen keskeisiä menetelmiä verkkoliikenteen analyysissä. Seuraavissa luvuissa käsitellään koneoppimispohjaisten luokittelumenetelmien käyttökohteita, kuten sovellusliikenteen tunnistamista, poikkeamien havaitsemista ja haitallisen liikenteen tunnistamista. Lopuksi tarkastellaan menetelmien rajoitteita ja käytännön haasteita sekä kootaan yhteen työn keskeiset johtopäätökset.

2 Internet-liikenteen luokittelu tietoturvassa

Verkkoliikenteen luokittelun tavoitteena on erottaa toisistaan erilaisia verkossa kulkevia yhteyksiä ja liikennetyyppejä. Samassa verkossa voi kulkea yhtä aikaa esimerkiksi selausliikennettä, videopalveluiden liikennettä, pilvipalveluiden yhteyksiä, tiedostonsiirtoa, IoT-laitteiden liikennettä sekä erilaisia taustalla toimivia palveluita. Luokittelun avulla pyritään ymmärtämään, mihin tarkoitukseen tietty liikenne liittyy ja millainen rooli sillä on verkon toiminnassa. Tietoturvan kannalta normaalin liikenteen tunteminen on tärkeää, sillä sen avulla voidaan helpommin tunnistaa myös normaalia poikkeavaa tai mahdollisesti haitallista toimintaa [2].

2.1 Perinteiset luokittelumenetelmät

Internet-liikenteen luokittelussa on perinteisesti käytetty menetelmiä, jotka perustuvat esimerkiksi porttinumeroihin tai pakettien sisällön tarkastamiseen. Porttipohjaisessa tunnistuksessa sovellus tai palvelu pyritään päättämään TCP- tai UDP-portin perusteella. Menetelmän lähtökohtana on, että tietyt sovellukset käyttävät niille vakiintuneita portteja. Porttien avulla tehtävä tunnistus on yksinkertaista ja kevyttä toteuttaa, mutta sen luotettavuus heikkenee, jos sovellus käyttää muuttuvia, epätyypillisiä tai tarkoituksella harhaanjohtavia porttinumeroita [3].

Toinen perinteinen tapa luokitella liikennettä on pakettien sisällön tarkastaminen eli deep packet inspection (DPI). Siinä liikennettä ei tunnisteta pelkän porttinumeron perusteella, vaan pakettien hyötykuormasta etsitään sovellukselle tai protokollalle tyypillisiä tunnisteita. Menetelmä voi tarjota porttipohjaista tunnistusta tarkempaa tietoa liikenteestä, mutta samalla se vaatii enemmän laskentatehoa ja ajantasaista tietoa sovellusten ja protokollien rakenteista. Salattu liikenne aiheuttaa DPI:lle erityisen ongelman, koska hyötykuorman sisältö ei tällöin ole suoraan tarkastavissa. Pakettien sisällön analysointia voivat lisäksi rajoittaa yksityisyyteen liittyvät vaatimukset [1, 3].

Näiden rajoitteiden vuoksi liikennettä on alettu tarkastella yhä enemmän sen käyttäytymistä kuvaavien piirteiden avulla. Tällaisia voivat olla esimerkiksi pakettien koot, yhteyksien kestot, pakettien väliset aikavälit ja liikenteen suunta. Tilastollisten piirteiden hyöty on siinä, että niitä voidaan muodostaa myös silloin, kun pakettien varsinaista sisältöä ei voida lukea. Samoja liikenteen ominaisuuksia voidaan siten hyödyntää sekä yleisessä liikenteen luokittelussa että salatun liikenteen analysoinnissa [3, 5].

Perinteiset menetelmät eivät kuitenkaan ole täysin käyttökeltottomia. Porttinumeroita voidaan edelleen käyttää yksin tai yhdessä muiden piirteiden kanssa, ja DPI voi toimia tilanteissa, joissa liikenne ei ole salattua ja sisällön tarkastaminen on mahdollista. Nykyisessä verkkoympäristössä niiden rinnalle tarvitaan kuitenkin menetelmiä, jotka eivät ole yhtä riippuvaisia kiinteistä porteista tai näkyvästä hyö-

tykuormasta. Erityisesti salauksen yleistymisen on lisännyt tarvetta tällaisille vaihtoehtoisille lähestymistavoille [2].

2.2 Salatun liikenteen tuomat haasteet

Salattu liikenne on yksi keskeinen syy siihen, miksi Internet-liikenteen luokittelu on muuttunut vaikeammaksi. Salaus suojaa käyttäjän tietoja ja yksityisyyttä, mutta samalla se piilottaa suuren osan tiedosta, jota liikenteen tunnistamisessa on perinteisesti hyödynnetty. Kun pakettien hyötykuorma ei ole nähtävissä, sovellusta tai toimintaa ei voida päätellä samalla tavalla paketin varsinaisesta sisällöstä. Tämä heikentää erityisesti DPI:n kaltaisten menetelmien käyttömahdollisuuksia [5].

Kaikki liikenteeseen liittyvä tieto ei kuitenkaan katoa salauksen seurauksena. Verkoliikenteestä voidaan edelleen tarkastella esimerkiksi pakettien kokoja, määrää ja suuntaa, yhteyksien kestoa sekä pakettien välisiä aikavälejä. Tällaiset piirteet eivät kerro suoraan, mitä viestin sisällä lukee, mutta ne voivat kuvata yhteyden käyttäytymistä ja auttaa erottamaan eri liikennetyyppejä toisistaan [5].

Koneoppiminen mahdollistaa näiden epäsuorien piirteiden hyödyntämisen luokittelussa. Mallille voidaan antaa esimerkiksi tilastollisia piirteitä, aikasarjoja, pakettisekvenssejä tai muita liikenteestä muodostettuja esityksiä. Tällöin luokittelu ei perustu paketin varsinaisen viestisisällön lukemiseen, vaan eri liikennetyyppien käyttäytymisessä havaittaviin eroihin. Sama periaate mahdollistaa sekä salaamattoman että salatun liikenteen käsittelyn [5].

Salatun liikenteen luokittelu ei tästä huolimatta ole yksinkertainen tehtävä. Eri sovellusten liikenne voi muistuttaa toisiaan, ja saman sovelluksen käyttäytyminen voi muuttua verkon, laitteen tai käyttötilanteen mukaan. Myös tutkimusaineistossa hyvin toimivan mallin suorituskyky voi heikentyä, jos sitä käytetään erilaisessa verkkoympäristössä kuin missä se on koulutettu. Salatun liikenteen analyysissä tarvitaan tämän vuoksi riittävän edustavaa aineistoa, huolellista arviointia ja menetelmiä, jotka pystyvät toimimaan myös verkkoympäristön muuttuessa [5].

3 Koneoppiminen verkkoliikenteen analyysissä

Koneoppimista voidaan hyödyntää verkkoliikenteen analyysissä silloin, kun liikenteestä halutaan tunnistaa toistuvia piirteitä, malleja tai poikkeavuuksia. Sen sijaan, että liikennettä luokiteltaisiin pelkästään käsin määriteltujen sääntöjen perusteella, malli voidaan opettaa käyttämään aiemmin kerättyä dataa. Verkkoliikenteen luokittelu muodostuu tällöin useasta vaiheesta. Aineiston keräämisen jälkeen liikenteestä muodostetaan tai valitaan mallille sopivia piirteitä, minkä jälkeen tehdään varsinainen luokittelu ja arvioidaan tulosten luotettavuutta [6].

Koneoppimismenetelmät eroavat toisistaan erityisesti siinä, millaista opetusaineistoa ne tarvitsevat ja millä tavalla ne oppivat datassa olevia rakenteita. Tässä luvussa tarkastellaan valvottua ja valvomatonta oppimista sekä syväoppimista. Lopuksi käsitellään piirteiden valintaa ja aineistoja, joilla on keskeinen vaikutus luokittelumallin toimintaan.

3.1 Valvottu oppiminen

Valvottu oppiminen perustuu valmiiksi merkittyyn opetusaineistoon. Jokaiselle opetuksessa käytettävälle havainnolle tunnetaan oikea luokka, jonka perusteella malli oppii yhteyden havaittujen piirteiden ja luokan välillä. Verkkoliikenteessä luokka voi tarkoittaa esimerkiksi tiettyä sovellusta, protokollaa tai jakoa normaaliin ja haitalliseen liikenteeseen. Opetuksen jälkeen tavoitteena on, että malli pystyy luokittelemaan oikein myös sellaista liikennettä, jota se ei ole aikaisemmin nähnyt [7].

Valvottuun liikenteen luokitteluun voidaan käyttää useita erilaisia koneoppimismenetelmiä. Yleisiä vaihtoehtoja ovat esimerkiksi tukivektorikoneet, lähimmän naapurin menetelmä, päätöspuut, satunnaismetsät, boosting-menetelmät ja neuroverkot. Menetelmät ratkaisevat luokitteluongelman eri tavoin, joten niiden soveltuvuus riippuu muun muassa käytettävistä piirteistä, aineiston koosta ja itse luokittelutetävästä. Sopivan mallin valintaa ei tämän vuoksi voida täysin erottaa käytettävän aineiston ja piirteiden ominaisuuksista [6, 8].

Tietoturvan näkökulmasta valvottu oppiminen on hyödyllinen erityisesti silloin, kun tunnistettavista liikenneluokista on saatavilla riittävästi esimerkkejä. Malli voidaan esimerkiksi kouluttaa erottamaan normaalia liikennettä tunnetuista haitallisista liikennetyypeistä tai tunnistamaan tiettyjä sovelluksia ja protokollia. Riippuvuus merkitystä opetusaineistosta on samalla menetelmän keskeinen rajoite. Jos tietyn hyökkäyksen kaltaisia esimerkkejä ei ole opetusaineistossa tai verkkoliikenne muuttuu merkittävästi, malli ei välttämättä tunnista uutta tilannetta luotettavasti. Lisäksi suuren verkkoliikenneaineiston tarkka luokittelu ja merkitseminen voi olla työlästä [7].

3.2 Valvomaton oppiminen

Valvomattomassa oppimisessä lähtökohta on erilainen, koska mallille ei anneta valmiiksi oikeita luokkia. Sen sijaan tarkoituksena on löytää datasta rakenteita, samankaltaisuuksia tai poikkeavuuksia. Verkkoliikenteen analyysissä tämä voi tarkoittaa esimerkiksi samankaltaisten liikennevirtojen ryhmittelyä tai normaalista käyttäytymisestä selvästi poikkeavien havaintojen etsimistä. Tietoturvan kannalta lähestymistapa on kiinnostava erityisesti silloin, kun kaikkia mahdollisia poikkeamia tai hyökkäystapoja ei tunneta etukäteen [7].

