Tekoäly oppii

Arno Solin

Aalto-yliopisto

Sisällys

Tekoäly oppii 3

Mistä on kyse? 3

Ohjaamaton oppiminen 4

Ohjattu oppiminen 5

Vahvistusoppiminen 5

Tekoäly muutosvoimana 5

Huomioitavaa päättäjille ja kansalaisille 6

Viitteet 7

Kainalolaatikko 8

Tekoäly oppii

Arno Solin, Aalto-yliopisto

Ihmiset ja eläimet omaksuvat asioita havainnoimalla ympäristöään sekä yrityksen ja erehdyksen kautta tai yleistämällä jo oppimiaan asioita. Tekoälyt lainaavat meiltä tapoja oppia, mutta yhtäläisyyksiä on vähemmän kuin äkkiseltään toivoisi. Tekoälyt kykenevät kuitenkin *mukautumaan* havaintoihin, *oppimaan* kokemuksistaan ja *selittämään* ympäröivää maailmaa.

Mistä on kyse?

Sana 'tekoäly' viittaa moniin erilaisiin menetelmiin ja eri osista koostuviin järjestelmiin. Jo tekoälyn määrittelyn hankaluus antaa vihjeen siitä, että tekoälyt eivät opi yhdellä selkeästi määritellyllä tavalla. Tapa omaksua tietoa riippuu menetelmien rakenteesta ja sovellusalueesta. Kaikki tekoälyt eivät opi, vaan niiden toiminta voi myös perustua ihmisen määrittelemään käytökseen. Tällaisia ovat esimerkiksi jotkin pelitekoälyt, jotka seuraavat niille syötettyjä lainalaisuuksia (pelin säännöt), tai esimerkiksi älykäs termostaatti, joka seuraa lämpötilaa ja säätelee lämmitystä annettujen havaintojen perusteella. Moni nykyisistä tekoälyjärjestelmistä perustuu kuitenkin tekoälyn kouluttamiseen havaintoaineiston (datan) avulla eli koneoppimiseen.

Koneoppiminen (engl. *machine learning*) viittaa järjestelmiin, jotka oppivat havainnoimalla tietoaineistoja ja seuraamaan löytämiään lainalaisuuksia. Koneoppimisessa ei tyypillisesti tavoitella ihmisen kaltaista yleistä älykkyyttä vaan koulutetaan menetelmiä, jotka kykenevät tekemään yksittäisiä samankaltaisia asioita, joita me älykkäät olennot kykenemme tekemään.

Voimme kuvitella kaksi ääripäätä, joista toisessa mallinnettavasta ilmiöstä tiedetään kaikki ja havaintoja ei tarvita (esim. shakin pelaamisessa säännöt) ja toisessa ääripäässä ei ennalta tiedetä mitään mutta kaikki voidaan havainnoida/mitata (esim. kaikkien kansalaisten syntymäaika). Kummassakaan ääripäässä koneoppimista ei tarvita, sillä kaikki mahdollinen joko tiedetään tai havainnoidaan. Koneoppimista tarvitaan näiden kahden ääripään välissä, jossa on tarpeen yhdistää ennakkotietoa ja havaintoja opittavasta ilmiöstä. Ennakkotieto voi koostua myös oletuksista. Oletusten ja havainnoinnin tasapaino vaihtelee sovelluksen mukaan; fysiikan lait on helppo kertoa tekoälylle matemaattisessa muodossa, kun taas kissan kuvailu kaavoilla on vaikeaa. Tämän takia jotkin ihmiselle helpot asiat ovat tekoälylle vaikeasti opittavia – ja toisin päin.

