TEKOÄLY HAVAINNOI – KONEAISTIT

Asiantuntijalausunto Eduskunnan tulevaisuusvaliokunnalle

Signaalinkäsittelyn professori Joni-Kristian Kämäräinen,
Tietotekniikan yksikkö, Tampereen yliopisto

Sisällys

Tekoäly havainnoi - tulin, näin, voitinko? 3

Koneaistit - tietokonenäkö tekoälyn veturina 3

Yksi menetelmä löytää heidät, se yksi heitä hallitsee 4

Tekoälyn nykytila - ” kun Suomi putos puusta” 5

Tutkimus siirtynyt yrityksiin - kaiken takana on data, joka ihanasti opettaa 5

Tehdään Suomesta taas suuri - oppia elokapinalta ja Grezkyltä 6

Eurooppalainen vastarintaliike - datakapina 6

Pienen Suomen pitää tehdä sitä missä se on hyvä eikä sitä mikä on kivaa 7

Minne kiekko menee seuraavaksi? Luonnollinen kieli ja robotiikka 8

Tekoäly havainnoi – tulin, näin, voitinko?

Tekoäly kattaa kaikki teknologiat, joita tarvitaan fyysisten töiden tekemiseen kykenevien robottien rakentamiseksi. Robotti havainnoi koneaisteilla kuten kuulo, näkö ja tunto, liikkuu raajoilla, telaketjuilla, pyörillä tai lentäen ja vaikuttaa ympäristöönsä koskettamalla ja tarttumalla. Tällaisen robotin pitäisiä kyetä tekemään ihmisen sille antamia tehtäviä ja vapauttaa ihmiset tärkeämpiin tehtäviin. Tärkeämpää voi olla vaikkapa toisen ihmisen, asukkaan, kohtaaminen senioritalossa.

Robotin ja muiden koneiden havainnoinnin toteuttamisessa tekoäly ja erityisesti koneoppimismenetelmät ovat olleet keskeisessä roolissa. Aisteista kenties tärkein, näkö, ja sen laskennallinen toteuttaminen, tietokonenäkö (eng. computer vision), on ollut erittäin merkittävässä roolissa. Nyt on sopiva hetki välitilinpäätökselle eli mistä tultiin, missä ollaan ja mihin pitäisi mennä seuraavaksi?

Koneaistit - tietokonenäkö tekoälyn veturina

Tekoälyä on tutkittu aktiivisesti jo 70-luvulta, jolloin sen ensimmäiset tieteelliset konferenssit ja lehdet aloittivat toimintansa. Ala houkutteli ja houkuttelee edelleen tutkijoita eri tieteenalojen piiristä, koska ihmisenkaltaisen laskennallisen älykkyyden kehittäminen on houkutteleva tutkimuskohde. Tietokonenäköä ja erityisesti kohteiden tunnistamista pidettiin yhtenä avaimena tekoälyyn ja 90-luvun lopusta alkaen tietokonenäkö onkin houkutellut valtavasti tutkijoita. Tietokonenäöstä tuli suurin ala tutkijoiden ja artikkeleiden määrällä mitattuna ja sen pääkonferenssi, CVPR, lähentelee viittausmäärillään sellaisia tieteen kärkilehtiä kuin Science ja Nature.

Tietokonenäössä tavoitteena on antaa kameralla varustetulle tietokoneelle ihmisenkaltainen näkökyky. Vastaavasti konekuulossa (eng. machine hearing) tavoitteena on antaa mikrofonilla varustetulle tietokoneelle ihmisenkaltainen kuulo.

Kehitys tietokonenäön ja muiden koneaistien alalla oli tasaista vuoteen 2012 asti ja sitä voidaankin pitää modernin tekoälyn syntymävuotena. Silloin brittiläiskanadalainen Toronton yliopiston professori Geoffrey Hintonin (s. 1947) tutkimusryhmä julkaisi läpimurtoartikkelin[1]. Artikkelissa esiteltiin uusi lähestymistapa, joka voitti tietokonenäön seuratuimman kilpailun: Pascal VOC. Hintonin ryhmän menetelmä perustui laskennalliseen rakenteeseen nimeltä *konvoluutioneuroverkko* (eng. convolutional neural network, CNN).