Yksi valvomattoman oppimisen keskeisistä menetelmistä on klusterointi, jossa havaintoja ryhmitellään niiden keskinäisen samankaltaisuuden perusteella. Verkkoliikenteessä ryhmittelyn avulla voidaan esimerkiksi muodostaa joukkoja samantyyppisistä yhteyksistä ilman, että niiden oikeita luokkia tunnetaan etukäteen. Samalla voidaan etsiä havaintoja, jotka poikkeavat selvästi muodostuneista ryhmistä tai liikenteen tavallisesta käyttäytymisestä. Tällainen poikkeama voi olla merkki tapahtumasta, joka vaatii tarkempaa tutkimista [7].

Poikkeava havainto ei kuitenkaan automaattisesti tarkoita haitallista toimintaa. Se voi johtua esimerkiksi harvinaisesta mutta normaalista käyttäytymisestä, uudesta sovelluksesta tai verkon tavallisesta vaihtelusta. Myös tulosten arviointi on valvomattomassa oppimisessä vaikeampaa kuin tilanteessa, jossa oikeat luokat tunnetaan valmiiksi. Mallin muodostamille ryhmille tai löytämille poikkeamille ei välttämättä ole suoraa vertailukohtaa, minkä vuoksi niiden merkitystä joudutaan usein arvioimaan yhdessä muun verkkotiedon ja asiantuntija-arvion kanssa [7].

3.3 Syväoppiminen

Syväoppiminen on koneoppimisen osa-alue, jossa hyödynnetään useista kerroksista koostuvia neuroverkkoja. Sen merkittävä ero moniin perinteisiin koneoppimismenetelmiin liittyy piirteiden muodostamiseen. Perinteisessä koneoppimisessä sopivia piirteitä voidaan joutua suunnittelemaan ja valitsemaan käsin, kun taas syväoppimismalli pystyy oppimaan osan tarvitsemistaan piirteistä suoraan sille annettusta datasta. Tämä vähentää riippuvuutta kokonaan käsin tehdystä piirteiden muodostamisesta ja mahdollistaa monimutkaisempien rakenteiden hyödyntämisen liikenteen luokittelussa [4].

Kaikissa tilanteissa liikenteen olennaisia piirteitä ei ole helppo määritellä etukäteen. Syväoppimisessä voidaan hyödyntää erilaisia neuroverkkorakenteita sen mukaan, millaisessa muodossa liikenne esitetään mallille. Konvoluutioneuroverkkoja (CNN) voidaan käyttää rakenteellisten esitysten käsittelyyn, kun taas toistuvat neuroverkot (RNN) ja erityisesti LSTM-mallit soveltuvat ajallisten riippuvuuksien käsittelyyn. Tämä on verkkoliikenteessä olennaista, koska yksittäisten pakettien lisäksi myös niiden järjestys, ajoitus ja yhteyden aikana tapahtuvat muutokset voivat sisältää luokittelun kannalta hyödyllistä tietoa [1].

Automaattinen piirteiden oppiminen ei kuitenkaan tarkoita, että syväoppiminen ratkaisisi kaikki liikenteen luokitteluun liittyvät ongelmat. Syväoppimismallit tarvitsevat usein runsaasti opetusaineistoa ja laskentaresursseja, ja laadukkaan merkityn

verkkoliikenneaineiston kerääminen voi olla vaikeaa. Lisäksi monimutkaisen mallin tekemän päätöksen perusteita ei aina ole helppo tulkita. Tämä on tietoturvasovelluksissa tärkeä rajoite, sillä haitalliseksi luokitellun liikenteen kohdalla olisi hyödyllistä pystyä arvioimaan myös sitä, millä perusteella malli päätyi tulokseensa [4, 5].

Salattu liikenne on yksi alue, jossa syväoppimisen ominaisuuksista voi olla erityistä hyötyä. Kun pakettien varsinaista sisältöä ei voida tarkastella, malli voi hyödyntää esimerkiksi pakettien kokoja, suuntia, järjestystä ja liikenteen ajallista käyttäytymistä. Syväoppiminen mahdollistaa myös raakaliikenteen tai siitä muodostettujen sekvenssien käyttämisen mallin syötteenä, jolloin kaikkia luokittelun kannalta hyödyllisiä piirteitä ei tarvitse määritellä etukäteen. Menetelmän toimivuus riippuu silti käytettävästä aineistosta, eikä yhdessä tutkimusympäristössä saavutettu hyvä suorituskky välttämättä siirry sellaisenaan toiseen verkkoon tai muuttuneeseen liikenneympäristöön [1, 5].

Syväoppimismallin suorituskky ei siten riipu pelkästään valitusta neuroverkko-rakenteesta. Myös sillä, millaisesta aineistosta malli oppii ja millaisessa muodossa verkkoliikenne sille esitetään, on keskeinen merkitys. Seuraavaksi tarkastellaan tarkemmin piirteiden valintaa ja käytettävien aineistojen merkitystä verkkoliikenteen luokittelussa.

3.4 Piirteiden valinta ja aineistot

Koneoppimismallin suorituskky riippuu olennaisesti siitä, millaista tietoa verkkoliikenteestä käytetään luokittelussa. Liikenteestä voidaan muodostaa piirteitä esimerkiksi pakettien koon, määrän ja suunnan, yhteyden keston sekä pakettien välisten aikavälien perusteella. Näiden lisäksi voidaan käyttää erilaisia tilastollisia tunnuslukuja, jotka kuvaavat liikennevirran käyttäytymistä kokonaisuutena. Sopivien piirteiden tavoitteena on tuoda esiin sellaisia eroja, joiden avulla eri sovellukset, protokollat tai liikennetyypit voidaan erottaa toisistaan. Erityisesti salatussa liikenteessä tällaiset piirteet ovat tärkeitä, koska pakettien varsinaista sisältöä ei yleensä voida käyttää luokittelun perustana [5, 6].

Piirteiden lisäksi myös käytettävän aineiston laatu vaikuttaa suoraan siihen, mitä koneoppimismalli pystyy oppimaan. Opetusaineiston tulisi sisältää riittävästi erilaisia liikennetyyppejä ja verkkotilanteita, jotta malli ei opi vain yhden rajatun ympäristön ominaisuuksia. Valvotussa oppimisessa myös luokkamerkintöjen oikeellisuus on tärkeää, sillä virheelliset tai puutteelliset merkinnät voivat heikentää mallin toimintaa. Jos aineisto ei vastaa riittävän hyvin todellista verkkoliikennettä, tutkimusaineistossa hyvin toimiva malli voi suoriutua selvästi heikommin uudessa verkkoympäristössä [4, 7].

Koneoppimismallin arviointia varten aineisto jaetaan yleensä erillisiin opetus-, validointi- ja testiaineistoihin. Opetusaineistoa käytetään mallin parametrien oppimiseen, kun taas validointiaineiston avulla voidaan arvioida mallin toimintaa koulutuksen aikana ja tehdä esimerkiksi mallin asetuksiin liittyviä valintoja. Lopullista suorituskkyä arvioidaan testiaineistolla, jota malli ei ole käyttänyt oppimiseen. Tällainen erottelu vähentää riskiä, että mallin suorituskky arvioitaisiin samalla

aineistolla, johon se on jo mukautunut [4].

Piirteiden ja aineiston valinta liittyvät siten tiiviisti toisiinsa. Hyväkään luokittelumenetelmä ei välttämättä toimi luotettavasti, jos käytetty aineisto on liian rajallinen tai jos sen ominaisuudet poikkeavat selvästi siitä liikenteestä, jota mallin pitäisi myöhemmin käsitellä. Verkkoliikenteessä tämä on erityisen tärkeää, koska sovellukset, protokollat ja liikenteen käyttäytyminen muuttuvat ajan myötä. Mallin suorituskykyä ei tämän vuoksi voida arvioida pelkästään yhden aineiston perusteella, vaan olennaista on myös se, kuinka hyvin menetelmä toimii sille ennestään tuntemattomassa liikenteessä [4, 5].

Luvussa käsitellyt koneoppimismenetelmät eroavat toisistaan erityisesti opetusaineiston, oppimistavan ja käyttökohteiden perusteella. Taulukkoon 1 on koottu valvotun oppimisen, valvomattoman oppimisen ja syväoppimisen keskeisiä ominaisuuksia verkkoliikenteen analyysin ja tietoturvan näkökulmasta [1, 7, 8].

Taulukko 1. Koneoppimismenetelmien vertailu verkkoliikenteen luokittelussa

Menetelmäryhmä	Opetusaineisto	Tyypillisiä menetelmiä	Soveltuvuus tietoturvaan	Keskeisiä rajoitteita
Valvottu oppiminen	Valmiiksi luokiteltu ja merkitty aineisto	SVM, päätöspuut, satunnaismetsä, k-NN, boosting	Tunnettujen liikenne- luokkien, sovellusten ja haitallisen liikenteen tunnistaminen	Riippuvuus laadukkaasta merkitystä aineistosta; uusien tai tuntemattomien hyökkäysten tunnistaminen voi olla vaikeaa
Valvoton oppiminen	Luokkamerkintöjä ei tarvita	Klusterointi ja poikkeamien tunnistaminen	Normaalista poikkeavan käyttäytymisen ja mahdollisesti tuntemattomien tapahtumien havaitseminen	Poikkeama ei välttämättä tarkoita hyökkäystä; tulosten arviointi ja tulkinta voivat olla vaikeita
Syväoppiminen	Usein suuri määrä opetusaineistoa; syöteenä voidaan käyttää myös raakaliikennettä tai sekvenssejä	CNN, RNN, LSTM, autoenkooderit	Salatun liikenteen, monimutkaisten liikennepiirteiden ja haitallisen liikenteen tunnistaminen	Suuri aineisto- ja laskentaresurssien tarve, tulkittavuus sekä yleistettävyyys uusiin verkkoympäristöihin

Menetelmäryhmien välillä ei ole yhtä vaihtoehtoa, joka olisi kaikissa verkkoliikenteen luokittelutehtävissä paras. Valvottu oppiminen soveltuu hyvin tilanteisiin, joissa liikenneluokat tunnetaan ja laadukasta merkittyä aineistoa on saatavilla. Valvomattoman oppimisen vahvuus puolestaan liittyy tilanteisiin, joissa kaikkia poikkeamia ei tunneta etukäteen. Syväoppiminen mahdollistaa monimutkaisten ja osittain automaattisesti opittujen piirteiden hyödyntämisen, mutta samalla sen käyttö lisää aineistoon, laskentaan ja mallin tulkintaan liittyviä vaatimuksia. Menetelmän valinta riippuu siten ennen kaikkea luokittelutehtävästä, käytettävästä aineistosta ja siitä, millaista verkkoliikennettä halutaan tunnistaa [1, 4, 7].

4 Koneoppimis pohjaiset käyttökohteet tietoturvassa

Koneoppimiseen perustuvaa verkkoliikenteen luokittelua voidaan hyödyntää useissa tietoturvaan liittyvissä tehtävissä. Liikenteestä opittujen piirteiden avulla voidaan tunnistaa verkossa käytettäviä sovelluksia ja palveluita, etsiä normaalista poikkeavaa käyttäytymistä sekä erottaa haitalliseksi arvioitua liikennettä tavallisesta verkkoliikenteestä. Tällaiset käyttökohteet liittyvät toisiinsa, sillä tieto liikenteen tyypistä ja normaalista käyttäytymisestä voi auttaa havaitsemaan myös tilanteita, joissa verkon toiminta muuttuu odottamattomalla tavalla [2, 7].

