

Joonas Tamminen

LL, erikoistuva lääkäri
Tampereen yliopisto ja Rovaniemen kaupunki
PSHP, Tays, EKA, Ensihoitokeskus
joonas.i.tamminen@tuni.fi

**Antti Kallonen**

FM, diplomi-insinööri
Tampereen yliopisto
antti.kallonen@tuni.fi

**Sanna Hoppu**

Dos., ylilääkäri
PSHP, Tays, EKA, Ensihoitokeskus
sanna.hoppu@pshp.fi



Tekoäly päätöksenteon työkaluna ensihoitopalvelussa

Sähköinen ensihoitokertomus avaa ovet hoitopäätöksiä ohjaavan tekoälyn rakentamiseksi. Paikalliseen ensihoitojärjestelmään räätälöity tekoäly auttaisi kliinikkoa tunnistamaan korkean ja keskiuuren riskin potilaat ilman, että päivystyspisteet ruuhkautuisivat.

Lähestyvän elottomuuden merkit näkyvät potilaan peruselintoimintojen häiriöinä tunteja ennen sydänpysähdystä tai kuolemaa (1,2). Sairaalapotilaiden uhkaavaa elottomuutta on ennustettu aikaisen varoituksen pisteytysjärjestelmällä, joista suomalaisille tutuin lienee NEWS (National Early Warning Score) (3). On osoitettu, että NEWS soveltuu käytettäväksi myös sairaalan ulkopuolella (4). NEWS kykenee ennustamaan kohtalaisesti potilaiden kuolleisuutta jopa 30 päivän aikajänteellä, kun NEWS-pisteet ovat korkeat (yli 6 pistettä) tai hyvin matalat (0 pistettä) (5).

NEWS-pisteytys on kuuteen fysiologiseen mittaukseen ja lisähapen tarpeeseen pohjautu-

va riskinarvion apuväline. Se täydentää, muttei korvaa potilaan huolellista kliinistä tutkimista, ja NEWS-pisteytyksellä on yleisesti tunnetut rajoituksensa. NEWS-pisteytystä ei esimerkiksi tule käyttää lasten ja raskaana olevien eikä selkäydinvamma potilaiden riskinarvion tai epäiltäessä akuuttia koronaarioireyhtymää tai suolistoverenvuotoa. Lisäksi kroonista ventilaatiovajausta sairastaville potilaille tulisi käyttää muokattuja saturaatioarvoja, mikä on huomioitu päivitetystä NEWS 2 -pisteytyksessä (6). Pediatriisille sekä obstetrisille potilaille on ehdotettu käytettäväksi omia aikaisen varoituksen pisteytyksiä (7,8). NEWS-pisteytystä tai muita aikaisen varoituksen pisteytysjärjestelmää lieneekin mahdotonta räätälöidä koskemaan yleisesti kaikkia ensihoi-

>>

don kohtaamia potilaita. Voisiko kuitenkin tekoäly taipua yleispäteväksi riskinarvion apuvälineeksi klinikon työkalupakkiin?

Tekoälyn anatomia

Tekoälyllä (artificial intelligence) tarkoitetaan tavallisesti jonkinlaista tietokonetta tai robottia, joka kykenee matkimaan ihmisen kognitiivista kapasiteettia vaativia toimintoja kuten ongelmanratkaisua ja oppimista. Tämän lavean kattotermin alle voidaan sijoittaa koneoppiminen (machine learning). Koneoppimisessa matemaattinen malli ja algoritmi kykenee oppimaan ja parantamaan suorituskyykyään. Kyky oppia eli löytää yhteyksiä, hahmoja ja rakenteita niille syötetyistä esimerkeistä erottaa koneoppivat mallit perinteisesti ohjelmoiduista koodinpätkistä. Koneoppivat mallit luovat itse oman ohjelmakoodinsa. Niille esitetään syöte, josta ne oppivat optimaalisen vastauksen esitettyyn kysymykseen tai ongelmaan. Toisin sanoen koneoppivat mallit ovat dataohjautuvia ja kykenevät tunnistamaan annetuista esimerkeistä entuudestaan tuntemattomia yhteyksiä ja riippuvuuksia. Tämä ominaisuus tekee niistä erityisen hyviä mallintamaan monitahoisia fysiologisia ilmiöitä, joita ei pystyttäisi tyhjentävästi kuvaamaan perinteisillä tilastollisilla menetelmillä.