Konvoluutioneuroverkko eli lyhyemmin konvoluutioverkko on uusi laskentaparadigma, jossa kehitys on korkeammalla abstraktiotasolla. Kehitystyössä täytyy yhdistellä erilaisia laskentakerroksia ongelmaan sopivaksi arkkitehtuuriksi ja loppu opitaan

[1]A. Krizhevsky, I. Sutskever, G.E. Hinton (2012): ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems (NeurIPS).

opetusaineistosta eli datasta. Pascal VOC -kilpailussa kuvat ja niiden luokat muodostavat opetusaineiston. Opittu tieto tallentuu kerrosten painoihin (eng. weights), joita Hintonin ryhmän arkkitehtuurissa on kaikkiaan 60 miljoonaa.



Kuva 1: Konvoluutioverkko koostuu kerroksista, joiden läpi kuva kulkee tuottaen kunkin luokan todennäköisyyden. Kuva vastaa Hintonin ryhmän vuonna 2012 esittelemää arkkitehtuuria (AlexNet)

Hintonin ryhmän esittelemä verkko, ”AlexNet”, on esitetty Kuvassa 1. Verkon painot opetetaan syöttämällä miljoonia opetuskuvia toistuvasti verkon läpi. Verkon toinen pää tuottaa kullekin tuhat todennäköisyyttä, joista jokainen edustaa yhtä luokkaa (”Auto”, ”Kissa”, ”Koira”). Ideaalisesti vain oikean luokan todennäköisyys on 100 prosenttia ja kaikkien muiden 0, mutta aluksi näin ei ole vaan ulostulon virhe takaisinohjataan verkon läpi ja jokaisen kerroksen painoja säädetään virheen pienentämiseksi.

Hintonin ryhmän AlexNet-verkko mullisti tekoälytutkimuksen ja nykyään kaikkien koneaistien alueilla parhaat tulokset saadaan konvoluutioverkoilla.

Yksi menetelmä löytää heidät, se yksi heitä hallitsee

Tekoälytutkimuksen kannalta on merkittävää, että ennen konvoluutioverkkoja tutkimus oli sirpaloitunut tutkimusyhteisöihin. Tietokonenäön, luonnollisen kielen käsittelyn ja äänen analysoinnin tutkijat kävivät eri konferensseissa, julkaisivat eri lehdissä ja heidän käyttämänsä menetelmät olivat erilaisia.

Vuoden 2012 jälkeen kaikki alat ovat siirtyneet käyttämään konvoluutioverkkoihin perustuvia arkkitehtuureja ja sen seurauksena alojen väliset erot ovat kaventuneet. Nyt tietyn alan tutkija voi ottaa nopeasti käyttöönsä parhaan mahdollisen menetelmän toiselta alalta ja jopa parantaa sitä. Vuosi vuodelta konvoluutioverkot laajentuvat yhä uusille tutkimus- ja sovellusalueille, kuten biologiaan ja lääketieteeseen[2].

Konvoluutioverkosta on tullut Taru Sormusten Herran Sormusten sormus - ”Yksi sormus löytää heidät, se yksi heitä hallitsee”.

[2]P. Gainza et al. (2020): Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. Nature Methods 17. DOI

Tekoälyn nykytila – ”kun Suomi putos puusta”

Vielä 90-luvulla Suomi oli tekoälyn kärkimaa ja täällä tehtiin juuri neuroverkkotutkimusta. Kiitos kuuluu akateemikko Teuvo Kohoselle (s. 1934) ja hänen kouluttamilleen tutkijasukupolville. 2000-luvulla neuroverkkotutkimuksesta laajalti luovuttiin ja keskityttiin muihin koneoppimisen menetelmiin.

2010-luvun lopulla Suomessa ja erityisesti Aalto-yliopistossa palattiin neuroverkkotutkimuksen pariin ja sitä on myös ansiokkaasti yhdistetty 2000-luvulla tehtyyn tutkimukseen (mm. Bayesilainen päättely). Uudessa nousussa on auttanut merkittävät panostukset Aalto-yliopiston tekoälytutkimukseen ja yliopisto onkin taas merkittävä tutkimuslaboratorio. Aalto-yliopistosta on syntynyt myös uusia avauksia, kuten konvoluutioverkkojen käyttö tietokonegrafiikassa.