4.1 Sovellus- ja protokollaliikenteen luokittelu

Sovellus- ja protokollaliikenteen luokittelussa pyritään tunnistamaan, mikä sovellus tai verkkoprotokolla havaitun liikennevirran taustalla on. Luokittelutietoa voidaan hyödyntää esimerkiksi verkon hallinnassa, resurssien kohdentamisessa, pääsynvalvonnassa sekä verkon valvonnan ja epäilyttävän liikenteen tunnistamisen tukena. Perinteiset porttinumeroihin tai pakettien sisältöön perustuvat menetelmät eivät kuitenkaan aina sovellu tähän tehtävään, sillä sovellukset voivat käyttää dynaamisia portteja ja salausta estää hyötykuorman suoran tarkastelun. Koneoppimisen avulla liikennettä voidaan tällöin luokitella sellaisista ominaisuuksista, jotka säilyvät havaittavina myös salatussa yhteydessä, kuten pakettien koosta, suunnasta, ajoituksesta ja liikennevirran tilastollisista piirteistä [9–11].

Salatun liikenteen sovellustunnistus perustuu usein siihen, että eri sovellukset tuottavat toisistaan poikkeavia liikenneprofileja, vaikka pakettien sisältö ei olisi nähtävissä. Esimerkiksi pakettien kokoihin, suuntiin ja liikennevirtojen tilastollisiin ominaisuuksiin perustuvilla malleilla on pystytty erottamaan toisistaan SSH- ja Skype-liikennettä ilman IP-osoitteiden, porttinumeroiden tai hyötykuorman käyttöä. Samankaltaista lähestymistapaa voidaan soveltaa myös mobiilisovelluksiin, joiden salatusta liikenteestä jää havaittavaksi sovellukselle tyypillisiä piirteitä. Näiden piirteiden avulla sovelluksia voidaan tunnistaa myös silloin, kun SSL- tai TLS-salaus estää varsinaisen viestisisällön tarkastamisen [12, 13].

Koneoppimiseen perustuva sovellusluokittelu voidaan toteuttaa sekä käsin muodostettujen liikennepiirteiden että automaattisesti opittujen esitysten avulla. Flow-pohjaisessa lähestymistavassa mallille voidaan antaa esimerkiksi yhteyden kestoa, pakettien ja tavujen määrää sekä muita liikennevirran ominaisuuksia kuvaavia piirteitä. Huolellinen piirteiden valinta voi parantaa luokittelua myös silloin, kun käytössä on perinteinen koneoppimismenetelmä, kuten satunnaismetsä. Syväoppimisessa vaihtoehtona on antaa mallin muodostaa hyödyllisiä piirteitä suoremmin liikennedatasta. Deep Packet -menetelmässä tähän käytetään konvoluutioneuroverkkoa ja

autoenkooderia, joiden avulla voidaan tehdä sekä sovellusten tunnistamista että liikennetyyppien luokittelua [11, 14].

Syväoppiminen ei kuitenkaan automaattisesti johda perinteisiä koneoppimismenetelmiä parempaan luokittelutulokseen. Menetelmien suorituskky riippuu myös käytettävästä aineistosta, syötteen esitystavasta ja siitä, kuinka hyvin valitut piirteet erottavat liikenneluokat toisistaan. Esimerkiksi flow-piirteisiin perustuva satunnaismetsä voi joissakin aineistoissa suoriutua CNN-malleja paremmin, vaikka syväoppimisen etuna on mahdollisuus oppia piirteitä automaattisesti. Myös mobiilisovellusten salatussa liikenteessä eri mallien väliset erot riippuvat luokittelutehtävästä, eikä yksi mallityyppi ole kaikissa tilanteissa selvästi paras [11, 15].

Sovellus- ja protokollaliikenteen tunnistamisen haasteena on lisäksi se, että liikenne ei pysy muuttumattomana ajan myötä. Sovellusten päivitykset, eri laitteet ja käyttöympäristöt voivat muuttaa niiden tuottamia liikennepiirteitä, jolloin aiemmin opittu malli ei välttämättä tunnista samaa sovellusta yhtä luotettavasti myöhemmin. Mobiilisovellusten tapauksessa sovellusversioiden ja laitteiden vaihtumisen on havaittu heikentävän tunnistustarkkuutta selvästi. Tämän vuoksi käytännössä toimivan luokittelujärjestelmän tulisi huomioida liikenteen muuttuminen ajan myötä, ja arvioinnissa olisi tarkasteltava myös sitä, kuinka hyvin malli toimii opetusaineistosta poikkeavissa tilanteissa [13, 15].

4.2 Poikkeavan liikenteen havaitseminen

Poikkeavan liikenteen havaitsemisessa tavoitteena on tunnistaa verkkoliikenteestä sellaisia käyttäytymismalleja, jotka eroavat verkon normaalista toiminnasta. Poikkeamat voivat liittyä esimerkiksi hyökkäyksiin, väärinkäytöksiin tai muihin epätavallisiin verkkotapahtumiin, mutta havaittu poikkeama ei itsessään tarkoita, että liikenne olisi haitallista. Koneoppimista voidaan hyödyntää muodostamalla aineiston perusteella käsitys normaalista ja poikkeavasta liikenteestä sekä luokittelemalla uusia liikennevirtoja tämän perusteella. Valvotussa oppimisessa mallia voidaan kouluttaa valmiiksi merkityillä normaali- ja hyökkäysnäytteillä, kun taas valvomattomassa oppimisessa poikkeamia voidaan etsiä ilman valmiita luokkamerkintöjä [7].

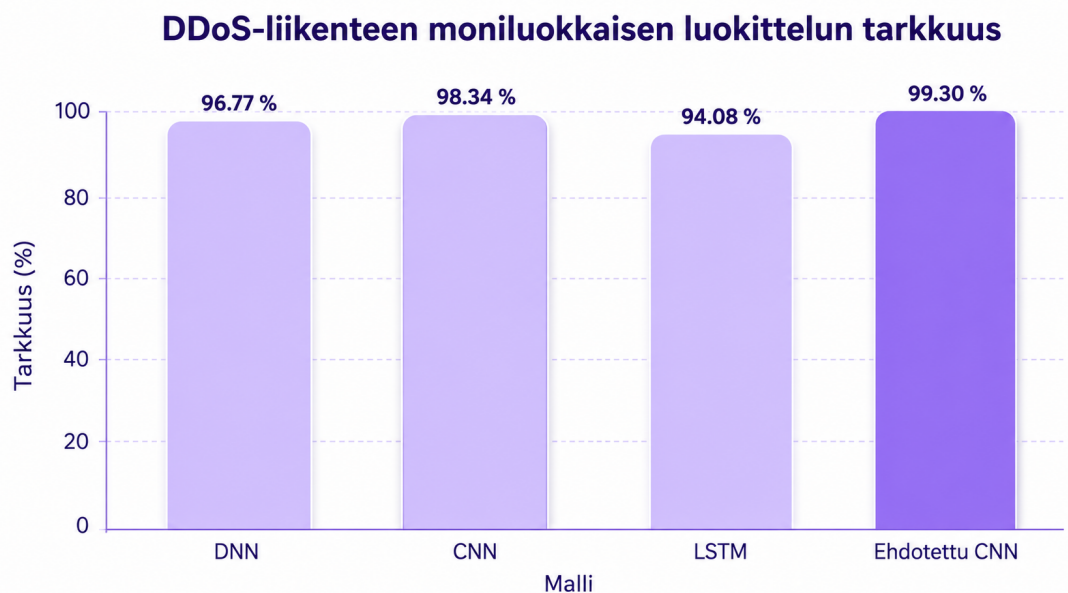
Poikkeavan liikenteen tunnistamista voidaan toteuttaa liikennevirtatasolla, jolloin yksittäisten pakettien sijaan tarkastellaan yhteyden kokonaisuutta kuvaavia ominaisuuksia. Tällaisia piirteitä voidaan käyttää normaalin liikenteen ja hyökkäysliikenteen erottamiseen sekä tarkempaan eri hyökkäystyyppien luokitteluun. CICIDS2017-aineistolla tehdyssä tutkimuksessa useita koneoppimismenetelmiä verrattiin sekä binäärisessä että moniluokkaisessa tunkeutumisen havaitsemisessa. Päätöspuihin perustuvat menetelmät saavuttivat tutkimuksessa korkean suorituskvyn, ja piirteiden määrää pystyttiin vähentämään siten, että mallin testaukseen tarvittava aika pieneni yli 30 prosenttia vain vähäisellä suorituskvyn heikkenemisellä [16].

Poikkeavan liikenteen havaitsemista vaikeuttaa se, että todellisessa verkkoliikenteessä normaaleja tapahtumia on yleensä huomattavasti enemmän kuin hyökkäyksiä. Tällöin opetusaineisto voi olla voimakkaasti epätasapainoinen, jolloin koneoppimismalli oppii tunnistamaan enemmistönä olevan normaalin liikenteen hyvin mutta

harvinaisemmat hyökkäysluokat heikommin. Ongelmaa voidaan lieventää esimerkiksi vähentämällä enemmistöluokan näytteitä tai lisäämällä vähemmistöluokkien näytteitä opetusaineistossa. NSL-KDD- ja CSE-CIC-IDS2018-aineistoilla tehdyssä tutkimuksessa luokkien tasapainottamista hyödynnettiin sekä perinteisten koneoppimismenetelmien että syväoppimismallien kanssa, jotta harvinaisempien hyökkäysluokkien tunnistamista voitiin parantaa [17].

DDoS-hyökkäysten tunnistaminen on yksi esimerkki poikkeavan liikenteen havaitsemisesta, jossa koneoppimista voidaan käyttää normaalin liikenteen ja useiden hyökkäystyyppien erottamiseen. CIC-DDoS2019-aineistolla tehdyssä tutkimuksessa verrattiin DNN-, CNN- ja LSTM-malleja sekä hyödynnettiin aineiston esikäsittelyä, kuten piirteiden valintaa, päällekkäisten näytteiden poistamista, luokkien tasapainottamista ja normalisointia. Parhaiten suoriutunut CNN-pohjainen malli saavutti tutkimusaineistossa 99,99 prosentin tarkkuuden binäärisessä luokittelussa ja 99,30 prosentin tarkkuuden moniluokkaisessa luokittelussa. Tulokset osoittavat, että syväoppimista voidaan hyödyntää DDoS-liikenteen tunnistamisessa, mutta korkeat tulokset liittyvät samalla käytettyyn aineistoon ja sen esikäsittelyyn [18].

