

NEUROVERKKO, FUNKTION APPROKSIMAATTORI

Tästä artikkelista

Neuroverkko on tekoälyn toiminnan perusta, eräänlainen laskentakehikko, joka laskee tekoälylle annetusta syötetiedosta tekoälyn tuottamat tulostiedot. Tämän artikkelin tavoite on antaa lukijalle ymmärrys siitä, että neuroverkon toiminta perustuu sen kykyyn approksimoida mielivaltaisen monimutkaisia epäsäännöllisiä funktioita tai toisin sanoen, mallintaa monimutkaisia tilastollisia riippuvuuksia.

Jos neuroverkot eivät ole lukijalle mitenkään entuudestaan tuttuja, olisi hyödyksi tutustua aiheeseen ensin vaikkapa perehtymällä kalvosarjaani ”[Mitä tekoäly on ja mitä se ei ole](#)”. Linkki kalvosarjaani ja muuhun tekoälymateriaaliini löytyy myös tekoälyisivultani <https://www.einouikkanen.fi/AI/>.

Neuroverkon toiminta funktion approksimoijana esitetään analogialla lineaarisen regressioanalyysin ja neuroverkkolaskennan kesken.

Artikkelissa esiintyvät matemaattiset käsitteet sisältyvät lukion matematiikan oppimäärään, mutta asian periaatteelliseen ymmärtämiseen riittää myös peruskoulun matematiikan oppimäärä.

Artikkelin on kirjoittanut Eino Uikkanen, joka vastaa mieluusti artikkelia koskeviin kysymyksiin osoitteessa eino.uikkanen@iki.fi.

Sisällysluettelo

| | |
|---|---|
| Funktio | 2 |
| Epäsäännöllinen funktio..... | 2 |
| Funktion approksimointi | 3 |
| Neuroverkko funktion approksimoijana | 4 |
| Vertailu lineaarinen regressioanalyysi vs. neuroverkko | 5 |
| Jos neuroverkko on vain tilastomatematiikkaa, miten selittyvät tekoälyn huikeat tulokset? | 6 |
| Yhteenveto | 7 |

NEUROVERKKO – FUNKTION APPROKSIMAATTORI

Funktio

Avain keinotekoisien neuroverkkojen toiminnan ymmärtämiseen on funktiokäsitteen ymmärtäminen. Funktio eli kuvaus kertoo asioiden välisistä riippuvuussuhteista. Funktio voidaan esittää sääntönä, joka liittyy funktion määrittelyjoukon (=lähtötiedot) jokaisen alkion täsmälleen yhteen maalijoukon (=tulostiedot) alkioon. Tällöin sanotaan, että funktion arvo määrittelyjoukon alkion arvolla eli argumentin arvolla (esim. x) on vastaava maalijoukon alkion arvo (esim. y). Usein tätä kuvataan merkinnällä $y = f(x)$, jossa x on funktion argumentti ja y on funktion arvo.

Esimerkkejä funktioista:

- Neliön pinta-ala (a) on neliön sivun pituuden (s) funktio kaavalla $a = s^2$ tai $y = f(x) = x^2$
- Tasaisella kiihtyvyydellä (a) kiihtyvän kappaleen kulkema matka (s) hetkellä (t) on lähtönopeuden (u), kiihtyvyyden (a) ja ajan (t) funktio kaavalla $s = ut + \frac{1}{2}at^2$ tai $y = f(u, a, t) = ut + \frac{1}{2}at^2$

Epäsäännöllinen funktio

Funktio voi noudattaa tiettyä sääntöä tai matemaattista kaavaa kuten yllä olevissa esimerkeissä, jolloin funktion argumentti määrittää säännön tai kaavan avulla täsmällisesti funktion arvon. Funktio voi olla myös epäsäännöllinen, jolloin taustalla voi olla esimerkiksi tilastollinen riippuvuus. Silloin funktiota ei voida kuvata millään säännöllä tai matemaattisella kaavalla, vaan ainoastaan luettelemalla joukko lähtötietoja (funktion argumentit) ja vastaavia tulostietoja (funktion arvot). Jos esimerkiksi kerätään päivittäiset sademäärät tietyltä ajalta, määrittää päivämäärä yksikäsitteisesti sademäärän eli sademäärä on päivämäärän funktio.

