

**Kohosen itseorganisoituvat verkot lääketieteellisen datan
luokitteluongelmassa**

Johanna Autio

Tampereen yliopisto
Tietojenkäsittelytieteiden laitos
Tietojenkäsittelyoppi
Pro gradu -tutkielma
Ohjaaja: Martti Juhola
Huhtikuu 2008

Tampereen yliopisto
Tietojenkäsittelytieteiden laitos
Tietojenkäsittelyoppi
Johanna Autio: Kohosen itseorganisoituvat verkot lääketieteellisen datan
luokitteluongelmassa
Pro gradu -tutkielma, 71 sivua, 7 liitesivua
Huhtikuu 2008

Tutkimuksessa tarkastellaan, kuinka Kohosen itseorganisoituvat verkot soveltuvat luokittelemaan vaikeasti diagnosoitavia huimauspotilaiden tauteja korvalääketieteen alalta. Aineistoa on luokiteltu aikaisemminkin neuroverkoilla käyttäen tähän pääosin monikerroksisia perceptron-verkkoja. Itseorganisoituvien verkkojen soveltuvuudesta aineiston luokitteluun kaivattiin lisätietoa ja oltiin myös kiinnostuneita, kuinka aineistossa olevien pienten tautiluokkien luokittelu niillä onnistuisi.

Tutkimuksessa todettiin itseorganisoituvien verkkojen soveltuvan hyvin tämän aineiston luokitteluun. Aineiston pääjoukon luokittelu onnistui sensitiivisyydellä 81 % ja kokonaistarkkuudella 91 %. Pientenkin tautiluokkien luokittelun mahdollisuudesta saatiin viitteitä, mutta tautiluokkien jakauman tulisi olla tasaisempi kaikki tautiluokat optimaalisesti luokittelevan mallin muodostamiseksi.

CR-luokat: I.5.1 [**Pattern Recognition**]: Models – *Neural nets*; I.5.2 [**Pattern Recognition**]: Design Methodology – *Classifier design and evaluation*.

Avainsanat ja -sanonnat: neuroverkko, itseorganisoituva kartta, SOM, luokittelu, lääketieteellinen data-analyysi, huimaus.

Sisällys

1.	Johdanto.....	1
2.	Kohosen itseorganisoituvat verkot	4
2.1.	Neuroverkoista yleisesti	4
2.1.1.	Yhteys aivotoimintaan.....	5
2.1.2.	Neuroverkkojen peruselementtejä.....	8
2.1.3.	Erilaisia neuroverkkomenetelmiä.....	9
2.2.	Ominaisuudet ja käyttöalueet	12
2.3.	Oppiminen	13
2.3.1.	Ohjaamattoman oppimisen eduista	13
2.3.2.	Neuroneiden painovektorit.....	14
2.3.3.	Painovektoreiden alustaminen.....	14
2.3.4.	Painovektoreiden muokkaaminen	15
2.3.5.	Oppimisen vaiheet ja opetusparametrit	16
2.4.	Kalibrointi	17
3.	Aineistosta ja tutkimusongelmasta	19
3.1.	Aineisto ja tutkimusongelma	19
3.2.	Aineiston osajoukot	22
3.3.	Aineiston esikäsittely	24
3.3.1.	Puuttuvat arvot	24
3.3.2.	Tulosluokkien jakauma	24
3.3.3.	Muuttujien normalisointi.....	25
3.4.	Aineiston jakaminen opetukseen ja testaukseen	25
4.	Testit ja tulokset	27
4.1.	Käytetyt laitteistot ja ohjelmistot	27
4.2.	Luokittelun onnistumisen arviointimenetelmät	29
4.3.	Verkkojen kalibrointi	31
4.4.	Opetusparametrit.....	33
4.4.1.	Verkon koko.....	34
4.4.2.	Verkon alustaminen	34
4.4.3.	Hilan muoto.....	35
4.4.4.	Naapurusto	35
4.4.5.	Opetuksen kesto	37
4.4.6.	Oppimisnopeus.....	37
4.5.	Opetus ja testaus aineiston osajoukoilla	38
4.5.1.	Pääjoukko	39
4.5.2.	Kolmas osajoukko.....	53

4.6. Testaus aineiston osajoukoilla	60
4.6.1. Toinen osajoukko	60
4.6.2. Neljäs osajoukko	62
5. Yhteenveto.....	64
Viiteluettelo	68
Liitteet	
Liite 1. Pääjoukon testien opetusparametritaulukoita.....	72
Liite 2. Kolmannen osajoukon testien opetusparametritaulukoita	77

1. Johdanto

Tässä tutkimuksessa luokiteltiin huimauspotilaiden dataa korvalääketieteen alalta Kohosen itseorganisoituvilla verkoilla. Aineisto koostuu vaikeasti toisistaan erotettavista taudeista. Tautien diagnosointiin on kehitetty myös asiantuntijajärjestelmä ONE [Auramo *et al.*, 1993; Kentala *et al.*, 1996] ja useiden eri menetelmien sopivuutta tautien luokitteluun on tutkittu [Juhola *et al.*, 1999].

Neuroverkot sopivat hyvin tällaiseen vaativaan luokittelutehtävään. Ne ovat oppivia ja älykkäitä järjestelmiä, joita käytetään monimutkaisten ongelmien ratkaisuun useilla aloilla, mm. lääketieteessä. Lääketieteessä ne ovat osoittautuneet hyödyllisiksi esimerkiksi diagnostiikassa, biolääketieteellisten signaalien prosessoinnissa ja analyysissä sekä lääketieteellisessä kuva-analyysissä [Begg *et al.*, 2006].