Akgun et al. [18] raportoivat CNN-pohjaisten mallien saavuttaneen DNN- ja LSTM-malleja korkeamman tarkkuuden moniluokkaisessa DDoS-liikenteen luokittelussa, kuten kuvassa 1 esitetään.



Kuva 1. DNN-, CNN-, LSTM- ja ehdotetun CNN-mallin moniluokkaisen DDoS-liikenteen luokittelutarkkuus CIC-DDoS2019-aineistolla. Kuva on muodostettu lähteen [18, s. 10] tulosten perusteella ja muokattu työn visuaaliseen ilmeeseen sopivaksi.

Kuvan 1 perusteella mallien välillä oli eroja, vaikka kaikkien tarkkuus oli tutkimuksessa korkea. Ehdotettu CNN-malli saavutti parhaan tuloksen 99,30 prosentin tarkkuudella, kun taas LSTM-mallin tarkkuus oli vertailun matalin, 94,08 prosenttia. Tuloksia tarkasteltaessa on kuitenkin huomioitava, että ne perustuvat yhteen aineistoon ja tutkimuksessa tehtiin esikäsittelyvaiheisiin, joten niiden perusteella ei voida päätellä CNN-mallien olevan yleisesti muita menetelmiä parempia [18].

Poikkeavan liikenteen havaitsemisessa korkea suorituskky tunnetulla testiaineistolla ei vielä takaa sitä, että järjestelmä tunnistaisi myös aiemmin näkemättömiä hyökkäyksiä. Valvotut menetelmät oppivat opetusaineistossa esiintyvien luokkien perusteella, minkä vuoksi kokonaan uudenlainen hyökkäys voi jäädä tunnistamatta. Valvomattomat menetelmät voivat tässä suhteessa olla hyödyllisiä, koska ne etsivät normaalista käyttäytymisestä poikkeavia havaintoja ilman valmiita hyökkäysluokkia. Niiden haasteena on kuitenkin se, että poikkeama ei välttämättä tarkoita hyökkäystä ja tulosten oikeellisuuden arviointi voi olla vaikeampaa. Poikkeamien havaitsemisen arvioinnissa tulisi tämän vuoksi huomioida luokittelutulosten lisäksi myös menetelmän kyky yleistyä uusiin ja muuttuviin verkkotilanteisiin [7, 16].

4.3 Haitallisen liikenteen tunnistaminen

Haitallisen liikenteen tunnistamisessa tavoitteena on erottaa esimerkiksi haittaohjelmien tuottama verkkoliikenne tavallisesta liikenteestä ja mahdollisuuksien mukaan tunnistaa myös liikenteen taustalla oleva haittaohjelmaperhe. Koneoppiminen soveltuu tähän tehtävään erityisesti silloin, kun liikenteestä voidaan muodostaa erottavia piirteitä ilman varsinaisen hyötykuorman tarkastamista. Tämä on tärkeää salatussa liikenteessä, jossa perinteinen sisältöön tai tunnettuihin allekirjoituksiin perustuva tunnistus voi vaikeutua. Luokittelussa voidaan hyödyntää esimerkiksi liikennevirran metatietoja, pakettien ominaisuuksia sekä niistä automaattisesti opittuja esityksiä [19, 20].

Haittaohjelmaliikenteen luokittelussa syväoppimisen avulla piirteitä voidaan oppia suoraan raakaliikenteestä ilman, että kaikki luokittelussa käytettävät ominaisuudet määritellään etukäteen käsin. Verkkoliikenne voidaan esimerkiksi muuntaa numeeriseksi kaksiulotteiseksi esitykseksi, jota CNN-malli käsittelee kuvan tavoin. USTC-TFC2016-aineistolla toteutetussa tutkimuksessa tällaista lähestymistapaa käytettiin sekä haitallisen ja normaalin liikenteen erottamiseen että tarkempaan liikenneluokkien tunnistamiseen. Tutkimuksessa tarkasteltiin binääristä luokittelua sekä 10 ja 20 luokan luokittelutehtäviä, ja eri kokeiden keskimääräiseksi tarkkuudeksi raportoitiin 99,41 prosenttia. Tulokset osoittavat, että automaattisesti opittuja esityksiä voidaan hyödyntää haittaohjelmaliikenteen tunnistamisessa ilman perinteistä käsin tehtävää piirteiden muodostamista [20].

Korkea luokittelutarkkuus ei kuitenkaan yksin osoita, että menetelmä toimisi yhtä hyvin todellisessa ja jatkuvasti muuttuvassa verkkoympäristössä. USTC-TFC2016-aineistoon perustuvassa tutkimuksessa liikennenäytteiden koko ja luokkien määrä olivat ennalta määriteltäviä, minkä vuoksi mallin yleistettävyyttä uudenlaiseen aineistoon ei voitu vielä varmistaa. Lisäksi CNN-malli hyödynsi pääasiassa liikenteen rakenteellisia ominaisuuksia eikä ottanut huomioon ajallisia piirteitä. Tutkimuksessa luokiteltiin myös vain tunnettuja haittaohjelmatyyppejä, joten menetelmän kyky tunnistaa kokonaan uusia haittaohjelmia jäi avoimeksi [20].

Salatun haittaohjelmaliikenteen tunnistamisessa koneoppiminen voi hyödyntää yhteydestä saatavaa metatietoa myös silloin, kun varsinaista viestisisältöä ei voida tarkastella. Käytännön ympäristössä luokittelua vaikeuttavat kuitenkin erityises-

ti virheelliset luokkamerkinnät ja verkkoliikenteen muuttuminen ajan myötä. Miljooniin TLS-salattuihin istuntoihin ja 12 kuukauden aineistoon perustuvassa tutkimuksessa satunnaismetsä suoriutui vertailtujen menetelmien joukossa hyvin, mutta vielä suurempi vaikutus luokittelutuloksiin oli käytettyjen piirteiden laadulla. Monipuolisempi, verkkoliikenteen ominaisuuksia paremmin kuvaava piirrejoukko paransi tuloksia enemmän kuin pelkkä koneoppimisalgoritmin vaihtaminen [19].

Haitallisen liikenteen luokittelussa syväoppiminen ei välttämättä tarjoa parempaa suorituskykyä kuin perinteiset koneoppimismenetelmät. Kahdella haitallista verkkoliikennettä sisältävällä aineistolla tehdyssä vertailussa tarkasteltiin sekä binääristä haittaohjelmien tunnistamista että haittaohjelmaperheiden ja aiemmin näkemättömien hyökkäysten luokittelua. Tuloksissa esimerkiksi satunnaismetsä ja päätöspuut saavuttivat useissa tehtävissä syväoppimismallien kanssa kilpailukykyisen tai paremman suorituskyvyn, vaikka niiden laskennalliset vaatimukset olivat pienemmät. Menetelmän valinnassa ei tämän vuoksi ole perusteltua olettaa monimutkaisemman mallin olevan automaattisesti paras, vaan myös aineiston ominaisuudet, luokittelutehtävä ja käytettävissä olevat laskentaresurssit vaikuttavat sopivan menetelmän valintaan [21].

Haittaohjelmaliikenteen tunnistamista on tutkittu myös IoT-ympäristöissä, joissa päätelaitteiden rajalliset resurssit voivat vaikeuttaa raskaiden tietoturvamenetelmien käyttöä suoraan laitteissa. Eräässä lähestymistavassa verkkoliikenne muunnettiin visuaaliseen muotoon ja luokiteltiin ResNet50-neuroverkolla normaaliksi tai haitalliseksi liikenteeksi. Tuhannesta PCAP-tiedostosta muodostetulla aineistolla malli saavutti 94,50 prosentin tarkkuuden, 95,78 prosentin precision-arvon, 94,02 prosentin recall-arvon ja 94,90 prosentin F1-arvon. Tulos osoittaa, että verkkoliikenteen visuaalista esitystä voidaan hyödyntää haitallisen IoT-liikenteen tunnistamisessa, mutta menetelmän toimivuutta on arvioitava myös käytetyn aineiston ulkopuolella ennen johtopäätöksiä sen yleisestä suorituskyvystä [22].

Haitallisen liikenteen luokittelussa yksittäisen tarkkuusluvun lisäksi on siten arvioitava mallin yleistettävyyttä, piirteiden laatua ja laskennallisia vaatimuksia [19–21].

Luvussa tarkastellut tutkimukset osoittavat, että koneoppimiseen perustuvaa verkkoliikenteen luokittelua voidaan soveltaa erilaisiin tietoturvatehtäviin ja hyvin erilaisilla menetelmillä. Taulukkoon 2 on koottu esimerkkejä sovellusliikenteen, poikkeavan liikenteen ja haitallisen liikenteen luokittelusta sekä tutkimuksissa havaituista keskeisistä tuloksista ja rajoitteista.

Taulukko 2. Luvussa 4 tarkasteltujen koneoppimispohjaisten verkkoliikenteen luokittelututkimusten vertailu

Tutkimus	Tehtävä ja aineisto	Menetelmä	Keskeinen havainto	Rajoite tai huomio
Taylor et al. [13]	Salattujen mobiilisovellusten tunnistaminen; 110 Android-sovellusta ja eri ajankohtina kerättyjä aineistoja	AppScanner ja liikennevirroista muodostetut piirteet	Sovelluksia pystyttiin parhaassa tapauksessa tunnistamaan kuusi kuukautta myöhemmin jopa 96 % tarkkuudella	Laitteen, käyttöjärjestelmän ja erityisesti sovellusversion muuttuminen heikensi tunnistamista
Barut et al. [11]	TLS-salatun sovellusliikenteen luokittelu; non-VPN2016-aineisto	Satunnaismetsä, k-NN sekä yksi- ja kaksiulotteiset CNN-mallit	Satunnaismetsä saavutti vertailussa CNN-malleja paremman luokittelutuloksen, ja piirteiden valinta paransi TLS-liikenteen luokittelua	TLS-salattujen näytteiden määrä oli aineistossa pieni, mikä rajoitti erityisesti syväoppimismallien arviointia
Rodríguez et al. [16]	Tunkeutumisen havaitseminen; CICIDS2017	PART, J48, satunnaismetsä ja muut ML-menetelmät sekä CFS-piirteiden valinta	Päätöspuupohjaiset menetelmät saavuttivat F1-arvoja yli 0,999. Piirteiden vähentäminen pienensi testausaikaa yli 30 % vain pienellä suorituskyvyn heikkenemisellä	Tulosten yleistettävyyys muihin IDS-aineistoihin jäi arvioitavaksi
Akgun et al. [18]	DDoS-liikenteen tunnistaminen; CIC-DDoS2019	DNN, CNN ja LSTM sekä aineiston esikäsittely	Ehdotettu CNN saavutti 99,99 % tarkkuuden binäärisessä ja 99,30 % moniluokkaisessa luokittelussa	Tulokset riippuvat käytetystä aineistosta ja esikäsittelystä eivätkä osoita CNN:n olevan yleisesti paras menetelmä

Jatkuu seuraavalla sivulla.