Muita esimerkkejä epäsäännöllisistä funktioista:

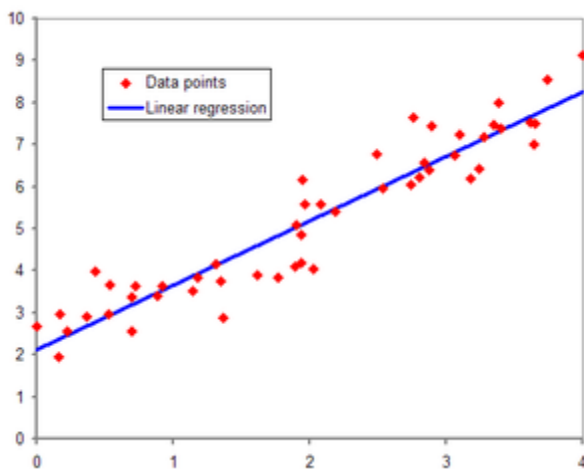
- Päivän keskilämpötila tietyllä alueella on päivämäärän funktio.
- Vehnän hehtaarisato (t/ha) on jollakin tarkkuudella kesän keskilämpötilan, sademäärän sekä maan muokkauksen, lannoituksen ja tuholaisten torjunnan funktio.
- Valitaan joukko äänitteitä linnun laulusta ja liitetään niihin linnun tiedetty laji. Tällöin laji on äänitteiden funktio.
- Valitaan joukko keuhkokuvia ja liitetään niihin kuvan perusteella tehty diagnoosi. Tällöin diagnoosi on keuhkokuvien funktio.

Funktion approksimointi

Koska epäsäännöllistä funktiota ei voida ilmaista yksikäsitteisellä säännöllä tai matemaattisella kaavalla, vaan esimerkiksi taulukkona, on sen takana olevan ilmiön ymmärtäminen ja hallinta on vaikeaa. Tätä voidaan auttaa etsimällä matemaattisena kaavana ilmaistavissa oleva funktio, joka antaa kuitenkin likimäärin oikeita arvoja riittävällä tarkkuudella. Tätä kutsutaan funktion approksimoimiseksi. Approksimoitavasta funktiosta riippuu, kuinka monimutkainen approksimaatio tarvitaan.

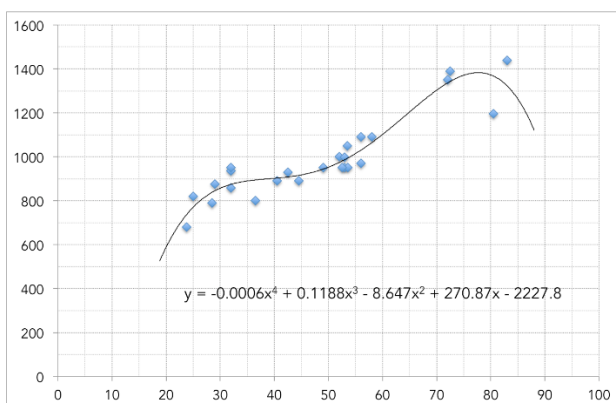
Yksinkertaisessa riippuvuudessa saattaa riittää, että riippuvuus kuvataan lineaarisella approksimaatiolla. Tämä tarkoittaa, että etsitään esimerkiksi lineaarisella regressioanalyysillä lineaarisen funktion $y = ax + b$ parametreille a ja b parhaat arvot siten, että funktio mahdollisimman hyvin approksimoi mallinnettavaa riippuvuutta.

Esimerkki lineaarisesta approksimaatiosta: (lähde: [Wikipedia, lineaarinen regressioanalyysi](#))



Kun riippuvuus on monimutkaisempi, joudutaan riippuvuutta kuvaavan funktion approksimoinnissa käyttämään vaativampia menetelmiä, esimerkiksi polynominen sovitus:

(lähde: <https://kisallioppiminen.fi/kurssit/mab4/luku3/>).



Pystyvin menetelmä monimutkaistenkin funktioiden approksimointiin on neuroverkkotekniikka. On osoitettu, että keinotekoisella neuroverkolla voidaan verkon tasoja lisäämällä ja syventämällä mallintaa mikä tahansa funktio.

