

Occamin partaveitsi ja oppivat koneet

Petri Myllymäki
Tietojenkäsittelytieteen laitos
Helsingin yliopisto

Kun ihminen halusi oppia rakentamaan lentäviä koneita, mallia haettiin ensin luonnosta lintujen ja muiden lentävien eläinten rakennetta ja toimintaa tarkkailemalla. Kun aerodynamiikan lait opittiin tuntemaan paremmin, havaittiin että luonnon orjallinen matkiminen ei ole paras tapa rakentaa lentäviä koneita, vaan olennaista on ymmärtää ilmiön perustavat lainalaisuudet, ja hyödyntää niitä järkevällä tavalla. Lopputulos onkin melko erilainen kuin esikuvansa: nykyiset jumbojetit eivät räpyttele siipiään, eivätkä pysty laskeutumaan puun oksalle, mutta ne voivat kuskata satoja ihmisiä toiselle puolelle maailmaa muutamassa tunnissa.

Yrittäessämme rakentaa älykkäitä koneita meidän olisi samalla tavalla pystyttävä identifioimaan aerodynamiikan lakeja vastaavat ”älykkyiden luonnonlait” – mitkä ovat älykästä toimintaa karakterisoivat perustavat lainalaisuudet? Ongelmaa on mietitty eri tieteissä pitkään ja hartaasti, mutta ikävä kyllä olemme vielä melko kaukana asian syvällisestä ymmärtämisestä. Yhdestä asiasta useimmat tutkijat ovat kuitenkin yhtä mieltä: adaptiivisuus, kyky oppia uusia asioita on olennainen osa älykkyyttä.

Koneoppiminen (machine learning) on tekoälytutkimuksen osa-alue, joka pyrkii kehittämään sellaisia oppivia koneita jotka kykenevät mukauttamaan toimintaansa tehtyjen havaintojen perusteella. Koneiden oppimiskyky on toistaiseksi hyvin rajallista siinä mielessä että ne pystyvät adaptoitumaan melko rajoitetuissa tilanteissa – ne voivat esimerkiksi oppia tunnistamaan ihmisiä valokuvista tai erottelemaan roskapostit asiallisista viesteistä – ja tällaisetkin koneet on ohjelmoitava ensin käsin siinä mielessä että niihin on sisäänrakennettava oppimista ohjaavat algoritmit. Joudumme kuitenkin jo tällaisten yksinkertaistenkin ongelmien algoritmisessa ratkaisemisessa ottamaan kantaa siihen mitä oppiminen on, ja kuinka karakterisoida sitä matemaattisesti.

Oppiminen ei selvästikään ole pelkkää muistamista – jos näin olisi, niin koneet olisivat tässä suhteessa paljon ”älykkäämpiä” kuin ihminen, voivathan ne toistaa täydellisesti kaiken muistivarastoihinsa talletetun sisällön. Sen sijaan oppimiseen tuntuu liittyvän havaintodatan taustalla piilevien lainalaisuuksien ymmärtäminen, kyky erottaa ja tiivistää datan sisältämä olennainen informaatio. Tällaisten taustailmiön lainalaisuuksia kuvaavien mallien järkevyyttä voidaan testata tutkimalla kuinka hyvin ne kykenevät yleistämään havaitsemaansa ja selittämään annetun datan ulkopuolisia havaintoja, esimerkiksi kuinka hyvin ne voivat ennustaa tulevia tapahtumia annettuna tämänhetkiset havaintomme.

Erilaisten mallien ennustamiskykyä tutkittaessa on tieteen alkuajoista saakka toistettu lukemattomia kertoja seuraava havainto: kahdesta havaintoihin yhtä hyvin sopivasta mallista (hypoteesista), yksinkertaisempi tuottaa yleensä luotettavampia ennustuksia. Tämä Occamin partaveitsekseksi nimetty empiirinen periaate näyttäisi siis muodostavan yhden tärkeän ”oppimisen luonnonlain”. Periaate sinänsä on kuitenkin hyvin epämääräinen: mitä tarkoittaa se, että hypoteesit sopivat havaintoihin ”yhtä hyvin”, ja onko mallien yksinkertaisuus

hyvin määriteltävissä? Jotta voisimme rakentaa oppivan koneen jossa Occamin partaveitsi on sisäänrakennettuna algoritmina, on meidän ensin pystyttävä formalisoimaan tämä periaate matemaattisesti.

