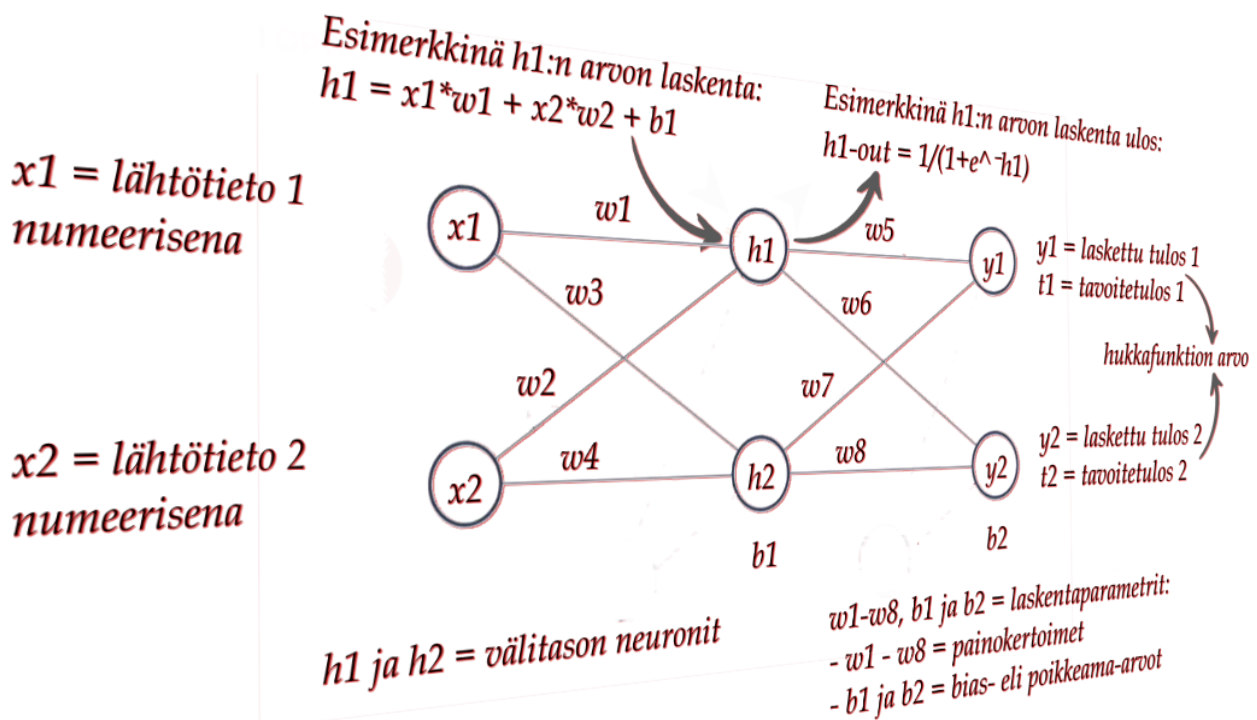


# LYHYT JOHDATUS NEUROVERKON TOIMINTAAN

## Sisällysluettelo

Tästä artikkelista	1
Funktio	1
Epäsäännöllinen funktio	2
Funktion approksimointi	2
Neuroverkko funktion approksimoijana	4
Vertailu lineaarinen regressioanalyysi vs. neuroverkko	5
Neuroverkko ja tekoäly	6
Jos neuroverkko on vain tilastomatematiikkaa, miten selittyvät tekoälyn huikeat tulokset?	6
Yhteenveto	7
Lisää tietoa	8
LIITE 1, VERTAILU LINEAARINEN REGRESSIOANALYYSI VS. NEUROVERKKOLASKENTA	9



# Lyhyt johdatus neuroverkon toimintaan

## Tästä artikkelista

Tämän artikkelin tarkoitus on antaa matematiikkaa jonkin verran hallitsevalle henkilölle nopeasti ja helposti ymmärrys tekoälyn pohjana olevien keinotekoisien neuroverkkojen toiminnasta. Kaikki tässä artikkelissa esiintyvät matemaattiset käsitteet sisältyvät lukion matematiikan oppimäärään, mutta asian ymmärtämiseen riittää hyvin myös peruskoulun matematiikan oppimäärä.

Tavoitteena on, että lukija ymmärtää tämän artikkelin luettuaan, ettei neuroverkko ole älykäs toimija, vaan matemaattinen funktio, jolla voidaan mallintaa monimutkaisia tilastollisia riippuvuuksia.

Artikkelin on kirjoittanut Eino Uikkanen, jolle voi lähettää mahdolliset kommentit osoitteeseen [eino.uikkanen@iki.fi](mailto:eino.uikkanen@iki.fi).