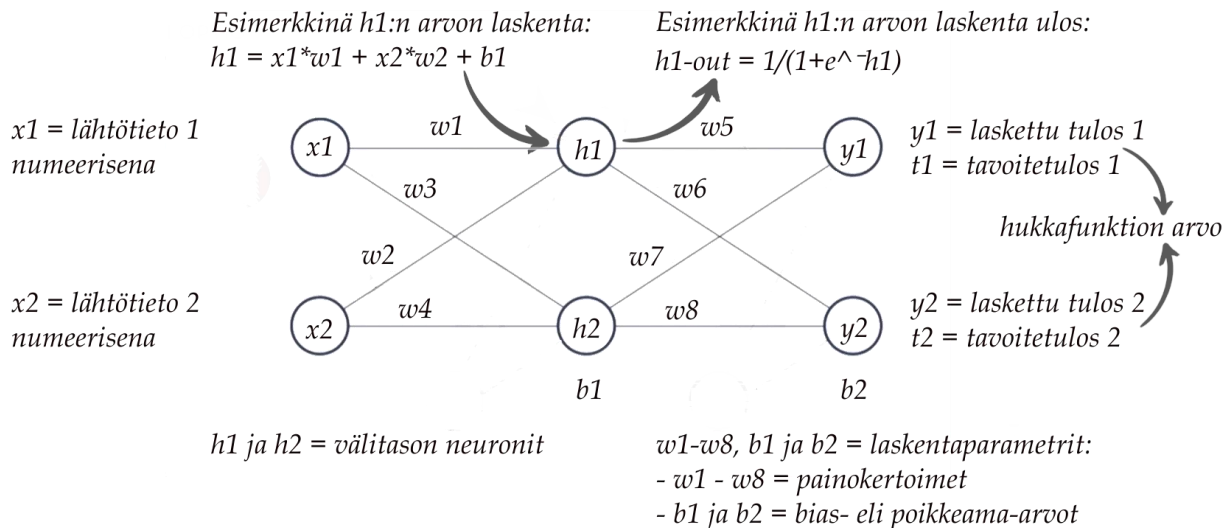
Neuroverkko funktion approksimoijana

Keinotekoinen neuroverkko koostuu neuroneista ja neuroneiden välisistä synapseista. Neuronit kuvataan palloina ja synapsit neuronien välisinä viivoina. Neuroverkon rakenne määrittää matemaattisen funktion, jossa lähtötietoina eli argumentteina ovat lähtöneuronien arvot ja tulostietoina eli funktion arvona tulosneuronien arvot.

Neuroverkon määrittämä funktio on sarja yksinkertaisia matemaattisia operaatioita, jotka on kuvattu alla olevassa mallissa. Laskennan lopputuloksen määräävät verkon laskentaparametrit, joita ovat synapsien painoarvot (esimerkissä $w1-w8$) ja ns. bias- eli poikkeama-arvot (esimerkissä $b1$ ja $b2$).

Neuroverkon laskenta etenee seuraavasti:

- Kuhunkin neuroniin summataan sitä edeltävien neuronien arvot kerrottuna vastaavan synapsin painoarvolla (esimerkissä $w1-w8$).
- Summaan lisätään neuronitasoa vastaava bias- eli poikkeama-arvo (esimerkissä $b1$ ja $b2$).
- Saatu luku sijoitetaan neuroverkolle valittuun ns. aktivaatiofunktioon ja saatu funktion arvo edustaa neuronin arvoa eteenpäin.
- Laskenta viedään eteenpäin läpi koko neuroverkon.



Verkon määrittämä funktio voidaan esittää myös suoraan lausekkeena. Alla yllä olevan neuroverkon määrittämä funktio lausekemuodossa. Käytännössä laskentaa ei kuitenkaan suoriteta lausekkeena, vaan kaavion määrittämänä algoritmina, joka toteuttaa laskentakaavan. Esitänkin lausekkeen lähinnä konkreettisenä näyttönä siitä, että neuroverkko on puhdas matemaattinen funktio:

$$y1 = 1/(e^{-1*((1/(e^{-1*(x1*w1+x2*w2+b1)})))*w5+(1/(e^{-1*(x1*w3+x2*w4+b1)})))*w7+b2}))$$

$$y2 = 1/(e^{-1*((1/(e^{-1*(x1*w1+x2*w2+b1)})))*w6+(1/(e^{-1*(x1*w3+x2*w4+b1)})))*w8+b2}))$$

Vertailu lineaarinen regressioanalyysi vs. neuroverkko

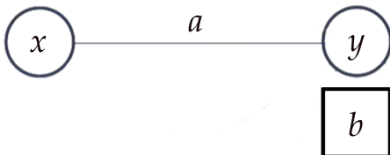
Lineaarinen regressioanalyysi ja neuroverkon mallintaminen edustavat tilastollisen mallintamisen ääripäitä; lineaarisen riippuvuuden ja mielivaltaisen moniulotteisen riippuvuuden mallintamista.