Itseorganisoituvat verkot ovat suomalaisen Teuvo Kohosen vuonna 1982 kehittämä ohjaamattomaan oppimiseen perustuva neuroverkkomenetelmä [Kohonen, 2001]. Ohjaamattomassa opettamisessa neuroverkoille ei kerrota, kuinka niiden tulisi vastata syötteisiin, vaan verkot päättävät siitä itse. Tästä on monia etuja, datasta esimerkiksi voidaan tällaisella menetelmällä löytää uusia ilmiöitä [Joutsiniemi *et al.*, 1995]. Neuroverkoilla yleensäkin mallinnettiin alun perin aivojen toimintaa, mutta vastaavuutta on lähinnä vain tiedon käsittelyn ja siirtymisen mielessä [Juhola, 2004]. Kohonen kuitenkin vei yhteyden biologiaan hiukan tätä lähemmäksi mallintaessaan verkoillaan itseorganisoitumalla muodostuvia aivokarttoja. Kohosen itseorganisoituvat verkot ovat nykyään yksi suosituimmista neuroverkkomenetelmistä [Oja ja Mannila, 2007]. Itseorganisoituvan verkon erityispiirre muihin luokittelumenetelmiin verrattuna on sen topologian säilytyskyky [Oja and Kaski, 1999], verkko tiivistää tietoa säilyttäen tärkeimmät datassa olevat suhteet [Kohonen, 2001]. Luokittelun lisäksi itseorganisoituvia verkkoja voidaan käyttää mm. klusteroinnissa, epälinearisessa projektiossa, aikasarja-analyysissä ja monidimensioisen datan visualisoinnissa [Cottrell and Verleysen, 2006].

Neuroverkoista eniten on tämän aineiston luokittelussa käytetty yleisestikin suosituinta [Swingler, 1996] neuroverkkomenetelmää monikerroksista perceptronia [Autio *et al.*, 2007; Juhola *et al.*, 1999; Juhola *et al.*, 2001; Juhola and Laurikkala, 2003; Juhola and Siermala, 2005; Siermala and Juhola, 2006; Siermala *et al.*, 2007], mutta myös tässä tutkimuksessa käytettyjä Kohosen itseorganisoituvia verkkoja on kokeiltu aineiston luokittelussa aiemmin [Juhola *et al.*, 2001]. Tuolloin ne osoittautuivat lupaavaksi menetelmäksi ja tämä tutkimus tehtiinkin, jotta verkkojen soveltuvuudesta aineiston

luokitteluun saataisiin lisätietoa. Nyt tarkoituksena oli etsiä tautien luokitteluun mahdollisimman hyvin soveltuvia verkkorakenteita. Parhaiden verkkorakenteiden valinnassa painotettiin luokittelun onnistumista kokonaisuutena, jolloin tautiluokkien erikokoisuuden vuoksi luokittelun onnistumista mitattiin tautitapausten määrillä painottaen. Siten parhaiden verkkorakenteiden valinnassa korostettiin suurimpien tautiluokkien luokittelun onnistumista. Tämän lisäksi kuitenkin tavoitteena oli selvittää pientenkin tautiluokkien luokittelun mahdollisuutta itseorganisoituvilla verkoilla.

Tutkimuksessa todettiin itseorganisoituvien verkkojen soveltuvan hyvin tämän aineiston luokitteluun ja myös pienten tautiluokkien luokittelun mahdollisuudesta saatiin viitteitä. Tautiluokkien jakauman tulisi kuitenkin olla nykyistä tasaisempi, jotta voitaisiin muodostaa kaikki taudit optimaalisesti luokitteleva malli.

Tutkielman toisessa luvussa perehdytään ensin tutkimuksessa luokitteluun käytettyyn menetelmään Kohosen itseorganisoituviiin verkkoihin. Tässä yhteydessä luodaan myös katsaus neuroverkkoihin yleisesti. Itseorganisoituvien verkkojen käsittelyssä keskitytään erityisesti tämän tutkimuksen kannalta olennaisiin seikkoihin, kuten verkkojen kalibrointiin, joka on tehtävä verkoille, jotta niitä voidaan käyttää luokittelussa.

Kolmannessa luvussa kerrotaan aineistosta ja tutkimusongelmasta, jolloin käsitellään mm. aineiston koostumusta, osajoukkoja, esikäsitelyä, jakamista opetukseen ja testaukseen sekä kuinka aineistoa on aikaisemmissa tutkimuksissa luokiteltu ja miten tämä nyt tehtiin.

Neljännessä luvussa kuvataan tutkimuksen empiirinen osuus eli tarkastellaan tehtyjä neuroverkkoja ja niillä saatuja tuloksia. Ensin kerrotaan käytetyistä laitteistoista ja ohjelmistoista sekä annetaan suoritusaikaesimerkkejä. Tämän jälkeen määritellään luokittelun onnistumisen arviointimenetelmät. Kun toisessa luvussa käsiteltiin yleisemmällä tasolla itseorganisoituvien verkkojen käyttöön liittyviä asioita, kuten verkkojen kalibrointia ja opetusparametreja, tarkennetaan niitä tässä luvussa ja kerrotaan perustellen, kuinka niiden suhteen on tässä tutkimuksessa menetelty. Luvun lopuksi selostetaan aineiston osajoukkojen luokitteluun parhaiten soveltuvien verkkorakenteiden kehitysprosessi ja pienten tautiluokkien luokittelun mahdollisuuden selvittämiseksi tehdyt testit, minkä yhteydessä myös analysoidaan saatuja tuloksia ja arvioidaan luokittelun onnistumista luvun alussa määritellyillä tunnusluvuilla. Tuntemattomista diagnooseista koostuvan osajoukon osalta kerrotaan, kuinka varmimmista diagnooseista koostuvalla osajoukolla opetettu verkko luokitteli diagnosoimattomat tapaukset.

Lopuksi viidennessä luvussa tehdään yhteenveto ja suhteutetaan tutkimus aikaisempaan.

2. Kohosen itseorganisoituvat verkot

Itseorganisoituva kartta (Self-Organizing Map, SOM, Self-Organizing Feature Map, SOFM, piirrekartta) on suomalaisen akateemikon Teuvo Kohosen vuonna 1982 kehittämä neuroverkkomenetelmä [Kohonen, 2001], joka on nykyään käytetyimpiä neuroverkkoalgoritmeja [Oja ja Mannila, 2007]. Laajennuksineen se on suosituin ohjaamattomaan oppimiseen, luokitteluun ja datan visualisointiin käytetty neuroverkkoalgoritmi [Cottrell and Verleysen, 2006].

Tässä luvussa käsitellään ensin neuroverkkoja yleisesti ja syvennyttään tämän jälkeen tarkemmin itseorganisoituvaan karttaan.