Tekoälyyn liittyvät sovellukset ovat pulpahtaneet pinnalle edeltävän vuosikymmenen aikana, koska tietokoneiden roimasti kohentunut laskentatehokapasiteetti on mahdollistanut mutkikkaan matemaattisen mallintamisen mielekkäässä ajassa. Tekoäly voidaan valjastaa suoratoistopalveluiden katsojakohtaisten suositusten laatimisen ohella lääketieteellisiin tarkoituksiin erityisesti diagnostisilla aloilla sekä tehohoitolääketieteessä. Hiljattain julkaistussa Duodecimin teemanumerossa käsiteltiin tekoälyn soveltamista radiologiassa, patologiassa, silmänpohjakuvien tulkitsemisessä sekä vastasyntyneiden teho-osastolla (9). Viime vuonna New England Journal of Medicine -lehdessä julkaistiin aiheesta kattava katsausartikkeli (10). Artikkelissa todettiin, että tuoreisiin tekoälytutkimustuloksiin voi suhtautua varovaisen

optimistisesti; toivottavasti jo lähitulevaisuudessa tekoäly avustaa hoitopäätösten tekemistä.

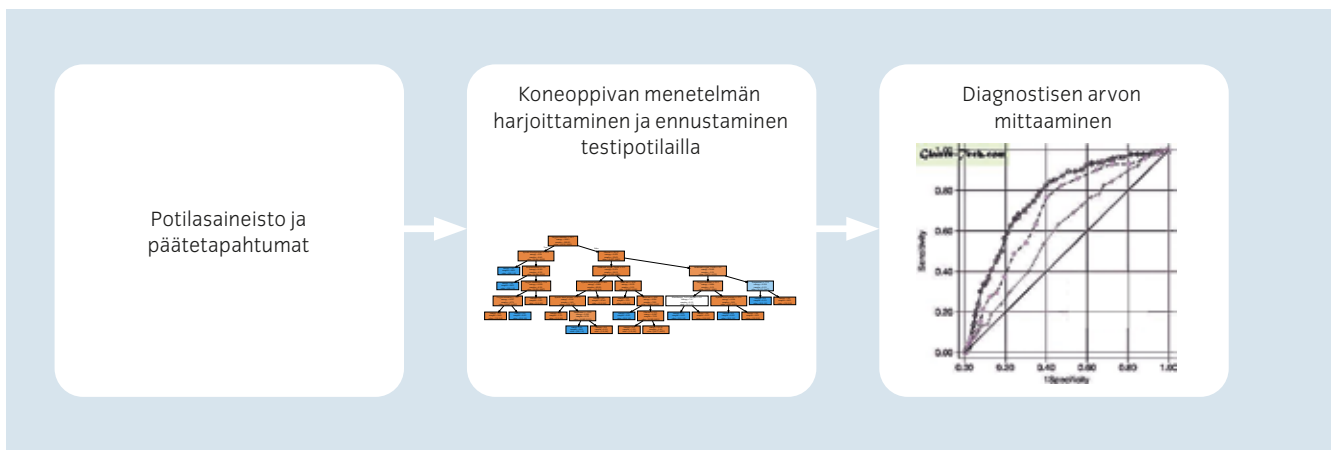
Koneoppivan mallin oppiminen voi olla valvottua (supervised) tai valvomattomaa (unsupervised). Valvotussa oppimisessa mallien koulutusvaiheessa algoritmille voidaan esittää tuhansia, miljoonia tai jopa miljardeja esimerkkejä kerättyjen havaintojen ominaisuuksista, ja sille kerrotaan kunkin havainnon mallivastaus (kuva 1). Algoritmille voidaan esimerkiksi esittää potilastapausten demografisia tietoja, pitkäaikaissairauksia, ajankohtaisia oireita, fysiologisia mittauksia ja laboratoriovastauksia. Mallille kerrottu oikea vastaus voi olla jokin päätetapahtuma kuten potilaan äkillinen menehtyminen. Koulutusvaiheen jälkeen valvotun koneoppivan mallin suorituskyykyä arvioidaan mallille esitetyillä uusilla havainnoilla, ja ennustusvoima voidaan raportoida ROC-käyrän alle jäävänä pinta-alana (AUC). Valvotut koneoppivat mallit siis antavat räätälöidyn ennusteen tai vastauksen niille esitettyyn kysymykseen. Vastaavasti valvomattomat koneoppivat mallit oppivat yrityksen ja erehdyksen kautta. Valvomattomat koneoppivat mallit koulutetaan simuloidussa ympäristössä ja ne saavat positiivista vahvistusta oikein tehdyistä suoritteista ja negatiivista vahvistusta vääristä vastauksista. Valvottoman koneoppiva malli voisi esimerkiksi olla tekoäly, joka suunnittelisi potilaille yksilöllisiä hoitopolkuja ottamalla huomioon useita hoidon laadullisia mittareita.