Koneaistien osalta Suomessa tehdään merkittävää tutkimusta ainakin tietokonenäössä ja konekuulossa. Tietokonenäön tutkimus on aktiivisinta Aalto-yliopistossa ja Tampereen ja Oulun yliopistoissa. Konekuulon osalta tutkimus on merkittävää Aalto-yliopistossa ja Tampereen yliopistossa. Muiden aistien osalta tutkimus Suomessa on vähäistä ja Suomeen olisikin hyvä saada laajempi teknologinen osaaminen koneaisteissa kuten tunto (eng. tactile sensing), joka on tärkeää robotiikalle.

Tekoälytutkimuksen tärkeimmältä osa-alueelta eli konvoluutioverkkojen osalta perustutkimus on vahvaa ja kriittisen massan ylittävää Aalto-yliopistossa, mutta muiden menetelmien ja koneaistien osalta useampi yliopisto on saavuttanut riittävän tutkijamassan (Helsingin yliopisto, Tampereen yliopisto ja Oulun yliopisto). Nämä yliopistot pystyvät palvelemaan Suomen koulutus-, tutkimus- ja tuotekehitystarpeita ja panostuksia on hyvä suunnata valmiiksi vahvoihin yliopistoihin. Muiden yliopistojen osalta on hyvä painottaa paikallisia tarpeita, koska vahvaa teollisuutta on laajemmin.

Tekoälyyn, koneoppimiseen ja koneaisteihin suuntautuvaa tutkimusta ja opetusta tulee vahvistaa, koska kaikki ennusteet viittaavat näiden alojen merkittävään rooliin tulevaisuudessa. Yliopistoja on parasta vahvistaa kunkin vahvuusalueella, koska tekoälyn perustutkimuksen nostamiseen kriittisen massan yli jokaisessa suomalaisessa yliopistossa on resurssien osalta mahdotonta.

Tutkimus siirtynyt yrityksiin – kaiken takana on data, joka ihanasti opettaa

Edellä on puhuttu vain menetelmistä ja erityisesti konvoluutioverkoista, mutta alalla on toinen paljon merkittävämpi muutos, jonka konvoluutioverkot ovat mahdollistaneet. Konvoluutioverkkomenetelmät toimivat vain niin hyvin kuin niiden opetuksessa käytetty data on määrällisesti ja laadullisesti. Tätä rajoitetta on yritetty kiertää mm. luomalla simuloitua dataa, mutta kokemukset tähän mennessä ovat olleet heikkoja. Näyttää siltä, että mikään ei voita aitoa ja laadukasta todellista dataa.

Koska datan merkitys on menetelmiä suurempaa ja koska datan keräämistä harvoin pidetään tieteellisenä työnä, niin luonnollisesti monen tutkimusongelman osalta paras osaaminen on siirtynyt yrityksiin. Yrityksillä on resursseja kerätä ja merkitä (eng. annotate tai label) valtavia määriä dataa (kuvia, ääntä, lääketieteellisiä näytteitä jne.)

On olemassa suuri vaara, että yritykset vaikuttavat siihen mitä tutkitaan, koska ne voivat jo pelkällä tutkijoiden määrällään ohjata sitä mitkä ovat pääongelmia ja mitkä kuuluvat marginaaliin. Tekoälyn tärkeimmissä konferensseissa merkittävä osa tiedeartikkeleista tulee suurimmista yrityksistä. Yksittäinen yritys pystyy kilpailemaan suurienkin maiden kanssa tiedeartikkelien määrässä ja näin ohjata tiedettä.

Tekoälytutkimuksesta ja erityisesti tietokonenäön ja konekuulemisen osalta suurin osa tutkimuksesta on siirtynyt jättiyrityksiin kuten Google, Facebook, Apple, Amazon, Microsoft, Baidu, Alibaba ja Huawei. Myös laskenta-alustojen yksipuolisuus eli näytönohjainteknologiaan perustuvat GPU-kortit, on nostanut Nvidian erittäin merkittävään asemaan. Tästä muutoksesta ja sirupulasta johtuen monilla suomalaisilla tutkijoilla on ollut vaikeuksia saada laskentaan soveltuvia GPU-kortteja.

Esimerkkinä datan voimasta toimii sähköautojen valmistaja Tesla, jonka itseajava ohjelmisto on johtava maailmassa. Teslan tekoälyjohtaja Andrej Karpathyn mukaan hänen johtamansa ryhmä käyttää 75 prosenttia ajastaan datan keräämiseen ja merkitsemiseen, mutta todellinen osuus saattaa olla jopa suurempi. Teslan menetelmänkehitys perustuu datan hankkimiseen ja vastaavia yrityksiä tulee lisää.