2.1. Neuroverkoista yleisesti

Neuroverkot (hermoverkko, neural network, NN, artificial neural network, ANN) ovat oppivia ja älykkäitä järjestelmiä. Niillä käsitellään monimutkaisia ongelmia, joita on vaikea tai jopa mahdotonta ratkaista perinteisin menetelmin, joissa ohjelmoidaan tarkat säännöt ongelman käsittelyyn. Neuroverkot sen sijaan oppivat esimerkeistä. Kun niille syötetään käsiteltävää ongelmaa kattavasti esittelevää dataa, ne käyttävät sen oppimiseen neuroverkkomenetelmänsä mukaista oppimisalgoritmiaan ja näin muodostavat sisäisen esityksen datasta, jonka perusteella ne pystyvät antamaan vastauksia myös uusiin samaa ongelmaa käsitteleviin syötteisiin.

Uusiin opetusdatassa esiintymättömiin syötteisiin järkevästi vastaaminen on yleistämistä [Haykin, 1994]. Jotta neuroverkot voivat yleistää, ne on täytynyt opettaa hyvälaatuisella ja kattavalla opetusdatalla. Hyvän datan kriteereitä käsitellään aineiston esikäsitteystä kerrottaessa kohdassa 3.3. Kattavuudella tarkoitetaan sitä, että aineiston tulisi ilmentää tutkittavan ongelman eri ilmiöitä mahdollisimman laajasti. Yleistämiskykynsä ansiosta neuroverkot kuitenkin pystyvät hiukan puutteellisellakin datalla opetuksen jälkeen vastaamaan syötteisiin alueilta, joita ei ole opetusdatassa esitetty [Swingler, 1996]. Tässä tulee kuitenkin huomioida, että interpolointi onnistuu neuroverkoilta paremmin kuin ekstrapolointi [Juhola, 2004], koska yleistäminen on opitusta datasta epälineaarista keskimääräistämistä [Swingler, 1996].

Puutteellisuuden lisäksi opetusaineistossa mahdollisesti esiintyvä kohina eli häiriö tuo haasteellisuutta neuroverkon oppimiseen. Neuroverkot ovat kuitenkin myös vikasietoisia eli datassa oleva kohina ei välttämättä pilaa koko oppimisprosessia ja

järjestelmän suorituskykyä. Yleistämiskykynsä ansiosta ne suoriutuvat kohinaisestakin datasta, ja useisiin neuroneihin hajautetun laskennan vuoksi yksittäisten neuroneiden epäonnistunut oppiminen ei välttämättä paljoakaan huononna järjestelmän suorituskykyä [Juhola, 2004].

Häiriöitä ei kuitenkaan saa olla liikaa, vaan datan täytyy olla hyvälaatuista, koska neuroverkotkaan eivät opi huonosta datasta. Lisäksi neuroverkkoa opettaessa on varottava sen ylisovittumista opetusaineistoon ja siinä esiintyvään kohinaan, sillä neuroverkon oppiessa kohinan ominaisuuksia sen yleistyskyky heikkenee kohinan ollessa satunnaisena ilmiönä epäennustettava. [Juhola, 2004] Tarkkuus kyllä tällöin paranee ja neuroverkon käyttötarkoituksesta riippuu, onko yleistyskyky vai tarkkuus tärkeämpää. Esimerkiksi data-analyysia tehtäessä tarvitaan ennen kaikkea tarkkuutta, mutta useimmiten, mm. luokittelutehtävissä, yleistyskyvyn säilyttäminen on erittäin tärkeää.

Vaikka ihanteellisinta olisi saada kehitettyä neuroverkko, joka olisi hyvin yleistämiskykyinen ja tarkka, niin käytännössä joudutaan usein tinkimään jommastakummasta [Swingler, 1996]. Yleensä kannattaa suosia yksinkertaisia neuroverkkoja niiden paremman yleistyskyvyn vuoksi. Swingler [1996] kehottaakin valitsemaan ongelman hyvin selittävästä neuroverkkomalleista yksinkertaisimman. Ylisovittumista voi valvoa neuroverkkoa opettaessa validointidatalla, joka on neuroverkolle opetuksessa esittämätöntä uutta dataa, ja siten siitä suoriutumista voidaan käyttää yleistyskyvyn mittarina [Juhola, 2004].

Neuroverkot sopivat siis ongelmiin, joiden ratkaisuun ei tiedetä sääntöjä, mutta joista voidaan neuroverkkojen opetusta varten kerätä riittävästi ongelmaa kuvaavaa dataa, joka on neurolaskentaa varten muunnettavissa numeeriseen muotoon [Swingler, 1996]. Neuroverkoille hyvin soveltuvia tehtäviä ovat luokittelu, ennustaminen, aikasarjaennustaminen ja tiedon segmentointi [Juhola, 2004].

2.1.1. Yhteys aivotoimintaan

Aivojen hermoverkot ovat tietoa käsitteleviä, lähettäviä ja vastaanottavia järjestelmiä ja ne muodostuvat toisiinsa yhteyksissä olevista hermosoluista eli neuroneista [Lyytinen ja Himberg, 1996]. Keinotekoisilla neuroverkoilla mallinnettiin alun perin aivojen toimintaa, mutta nykyisin neuroverkkojen kehityksen myötä ne eivät ole enää paljoakaan aivojen kaltaisia. Vastaavuutta aivojen toimintaan on yleisesti ottaen lähinnä vain tiedon käsittelyn ja siirtymisen mielessä. Aivojen hermoverkoissa tieto siirtyy

neuroneista toisiin niiden välisten yhteyksien kautta, ja myös neuroverkoissa siirretään tietoa neuroneiden välillä ja näin saadaan vasteita syötteisiin. [Juhola, 2004]

Keinotekkoisten neuroverkkojen teoria on edistänyt myös aivotutkimusta selittäessään useita aivojen toimintoja kuten itseorganisoitumista, piirretunnistimia ja muistia [Ilmoniemi, 2008].

2.1.1.1. Plastisiteetti

Ihmisen aivotoiminnalle ominaista on plastisiteetti ja oppimiskyky. Plastisiteetilla tarkoitetaan aivojen mukautumiskykyä niin normaalin oppimisen kuin vaurioista toipumisen yhteydessä. Oppiessaan aivot uudelleen organisoivat aivotoimintoja, ja vaurioiden jälkeen vahingoittumattomat aivoalueet voivat korvata vahingoittuneita alueita. [Lyytinen ja Himberg, 1996] Neuroverkoilla on yhteneväisyyttä näihin aivojen ominaisuuksiin, koska kuten jo aiemmin mainittiin, nekin ovat oppimiskykyisiä ja vikasietoisia.