Lineaarisen regressioanalyysin ja neuroverkon mallintamisen välillä on kuitenkin analogia:

- Molemmissa mallinnetaan tilastollinen riippuvuus.
- Molemmissa laskennan lähtötiedot ovat numeerisia.
 - Neuroverkoilla käsitellään kuvaa, ääntä ym. ei numeerista tietoa, mutta ne muunnetaan aina numeerisiksi ennen neuroverkkoon syöttämistä.
- Molemmissa mallinnus tehdään etsimällä optimaaliset arvot funktion parametreille:
 - Lineaarisisessa regressioanalyysissä funktio on $y = a + bx$, ja estimoitavat parametrit tämän funktion parametrit a ja b .
 - Neuroverkossa funktio on itse neuroverkko ja estimoitavat parametrit synapsien painoarvot ja bias- eli poikkeama-arvot.
- Molemmissa määritellään virheen suuruutta kuvaava funktio ja laskennassa etsitään tämän funktion minimiä.
- Molemmissa on lopputuloksena puhdas matemaattinen funktio.

Erona se, että lineaarisessa regressioanalyysissä voidaan mallintaa vain yksinkertainen lineaarinen riippuvuus, mutta neuroverkolla mielivaltaisen monimutkainen funktio. Lisäksi lineaarisessa regressioanalyysissä puhutaan funktion tuntemattomien parametrien estimoinnista, neuroverkkojen mallintamisessa verkon kouluttamisesta. Kuitenkin myös neuroverkon mallintamisessa on kysymys funktion eli suunnitellun neuroverkon tuntemattomien parametrien estimoinnista.

Todettakoon vielä analogiasta, että alla oleva yksinkertainen mahdollinen neuroverkko toteuttaa lineaarisessa approksimaatiossa käytetyn kaavan $y = ax + b$.



Jos neuroverkko on vain tilastomatematiikkaa, miten selittyvät tekoälyn huikeat tulokset?

Nykyiset tekoälysovellukset perustuvat pääosin neuroverkkoihin. Tekoäly pohjautuu siis puhtaaseen matemaattiseen menetelmään, johon ei liity ihmisen älyn kaltaista älyä, arvovalintoja tai muuta vastaavaa. Kysymys on vain varsin monimutkaisten tilastollisten riippuvuuksien mallintamisesta tunnetulla aineistolla siten, että mallilla voidaan riittävän luotettavasti ennustaa tai estimoida tuloksia myös tuntemattomalle aineistolle.

Neuroverkko on tosiaan periaatteeltaan kohtuullisen yksinkertainen matemaattinen funktio; yhteenlaskua, kertolaskua ja sijoitus valittuun aktivaatiofunktioon. Yksinkertaisuus ei tässä kuitenkaan merkitse heikkoutta ja huonoutta, vaan nerokkuutta, sillä neuroverkolla voidaan approksimoida mielivaltaisen monimutkaisia funktioita ja sitä kautta mallintaa mielivaltaisen monimutkaisia tai paremminkin moniulotteisia tilastollisia riippuvuuksia.

Tekoälysovellukset taas perustuvat tilastollisten riippuvuuksien mallintamiseen. Tästä esimerkkejä alla.

- Keuhkokuvista diagnooseja tekevä tekoälysovellus pohjautuu neuroverkkoon, joka approksimoi funktiota, jossa lähtötietona eli funktion argumenttina ovat keuhkokuvan pisteet ja tulostietona eli funktion arvona diagnoosit.
- Keskustelevat tekoälysovellukset, kuten ChatGPT ja Gemini, perustuvat tilastolliseen todennäköisyyteen sille, millä sanalla syötteenä annettu teksti jatkuu. Ihmismäiset vasteet luovat illuusion älykkyydestä, mutta taustalla on silti vain tilastomatematiikkaa.
- Käsin kirjoitetun numeron tulkitseva tekoälysovellus perustuu neuroverkkoon, joka on mallinnettu suurella määrällä käsin kirjoitettujen numeroiden kuvia, joista tunnetaan tarkoitettu numero.

Näitä tekoälysovelluksia mahdollistavia neuroverkkomalleja ei olisi voitu toteuttaa ilman tietokoneiden suureksi kasvanutta suorituskykyä ja suurta määrää verkosta saatavaa materiaalia, joiden perusteella tilastollinen mallinnus on voitu tehdä. Suorituskykyä vaaditaan, koska neuroverkot voivat olla kooltaan varsin suuria; esimerkin neuroverkossa on kymmenen parametria, mutta suurimmissa käytössä olevissa neuroverkoissa on miljardeja parametreja.