Taulukko 2 jatkuu edelliseltä sivulta.

Tutkimus	Tehtävä ja aineisto	Menetelmä	Keskeinen havainto	Rajoite tai huomio
Wang et al. [20]	Haaitaohjelma- liikenteen tunnistaminen; USTC-TFC2016	CNN ja raa- kaliikenteestä automaattisesti opittu esitys	Eri kokeiden keski- määräiseksi tark- kuudeksi raportoi- tiin 99,41 %, eikä piirteitä tarvinnut muodostaa koko- naan käsin	Tutkimus käsitteli tunnettuja hait- taohjelmia, ajalliset piirteet jäivät pois ja yleistettävyyys vaati lisäarviointia
Anderson ja McGrew [19]	Salatun haaitaohjelma- liikenteen tunnistaminen; miljoonia TLS- istuntoja 12 kuu- kauden ajalta	Kuusi ML- menetelmää ja TLS- metatiedoista muodostetut piirteet	Satunnaismetsä suo- riutui hyvin, mutta piirteiden laadulla oli algoritmin valin- taa suurempi vaiku- tus tuloksiin	Virheelliset luok- kamerinnät ja liikenteen muuttu- minen ajan myötä vaikeuttivat luokit- telua
Lichy et al. [21]	Salatun haaitaohjelma- liikenteen, hait- taohjelma-perhei- den ja aiemmin näkemättömien hyökkäysten luo- kittelu	Satunnais- metsä ja pää- töspuut se- kä CNN- ja GRU-pohjaiset mallit	Perinteiset ML- menetelmät olivat useissa tehtävissä kilpailukykyisiä tai syväoppimista pa- rempia pienemmillä laskennallisilla vaati- muksilla	Mallien suorituskyy- ky vaihteli aineis- ton ja luokittelu- tehtävän mukaan, joten yhtä yleisesti parasta menetel- mää ei voitu osoit- taa

Taulukon 2 perusteella koneoppimismenetelmien välillä ei muodostu yksiselitteistä suorituskyykyjärjestystä. Syväoppimismalleilla on saavutettu erittäin korkeita tulok- sia esimerkiksi DDoS- ja haaitaohjelmaliikenteen luokittelussa, mutta perinteiset ko- neoppimismenetelmät ovat useissa tutkimuksissa olleet niiden kanssa kilpailukykyi- siä tai jopa parempia [11,21]. Tulosten vertailua vaikeuttavat lisäksi erot käytetyissä aineistoissa, piirteissä, luokittelutehtävissä ja arviointimenetelmissä [16,18,19]. Käy- tännön tietoturvasovelluksessa menetelmän valinnassa on tämän vuoksi huomioitava luokittelutarkkuuden lisäksi myös yleistettävyyys, aineiston edustavuus ja laskennal- liset vaatimukset.

5 Menetelmien rajoitteet ja käytännön haasteet

Koneoppimiseen perustuvilla verkkoliikenteen luokittelumenetelmillä voidaan saavuttaa tutkimusaineistoissa korkea suorituskky, mutta todellisissa verkoissa niiden toimintaan vaikuttavat aineiston laatu, liikenteen ja hyökkäystapojen muuttuminen, salaus, mallien tulkittavuus sekä reaaliaikaisen käsittelyn laskennalliset vaatimukset. Siksi korkea luokittelutarkkuus yhdessä aineistossa ei vielä osoita menetelmän toimivan luotettavasti muuttuvassa verkkoympäristössä [5, 7, 19].

5.1 Aineiston laatu ja yleistettävyyys

Aineiston laatuun liittyvät ongelmat voivat heikentää liikenteen luokittelua jo ennen kuin mallia käytetään uudessa verkkoympäristössä. Verkkoliikenneaineistossa voi esiintyä esimerkiksi puuttuvia arvoja, päällekkäisiä havaintoja, virheellisiä mitauksia tai epätarkkoja luokkamerkintöjä. Tietoturvasovelluksissa luokkamerkintöjen oikeellisuus on erityisen tärkeää, koska aina ei voida varmasti tietää, onko yksittäinen liikennevirta normaalia vai haitallista. Lisäksi normaalin liikenteen määrä on usein selvästi haitallista liikennettä suurempi, jolloin harvinaiset hyökkäysluokat voivat jäädä opetusaineistossa aliedustetuiksi. Tällöin malli voi oppia enemmistönä olevan normaalin liikenteen hyvin mutta suoriutua heikommin juuri niissä harvinaisissa liikenneluokissa, joiden tunnistaminen olisi tietoturvan kannalta tärkeää [17, 19].

Verkkoliikenteen muuttuminen ajan myötä voi johtaa tilanteeseen, jossa opetusaineiston ja myöhemmin luokiteltavan liikenteen jakaumat eivät enää vastaa toisiaan. Tällaista muutosta voidaan kuvata data drift -ilmiöllä, jossa mallin syötteenä käytettävän datan jakauma muuttuu ajan kuluessa. Concept drift puolestaan kuvaa tilannetta, jossa mallin syötteiden ja ennustettavan luokan välinen suhde muuttuu ajan myötä. Internet-liikenteessä muutoksia voivat aiheuttaa esimerkiksi uusien protokollien käyttöönotto, sovellusten päivitykset ja liikenteen käyttäytymisen muuttuminen. Tällöin aiemmin opitut luokittelusäännöt eivät välttämättä enää vastaa uutta liikennettä, mikä heikentää mallin yleistettävyyttä [23, 24].

Data driftin vaikutus näkyy konkreettisesti salatun verkkoliikenteen luokittelussa. Yli kahden vuoden ajalta kerätyillä todellisen mobiilioperaattorin verkkoliikenneaineistoilla havaittiin, että vanhemmalla aineistolla opetettujen mallien suorituskky heikkeni selvästi uudemman liikenteen luokittelussa. Kun noin kaksi vuotta vanhalla aineistolla opetettua mallia käytettiin myöhemmin kerätyn liikenteen luokitteluun, UW-H-mallin tarkkuus heikkeni 43,9 prosenttia ja toisen tarkastellun CNN-mallin 49,6 prosenttia. Lyhyemmällä aikavälillä suorituskyyvyn heikkeneminen oli pienempää, mikä osoittaa, että opetusaineiston ajantasaisuudella voi olla suuri merkitys mallin yleistettävyydelle. Tietoturvan näkökulmasta tulos korostaa sitä, että kerän koulutettua liikenteen luokittelumallia ei voida välttämättä käyttää muuttumattomana pitkään, jos verkossa esiintyvät protokollat, sovellukset ja liikennepro-

fiilit muuttuvat [24].

Mallin yleistettävyyttä arvioitaessa on tärkeää, että testiaineisto vastaa mahdollisimman hyvin sellaista liikennettä, jota malli kohtaisi käytännössä. Jos saman ajanjakson liikenne jaetaan satunnaisesti opetus- ja testiaineistoon, molemmissa osissa voi esiintyä hyvin samankaltaisia liikennepiirteitä, jolloin saavutettu suorituskky voi antaa liian optimistisen kuvan mallin toimivuudesta myöhemmin kerätyssä liikenteessä. Ajallisesti johdonmukaisessa arvioinnissa opetusaineisto kerätään ennen testiaineistoa, jolloin voidaan paremmin tarkastella sitä, kuinka hyvin malli toimii liikenteen muuttuessa ajan myötä. Tietoturvan näkökulmasta tällainen arviointi on tärkeää, koska käytännössä luokittelumallia sovelletaan koulutuksen jälkeen uuteen ja osittain ennestään tuntemattomaan verkkoliikenteeseen [25].

Koneoppimiseen perustuvan liikenteen luokittelun yleistettävyys riippuu algoritmin lisäksi aineiston laadusta, luokkajakaumasta ja arviointitavasta. Koska verkko-liikenne muuttuu jatkuvasti, käytössä olevan mallin suorituskky on seurattava ja mallia tarvittaessa päivitettävä uudemmalla aineistolla [19, 24, 25].

5.2 Muuttuvat hyökkäystavat ja salattu liikenne

Tietoturvakäytössä Internet-liikenteen luokittelua vaikeuttaa se, että haitallinen liikenne ja hyökkäystavat muuttuvat jatkuvasti. Erityisesti valvottuun oppimiseen perustuva luokittelumalli oppii opetusaineistossa esiintyvien liikenneluokkien ja piirteiden välisiä yhteyksiä, minkä vuoksi kokonaan uudenlainen hyökkäys ei välttämättä vastaa mitään mallin aiemmin oppimaa luokkaa. Tämä on ongelmallista esimerkiksi zero-day-hyökkäysten kohdalla, koska niiden liikennepiirteitä ei välttämättä ole ollut saatavilla mallin koulutusvaiheessa. Tällöin haitallinen liikenne voi jäädä tunnistamatta tai tulla luokitelluksi johonkin tunnetuista liikenneluokista. Poikkeamien havaitsemiseen perustuvat menetelmät voivat helpottaa uusien uhkien tunnistamista, mutta normaalin verkkoliikenteen suuren vaihtelun vuoksi myös niiden on vaikea erottaa luotettavasti aidosti haitallista toimintaa tavallisesta liikenteen vaihtelusta [26].

Hyökkääjä voi myös tarkoituksellisesti muuttaa verkkoliikenteen piirteitä siten, että koneoppimismalli tekee väärän luokituksen. Tällaisissa väistöhyökkäyksissä (evasion attacks) tavoitteena ei ole välttämättä muuttaa haitallisen toiminnan varsinaista tarkoitusta, vaan muokata liikenteen havaittavia ominaisuuksia niin, että se muistuttaa mallin näkökulmasta enemmän normaalia liikennettä. Muutoksia voidaan tehdä esimerkiksi pakettien määrään, ajoitukseen tai muihin liikennevirrasta muodostettaviin piirteisiin. Tietoturvan kannalta tämä on erityisen merkittävä ongelma, koska hyvin toimiva luokittelumalli voi joutua hyökkääjän aktiivisen manipuloinnin kohteeksi eikä sen suorituskkyä voida arvioida vain tavallisella testiaineistolla [26, 27].

Väistöhyökkäysten vaikutus on osoitettu myös käytännön verkkoliikenteellä tehdyissä kokeissa. Yan et al. [27] tarkastelivat tilannetta, jossa hyökkääjällä ei ollut suoraa pääsyä kohdemallin rakenteeseen tai parametreihin, vaan hän pystyi ainoastaan havainnoimaan, hyväksyikö vai hylkäsikö tunnistusjärjestelmä muokatun liikenteen.

Tästä rajoitetusta tiedosta huolimatta hyökkäys onnistui väistämään useita erilaisia koneoppimiseen perustuvia tunkeutumisen tunnistusmalleja, ja keskimääräinen onnistumisaste oli yli 75 prosenttia. Tulokset osoittavat, että liikenteen luokittelumallin haavoittuvuus ei edellytä sitä, että hyökkääjä tuntee mallin yksityiskohtaisesti. Tietoturvasovelluksessa tämä korostaa tarvetta arvioida luokittelumenetelmän tarkkuuden lisäksi myös sen kestävyyttä tarkoituksellisesti muokattua verkkoliikennettä vastaan.