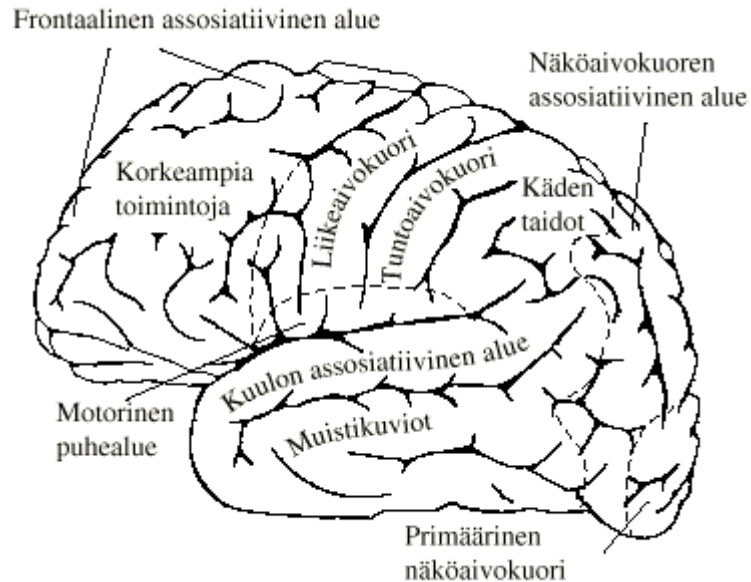
2.1.1.2. Epälineaarinen hajautettu järjestelmä

Aivot pystyvät suorittamaan hyvin monimutkaisia toimintoja kuten hahmontunnistusta erittäin nopeasti massiivisen rinnakkaisprosessointinsa ansiosta [Ilmoniemi, 2008]. Neurolaskennassa mallinetaan tätä rinnakkaisuutta. Keinotekoiset neuronit prosessoivat hajautetusti ongelmaa, ja lopputuloksena saadaan niiden yhdessä työstämä ratkaisu. Luonnollisestikin kun neurolaskentaa suoritetaan prosessorimääriltään hyvin rajoittuneilla tietokoneilla, kyse on vain simuloinnista aidon rinnakkaisuuden sijaan. [Juhola, 2004] Lisäksi sekä aivot että keinotekoiset neuroverkot käsittelevät tietoa epälineaarisesti ja ovat siten epälineaarinen hajautettu järjestelmä [Haykin, 1994]. Epälineaarisuus mahdollistaa monimutkaisten riippuvuussuhteiden käsittelyn.

2.1.1.3. Aivokartat – itseorganisoituvat kartat

SOM:lla on muita neuroverkkoja enemmän yhteyttä biologiaan. Se mallintaa aivokarttoja, jotka ovat aivoissa ja erityisesti aivokuorella eri toimintoihin kuten eri aistitoimintoihin erikoistuneita alueita ja niiden välisiä assosiatiivisia alueita. Nämä alueet muodostuvat itseorganisoitumalla pääosin aistininformaation ohjaamana ja aivojen plastisiteetin mahdollistamana. Aistinsolut projisioituvat aivokuoreen eli niitä

edustetaan aivokuorella aistinelimien mukaisessa topografisessa järjestyksessä. [Kohonen, 2001] Kuvassa 1 esitetään joitakin näistä aivokuoren eri toiminnoista huolehtivista alueista.



Kuva 1. Aivokuoren alueita [Ilmoniemi, 2008].

Topografisesti järjestyneiden karttojen muodostumiseen johtava itseorganisaatioprosessi on Kohosen [2001] mukaan määriteltävissä kolmella osaprosessilla: syötteen levittämällä kaikkiin neuroneihin, neuronien kilpailulla siitä, mikä niistä syötettä eniten vastaavana saa aktivoitua syötteestä eniten sekä kyseisen voittajan aktivoinnin yhteydessä myös sen naapuruston aktivoinnilla. Nämä ovatkin Kohosen kehittämän SOM-algoritmin perusteet, ja näitä käsitellään tarkemmin kohdassa 2.3.

Aivotointojen sijainnillisen järjestyneisyyden mallintaminen tuo monia etuja. Se esimerkiksi vähentää tarvittavien yhteyksien määrää tuodessaan toisiinsa liittyvät toiminnot lähelle toisiaan ja mahdollistaa loogisemman ja robustimman arkkitehtuurin. Kohonen [2001] ihmettelee, miksi tämä useimmiten jätetään neurolaskennassa täysin huomioimatta.

2.1.2. Neuroverkkojen peruselementtejä

Neuroverkoissa neuronit ovat järjestettyinä kerroksiin ja niiden välillä on yhteyksiä. Käsitellään seuraavaksi näitä neuroverkkojen peruselementtejä.

2.1.2.1. Neuronit

Biologinen neuroni muodostuu tietoa vastaanottavista osista eli dendriiteistä, tuman sisältävästä solukeskuksesta eli soomasta ja tietoa vievästä haarakkeesta eli aksonista. Tieto neuronista toiseen siirtyy lähettävän solun aksonin ja vastaanottavan solun dendriitin yhtymäkohdan synapsin kautta, jossa peräkkäiset neuronit eivät kuitenkaan aivan kosketa toisiaan, vaan välissä on synapsirako. Tiedon siirtyminen tapahtuu solun sisällä sähköisesti ja solujen välillä synapsissa kemiallisesti. [Lyytinen ja Himberg, 1996] Neuroniiin saapuvat syötteet lasketaan yhteen, ja summan ylittäessä määrätyn kynnyksarvon tulos välitetään eteenpäin synapseista, joiden informaation läpipäästämisen vahvuus vaihtelee [Juhola, 2004].