Koneoppimismalli voi olla musta laatikko (black box). Tällä viitataan mallien monimutkaiseen ja joskus epäintuitiiviseen tapaan yhdistellä

eri muuttujia: vertauskuvaunnollisesti tekoälyn mustaan laatikkoon ahdetaan dataa, hämärässä laatikossa algoritmi tarkastelee miljoonia yksittäisiä heikkoja signaaleja ja lopulta laatikosta putkahtaa ulos lopputulos. Vaikka mallin sisältämän matematiikan kirjoittaisi auki paperille, yksittäiselle kertoimelle ei välttämättä ole mielekästä

Toivottavasti jo lähitulevaisuudessa tekoäly avustaa hoitopäätösten tekemistä.

kliinistä merkitystä toisin kuin vaikkapa regressiomallien kertoimille. Tämä onkin tehokkaiden koneoppimismenetelmien kuten hermoverkkojen kääntöpuoli. Mitä monimutkaisempia ilmiöitä ja yhteyksiä malli kykenee oppimaan, sitä vaikeam-



Kuva 1. Esimerkki valvotusta oppimisesta.

pi on täydellisesti ymmärtää, mitä malli on oppinut sille syötetyistä esimerkeistä. Mustan laatikon vahvuus piilee kuitenkin siinä, että harjoitetun koneoppimismallin ennuste on aina deterministinen eli sen suoriutuminen voidaan varmistaa erikseen kerätyllä edustavalla potilasaineistolla. Kun algoritmi on ulkoisesti validoitu, myös skeptisimmänkin klinikon kannattaa pysähtyä pohtimaan ennusteen merkitystä, vaikkei siihen ensisilmäyksellä uskoisikaan.

Kone päihittää kynän ja paperin

NEWS-pisteytyksen kiistaton vahvuus on sen yksinkertaisuus. NEWS-pisteiden laskeminen onnistuu helposti ja pisteiden tulkinta on suoraviivaista. NEWS-pisteytyksessä ja muut aikaisen varoituksen pisteytysjärjestelmät perustuvat kuitenkin perinteiseen tilastotieteeseen eli lineaariseen logistiseen regressioon. Tässä piilee muutama itse tilastolliseen menetelmään liittyvä sudenkuoppa, vaikka NEWS-pisteytystä käytettäisiin oikeaoppisesti siihen soveltuvien potilaiden riskinarvioon.

Jokainen seitsemästä NEWS-muuttujista tulisi arvioida, jotta kokonaispistemäärä voitaisiin laskea ja tehdä luotettavia johtopäätöksiä. Käytännössä kaikkien NEWS-muuttujien mittaaminen ei ole välttämätöntä jokaisen ensihoitopotila-

aan kohdalla. Nilkkansa nyrjäyttäneelle ei ole tarpeen arvioida NEWS-pisteitä. Toisaalta kyseessä voi olla hätätilapotilas, jolta ensihoitajat eivät yksinkertaisesti ehdi mittaamaan kaikkia NEWS-muuttujia, ja kokonaispisteet jäävät valheellisen mataliksi. Puuttuvat arvot muodostavat siis merkittävän rajoitteen, mutta koneoppiville menetelmille tämä ei koidu ongelmaksi, ja mallit voidaan kouluttaa ”reikäisellä” aineistolla. Koneoppimalli pystyy siis antamaan luotettavan ennusteen, vaikka potilaan hengitysfrekvenssi ja lämpötila olisivat vielä mittaamatta. On myös mahdollista, että joissakin hätätilanteissa tietyt mittaukset jäävät systemaattisesti tekemättä, ja koneoppiva malli pystyisi tunnistamaan puuttuvien mittausten sormenjäljen ja hyödyntämään tätä informaatiota ennusteessaan.

Puuttuvien arvojen lisäksi lineaarinen logistinen regressio arvioi jokaisen muuttujan erikseen eikä huomioi mahdollisia peruselintoimintojen välisiä interaktioita. NEWS-pisteytyksessä hengitysfrekvenssin normaaliarvot on nauhattu lukemiin 12–20/min riippumatta esimerkiksi syketaajuuden muutoksista. Modernit koneoppivat menetelmät kykenevät

kiertämään tämän ongelman ja voivat löytää epälineaarisia yhteyksiä muuttujien välille. Koneoppivat menetelmät voivat myös käsitellä valtavan määrän muuttujia, kunhan opetusaineistossa on riittävästi havaintoja eli malliesimerkkejä. Eräas-

Koneoppivat mallit toimivat myös ”reikäisellä” aineistolla.