Salatun haittaohjelmaliikenteen kohdalla luokittelua voidaan pyrkiä väistämään muuttamalla sellaisia liikennevirran ominaisuuksia, joita koneoppimismalli käyttää päätöksenteossa. Esimerkiksi haittaohjelman komento- ja ohjausliikenteessä paketteja voidaan lisätä tai niiden ajoitusta viivästyttää siten, että liikennevirran tilastolliset piirteet muuttuvat ilman, että haittaohjelman varsinainen toiminta estyy. Tällöin luokittelumalli voi tulkita muokatun liikenteen eri tavalla kuin alkuperäisen haitallisen liikenteen. Haaste on merkittävä erityisesti silloin, kun salauksen vuoksi hyötykuorman sisältö ei ole käytettävissä ja luokittelu perustuu suurelta osin juuri liikennevirran ominaisuuksiin [28].

Salauksen yleistymisen rajoittaa niitä tietoja, joihin koneoppimiseen perustuva Internet-liikenteen luokittelu voi tukeutua. TLS-salatussa liikenteessä sovellustason hyötykuorma ei ole suoraan tarkasteltavissa, joten luokittelussa hyödynnetään esimerkiksi pakettien kokoja ja suuntia, saapumisvälejä, liikennevirran tilastollisia ominaisuuksia sekä salausprotokollan näkyviin jääviä kättelytietoja. Näiden piirteiden avulla liikennettä voidaan luokitella ilman salauksen purkamista, mutta samalla malli voi tulla riippuvaiseksi yksittäisistä protokollakentistä tai tietyistä liikennekuvioista. Jos tällainen piirre myöhemmin salataan tai protokollan rakenne muuttuu, mallin aiemmin oppima erottelutapa ei välttämättä enää toimi yhtä luotettavasti [29, 30].

Hyvä esimerkki salauksen kehittymisen vaikutuksesta on TLS-yhteyden muodostamisessa käytettävä Server Name Indication (SNI), joka ilmaisee palvelimen verkko-tunnuksen ja on perinteisesti ollut näkyvissä salaamattomana. SNI voi tarjota liikenteen luokittelulle hyvin erottuvan piirteen, mutta Encrypted ClientHello (ECH) pyrkii salaamaan myös tämän tiedon käyttäjän yksityisyyden parantamiseksi. ECH:n käyttöönoton on havaittu heikentävän sellaisten liikenteen luokittelumenetelmien suorituskykyä, jotka hyödyntävät TLS-kättelyn näkyviä tietoja, mikä osoittaa mallien riippuvuuden käytettävissä olevista protokollapiirteistä. Tietoturvan näkökulmasta kehitys muodostaa kaksijakoisen tilanteen: vahvempi salaus parantaa viestinnän yksityisyyttä, mutta samalla se vähentää verkon valvonnassa ja haitallisen liikenteen luokittelussa käytettävissä olevaa tietoa [29, 30].

QUIC vahvistaa samaa kehitystä, sillä protokollassa liikenteen sisältö on salattu ja myös yhteyden muodostamiseen liittyvää tietoa on vaikeampi hyödyntää suoraan luokittelussa. QUIC-liikenteessä esimerkiksi palvelun tunnistamiseen hyödyllinen SNI ei ole suoraan nähtävissä ilman kättelytietojen purkamista, minkä vuoksi koneoppimismallien on tukeuduttava enemmän liikennevirran metatietoon, pakettien kokoihin, ajoitukseen ja muihin epäsuoriin piirteisiin. Kättelytietojen purkaminen voi tarjota lisää tietoa, mutta se kasvattaa verkkovalvonnan laskennallista kuormaa ja edellyttää järjestelmän päivittämistä protokollaversioiden muuttuessa.

Tämän vuoksi uusien salattujen protokollien käyttöönotto ei ainoastaan vähennä näkyvää tietoa, vaan voi myös muuttaa sitä, mitkä piirteet soveltuvat koneoppimis-pohjaiseen liikenteen luokitteluun [29, 31].

Edellä käsitelty data drift voi näkyä salatun liikenteen luokittelussa myös lyhyellä aikavälillä, jos luokittelussa hyödynnettävät metatiedot muuttuvat. Luxemburk et al. [31] havaitsivat todellisella QUIC-liikenteellä tehdyssä kuukauden mittaisessa arvioinnissa, että Google-palvelujen recall-arvo laski noin 30 prosenttia jo toisen viikon aikana. Tutkimuksessa suorituskyvyn heikkenemisen arvioitiin liittyvän Googlen TLS-palvelinvarmenteen muutokseen, joka vaikutti kättelypaketeista saataviin piirteisiin. Kättelypaketteihin vahvemmin tukeutunut CNN-malli oli muutokselle herkempi, kun taas ilman kättelypaketteja opetetut mallit olivat vakaampia, mutta niiden luokittelukyky oli samalla heikompi. Tulos havainnollistaa käytännön kompromissia: helposti erottuvat protokollapiirteet voivat parantaa luokittelua, mutta niihin voimakkaasti tukeutuva malli voi menettää suorituskyykyään nopeasti, jos nämä piirteet muuttuvat.

Muuttuvat hyökkäystavat ja salaus merkitsevät, että sekä luokiteltava liikenne että siitä saatavat piirteet voivat muuttua. Uudet hyökkäykset ja väistöyritykset haastavat mallin kestävyyttä, kun taas TLS:n, ECH:n ja QUICin kehitys voi tehdä aiemmin hyödyllisistä protokollapiirteistä käyttökeltvottomia. Mallia on siksi arvioitava myös uudenlaista, tarkoituksellisesti muokattua ja ajallisesti muuttuvaa liikennettä vastaan [26, 27, 29, 31].

5.3 Selitettävyys, reaaliaikaisuus ja käyttöönotto

Koneoppimiseen perustuvan Internet-liikenteen luokittelun käytännön soveltuvuuteen vaikuttaa luokittelutarkkuuden lisäksi se, kuinka hyvin mallin tekemät päätökset voidaan ymmärtää. Tämä korostuu tietoturvasovelluksissa, joissa luokittelun tuloksen perusteella voidaan joutua arvioimaan, onko havaittu liikenne todella haitallista vai esimerkiksi väärä hälytys. Erityisesti syväoppimismallit voivat toimia niin sanottuina mustina laatikoina, jolloin yksittäisen luokittelupäätöksen muodostumista on vaikea tulkita suoraan mallin rakenteesta. Tällöin korkea tarkkuus ei yksin riitä, jos verkon ylläpitäjä ei pysty arvioimaan, mihin liikennepiirteisiin havainto perustuu ja voidaanko mallin tuottamaan tulokseen luottaa [5, 26].

Selitettävyyteen vaikuttaa myös se, millaisia piirteitä ja malleja liikenteen luokittelussa käytetään. Esimerkiksi päätöspuiden toimintaa ja selkeästi määriteltyihin tilastollisiin piirteisiin perustuvia luokittelupäätöksiä voidaan usein tulkita helpommin kuin syväoppimismallien sisäisiä esityksiä. Syväoppimisessa piirteitä voidaan oppia automaattisesti suoraan liikenneaineistosta, mutta tällöin muodostuneiden sisäisten esitysten merkitystä voi olla vaikeampi tulkita. Monimutkaisten mallien päätöksiä voidaan jälkikäteen selittää esimerkiksi selitettävän koneoppimisen menetelmillä, mutta pelkkä tieto yksittäisen piirteen vaikutuksesta ei aina vielä kerro, miksi havaittu liikenne on tietoturvan kannalta merkityksellistä. Käytännön verkkovalvonnassa mallin suorituskyvyn rinnalla onkin huomioitava, pystyykö sen tuottamaa luokitusta perustelemaan tavalla, joka auttaa asiantuntijaa arvioimaan havaintoa ja

tekemään sen perusteella päätöksiä [5, 26].

Reaaliaikainen käyttö asettaa verkkoliikenteen luokittelulle vaatimuksia sekä käsittelyn nopeuden että käytettävien piirteiden suhteen. Osa liikennevirran tilastollisista piirteistä voidaan muodostaa vasta, kun riittävä määrä paketteja tai koko liikennevirta on havaittu, mikä lisää luokitteluun kuluva aikaa. Toisaalta päätös voidaan tehdä aikaisemmin vain muutaman ensimmäisen paketin perusteella, mutta tällöin mallilla on käytettävissään vähemmän tietoa liikenteestä. Lisäksi piirteiden muodostaminen ja koneoppimismallin käyttäminen luokitteluun vaativat laskentatehoa ja muistia, mikä voi muodostua ongelmaksi suurissa ja nopeasti liikennöidyissä verkoissa. Tietoturvasovelluksessa luokittelumenetelmän on siten tasapainotettava tarkkuus ja käsittelyaika siten, että epäilyttävä liikenne voidaan tunnistaa riittävän nopeasti käytännön toimenpiteitä varten [5, 6].

Käsittelyn nopeuden lisäksi käytännön verkkoympäristössä luokittelumallin on pysyttävä käsittelemään suuri määrä samanaikaisia liikennevirtoja ilman, että analyysi muodostuu verkon valvonnan pullonkaulaksi. Laskennalliseen kuormaan vaikuttavat itse koneoppimismallin lisäksi esimerkiksi liikennepiirteiden muodostaminen, käsiteltävien pakettien määrä ja mallin tarvitsema muisti. Suurissa verkoissa kaikkien pakettien ja piirteiden käsittely ei välttämättä ole mahdollista riittävän nopeasti, jolloin analyysia voidaan joutua keventämään esimerkiksi vähentämällä käytettäviä piirteitä tai hyödyntämällä liikenteen näytteenottoa. Tämä voi nopeuttaa luokittelua, mutta samalla osa luokkien erottamiseen hyödyllisestä tiedosta voi jäädä pois. Tietoturvasovelluksessa menetelmän käytännön soveltuvuutta onkin arvioitava luokittelutarkkuuden lisäksi sen perusteella, kuinka paljon laskenta- ja muistiresursseja liikenteen jatkuva analysointi vaatii [6, 8].

Koneoppimismallin käyttöönotto todellisessa verkkoympäristössä ei pääty mallin kouluttamiseen, vaan sen toimintaa on ylläpidettävä liikenteen muuttuessa. Uudella aineistolla tehtävä päivittäminen tai uudelleen kouluttaminen aiheuttaa kuitenkin viivettä ja laskennallista kuormaa. Jos päivitetty malli ei ole käytettävissä riittävän nopeasti, saapuvaa verkkoliikennettä joudutaan edelleen luokittelemaan vanhemmalla mallilla, vaikka liikenteen ominaisuudet olisivat jo muuttuneet. Päivitysten nopeuttamiseen voidaan hyödyntää esimerkiksi siirto-oppimista ja mallin hienosäätöä, jolloin kaikkea oppimista ei tarvitse aloittaa alusta. Tietoturvakäytössä tämä tarkoittaa, että luokittelujärjestelmän käytännön toimivuuteen vaikuttavat mallin tarkkuuden lisäksi myös se, kuinka nopeasti muutokset verkkoliikenteessä voidaan havaita ja kuinka tehokkaasti malli voidaan mukauttaa niihin [25].