Suomen ja Euroopan sinisilmäinen suhtautuminen dataan on johtanut data-ajan kilpailukyvyn menetykseen. Tähän on yritetty vaikuttaa EU-tason päätöksillä, mutta ne eivät ole jarruttaneet datan kulkeutumista suuryrityksille. Joidenkin päätösten voidaan kärjistäen sanoa olleen idealistisen herttaisia, mutta käytännössä hyödyttömiä. Eurooppa tarvitsee vähemmän herttaisia, mutta tehokkaita ja itsekkäitä päätöksiä.

Suomen ja Euroopan pitää tehdä toimia eurooppalaisen datan suojelemiseksi niin, että eurooppalaisilla toimijoilla, yrityksillä ja yhteiskunnalla, on siihen suorempi oikeus kuin kansainvälisillä toimijoilla. Lisäksi on syytä ylläpitää osaamista laskenta-alustojen tutkimuksessa ja tuotekehityksessä (esim. System on Chip, SoC, -teknologiat), koska on nähtävissä, että tietoturvalliset tiettyyn tarkoitukseen ja rajoitteisiin pystyvät sirut näyttelevät merkittävää roolia tulevaisuuden autonomissa laitteissa. Esimerkiksi Tesla on päätynyt tekemään omia sirujaan konvoluutioneuroverkkolaskentaan.

Tehdään Suomesta taas suuri - oppia elokapinalta ja Gretzkyltä

Tekoälyssä kilpailu maiden ja yritysten välillä lähti voimakkaaseen kasvuun 2010-luvun jälkipuoliskolla ja 2020-luvulla kilpailu on erittäin kovaa. Lähes kaikilla mailla on oma tekoälystrategiansa ja kaikki maat tekevät kykyjensä mukaan massiivisia investointeja tekoälyn tutkimukseen ja tuotekehitykseen. Tekoälylle, kuten kaikille digitaaliajan yrityksille, on ominaista ns. Voittaja saa kaiken -periaate eli lopulta vain yksi tai muutama jäävät jäljelle ja hallitsevat markkinoita. Näin syntyy monopoliasema, joka on haitaksi muille yrityksille ja maille.

Eurooppalainen vastarintaliike - datakapina

Seuraavassa kuvataan lyhyesti suosituksia Euroopan tasolla.

Suositus 1: Datamonopolien tuhoaminen suuryrityksiä pilkkomalla

Suomi ja Eurooppa ovat datateollisuuden kehitysmaita ja siksi on vaikeaa saada jalansijaa kuluttajadataa hyödyntävillä markkinoilla. Tämä johtuu pääasiassa siitä, että tietyt kansainväliset yhtiöt ovat kasvaneet liian suuriksi[3]. Euroopan pitäisi vaatia näiden yhtiöiden pilkkomista niin, että dataa tuottavat osat (esim. Googlen hakukone) eriytetään omaksi liiketoiminnaksi, joka tarjoaa palveluitaan kaikille yrityksille samoilla kaupallisilla ehdoilla. Näin estetään monopoliaseman väärinkäyttö.

Suositus 2: Eurooppalainen kuolontähti (datamonopoli)

Jos Eurooppa ei kykene painostamaan suuryritysten pilkkomiseen, niin toinen vaihtoehto on ottaa Kiinan ja Venäjän malli ja rakentaa omat jätit Eurooppaan. Kiina ja Venäjä lähtivät jo varhaisessa vaiheessa kilpailemaan yhdysvaltalaisia yrityksiä vastaan ja tuloksena oli omat hakupalvelut (Baidu ja Yandex), jotka ovat kasvaneet suuryrityksiksi ja ekosysteemeiksi. Eurooppa voisi rakentaa yhteisellä rahoituksella omat kilpailevat yrityksensä ja vielä vaikuttaa markkinoihin. Euroopan etuna ovat luotettavuus ja avoimuus. ”Euroopan google” olisi myös Eurooppaa yhdistävä toimi, joka houkuttelisi osaajia ja investointeja Eurooppaan. Toinen mahdollisuus on kopioida yhteiseen hyvään pyrkivän OpenAI-yrityksen konsepti Eurooppaan. Yhtiö on ideologiallaan houkutellut monia hyviä tutkijoita, mutta myös tuottanut useampia läpimurtoja ja on odotettavissa, että sillä on lopulta myös kaupallista merkitystä.