Koneoppiminen on käytännössä kolmen osatekijän summa: mallin, havaintoaineiston ja koulutusmenetelmän. Ensinnä tarvitaan matemaattinen **malli**, joka sisältää mallioletukset ja jonka rakenne pitkälti määrää, mitä tekoäly kykenee oppimaan. Käytännössä malli sisältää vapaita parametreja. Mitä enemmän vapaita parametreja, sitä joustavampi malli yleensä on ja sitä monimutkaisempiin opittaviin asioihin se kykenee mu-



*Kuva 1: Vasemmalla: **Ohjaamattomasti** eläinkuvien avulla opetetun tekoälyn luomia kuvia eläimistä. Keskellä: **Ohjatusti** opetetun tekoälyn nimeämiä asioita ja esineitä kuvassa (henkilö, solmio, kuppi). Oikealla: **Vahvistusoppimalla** toimiva robotti, joka opettelee kävelemään yrityksen ja erehdyksen kautta.*

kautumaan. Mallioletukset voivat olla joko eksplisiittisiä eli esimerkiksi luonnonlaeista, säännöistä tai toiveista koostuvia tai implisiittisiä eli mallin rakenteesta johtuvia. Toiseksi tarvitaan **havaintoaineisto**, joka kuvaa opittavaa tehtävää. Kolmanneksi tarvitaan **menetelmä** (algoritmi), joka kykenee yhdistämään mallin ja aineiston. Käytännössä malli koulutetaan selittämään annettu havaintoaineiston löytämällä parametreille arvot, jotka mahdollistavat tämän. Tavoitteena on, että koulutettu malli pystyy tämän jälkeen päättelemään asioita myös sellaisesta datasta, jota se ei nähnyt opetuksen aikana, eli yleistämään oppimiaan asioita.

Ohjaamaton oppiminen

Ohjaamattomassa (engl. *unsupervised learning*) oppimisessa tekoälylle esitetään tietoaineisto sellaisenaan. Koulutusvaiheen tarkoitus on antaa tekoälyn omatoimisesti löytää rakennetta annetusta aineistosta. Esimerkiksi, jos tavoitteena on kouluttaa tekoäly, joka kykenee luomaan kuvitteellisia ihmisten kasvokuvia, tyypillinen tapa kouluttaa tällainen malli on esittää tekoälyllä tuhansia kuvia ihmiskasvoista ja antaa tekoälyn oppia mitä sen odotetaan pystyvän tuottamaan. Tällöin tekoälylle ei kerrota kasvojen rakenteesta (missä nenän kuuluu olla tai että silmiä on yleensä kaksi) vaan sen annetaan itse löytää lainalaisuudet ja rakenne aineistosta. Oheisessa kuvassa on hyödynnetty ohjaamatonta oppimista kuvitteellisten eläinkuvien luomiseen [1].

Ohjaamatonta oppimista käytetään myös silloin, kun halutaan tarkastella monimutkaista tietoaineistoa ja pyrkiä löytämään siitä lainalaisuuksia ja säännönmukaisuuksia. Yksinkertaisena esimerkkinä tästä on vaalikonevastauksille tehtävä peruskomponenttianalyysi (esim. vaalikonevastauksien visualisointi nelikenttänä sanomalehdissä), joka yleensä löytää liberaali–konservatiivi- ja vasemmisto–oikeisto-akselit vastauksista ilman, että näistä olisi kerrottu menetelmälle etukäteen. Ohjaamaton oppiminen tarjoaa siis työkaluja sekä monimutkaisten asioiden yksinkertaistamiseen (selittäminen) tai havaintojen generointiin, muokkaukseen tai tulkintaan.

Ohjattu oppiminen

Ohjatussa oppimisessa (engl. *supervised learning*) tekoäly koulutetaan näyttämällä sille esimerkkejä sekä syötteistä (mitä menee sisään) ja toivotuista tuloksista (mitä tulee mallista ulos). Tavoitteena on kouluttaa tekoäly, joka kykenee opetuksen jälkeen yleistämään opittuja asioita ja päättelemään, mitä tuloksia tuntemattomilla syötteillä tulisi saada aikaan. Ohjattu oppiminen on yleisesti käytetty tapa opettaa regressio- ja luokittelumalleja. Regressiomalli pyrkii ennustamaan tai selittämään jatkuvia suureita (esim. lapsen pituus aikuisena, kun vanhempien pituudet tunnetaan). Luokittelumalli pyrkii nimensä mukaan luokittelemaan asioita joko kahteen tai useampaan luokkaan (esim. onko kuvassa kissa vai koira).