Käytännön tietoturvaratkaisussa menetelmän valinta edellyttää siten kompromissia luokittelutarkkuuden, selitettävyyden, käsittelyn nopeuden, resurssivaatimusten ja ylläpidettävyyden välillä [5, 6, 25].

Luvussa tarkastellut rajoitteet vaikuttavat koneoppimis pohjaiseen Internet-liikenteen luokitteluun eri vaiheissa aineiston muodostamisesta mallin käytännön käyttöönottoon. Taulukkoon 3 on koottu keskeiset haasteet sekä niiden vaikutukset liikenteen luokitteluun ja tietoturvakäyttöön.

Taulukko 3. Koneoppimispohjaisen Internet-liikenteen luokittelun keskeiset rajoitteet tietoturvakäytössä

Rajoite	Vaikutus liikenteen luokitteluun	Merkitys tietoturvassa
Aineiston laatu ja luokkien epätasapaino	Virheelliset luokkamerkinnät, puuttuvat arvot ja harvinaisten liikenneluokkien vähäinen määrä voivat vääristää mallin oppimista.	Harvinaiset mutta tärkeät hyökkäykset voivat jäädä tunnistamatta tai tulla luokitelluiksi normaaliksi liikenteeksi.
Data drift ja concept drift	Opetusaineiston ja uuden verkkoliikenteen jakaumat tai piirteiden ja luokkien väliset suhteet voivat muuttua ajan myötä.	Aiemmin tarkka malli voi menettää suorituskyykyään muuttuvassa verkkoympäristössä, ellei sen toimintaa seurata ja päivitetä.
Uudet ja aiemmin tuntemattomat hyökkäykset	Valvottu luokittelumalli ei välttämättä tunnista liikennettä, joka poikkeaa opetusaineiston tunnetuista luokista.	Zero-day-hyökkäys voi jäädä havaitsematta tai tulla luokitelluksi väärään tunnettuun luokkaan.
Väistöhyökkäykset	Hyökkääjä voi muuttaa esimerkiksi pakettien määrää, ajoitusta tai muita mallin käyttämiä liikennepiirteitä väärän luokituksen aikaansaamiseksi.	Haitallinen liikenne voi päästä tunnistusjärjestelmän ohi, vaikka malli toimisi tavallisella testiaineistolla hyvin.
Salaus ja protokollien kehittyminen	TLS:n, ECH:n ja QUICin kaltaiset teknologiat vähentävät näkyviä protokolla- ja hyötykuormatietoja, jolloin luokittelu perustuu enemmän epäsuoriin piirteisiin.	Haitallisen salatun liikenteen tunnistaminen vaikeutuu ja aiemmin hyödyllisiin protokollapiirteisiin perustuvat mallit voivat menettää suorituskyykyään.
Selitettävyyys	Erityisesti monimutkaisten syväoppimismallien päätösten muodostumista voi olla vaikea tulkitella.	Asiantuntijan voi olla vaikea arvioida, miksi liikenne luokiteltiin haitalliseksi ja onko havainto luotettava vai väärä hälytys.
Reaaliaikaisuus ja laskennalliset vaatimukset	Piirteiden muodostaminen, suuren liikennemäärän käsittely ja monimutkaisten mallien käyttö voivat kasvattaa viivettä sekä laskenta- ja muistitarvetta.	Liian myöhään valmistuva luokittelu voi heikentää mahdollisuutta reagoida haitalliseen liikenteeseen ajoissa.

Jatkuu seuraavalla sivulla.

Taulukko 3 jatkuu edelliseltä sivulta.

Rajoite	Vaikutus liikenteen luokitteluun	Merkitys tietoturvassa
Mallin käyt- töönotto ja yl- läpito	Muuttuva verkkoliikenne voi edellyttää mallin jatkuvaa seuranta, päivittämistä ja uudelleen kouluttamista.	Vanhentunut malli voi tehdä epäluotettavia päätöksiä juuri silloin, kun verkkoympäristö tai hyökkäystavat ovat muuttuneet.

Taulukon 3 perusteella koneoppimispohjaisen liikenteen luokittelun käytännön haasteet liittyvät toisiinsa eikä niitä voida tarkastella täysin erillisinä ongelmina. Esimerkiksi muuttuva verkkoliikenne voi heikentää mallin yleistettävyyttä, mutta samalla se lisää mallin päivittämisen tarvetta ja siten myös laskennallisia vaatimuksia. Salaus puolestaan vähentää käytettävissä olevia piirteitä, jolloin mallin riippuvuus jäljelle jäävistä liikenneominaisuuksista voi kasvaa. Tietoturvakäytössä luokittelumenetelmän arvioinnissa on tämän vuoksi huomioitava suorituskyvyn lisäksi myös mallin kestävyys, selitettävyys, käsittelyn nopeus ja kyky mukautua muuttuvaan verkkoliikenteeseen.

6 Yhteenveto ja johtopäätökset

Tässä kandidaatintyössä tarkasteltiin koneoppimiseen perustuvaa Internet-liikenteen luokittelua tietoturvan näkökulmasta. Työn perusteella menetelmien käytännön toimivuus riippuu luokittelutarkkuuden lisäksi käytettävästä aineistosta, liikenteestä muodostettavista piirteistä, verkkoympäristön muutoksista ja käyttöönoton vaatimuksista.

6.1 Vastaukset tutkimuskysymyksiin

Ensimmäinen tutkimuskysymys käsitteli sitä, miten koneoppimista käytetään Internet-liikenteen luokittelussa. Luokittelussa hyödynnetään esimerkiksi pakettien kokoja ja suuntia, yhteyksien kestoa, pakettien välisiä aikavälejä ja muita liikennevirran tilastollisia ominaisuuksia. Valvottu oppiminen käyttää merkittyä aineistoa, valvomaton oppiminen etsii ryhmiä ja poikkeamia, ja syväoppiminen voi oppia piirteitä osittain suoraan raakaliikenteestä tai pakettisekvensseistä. Näin liikennettä voidaan luokitella myös silloin, kun salaus estää hyötykuorman suoran tarkastelun.

Toinen tutkimuskysymys käsitteli liikenteen luokittelun hyödyntämistä tietoturvassa. Sovellus- ja protokollaliikenteen tunnistaminen tukee verkon valvontaa. Poikkeamien havaitsemisella voidaan tunnistaa normaalista käyttäytymisestä eroavia havaintoja, ja haitallisen liikenteen luokittelulla voidaan erottaa esimerkiksi hyökkäys- tai haittaohjelmaliikenne tavallisesta liikenteestä. Salatussa liikenteessä hyödyllistä tietoa voidaan saada liikennevirran käyttäytymisestä, vaikka sisältöä ei voida tarkastella suoraan.

Kolmas tutkimuskysymys käsitteli menetelmien haasteita ja rajoitteita. Keskeisiä ongelmia ovat aineiston laatu ja edustavuus, verkkoliikenteen muuttuminen, uudet ja tarkoituksellisesti muokatut hyökkäykset, salauksen ja protokollien kehittyminen sekä selitettävyyden, reaaliaikaisuuden ja ylläpidon vaatimukset. Siksi luotettavuutta ei voida arvioida pelkän luokittelutarkkuuden perusteella, vaan myös yleistettävyys, kestävyys, käsittelynopeus ja ylläpidettävyys on huomioitava.

6.2 Keskeiset havainnot

Työn keskeinen havainto on, ettei yhtä yleisesti parasta luokittelumenetelmää voida nimetä. Syväoppiminen voi vähentää käsin tehtävää piirteiden muodostamista, mutta hyvin valituilla piirteillä myös perinteiset koneoppimismenetelmät voivat olla kilpailukykyisiä ja laskennallisesti kevyempiä. Menetelmän soveltuvuus riippuu siten tehtävästä, aineistosta ja käytettävistä piirteistä.

Toinen keskeinen havainto on, että tutkimusaineistossa saavutettu korkea tarkkuus ei takaa toimintaa muuttuvassa verkossa. Sovellukset, protokollat, palvelut ja hyökkäystavat muuttuvat, jolloin aiemmin hyvin toiminut malli voi menettää suori-

tuskykyään. Käytännön arvioinnissa on siksi tarkasteltava myös ajallista yleistettävyyttä ja toimintaa ennestään tuntemattomassa liikenteessä.

Käyttöönotossa nämä havainnot muodostavat kompromissin useiden vaatimusten välillä. Tarkkuuden lisäksi tarvitaan riittävää käsittelynopeutta, päätösten selitettävyyttä ja järjestelmän ylläpidettävyyttä. Paras ratkaisu ei siksi välttämättä ole yksittäisessä testissä tarkin malli, vaan menetelmä, joka säilyttää riittävän toimintakyvyn muuttuvassa verkkoympäristössä.

6.3 Jatkotutkimusmahdollisuudet

Jatkotutkimuksessa keskeinen suunta on kehittää liikenteen luokittelumenetelmiä, jotka pystyvät mukautumaan verkkoliikenteen muutoksiin ilman jatkuvaa täydellistä uudelleen kouluttamista. Erityisesti data driftin ja concept driftin vaikutusten tunnistaminen sekä mallien tehokas päivittäminen ovat tärkeitä, jos koneoppimis-pohjaista luokittelua halutaan käyttää pitkäaikaisesti todellisissa verkoissa. Samalla arviointimenetelmissä olisi hyödyllistä painottaa ajallisesti muuttuvaa liikennettä ja eri verkkoympäristöihin siirtymistä yksittäisellä aineistolla saavutetun tarkkuuden lisäksi. Näin voitaisiin paremmin arvioida, säilyttääkö luokittelumalli toimintakykynsä myös silloin, kun sovellukset, palvelut ja liikenneprofiilit muuttuvat.

Salatun liikenteen luokittelussa jatkotutkimusta tarvitaan erityisesti tilanteissa, joissa myös aiemmin näkyvissä olleet protokollatiedot vähenevät. TLS:n, ECH:n ja QUICin kaltaisten teknologioiden kehittyminen voi vähentää luokittelussa suoraan hyödynnettävissä olevaa protokollatietoa ja lisätä riippuvuutta muista liikenteestä havaittavista ominaisuuksista. Tämän vuoksi jatkossa olisi hyödyllistä kehittää menetelmiä, jotka tukeutuvat mahdollisimman yleisiin liikenteen käyttäytymistä kuvaaviin piirteisiin, kuten pakettien kokoihin, suuntiin ja ajoitukseen, eivätkä ole liian riippuvaisia yhden protokollaversion rakenteesta. Samalla olisi tärkeää arvioida, kuinka hyvin tällaiset menetelmät säilyttävät luokittelukykyä salausprotokollien ja verkkopalveluiden muuttuessa.