Kuten biologisissa neuroverkoissa, myös keinotekoisissa neuroverkoissa on neuroneita, joilla on yhteyksiä toisiinsa. Keinotekoisien neuroverkkojen neuroni on yksinkertainen malli biologisesta neuronista. Neuroniiin saapuu syötteitä, joista lasketaan niiden yhteyksissä synapsin yhdistämistehokkuutta kuvaavilla kertoimilla eli painoilla painotettu summa. Summa kynnystetään neuronin kynnysfunktioilla, jonka ollessa epälineaarinen neuroverkolla on mahdollista ratkaista epälineaarisia ongelmia. Kynnystetty arvo välitetään edelleen toisiin neuroneihin. [Juhola, 2004]

2.1.2.2. Kerrokset

Neuroverkoissa neuronit järjestetään kerroksiin, joista *syötekerroksen* (input layer) kautta saapuvat järjestelmän syötteet, jotka välitetään neuroneiden välisten yhteyksien kautta seuraaviin kerroksiin. Kun puhutaan yksi- ja monikerroksisista neuroverkoista, kerrosten määrällä viitataan usein syötekerroksen lisäksi verkossa olevien kerrosten määrään, koska syötekerroksessa ei tapahdu laskentaa. [Haykin, 1994] Juuri tästä samasta syystä monikerroksisten verkkojen tarkoista kerrosmääristä puhuttaessa esiintyy epäyhtenäistä käytäntöä siinä, sisällytetäänkö ilmoitettavaan kerrosmäärään syötekerrosta, mikä olisi suositeltavaa verkon rakennetta paremmin kuvaavana [Juhola, 2004].

Verkon viimeistä kerrosta kutsutaan *tuloskerrokseksi* (output layer) ja monikerroksisilla verkoilla syöte- ja tuloskerroksen välissä olevat kerrokset ovat *piilokerroksia* (hidden layer). Kerroksissa sijaitsevia neuroneita kutsutaan kerroksiaan vastaavasti syöte-, piilo- ja tulosneuroneiksi. [Haykin, 1994]

2.1.3. Erilaisia neuroverkkomenetelmiä

On olemassa useita neuroverkkomenetelmiä, jotka eroavat toisistaan niin oppimisalgoritmeiltaan kuin arkkitehtuureiltaan. Neuroverkolle käytettävästä oppimisalgoritmista riippuu, millainen arkkitehtuuri sille sopii parhaiten [Haykin, 1994]. Tarkastellaan seuraavaksi, kuinka neuroverkkomenetelmiä voidaan jaotella oppimisalgoritmeiltaan ja rakenteiltaan.

2.1.3.1. Oppimisalgoritmit

Oppimisalgoritmi määrittää, kuinka datasta oppiminen tapahtuu eli kuinka neuroverkon painoarvoja muokataan. Neuroverkot voidaan jakaa oppimisalgoritmeiltaan kahteen pääryhmään: ne oppivat joko *ohjatusti* (supervised learning) tai *ohjaamattomasti* (unsupervised learning). Ohjatussa oppimisessa neuroverkolle opetetaan, kuinka sen tulisi vastata annettuihin syötteisiin, kun taas ohjaamattomasti oppiessaan se itse päättää, kuinka se vastaa syötteisiin.

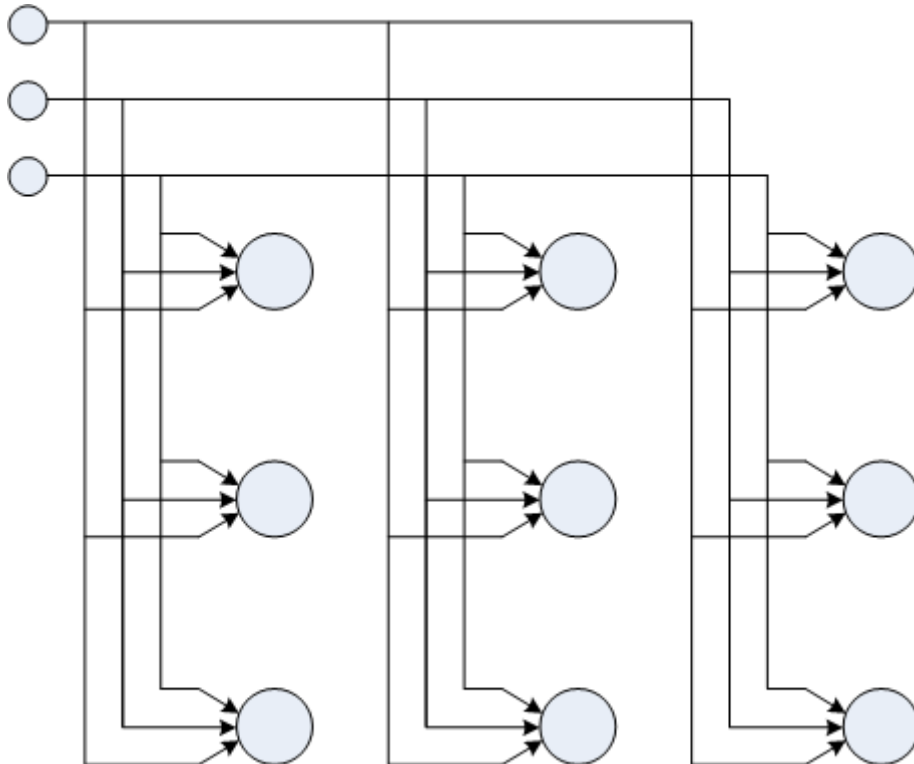
Ohjattu oppiminen voidaan edelleen jakaa *vahvistavaan oppimiseen* (reinforcement learning) ja *korjaavaan oppimiseen* (corrective learning), joista ensimmäisessä oikeita tuloksia vahvistetaan ja jälkimmäisessä virhe yhdessä syötteen kanssa määrittää, kuinka paljon painoja korjataan. Ohjaamaton oppiminen jakautuu myös alalajeihin: vahvistavaan oppimiseen ja *kilpailevaan oppimiseen* (competitive learning). [Rojas, 1996] Vahvistusoppiminen voi siis oppimisalgoritmista riippuen olla joko ohjattua tai ohjaamatonta oppimista.