Ohjattu oppiminen auttaa tyypillisesti tekoälyä oppimaan sellaisia tehtäviä, joista on kertynyt havaintoaineistoja, joiden lopputulos tiedetään tai joihin on liitetty ihmisen tavoittelema päämäärä. Tällainen tietoaineisto voi olla esimerkiksi kokoelma kuvia esineistä, jotka on nimetty ihmisen toimesta. Näin voidaan kouluttaa malli, joka kykenee tunnistamaan esineitä ja nimeämään ne. Ohjattu oppiminen muodostaa pitkälti monen nykyisen tekoälyjärjestelmän selkärangan.

Vahvistusoppiminen

Vahvistusoppiminen (engl. *reinforcement learning*) sopii käyttökohteeseen, jossa on helppo määritellä päämäärä oppimiselle (esim. pelin voittaminen tai robotin liikkuminen eteenpäin) mutta jota on hankala kuvailla matemaattisessa muodossa tai josta on vaikea kerätä sopiva opetusaineisto. Esimerkkinä vahvistusoppimisesta on robotin opettaminen kävelemään. On hankala määritellä, miten raajojen nivelten pitää käyttäytyä, jotta tasapaino säilyy ja samalla päästään eteenpäin – etenkin epätasaisessa maastossa. Sen sijaan on helppo määritellä, milloin oppiminen on johtanut tavoiteltuun lopputulokseen eli on päästy etenemään.

Vahvistusoppimisessa tekoäly tasapainoilee eksploitaation (jo opittujen tietojen ja taitojen hyödyntäminen) ja eksploraation (uuden tutkimisen, yrittämisen) välillä. Tähän törmäämme me ihmisetkin arjessa: tilaanko ravintolassa vanhan tutun annoksen (eksploitaatio) vai kokeilenko listalta jotain uutta (eksploraatio)? Tässä on nähtävissä analogia ihmislapsen liikkumiseen: alussa lapsi ei osaa hallita raajojaan, mutta *yrityksen ja erehdyksen* kautta oppii kääntymään, konttaamaan ja lopulta kävelemään. Vahvistusoppiminen on nouseva tekoälyn osa-alue, jossa läpimurtoja on tapahtunut aivan viime vuosina. Google DeepMindin AlphaZero-tekoäly [2], joka ensimmäisenä päihitti parhaat ihmispelaajat shakissa, gossa ja shogissa, perustui vahvistusoppimiseen, jossa tekoäly oli harjoitellut pelistrategioita pelaamalla itseään vastaan.

Tekoäly muutosvoimana

Tekoälyt ovat nopeasti siirtyneet akateemisesta tutkimuksesta jokapäiväiseen käyttöömme internetin, mobiililaitteiden ja kuluttajatuotteiden kautta. Oppivien tekoälyjen rivakasta kehitystä on ajanut kolme tekijää:

- Perustutkimus tekoälymenetelmien kehittämiseksi, joka on mahdollistanut sekä monipuolisemmat mallit että tehokkaammat opetusmenetelmät.
- Tarjolla oleva data tekoälymallien opettamiseksi (esim. avoimet kuva- ja tekstiaineistot).
- Kasvanut laskentakapasiteetti sekä päätelaitteiden yleistyminen (esim. älypuhelimet, älysormukset, ym.), jotka tarjoavat ekosysteemit kaupallisen toiminnan mahdollistamiseksi.

Tulevat kehitysaskeleet näiden kaikkien kolmen osa-alueen alla tulevat tukemaan tulevien innovaatioiden ja uusien avausten mahdollistamista tekoälyn alla.

Muutos on ollut ripeää ja monien seikkojen huomioiminen käytännön näkökulmista on seurannut perässä. Koneoppimismalleissa mahdollisia kompastuskohtia voi olla niin opetusaineistossa, mallioletuksissa kuin mallin opettamisessa. Pahimmillaan opetusaineisto sisältää sinne kuulumatonta tai ei-toivottua dataa (esim. vihapuhe tekstiaineistossa). Aineisto ei myöskään välttämättä ole kattavaa (esim. tekstiaineistosta puuttuvat murresanat). Toisaalta mallioletukset saattavat olla vääriä tai ne eivät huomioi sovelluksen erityispiirteitä. Opetus puolestaan saattaa olla tehotonta tai epäonnistua kokonaan. Näiden seikkojen huomioiminen on korostunut viimeaikaisessa tutkimuksessa samoin kuin tekoälyjen selitettävyyden ja tulkittavuuden parantaminen: Mitä tekoäly oppi tai miten se vetää johtopäätöksensä?