## 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.

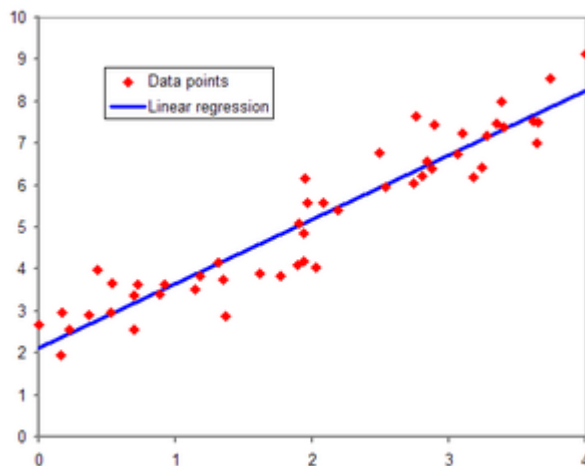
- Valitaan joukko äänitteitä linnun laulusta ja liitetään niihin linnun tiedetty laji. Tällöin laji on äänitteiden funktio.
- Valitaan joukko keuhkokuvia ja liitetään niihin kuvan perusteella tehty diagnoosi. Tällöin diagnoosi on keuhkokuvien funktio.

### 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.

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

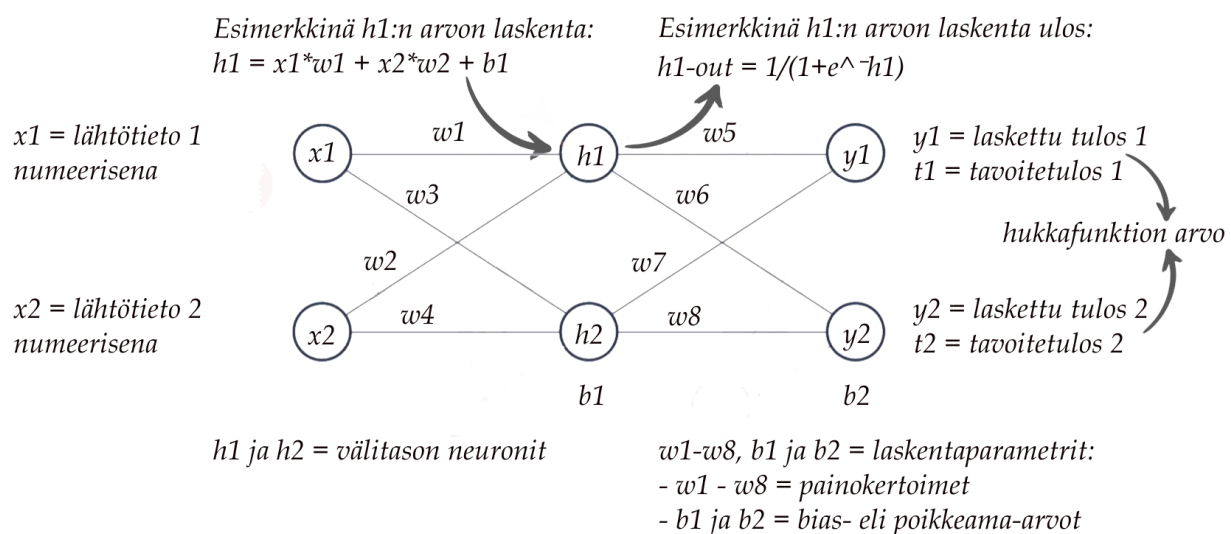
## Neuroverkko funktion approksimoijana

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 lopputulos riippuu verkon laskentaparametreista, joita ovat synapsien painoarvot (esimerkissä  $w_1-w_8$ ) ja ns. bias- eli poikkeama-arvot (esimerkissä  $b_1$  ja  $b_2$ ).

### Neuroverkon laskenta etenee suoraviivaisesti:

- Kuhunkin neuroniin summataan sitä edeltävien neuronien arvot kerrottuna vastaavan synapsin painoarvolla (esimerkissä  $w_1-w_8$ ).
- Summaan lisätään neuronitasoa vastaava bias- eli poikkeama-arvo (esimerkissä  $b_1$  ja  $b_2$ ).
- 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:

$$y_1 = 1/(e^{-1*((1/(e^{-1*(x_1*w_1+x_2*w_2+b_1)})) * w_5 + (1/(e^{-1*(x_1*w_3+x_2*w_4+b_1)})) * w_7 + b_2))})$$

$$y_2 = 1/(e^{-1*((1/(e^{-1*(x_1*w_1+x_2*w_2+b_1)})) * w_6 + (1/(e^{-1*(x_1*w_3+x_2*w_4+b_1)})) * w_8 + b_2))})$$