>>

sä tekoälytutkimuksessa arvioitiin koneoppivan menetelmän kykyä ennustaa päivystyspotilaiden osastohoitoa (11). Tutkimusaineisto koostui yli puolesta miljoonasta potilasta ja malliin sisällytettiin 972 muuttujaa.

Ruotsalaisten viime keväällä julkaisemassa tutkimuksessa verrattiin perinteistä NEWS-pisteytystä XGBoost-algoritmiin ensihoidon kohtaamien potilaiden kahden vuorokauden kuolleisuuden ennustamisessa (12). Tutkimusaineisto koostui 38 203 potilaasta, jotka kuljetettiin ambulanssilla päivystykseen ja joilla oli enintään kaksi puuttuvaa NEWS-muuttujan mittausta. Kolmekymmentä muuttujaa huomioiva algoritmi kykeni ennustamaan kuolleisuutta paremmin kuin tavallinen NEWS-pisteytys (XGBoost AUC, 0.89 [0.87–0.91] vs NEWS AUC, 0.85 [0.83–0.87]). Potilaiden valikoitumisen ohella tutkimuksen toinen merkittävä heikkous liittyy koneoppivan menetelmän valintaan. Ei ole ollenkaan selvää, mikä koneoppiva menetelmä soveltuisi parhaiten ensihoitopotilaiden kuolleisuuden ennustamiseen. Osastohoitopotilaisiin keskittyneessä monikeskustutkimuksessa päätöspuihin perustuva random forest -algoritmi osoittautui ennustevoimaltaan parhaimmaksi lukuisista testatuista koneoppivista menetelmistä (13).

Algoritmi opastaa keltanokkaa

Tekoälyn antamien ennusteiden integraatio nykyiseen ensihoitopalvelujärjestelmään olisi luonnollinen jatke sähköisille ensihoitokertomuksille. Enemmistössä sairaanhoitopiirejä osa NEWS-muuttujien mittauksista tallentuu automaattisesti sähköiseen potilasjärjestelmään, kun mittausanturit on saatu kytkettyä paikoilleen. Ensihoitajien täytyy kuitenkin lisätä ensihoitokertomukseen potilaan hengitysfrekvenssi, lämpötila ja tajunnantaso, jolloin nämä arvot jäävät harmillisen usein kirjaamatta kentällä (14). Ensihoitajien työtaakkaa ei ole kuitenkaan tarpeen lisätä kaikkien NEWS-muuttujien rutiininomaisella mittaamisella, koska koneoppivat mallit toimivat myös ”reikäisellä” aineistolla. NEWS-muuttujien lisäksi ensihoitokertomukseen tallentuu automaattisesti lukuisia

muita muuttujia, jotka voitaisiin sisällyttää halutun päätetapahtuman ennustamiseen.

Hienostuneimpien ja kehittyneimpien algoritmien numeeriset ennusteet potilaan kuolemanriskistä eivät korvaisi kliinistä päätöksentekoa. Ennusteet kuitenkin tarjoaisivat erinomaisen apuvälineen vähän työkokemusta omaaville ensihoitajille tai lääkäreille riskinarvioon ja epävarmuuden hallintaan. Ennuste voisi auttaa tunnistamaan korkean ja matalan riskin potilaat keskisuuren riskin potilaiden keskuudesta ja siten auttaa potilaan siirtoa asianmukaiseen terveydenhuollon yksikköön. Parhaimmillaan tekoälyn avulla välttyttäisiin päivystyspisteiden ruuhkautumiselta vaarantamatta potilasturvallisuutta. Yksi käytännön sovellus voisikin olla ensihoidon puhelinkonsultaatiot kohteeseen jättämisestä, jolloin konsultoitava lääkäri saattaa olla vastikään lisenssiaatin tutkinnon suorittanut terveystieteiden lääkäri.

Puheensorinaa ja kuolonkorinaa

Ensihoitopotilaan riskinarvio alkaa kontaktista hätäkeskukseen. Hätäpuhelu sisältää monenlaista mielenkiintoista informaatiota, jota hyödyntämällä voitaisiin täydentää hätäkeskuspäivystäjän tekemää riskinarviota. Riskinarvio voitaisiin muodostaa muutamalla eri tavalla kuten käyttämällä esimerkiksi puheentunnistusta. Puheentunnistukseen erikoistunut tekoäly litteroisi puheensorinaa tekstiksi, jonka pohjalta toinen tekoäly antaisi reaaliaikaisen ennusteen sydänpysähdyksen mahdollisuudesta. Vaihtoehtoisesti voitaisiin käyttää vain yhtä koneoppivaa mallia, joka olisi koulutettu tulkitsemaan hätäpuhelunauhoituksen jalostamatonta äänisignaalia. Raakadatahan sisältäisi kaiken ympäristön kohinan ja algoritmi pystyisi kenties tunnistamaan sydänpysähdyspotilaan agonaalisen hengityksen taustamelusta.