Kilpailevassa oppimisessa verkon neuronit kilpailevat siitä, mikä niistä saa vastata syötteeseen. Voittajan vastatessa muita neuroneita samanaikaisesti estetään. [Rojas, 1996] Tässä tutkimuksessa käytettävä itseorganisoituva kartta käyttää ohjaamatonta oppimista ja on sen alalajeista lähimpänä kilpailevaa oppimista. Tavanomaisesta kilpailevasta oppimisesta se kuitenkin poikkeaa siten, että yhden syötteen yhteydessä päivitetään voittajasolmun lisäksi myös tämän neuronin naapurustoa, minkä ansiosta viereiset neuronit oppivat vastaamaan samantapaisiin syötteisiin. [Matlab, 2007]

2.1.3.2. Arkkitehtuurit

Neuroverkot voidaan jakaa arkkitehtuureiltaan eli rakenteiltaan *yksikerroksisiin eteenpäin syöttäviin verkkoihin* (single-layer feedforward network), *monikerroksisiin eteenpäin syöttäviin verkkoihin* (multi-layer feedforward network), *rekursiivisiin verkkoihin* (recurrent network) ja *hilarakenteisiin verkkoihin* (lattice network) [Haykin, 1994].

Jos tietoa kerrosten välillä kulkee vain eteenpäin, neuroverkkoa kutsutaan eteenpäin syöttäväksi verkoksi. Jos neuroverkolla on vähintään yksi yhteys taaksepäin joko piilotaivulosneuroneistaan, on kyse rekursiivisesta verkosta. Hilarakenteinen verkko on eteenpäin syöttävä verkko, jossa jokaisesta syötöneuronista on yhteys kaikkiin joko yksi- tai useampiulotteisen hilan neuroneihin eli tulosneuroneihin, jotka ovat tuloskerroksella järjestettyinä riveihin ja sarakkeisiin. [Haykin, 1994] Kuvassa 2 esitetään esimerkki kaksiulotteisesta hilarakenteisesta neuroverkosta, jossa jokaisesta kolmesta syötöneuronista on yhteys kaikkiin hilan yhdeksään neuroniin.



Kuva 2. Kaksiulotteinen hilarakenteinen neuroverkko. Kuva on piirretty Haykinia [1994] mukaillen.

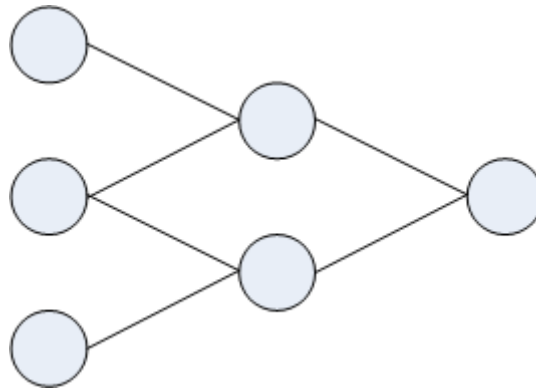
Tässä tutkimuksessa käytettävä itseorganisoituva kartta on hilarakenteinen neuroverkko. Hila on yleensä kaksiulotteinen mutta voi olla myös yksi- tai useampiulotteinen ja muodoltaan se on suorakulmainen, kuusikulmainen eli heksagonaalinen tai epäsäännöllinen. Syötevektorista on yksinkertaisimmassa tapauksessa yhteys kaikkiin verkon neuroneihin vaihtelevien skalaariarvoisten painojen kautta. [Kohonen, 2001]

Alakohdassa 2.1.3.3 esitellään lyhyesti monikerroksista perceptron -verkkoa, jolloin kerrotaan myös sen arkkitehtuurista.

2.1.3.3. Monikerroksinen perceptron

Käsitellään lyhyesti neuroverkoista yleisimmin käytettyä [Swingler, 1996] monikerroksista perceptronia (multi-layer perceptron, MLP), joka soveltuu monenlaisten vaikeiden ongelmien ratkaisuun [Haykin, 1994] ja jota on käytetty tämänkin tutkimuksen aineiston luokitteluun, kuten kohdassa 3.1 kerrotaan.

MLP on monikerroksinen eteenpäinsyöttävä verkko, joka käyttää virheen *takaisinlevittävää* (error backpropagation) algoritmia [Haykin, 1994]. MLP-verkossa neuroneiden kynnsfunktiot ovat epälineaarisia, jatkuvia ja derivoituvia ja ne on järjestetty syöte-, piilo- ja tuloskerroksiin, joista piilokerrokset toimivat piirteiden tunnistajina. Kolmogorovin teoreeman mukaan mikä tahansa tehtävä on suoritettavissa kolmikerroksisella eli syöte- ja tuloskerroksen lisäksi yhden piilokerroksen sisältävällä verkolla. [Juhola, 2004] Kuvassa 3 esitetään esimerkki tällaisesta kolmikerroksisesta MLP-verkosta. Kolmogorovin teoreema ei kuitenkaan määrittele, millainen verkon tulisi muilta osin olla minkä tahansa ongelman ratkaisun mahdollistamiseksi eli se ei ota kantaa neuroneiden määrään eikä niiden välisiin yhteyksiin ja yhteyksien painoihin [Juhola, 2004].



Kuva 3. Esimerkki kolmikerroksisesta MLP-verkosta. Ensimmäinen kerros on syötekerros, toinen kerros on piilokerros ja kolmas on tuloskerros.

Takaisinlevityksellä tarkoitetaan tapaa, jolla verkko oppii eli kuinka sen neuroneiden painoarvoja säädetään. MLP-verkkoa opetetaan ohjatusti ja siten tiedetään halutut tulokset, joiden suuntaan verkon kaikkien kerrosten neuroneiden painoarvoja korjataan. Tämä tapahtuu laskemalla syötteille virhefunktiot, jotka levitetään takaisin edeltäviin kerroksiin, joiden neuroneiden painoarvoja säädetään virhefunktion arvoa vähentäen. [Juhola, 2004]

2.2. Ominaisuudet ja käyttöalueet

SOM pystyy löytämään monimutkaisestakin vaikeasti hahmotettavasta datasta olennaisimman. Se tiivistää tietoa säilyttäen tärkeimmät datassa olevat suhteet ja esittää moniulotteisen datan omassa pienemmän dimensionalisuuden rakenteessaan, jolloin epälineaarisista tilastollisista suhteista muodostetaan yksinkertaisia geometrisia suhteita. [Kohonen, 2001] Tiedon suhteista säilytetään ja esitetään alueelliset tai topologiset suhteet [Juhola, 2004]. Tiedon tiivistäminen pienempidimensioiseksi muistuttaa klassista vektorikvantisointia, mutta on tästä poiketen järjestynyt ja syötedatan jakaamaa kuvaava [Kohonen, 2001].

