Blogi

Tekoälyn eettiset valinnat

Digitalisaatio | 13.11.2018 | Pertti Ukkonen

KIRJOITTAJA



Pertti Ukkonen
Vanhempi asiantuntija

Voiko tekoäly oppia tekemään eettisiä valintoja?

Yhdenvertaisuus- ja tasa-arvovaltuutetun esille ottama tapaus, jossa rahoituslaitos oli evännyt lainahakijan toivoman lainan, sai viime keväänä jonkin verran julkisuutta. Taustalla oli tekoäly, joka oli käyttänyt asiakkaan ominaisuuksia päätöksen tekemiseen. Voiko tekoäly oppia tekemään eettisiä valintoja?

Asiakkaan arviointi kuuluu lainan normaaliin myöntämisprosessiin. Siinä pyritään selvittämään asiakkaan kykyä ja todennäköisyyksiä maksaa laina korkoineen takaisin. Ongelmaksi muodostui se, että kielteinen lainapäätös tehtiin puhtaasti automatiikalla ja perusteltiin ainoastaan analysointisovelluksen antamalla tuloksella, siis profilointina. Osa sovelluksen käyttämistä tiedoista oli sellaisia, jotka ovat Suomen lainsäädännön mukaan syrjiviä ja joita ei siksi saa käyttää asiakkaiden erottelun perusteena muuten kuin laissa mainituin poikkeuksin.

Miten tämä sitten eroaa normaalista luottopäätöksestä? Kyllähän siinäkin pankin tai muun lainanantajan asiantuntija haluaa kuulla ja nähdä mahdollisesta asiakkaasta tosiasioita, joiden perusteella hän muodostaa arvion hakijan maksukyvystä ja ehkä jopa maksuhalustakin.

Ongelmallisia kohtia oli kaksi. Toinen liittyi siihen, että tasa-arvoisen kohtelun eettiseen ideaaliin perustuva syrjimättömyyden vaatimus oli unohtunut opettaa arviointialgoritmille. Tekoäly ei siis "tiennyt", että Suomessa palvelun tarjoaja ei saa asettaa palvelun ostajia keskenään eriarvoiseen asemaan esimerkiksi sukupuolen tai kielellisen tai etnisen taustan perusteella.

Toinen ongelmallinen kohta puolestaan oli se, että luoton epäämisestä "päättänyt" algoritmi oli kone, eikä sen kanssa ollut mahdollista keskustella tarkemmin päätöksestä ja sen perusteista. Näin ollen ei ollut mahdollista tuoda esille sellaisia seikkoja, jotka olisivat mahdollisesti lisänneet

takaisinmaksun todennäköisyyttä kyseisellä lainan hakijalla.

Vastaavalla tavalla lääketieteellinen diagnostiikka on eettistä päätöksentekoa sivuava alue, jolla tekoälyn odotetaan parantavan työn tuottavuutta. Myös siinä törmätään ainakin maallikon näkökulmasta helposti tilanteeseen, jossa oikeakaan diagnoosi ei pelkkänä diagnoosina ja hoito-ohjeina riitä, vaan tarvitaan mukaan ihminen ja inhimillinen ajattelu. Tekoälysovellukselta voi toki kysyä tarkentavia kysymyksiä, mutta miten terveydestään huolestunut ihminen voi luottaa pelkän koneen tekemään päätökseen? Mikä varmistaa, että tekoäly pystyisi näkemään asiat kontekstissa ja tunnistamaan kysymykset, joihin sille ei ole opetettu vastausta?

Käytännön esimerkkien valossa päätyykin nopeasti johtopäätökseen, että keskeinen ihmiselle kuuluva tehtävä, jota ei voi luovuttaa koneille, on kyky ottaa vastuuta ja kantaa vastuuta. Tekoäly on ihmistä parempi tunnistamaan yksityiskohtia suuresta tietomassasta, ja se on nopeampi laskemaan lukemattomia vaihtoehtoja, mutta siltä puuttuu kyky ihmetellä, miksi tulokset ovat inhimillisen intuition vastaisia. Tekoäly osaa ratkaista tehokkaimman reitin ja laskea tuottavimman vaihtoehdon, mutta oikeaa se ei ainakaan vielä osaa erottaa väärästä eikä hyvää pahasta.

Tekoälyyn sopii sama viisaus kuin erääseen toiseen suureen, ihmiskunnan historian muuttaneeseen innovaatioon: myös tekoäly on hyvä renki vaikkei isäntänä kummoinen. Siksi jokaisen kannattaa oppia perusasiat ja tutustua esimerkiksi Helsingin yliopiston ja Reaktor Group Oy:n laatimaan tekoälyn verkkokoulutuspakettiin, joka on vastikään julkaistu myös suomeksi.

Asiasanat

etiikka, luotot, tekoäly