Algoritmi pystyisi tunnistamaan sydänpysähdyspotilaan agonaalisen hengityksen taustamelusta.

Idea hätäpuheluiden litteroinnista ei ole uusi. Tanskalaiset ovat kehittäneet hätäkeskuspäivystäjälle apuvälineen sydänpysähdyksen tunnistamiseksi (15). Lupavalta vaikuttava Corti AI on koulutettua hätäkeskuspäivystäjää herkempi (84,1 % vs 72,5 %), mutta se häviää tarkkuudessa ihmiselle (98,8

vs 97,3 %). Yleisradion viime kesänä aloittaman Lahjoita puhetta -kampanjan tavoitteena on kerätä 10 000 tuntia suomenkielistä puhetta puheentunnistukseen kykenevän tekoälyn kehittämiseksi. Ehkäpä lähitulevaisuudessa meidänkin hätäkeskuspäivystäjämme saavat käyttöönsä oman suomenkielisen tekoälyn.

Tulevaisuuden työkaveri

Tekoälyn tuleminen osaksi potilaan ensihoitoketjua on vääjäämätöntä ja samalla luontaista jatkumoa sähköistyville sairauskertomusten tallentamiselle. Ei ole ollenkaan yllättävää, että tekoäly kykenee parempaan potilasluokitteluun kuin perinteiset aikaisen varoituksen järjestelmät niin osastolla kuin kentällä. Koneoppivan mallin opetusaineisto voi hyvinkin sisältää enemmän potilaskertomuksia kuin yksittäinen lääkäri kohtaa koko työuransa aikana. Tekoäly ei kuitenkaan ole vielä korvaamassa ihmistä, jonka tehtävänä on tulkita ennusteita yksilöllisesti kunkin potilaan kohdalla. Mustista laatikoista löytyy tässä vaiheessa vain jotain kättä pidempää. ■

Viitteet

1. Schein RMH, Hazday N, Pena M, Ruben BH, Sprung CL. Clinical antecedents to in-hospital cardiopulmonary arrest. *Chest*. 1990;98:1388–92.
2. Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL, et al. Duration of life-threatening antecedents prior to intensive care admission. *Intensive Care Med*. 2002;28:1629–34.
3. Royal College of Physicians. National Early Warning Score (NEWS) - Standardising the assessment of acute-illness severity in the NHS. 2012.
4. Silcock DJ, Corfield AR, Gowens PA, Rooney KD. Validation of the National Early Warning Score in the prehospital setting. *Resuscitation*. 2015;89:31–5.
5. Patel R, Nugawela MD, Edwards HB, Richards A, Le Roux H, Pullyblank A, et al. Can early warning scores identify deteriorating patients in pre-hospital settings? A systematic review. *Resuscitation*. 2018;132:101–11.
6. Royal College of Physicians. National Early Warning Score (NEWS) 2 - Standardising the assessment of acute-illness severity in the NHS. 2017.
7. Akre M, Finkelstein M, Erickson M, Liu M, Vanderbilt L, Billman G. Sensitivity of the pediatric early warning score to identify patient deterioration. *Pediatrics*. 2010;125:e763.
8. Singh S, McGlennan A, England A, Simons R. A validation study of the CEMACH recommended modified early obstetric warning system (MEOWS). *Anaesthesia*. 2012;67:12–8.
9. Mirtti T. Tekoälyä ja älytekoja. 2020;:1945–7.
10. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380:1347–58.
11. Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. *PLoS One*. 2018;13:1–13.
12. Spangler D, Hermansson T, Smekal D, Blomberg H. A validation of machine learning-based risk scores in the prehospital setting. *PLoS One*. 2019;14:1–18.
13. Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards. *Crit Care Med*. 2016;44:368–74.
14. Pirneskoski J, Kuisma M, Olkkola KT, Nurmi J. Prehospital National Early Warning Score predicts early mortality. *Acta Anaesthesiol Scand*. 2019;63:1–8.
15. Blomberg SN, Folke F, Ersbøll AK, Christensen HC, Torp-Pedersen C, Sayre MR, et al. Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. *Resuscitation*. 2019;138:322–9.