SOM:n käyttöalueet ovat hyvin monipuoliset. Sitä voidaan soveltaa mm. konenäössä ja kuva-analyysissä, puheanalyysissä ja puheen tunnistuksessa, signaalinkäsittelyssä, tietoliikenteessä, teollisissa ja muissa mittauksissa, prosessin ohjauksessa, robotiikassa, elektroniikkapiirien suunnittelussa, fysiikassa, kemiassa, biolääketieteellisissä sovelluksissa, neurofysiologisessa tutkimuksessa, tiedonkäsittelyssä ja tietoanalyysissä, kielitieteessä sekä mm. tekoälyyn ja matematiikkaan liittyvissä ongelmissa [Kohonen, 2001]. Cottrell ja Verleysen [2006] mainitsevat SOM:ia käytettävään tiedonlouhintatehtävissä kuten luokittelussa, klusteroinnissa, epälineaarissa

projektiossa, aikasarja-analyysissa ja monidimensioisen datan visualisoinnissa. SOM on suosittu kaupallisissakin projekteissa, joissa sitä käytetään vaikeiden ongelmien ratkaisuun. Käyttöalueet laajenevat koko ajan, viime aikoina SOM:ia on alettu soveltaa talous- ja johtamistieteissäkin. [Cottrell and Verleysen, 2006]

2.3. Oppiminen

Tässä kohdassa käsitellään SOM-algoritmin periaatteita alkuperäisen inkrementaalisen algoritmin kautta, jota tässäkin tutkimuksessa käytetään. Inkrementaalisuus viittaa verkon opetustapaan ja se tarkoittaa, että painoja päivitetään jokaisen syötevektorin jälkeen, kun taas toista tyyliä edustaa eräopetus (batch training), jossa päivitykset tehdään vasta kaikkien syötteiden esittämisen jälkeen [Matlab, 2007]. SOM:sta on olemassa myös tällainen eräopetusversio (batch map), joka onkin paljon inkrementaalista versiota nopeampi [Kohonen, 2005].

Oppimisalgoritmeja käsiteltiin alakohdassa 2.1.3.1 ja todettiin SOM:n käyttävän ohjaamatonta oppimista. SOM:sta on kuitenkin olemassa myös ohjattua oppimista käyttävä versio (supervised SOM) [Kohonen, 2001], mutta tässä tutkimuksessa käytetään ohjaamatonta oppimista käyttävää SOM:ia.

Näiden mainittujen variaatioiden lisäksi SOM:sta on kehitetty lukuisia muitakin versioita, mutta niitä ei käsitellä tässä tutkimuksessa.

2.3.1. Ohjaamattoman oppimisen eduista

Ohjaamattomaan oppimiseen perustuvien menetelmien käyttäminen on ohjattuja menetelmiä vaivattomampaa, koska kaikkea opetukseen käytettävää dataa ei tarvitse etukäteen analysoida, vaan tämä riittää tehdä verkon kalibrointiin käytettävälle datalle. Joissakin tilanteissa tällainen etukäteen tehtävä syötteiden luokittelu voi olla myös hyvin vaikeaa, koska yksittäiset tapaukset saattavat ilmentää useita tutkittavia luokkia. Tällaista luokkien päällekkäisyyttä ilmenee esimerkiksi EEG-datassa. Ohjaamattomalla opetuksella on myös mahdollista löytää tutkittavasta datasta uusia ilmiöitä. [Joutsiniemi *et al.*, 1995]

2.3.2. Neuroneiden painovektorit

Jokaiseen neuroniin liittyy painovektori $m_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{in}]^T \in \mathfrak{R}^n$, joka toimii neuronin *mallivektorina* (model vector) eli *mallina*. Näitä painovektoreita kutsutaan myös *referenssivektoreiksi* (reference vector) ja myös *koodivektoreiksi* (code vector, codebook vector), jolloin niiden muodostamaa hilaa voidaan sanoa *koodikirjaksi*. [Kohonen, 2001]

2.3.3. Painovektoreiden alustaminen

Ennen opettamista neuroneiden mallivektorit täytyy alustaa. Alustaminen on mahdollista tehdä jopa satunnaisilla arvoilla. Tätä tapaa aluksi käytettiinkin SOM:n vahvan itseorganisoitumistaipumuksen osoittamiseen, eli jopa tällaisesta alkuasetelmasta mallit lopulta pystyvät järjestymään. Käytännössä SOM:ia sovellettaessa tämä ei kuitenkaan ole paras eikä edes suositeltava tapa verkon alustamiseen. [Kohonen, 2001; Kohonen, 2005]

Satunnaisilla arvoilla alustetut painovektorit voivat myös olla suorastaan ongelmallisia, koska verkko saattaa olla suppenematon tai erittäin hidas oppimaan ja lisäksi mikäli monet painovektoreista ovat alustuneet hyvin erisuuntaisiksi kuin suurin osa opetusdatan syötevektoreista, voi verkolle opetuksessa tulla hyvin huono neuronien käyttösuhde ja sen myötä heikko luokituskyky [Juhola, 2004].

Järkevämpää on käyttää jotain järjestettyä alustusta. Esimerkiksi lineaarinen alustus (linear initialization) on todettu hyväksi menetelmäksi. Tässä menetelmässä alustus tehdään syötedatan kahden suurimman pääkomponentin ominaisvektoreita käyttäen ja kartta saadaan jo tässä vaiheessa melko hyvin organisoituneeksi, mikä nopeuttaa koko prosessia ja mahdollistaa kapeamman naapurustofunktion ja pienemmän oppimisnopeuden käyttämisen. [Kohonen, 2001; Kohonen, 2005]

Hyvä alustusmenetelmä on myös painovektoreiden asettaminen samoiksi arvoiksi. Tämä menetelmä mahdollistaa kaikille neuroneille aluksi yhtä suuren todennäköisyyden tulla käytetyksi opetusprosessissa eikä aseta solmuja epäoikeudenmukaiseen asemaan, kuten satunnaisessa alustuksessa on mahdollista käydä [Juhola, 2004].