## Neuroverkon toiminnan vaiheet ovat seuraavat

1. Suunnittelu: Neuroverkon rakenne suunnitellaan ja siinä käytetyt laskentakaavat valitaan verkon tehtävän mukaisesti.
2. Alustus: Verkon laskentaparametrit (esimerkissä  $w_1$ - $w_8$  ja  $b_1$  ja  $b_2$ ) alustetaan antamalla niille lähtöarvot. Lähtöarvot alustetaan tyypillisesti täysin satunnaisilla arvoilla, koska parempaa perustetta lähtöarvoille ei ole.
3. Oppiminen: Neuroverkon läpi viedään suuri määrä opetusaineistoksi kutsuttua tietoa, josta tunnetaan sekä lähtötiedot että tavoitellut tulokset. Neuroverkko optimoidaan päivittämällä verkon laskentaparametreja matemaattisesti niin, että se tuottaa opetusaineistosta tuloksia, jotka poikkeavat mahdollisimman vähän tavoitelluista tuloksista.
4. Testaus: Neuroverkkoa testataan aineistolla, josta tunnetaan sekä lähtötiedot että tavoitellut tulokset, mutta joka ei ole ollut mukana opetusaineistossa.
5. Käyttö: Lopputuloksena neuroverkko tai malli, kuten koulutettua verkkoa usein kutsutaan, osaa tuottaa hyviä ja osuvia tuloksia myös uudesta ja tuntemattomasta aineistosta.

## Vertailu lineaarinen regressioanalyysi vs. neuroverkko

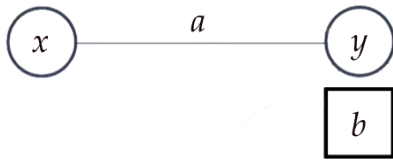
Lineaarinen regressioanalyysi ja neuroverkon mallintaminen edustavat tilastollisen mallintamisen ääripäitä; lineaarisen riippuvuuden ja mielivaltaisen moniulotteisen riippuvuuden mallintamista.

Lineaarisen regressioanalyysin ja neuroverkon mallintamisen välillä on kuitenkin analogia:

- Molemmassa mallinnetaan tilastollinen riippuvuus.
- Molemmassa laskennan lähtötiedot ovat numeerisia.
  - o Neuroverkoilla käsitellään kuvaa, ääntä ym. ei numeerista tietoa, mutta ne muunnetaan aina numeerisiksi ennen neuroverkkoon syöttämistä.
- Molemmassa etsitään optimaaliset arvot valmiiksi määritellyn funktion vakioparametrien arvoille.
  - o Linearisessa regressioanalyysissä funktio on  $y = a + bx$ , ja estimoitavat parametrit tämän funktion parametrit  $a$  ja  $b$ .
  - o Neuroverkossa funktio on itse neuroverkko ja estimoitavat parametrit synapsien painoarvot ja bias- eli poikkeama-arvot.
- Molemmassa määritellään virheen suuruutta kuvaava funktio ja laskennassa etsitään tämän funktion minimiä.
- Molemmassa 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  $y = ax + b$ .



Tarkempi lineaarisen regressioanalyysin ja neuroverkkolaskennan vertailutaulukko liitteessä 1, [“VERTAILU LINEAARINEN REGRESSIOANALYYSI VS. NEUROVERKKOLASKENTA”](#).

## Neuroverkko ja tekoäly

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.

## Jos neuroverkko on vain tilastomatematiikkaa, miten selittyvät tekoälyn huikat tulokset?

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.

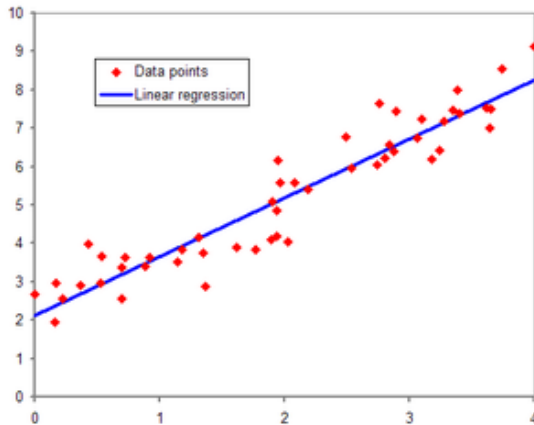
- Keuhkokuvista diagnooseja tekevä tekoälysovellus pohjautuu neuroverkkoon, joka approksimoi funktiota, jossa lähtötietona eli funktion argumenttina ovat keuhkokuvan pisteet ja tulostietona eli funktion arvona diagnoosit.
- Keskustelevat tekoälysovellukset, kuten ChatGPT ja Gemini, perustuvat tilastolliseen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