Huomioitavaa päättäjille ja kansalaisille

Ihmisen oppiminen on jatkuvaa ja tahatonta. Tekoälyt sen sijaan eivät tyypillisesti opi jatkuvasti, vaan niiden kouluttaminen ja käyttäminen ovat erillisiä prosesseja. Jotta tekoäly oppii uutta, on tekoäly käytännössä opetettava uudelleen ja otettava uudelleen käyttöön. Tämä myös tarkoittaa, että tekoäly ei tyypillisesti opi omista virheistään, vaan jos se erehtyy kerran, tulee se erehtymään jatkossakin samassa tilanteessa, jos uudelleenkoulutusta ei tehdä.

Ihminen kykenee huomioimaan epävarmuutta ja tietyissä rajoissa ilmaisemaan oman ymmärryksensä rajat. Tekoälyille epävarmuuksien hallinta ei tule luonnostaan, vaan ne saattavat olla ylitsevuotavan varmoja asioista, joista eivät tiedä opetusaineiston perusteella mitään (esim. tekoäly, joka on opetettu vain luokittelemaan kissoja ja koiria kuvista, saattaa olla täysin varma, että autoa esittävä kuva on koira). Epävarmuuksien hallinta vaatii erityistä huomioimista mallioletuksissa, opetusvaiheessa ja menetelmien käytössä.

Oppivien tekoälyjen käytettävyyteen vaikuttaa tarjolla oleva opetusaineisto. Tynnyrissä kasvanut tekoäly ei ymmärrä maailmaa. Pienenä kielialueena ja yhteiskuntana, jolla on omat kansalliset erityispiirteensä, meidän tulee tukea ja turvata kotimaisilla erityispiirteillä olevien opetusaineistojen saatavuus. Tämä koskee kielellisiä aineistoja sekä esimerkiksi Suomen luontoa, geeni- ja terveystietoja, tieverkoston erityispiirteitä, elinkeinorakennetta ja lainsäädäntö.

Tekoäly ei itsessään päätä mitä oppii, vaan mukautuu aineiston rakenteeseen ja oppii peilaamaan opetusaineistoaan mallioletustensa antamin reunaehdoin. Tästä syystä on tärkeää, että mallien rakentajat ymmärtävät mitä ovat tekemässä, mikä puolestaan vaatii resursointeja koulutukseen.

Viitteet

[1] Karras, Tero, ym. “Alias-free generative adversarial networks”. Advances in Neural Information Processing Systems (2021).

[2] Silver, David, ym. “A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play.” Science 362.6419 (2018): 1140–1144.

Kainalolaatikko

Havaintoaineistosta oppivat tekoälyt ovat yleisessä käytössä päätöksenteon ja havainnoinnin tukena, data-analytiikkatyökaluina ja tekemässä suosituksia. Koneoppimismenetelmät täydentävät monilla aloilla perinteisempään tilastotieteeseen pohjautuvia analyysimenetelmiä. Ne ovat mahdollistaneet uusia teknisiä avauksia palvelujen ja kuluttajatuotteiden osina. Suomalaisessa teollisuudessa monia näistä menetelmistä käytetään täydentämään jo olemassa olevia järjestelmiä ja tehostamaan toimintaa. Samalla kuitenkin kehitysaskeleet esimerkiksi havainnoinnin, kuten konenäön, parissa mahdollistavat uusia avauksia sekä entistä autonomisempia ja automatisoidumpia ratkaisuja (esim. autonomiset työkoneet, kuljettajaa helpottavat avustimet autoissa, robottivarastot tai tuotantoprosessien automaattinen valvonta). Oppivia tekoälyjä hyödynnetään myös konekääntämisessä (esim. Google-kääntäjä), asiakaspalvelun analysoinnissa (esim. Kansaneläkelaitos), mainosten ja palveluiden kohdentamisessa sekä lääketieteellisissä sovelluksissa.