Pienen Suomen pitää tehdä sitä missä se on hyvä eikä sitä mikä on kivaa

Suositus 3: Tukea datan keräämiselle

Suomalaisia ja eurooppalaisia yrityksiä pitäisi tukea siinä, että ne tunnistavat mitä mahdollisesti kilpailuetua tuovaa dataa niillä on tai mitä ne voisivat helposti kerätä. Suomen akatemian (SA) ja Business Finlandin (BF) rahoitusinstrumentit eivät tue datan keräämistä, koska tätä ei nähdä perustutkimuksena (SA) tai sen tuomaa arvoa ei osata mitata (BF). Tulisi luoda tutkimus- ja tuotekehitysinstrumentteja, joiden ensisijainen ehto olisi maailman mittakaavassa erityisen (laadullisesti ja määrällisesti) dataaineiston kerääminen.

Suositus 4: raskas teollisuus - rikas, rämä, mutta oma

Suomen kokoisen maan ei kannata kilpailla kaikilla tekoälyn aloilla. Suomen vahvuuksia ovat raskas teollisuus (esim. raskaat työkoneet), haastava ilmasto (autonomiset laitteet ääriolosuhteissa) ja korkea sosiaalipalveluiden taso (terveydenhuolto ja palvelut lapsille ja senioreille). Tutkimus ja tuotekehitys näillä aloilla vaatii tietokoneiden lisäksi laitteita, laboratorioita ja osaamista, jotka pudottavat kisasta heikoimmat kilpailijamaat. Suomen ei kannata kilpailla pelkässä datan prosessoinnissa, koska se vaatii infrastruktuurina vain tietokoneita ja on siksi mahdollista ja houkuttelevaa köyhimmillekin maille. Määrä voi voittaa kyvykkyyden. On paljon hyödyllisempää kehittää kalliita

[3]S. Galloway (2017): The Four: The Hidden DNA of Amazon, Apple, Facebook, and Google.

ja raskaita laitteita, jotka pudottavat monia kilpailijoita pois markkinoilta. Tällaisia ovat työkoneet, laivat, lentoalukset, tehdashallien ja satamien nosturit ja hissit ja raskas robotiikka. Keskittymällä näitä aloja palvelevaan tekoälyn ja koneaistien tutkimukseen suomalainen insinöörityö ja kokemus pystyy kilpailemaan kansainvälisesti.

Minne kiekko menee seuraavaksi? Luonnollinen kieli ja robotiikka

Tekoälyn murros alkoi tietokonenäöstä, mutta nyt tämä ala on saturoitunut valtavasta yritystutkimuksen määrästä. Saatta olla, että seuraava ”tekoälyn talvi” (eng. AI Winter) alkaa juuri tietokonenäöstä. Uudet teknologiat on kuitenkin nopeasti omaksuttu myös muilla aloilla ja onkin ehkä syytä seurata jääkiekkoilija Wayne Crezkyn isän neuvoa, että ei kannata mennä sinne missä kiekko on nyt vaan sinne mihin se menee seuraavaksi. Kaksi mahdollista ja nopeasti nousevaa alaa ovat

1. robotiikka ja
2. luonnollinen kieli.

Robotit ovat odottaneet tulemistaan koteihin ja pienyrityksiin jo vuosikymmeniä, mutta nyt on odotettavissa, että tekoälyn avustamana murros viimein tapahtuu. Suomessa merkittävää robotiikkatutkimusta tehdään käytännössä enää Aalto-yliopistossa ja Tampereen yliopistossa ja näissäkin ollaan Ruotsin ja Tanskan jäljessä. Näiden yliopistojen tutkimusta robotiikassa tulisi vahvistaa ja harkita jopa robotiikan yksiköiden ja koulutusohjelmien perustamista.

Luonnollisen kielen prosessointi (eng. natural language processing, NLP) oli tärkeässä roolissa tekoälyn alkuvaiheessa, mutta on 2000-luvulla ajautunut marginaaliin ja sen tutkimus ja koulutus hävinneet monista yliopistoista. Konvoluutioneuroverkot ovat kuitenkin nostaneet kielen prosessoinnin uuteen kukoistukseen ja nyt osaajista on valtava pula. Tämän alan tutkimuksella on pitkät perinteet ainakin Helsingin yliopistossa, mutta myös muualla ja välittömästi alan koulutukseen ja tutkimukseen investoimalla saataisiin osaajia yliopistoihin ja yrityksiin.