## Yhteenveto

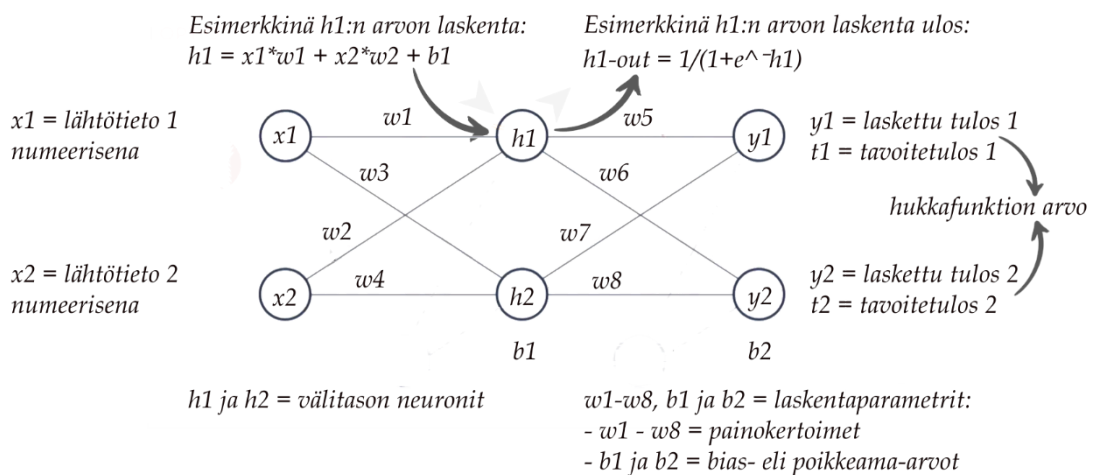
Kirjoituksen tavoite on esittää analogia lineaarisen regressioanalyysin ja neuroverkkotekniikan välillä ja tätä kautta selventää neuroverkon ideaa ja hälventää hämärtävää savuverhoa neuroverkon ympäriltä.

**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 työläämpiä 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.

## Lisää tietoa

Lisää tietoa neuroverkoista ja tekoälysovelluksista on saatavilla tekoälyesitelmäni ”Mitä tekoäly on ja mitä se ei ole” kalvosarjasta:

- Google Slides -muodossa avautuva Microsoft Powerpoint-versio (Ctrl+F5 avaa täyden esitysmuodon): [Mitä tekoäly on ja mitä se ei ole?](#)
- PDF-muotoinen versio: [Mitä tekoäly on ja mitä se ei ole?](#)

Linkit esitelmäni kalvosarjaan löytyvät myös tekoälysisivultani:

- Mitä tekoäly on ja mitä se ei ole: <https://www.einouikkanen.fi/AI/>



**LIITE 1, VERTAILU LINEAARINEN REGRESSIOANALYYSI VS. NEUROVERKKOLASKENTA**

	Lineaarinen regressio	Neuroverkko	Huomattavaa
<b>Laskennan lähtötiedot</b>	Joukko tunnettuja lukupareja (selittävä muuttuja, selitettävä muuttuja)	Joukko tunnettuja vektoripareja (lähtötiedot, tulostiedot)	Molemmissa laskennan lähtötiedot ovat numeerisia.
<b>Mikä on laskennan tavoite</b>	Selvitetään tilastollinen riippuvuus tunnettujen selittävien ja selitettävien tietojen välillä, jotta riippuvuutta voidaan käyttää selitettävän tiedon arvon estimointiin tapauksessa, jossa vain selittävä muuttuja tunnetaan.	Selvitetään tilastollinen riippuvuus tunnettujen lähtö- ja tulostietojen välillä, jotta riippuvuutta voidaan käyttää tulostietojen ennustamiseen tapauksessa, jossa vain lähtötiedot tunnetaan.	Molemmissa mallinnetaan tilastollinen riippuvuus.
<b>Mitä lasketaan</b>	Yhtälön $y = a + bx$ parametrit $a$ ja $b$ .	Neuroverkon laskentaparametrit eli synapsien painoarvot ja bias-arvot.	Molemmissa etsitään optimaaliset arvot valmiiksi määritellyn funktion vakioparametrien arvoille.  Molemmissa määritellään virheen suuruutta kuvaava funktio ja laskennassa etsitään tämän funktion minimiä.
<b>Mitä saadaan lopputulokseksi</b>	Yhtälö $y = a + bx$ , joka antaa hyviä estimaatteja selitettävän muuttujan arvolle, kun siihen sijoitetaan tuntemattoman tapauksen selittävän muuttujan arvo.	Neuroverkkomalli, joka antaa hyviä ennusteita tulostiedoiksi, kun neuroverkkoon syötetään tuntemattoman tapauksen lähtötiedot.	Molemmissa on lopputuloksena puhdas matemaattinen funktio.  Kummassakaan ei ole sijaa älylle, tahdolle, päättelylle tms.