Jatkotutkimusta tarvitaan myös mallien kestävyiden, selitettävyyden ja reaaliaikaisen käytön osalta. Tietoturvasovelluksissa luokittelumallien olisi pystyttävä toimimaan myös tarkoituksellisesti muokattua liikennettä vastaan, joten niiden kestävyttä väistöhyökkäyksille olisi tärkeää arvioida erilaisissa verkkoympäristöissä. Samalla mallien päätösten tulkittavuutta olisi hyödyllistä kehittää niin, että tietoturvasiantuntija pystyy arvioimaan, mihin havaittu luokittelu perustuu. Reaaliaikaisessa verkkovalvonnassa myös käsittelynopeus ja resurssien käyttö asettavat rajoitteita, minkä vuoksi jatkotutkimuksessa olisi hyödyllistä tarkastella kevyempiä malleja ja piirrejoukkoja, jotka säilyttävät riittävän luokittelukyvyn ilman tarpeettoman suurta laskennallista kuormaa. Näiden tavoitteiden yhdistäminen voisi parantaa koneoppimis-pohjaisen Internet-liikenteen luokittelun soveltuvuutta todellisiin tietoturvajärjestelmiin.

Koneoppimis-pohjaisen liikenteen luokittelun käytännön arvo riippuu lopulta siitä, kuinka luotettavasti ja tehokkaasti menetelmä säilyttää toimintakykyä verkkoliikenteen muuttuessa.

Lähteet

- [1] M. Abbasi, A. Shahraki, and A. Taherkordi, “Deep Learning for Network Traffic Monitoring and Analysis (NTMA): A Survey,” *Computer Communications*, vol. 170, pp. 19–41, 2021, doi: 10.1016/j.comcom.2021.01.021.
- [2] F. Pacheco, E. Expósito, M. Gineste, C. Baudoin, and J. Aguilar, “Towards the Deployment of Machine Learning Solutions in Network Traffic Classification: A Systematic Survey,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1988–2014, 2019, doi: 10.1109/COMST.2018.2883147.
- [3] T. T. T. Nguyen and G. Armitage, “A Survey of Techniques for Internet Traffic Classification Using Machine Learning,” *IEEE Communications Surveys & Tutorials*, vol. 10, no. 4, pp. 56–76, 2008, doi: 10.1109/SURV.2008.080406.
- [4] S. Rezaei and X. Liu, “Deep Learning for Encrypted Traffic Classification: An Overview,” *IEEE Communications Magazine*, vol. 57, no. 5, pp. 76–81, 2019, doi: 10.1109/MCOM.2019.1800819.
- [5] M. Shen, K. Ye, X. Liu, L. Zhu, J. Kang, S. Yu, Q. Li, and K. Xu, “Machine Learning-Powered Encrypted Network Traffic Analysis: A Comprehensive Survey,” *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 791–824, 2023, doi: 10.1109/COMST.2022.3208196.
- [6] M. S. Sheikh and Y. Peng, “Procedures, Criteria, and Machine Learning Techniques for Network Traffic Classification: A Survey,” *IEEE Access*, vol. 10, pp. 61135–61158, 2022, doi: 10.1109/ACCESS.2022.3181135.
- [7] S. Wang, J. F. Balarezo, S. Kandeepan, A. Al-Hourani, K. G. Chavez, and B. I. P. Rubinstein, “Machine Learning in Network Anomaly Detection: A Survey,” *IEEE Access*, vol. 9, pp. 152379–152396, 2021, doi: 10.1109/ACCESS.2021.3126834.
- [8] A. I. Getman and M. K. Ikonnikova, “A Survey of Network Traffic Classification Methods Using Machine Learning,” *Programming and Computer Software*, vol. 48, no. 7, pp. 413–423, 2022, doi: 10.1134/S0361768822070052.
- [9] S. Zander, T. T. T. Nguyen, and G. Armitage, “Automated Traffic Classification and Application Identification Using Machine Learning,” in *Proceedings of the 30th Annual IEEE Conference on Local Computer Networks (LCN 2005)*, pp. 250–257, 2005, doi: 10.1109/LCN.2005.35.

- [10] C. V. Wright, F. Monroe, and G. M. Masson, “On Inferring Application Protocol Behaviors in Encrypted Network Traffic,” *Journal of Machine Learning Research*, vol. 7, pp. 2745–2769, 2006.
- [11] O. Barut, R. Zhu, Y. Luo, and T. Zhang, “TLS Encrypted Application Classification Using Machine Learning with Flow Feature Engineering,” in *2020 the 10th International Conference on Communication and Network Security (ICCNS 2020)*, pp. 32–41, 2020, doi: 10.1145/3442520.3442529.
- [12] R. Alshammari and A. N. Zincir-Heywood, “Machine Learning Based Encrypted Traffic Classification: Identifying SSH and Skype,” in *2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA 2009)*, pp. 1–8, 2009, doi: 10.1109/CISDA.2009.5356534.
- [13] V. F. Taylor, R. Spolaor, M. Conti, and I. Martinovic, “Robust Smartphone App Identification Via Encrypted Network Traffic Analysis,” *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 63–78, 2018, doi: 10.1109/TIFS.2017.2737970.
- [14] M. Lotfollahi, M. Jafari Siavoshani, R. Shirali Hossein Zade, and M. Sabeirani, “Deep Packet: A Novel Approach for Encrypted Traffic Classification Using Deep Learning,” *Soft Computing*, vol. 24, no. 3, pp. 1999–2012, 2020, doi: 10.1007/s00500-019-04030-2.
- [15] G. Aceto, D. Ciunzo, A. Montieri, and A. Pescapé, “Mobile Encrypted Traffic Classification Using Deep Learning: Experimental Evaluation, Lessons Learned, and Challenges,” *IEEE Transactions on Network and Service Management*, vol. 16, no. 2, pp. 445–458, 2019, doi: 10.1109/TNSM.2019.2899085.
- [16] M. Rodríguez, Á. Alesanco, L. Mehavilla, and J. García, “Evaluation of Machine Learning Techniques for Traffic Flow-Based Intrusion Detection,” *Sensors*, vol. 22, no. 23, Art. no. 9326, 2022, doi: 10.3390/s22239326.
- [17] L. Liu, P. Wang, J. Lin, and L. Liu, “Intrusion Detection of Imbalanced Network Traffic Based on Machine Learning and Deep Learning,” *IEEE Access*, vol. 9, pp. 7550–7563, 2021, doi: 10.1109/ACCESS.2020.3048198.
- [18] D. Akgun, S. Hizal, and U. Cavusoglu, “A New DDoS Attacks Intrusion Detection Model Based on Deep Learning for Cybersecurity,” *Computers & Security*, vol. 118, Art. no. 102748, 2022, doi: 10.1016/j.cose.2022.102748.
- [19] B. Anderson and D. McGrew, “Machine Learning for Encrypted Malware Traffic Classification: Accounting for Noisy Labels and Non-Stationarity,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’17)*, pp. 1723–1732, 2017, doi: 10.1145/3097983.3098163.

- [20] W. Wang, M. Zhu, X. Zeng, X. Ye, and Y. Sheng, “Malware Traffic Classification Using Convolutional Neural Network for Representation Learning,” in *2017 International Conference on Information Networking (ICOIN)*, pp. 712–717, 2017, doi: 10.1109/ICOIN.2017.7899588.
- [21] A. Lichy, O. Bader, R. Dubin, A. Dvir, and C. Hajaj, “When a RF Beats a CNN and GRU, Together—A Comparison of Deep Learning and Classical Machine Learning Approaches for Encrypted Malware Traffic Classification,” *Computers & Security*, vol. 124, Art. no. 103000, 2023, doi: 10.1016/j.cose.2022.103000.
- [22] G. Bendiab, S. Shiaeles, A. Alruban, and N. Kolokotronis, “IoT Malware Network Traffic Classification Using Visual Representation and Deep Learning,” in *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, pp. 444–449, 2020, doi: 10.1109/NetSoft48620.2020.9165381.
- [23] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang, “Learning under Concept Drift: A Review,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 12, pp. 2346–2363, 2019, doi: 10.1109/TKDE.2018.2876857.
- [24] N. Malekghaini, E. Akbari, M. A. Salahuddin, N. Limam, R. Boutaba, B. Mathieu, S. Moteau, and S. Tuffin, “Deep Learning for Encrypted Traffic Classification in the Face of Data Drift: An Empirical Study,” *Computer Networks*, vol. 225, Art. no. 109648, 2023, doi: 10.1016/j.comnet.2023.109648.
- [25] G. Andresini, F. Pendlebury, F. Pierazzi, C. Loglisci, A. Appice, and L. Cavallaro, “INSOMNIA: Towards Concept-Drift Robustness in Network Intrusion Detection,” in *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security (AISec ’21)*, pp. 111–122, 2021, doi: 10.1145/3474369.3486864.
- [26] K. He, D. D. Kim, and M. R. Asghar, “Adversarial Machine Learning for Network Intrusion Detection Systems: A Comprehensive Survey,” *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 538–566, 2023, doi: 10.1109/COMST.2022.3233793.
- [27] H. Yan, X. Li, W. Zhang, R. Wang, H. Li, X. Zhao, F. Li, and X. Lin, “Automatic Evasion of Machine Learning-Based Network Intrusion Detection Systems,” *IEEE Transactions on Dependable and Secure Computing*, vol. 21, no. 1, pp. 153–167, 2024, doi: 10.1109/TDSC.2023.3247585.
- [28] C. Novo and R. Morla, “Flow-Based Detection and Proxy-Based Evasion of Encrypted Malware C2 Traffic,” in *Proceedings of the 13th ACM Workshop on Artificial Intelligence and Security (AISec ’20)*, pp. 83–91, 2020, doi: 10.1145/3411508.3421379.
- [29] I. Akbari, M. A. Salahuddin, L. Aniva, N. Limam, R. Boutaba, B. Mathieu, S. Moteau, and S. Tuffin, “A Look Behind the Curtain: Traffic Classification in an Increasingly Encrypted Web,” *Proceedings of the ACM on Measure-*

ment and Analysis of Computing Systems, vol. 5, no. 1, Art. no. 4, 2021, doi: 10.1145/3447382.

- [30] D. Shamsimukhametov, A. Kurapov, M. Liubogoshchev, and E. Khorov, “Is Encrypted ClientHello a Challenge for Traffic Classification?” *IEEE Access*, vol. 10, pp. 77883–77897, 2022, doi: 10.1109/ACCESS.2022.3191431.
- [31] J. Luxemburk, K. Hynek, and T. Čejka, “Encrypted Traffic Classification: The QUIC Case,” in *2023 7th Network Traffic Measurement and Analysis Conference (TMA)*, pp. 1–10, 2023, doi: 10.23919/TMA58422.2023.10199052.