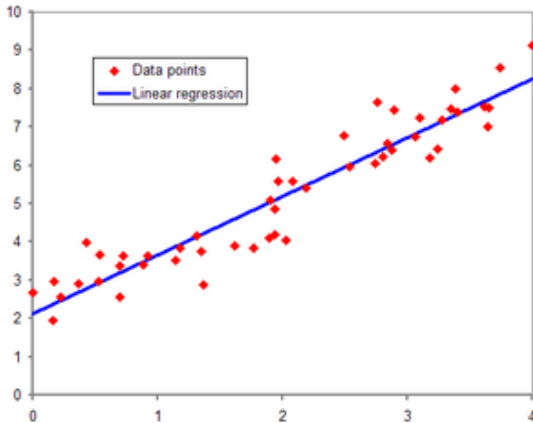
Tekoälyn huikeat tulokset perustuvat mm. seuraaviin asioihin:

- Neuroverkkojen kykyyn mallintaa mielivaltaisen monimutkaisia tilastollisia riippuvuuksia.
- Tietoverkoista saatavan neuroverkkojen opetuksessa käytettävän tiedon suureen määrään.
- Tietokoneiden kykyyn suorittaa neuroverkkojen vaatimaa suuria tehoja vaativaa laskentaa.

Yhteenveto

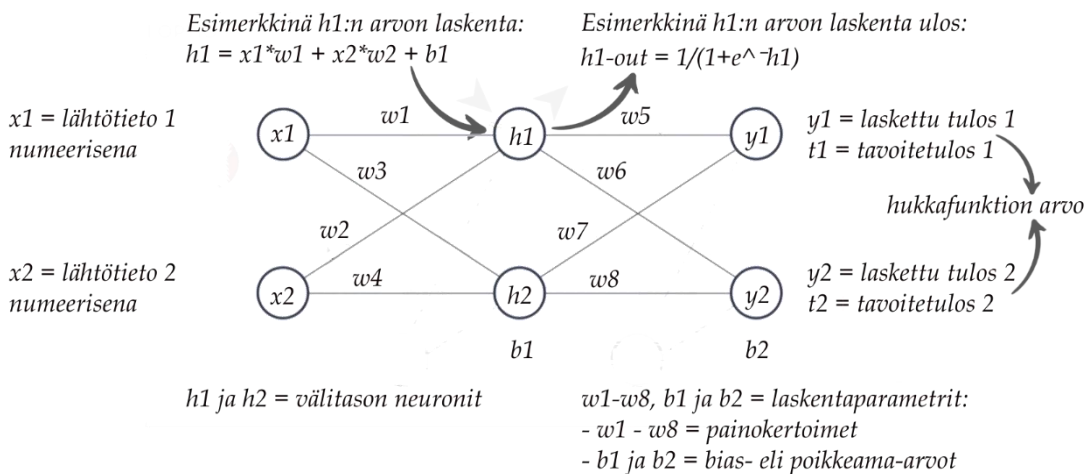
Kirjoituksen tavoite on esittää analogia lineaarisen regressioanalyysin ja neuroverkkotekniikan välillä ja tätä kautta selventää neuroverkon ideaa ja hälventää hämärtävää savuverhoa neuroverkon ympäriltä.

Lineaarisen regressioanalyysin idea on löytää yhtälön $y = ax + b$ parametreille a ja b arvot, joilla yhtälö parhaiten sovituu joukkoon tunnettuja pistepareja:



Neuroverkon idea on täsmälleen sama. Siinä funktio on vain pidempi ja parametreja enemmän. Alla pienen neuroverkon kaavakuva ja sama lausekkeena. Oikeat neuroverkot ovat yleensä laajempia, jopa miljardeja parametreja, mutta eivät monimutkaisempia - vain työlämpiä laskea.

Neuroverkkoesimerkki kaavakuvana:



Sama neuroverkkoesimerkki matemaattisena lausekkeena:

$$y1 = 1/(e^{(-1*((1/(e^{(-1*(x1*w1+x2*w2+b1))}))*w5+(1/(e^{(-1*(x1*w3+x2*w4+b1))}))*w7+b2))})$$

$$y2 = 1/(e^{(-1*((1/(e^{(-1*(x1*w1+x2*w2+b1))}))*w6+(1/(e^{(-1*(x1*w3+x2*w4+b1))}))*w8+b2))})$$

Neuroverkko on siis selkeä, suoraviivainen ja nerokas tilastomatemattinen menetelmä mallintaa mielivaltaisen monimutkaisia tilastollisia riippuvuuksia, mutta se ei ole esimerkiksi ihmisen älyn kaltaista tai muutenkaan älyä puhumattakaan omasta tahdosta tai pyrkimyksistä.