Yhden matemaattisesti konsistentin teoreettisen kehikon (kone)oppimiselle tarjoaa informaatioteoria: tiedon tiivistäminenhan perustuu nimenomaan olennaisen informaation tunnistamiseen ja datassa usein esiintyvien hahmojen ja lainalaisuuksien hyödyntämiseen, ja eri mallien voi siis ajatella sopivan havaintoihin sitä paremmin, mitä paremmin ne tarjoavat mahdollisuuden datan tiivistämiseen. Tiivistetyn datan purkaminen edellyttää toisaalta myös itse tiivistävän mallin koodaamista, ja Jorma Rissanen kehittämän lyhimmän kuvaus-pituuden (Minimum Description Length, MDL) periaatteen mukaan paras malli havainnolle on sellainen, joka minimoi tiivistetyn datan ja käytetyn tiivistäjän (mallin) yhteiskoodinpituuksien. MDL-periaatteen voidaan katsoa nyt formalisoivan Occamin partaveitsi -periaatteen: monimutkainen malli tiivistää datan tehokkaasti, mutta itse malli vaatii monimutkaisen koodin, kun taas yksinkertainen malli voidaan kuvata tehokkaasti, mutta tällöin datan vaatima koodinpituus pitenee. MDL-optimaalinen malli etsii tasapainon toisaalta mallin kompleksisuuden ja toisaalta mallin ja havaintojen yhteensopivuuden välillä. Periaate tarjoaa myös selityksen sille miksi Occamin partaveitsi toimii käytännössä: koska vain olennaisimmat (useimmin esiintyvät) lainalaisuudet kannattaa mallintaa (tarkasti), MDL-optimaalinen malli toimii hyvin nimenomaan kaikkien todennäköisimpien havaintojen tapauksessa, ja ennustaa siksi tulevia havaintoja (yleensä) hyvin.

Koodinpituuksien ja todennäköisyyksien välillä on suora, perustavanlaatuinen yhteys, ja informaatioteoria ja todennäköisyyslaskennan teoria ovatkin syvästi linkitettyjä toisiinsa. Intuitiivisesti tämä ei liene yllättävää, koska onhan esimerkiksi selvää että dataa pakatessa lyhin koodinpituus kannattaa antaa usein toistuville hahmoille, eli datasekvensseille joiden todennäköisyys on suurin. Käytännössä informaatioteoreettinen koneoppiminen onkin probabilistista mallintamista, eli datassa esiintyvien ilmiöiden tunnistamista ja niiden todennäköisyyksien arviointia. Koneoppiminen voidaan toki formalisoida myös muilla (ei-probabilistisilla) tavoilla, mutta informaatioteoreettinen (probabilistinen) lähestymistapa tarjoaa tälle tutkimusalueelle universaalin ja varsin elegantin teoreettisen kehikon.

Työ on kuitenkin vasta alussa: koneet osaavat jo etsiä hyviä probabilistisia malleja tilanteissa joissa ihminen osallistuu vahvasti prosessiin mm. esikäsittelemällä dataa ja valitsemalla ongelma-alueelle sopivia malliperheitä, ja tämä työ on tuottanut monia hyödyllisiä sovelluksia, kuten adaptiiviset roskapostisuodattimet tai puheentunnistusmenetelmät. Olemme kuitenkin vielä melko kaukana tilanteesta jossa koneet pystyisivät täysin itsenäisesti ja automaattisesti havainnoimaan, keräämään ja prosessoimaan heterogeenistä, strukturoimatonta dataa, ja kehittämään ja valitsemaan itse sopivat datan esitysmuodot, piirteet, käsitteet ja malliperheet, ja soveltaisivat Occamin partaveitsi -periaatetta ihmisten tapaan älykkäiden ja innovatiivisten hypoteesien ja johtopäätösten tekemisessä. Saattaa toisaalta myös hyvin olla niin että näin syntyvä älykkyys tulee muistuttamaan esikuvana käytettävää ihmisälyä melko vähän, yhtä vähän kuin lentokoneet, helikopterit, avaruussukkulat ja muut lentävät koneet muistuttavat lintuja, perhosia tai kimalaisia.