2.3.4. Painovektoreiden muokkaaminen

Käsitellään seuraavaksi painovektoreiden muokkaamista Kohosen [2001] mukaan.

Opetuksessa verkolle syötetään opetusaineiston syötevektorit ja jokaista opetuksessa käytettävää syötevektoria $x = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]^T \in \mathfrak{R}^n$ verrataan kaikkiin mallivektoreihin ja kullekin syötevektorille etsitään sitä eniten muistuttava malli. Neuronia, johon tämä mallivektori liittyy, kutsutaan voittajasolmuksi. Vertailu voidaan tehdä millä tahansa menetelmällä, mutta usein voittajasolmu määritetään syötevektorin ja mallivektorin välisellä euklidisella etäisyydellä $\|x - m_i\|$, jolloin voittajasolmu c on se solmu, jonka mallivektorin euklidinen etäisyys syötevektorista on pienin:

$$c = \arg \min_i \{\|x - m_i\|\} \text{ eli}$$

$$\|x - m_c\| = \min_i \{\|x - m_i\|\}.$$

Vaihtoehtoisesti voidaan käyttää esimerkiksi fysiologisia prosesseja muistuttavaa syötevektorin ja mallivektorin pistetuloon perustuvaa vertailujärjestelmää.

Kun voittajasolmu on löydetty, sen painovektoria muutetaan vielä enemmän syötevektorin kaltaiseksi ja näin tehdään myös voittajasolmun naapurustolle, ts. hilassa solmun ympärillä tietyn välimatkan päässä sijaitseville solmuille. Painovektorin päivitys tapahtuu:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)]$$

jossa

$t = 0, 1, 2, \dots$ on aikaa kuvaava kokonaislukuarvoja saava muuttuja ja

$h_{ci}(t)$ on naapurustofunktio, joka määritetään yleensä solmun i ja voittajasolmun c sijainnillisen etäisyyden sekä ajan funktiona:

$$h_{ci}(t) = h(\|r_c - r_i\|, t)$$

jossa

$r_c \in \mathfrak{R}^2$ on voittajasolmun c sijaintivektori hilassa

$r_i \in \mathfrak{R}^2$ on solmun i sijaintivektori hilassa.

Mitä kauempana solmu sijaitsee voittajasolmusta sitä pienemmän arvon naapurustofunktio saa eli kun $\|r_c - r_i\|$ suurenee, $h_{ci} \rightarrow 0$. Naapurustofunktiota on pienennettävä myös opetuksen aikana: $h_{ci}(t) \rightarrow 0$ kun $t \rightarrow \infty$.

Yleisesti käytettyjä naapurustofunktioita ovat yksinkertainen naapurijoukko ja Gaussin funktio. Esitellään seuraavaksi näistä ensimmäistä, koska sitä käytetään myös tässä tutkimuksessa. Naapurijoukko eli naapurusto käsittää hilassa voittajasolmun vieressä tietyllä säteellä sijaitsevat solmut. Yleensä tätä sädettä pienennetään monotonisesti ajan

kuluessa ja kun naapurustoa kuvataan symbolilla N_c , voidaan määrittää: $N_c = N_c(t)$. Solmun i kuuluessa tähän naapurustoon $N_c(t)$ käytetään sille naapurustofunktiona:

$$h_{ci}(t) = \alpha(t)$$

jossa

$\alpha(t)$ oppimismuutos ja saa arvoja väliltä $]0,1[$. Yleensä arvoja pienennetään monotonisesti ajan kuluessa.

Mikäli solmu i ei kuulu naapurustoon $N_c(t)$, $h_{ci}(t) = 0$ eli tällöin solmun painovektoria ei päivitetä kyseisellä hetkellä t esitetyn syötevektorin tapauksessa.

Topografisesti läheisten neuroneiden muuttaminen samaan suuntaan aiheuttaa verkon pinnassa paikallista tasoittumista eli relaksaatiota. Kun opetusprosessin aikana näitä paikallisia tasoittumisia tapahtuu verkon pinnan eri paikoissa, johtaa tämä lopulta koko verkon laajuiseen järjestykseen. Neuronit järjestyvät kuvaamaan syötteitä niin, että lähellä toisiaan sijaitsevat neuronit edustavat syötedatassa toisiaan muistuttavia syötevektoreita. Naapurustofunktiolla $h_{ci}(t)$ on suuri merkitys verkon relaksaatiolle. Sen keskimääräinen leveys ja muoto määrittävät verkon elastisuuden eli sen, kuinka jäykästi neuroneiden painovektorit soveltuvat syötedatapisteisiin.

2.3.5. Oppimisen vaiheet ja opetusparametrit

SOM:n opetuksessa on erotettavissa kaksi toisistaan eroavaa vaihetta. Ensin on karkeampi *järjestymisvaihe* (ordering phase) ja sitä seuraa *suppenemisvaihe* (final convergence phase) [Kohonen, 2001], jota voidaan kutsua myös *hienosäätövaiheeksi* (tuning phase) [Matlab, 2007].

Järjestymisvaiheen aikana [Matlab, 2007] neuroneiden painovektorit järjestyvät approksimoimaan syötedatan todennäköisyysfunktioita [Kohonen, 2001]. Syötedatan jakauma siis vaikuttaa neuroneiden järjestykseen kartalla. Jos syötevektorit täyttävät tasaisesti syöteavaruuden, myös neuronit asettuvat edustamaan niitä tasaisin etäisyyksin toisistaan. Vastaavasti syötedatan ollessa epätasaisesti jakautunutta neuronit keskittyvät alueille, joissa syötteitä esiintyy enemmän. [Matlab, 2007] Järjestymisvaihe voi kestää esimerkiksi 1000 askelta [Kohonen, 2001], kun yhdellä askeleella tarkoitetaan yhden syötevektorin esittämistä verkolle. Yhtä kokonaisen syötejoukon esittämistä verkolle ja painovektoreiden päivittämistä tämän perusteella kutsutaan *epookiksi* (epoch) [Matlab, 2007].

Hienosäätövaiheessa painovektorit levittäytyvät tasaisemmin mutta säilyttävät järjestyksenvaiheen aikana muodostetun topologisen järjestyksen [Matlab, 2007]. Tämän vaiheen tulisi kestää melko kauan. Askeleita on tehtävä vähintään 500 kertaa verkon neuroneiden lukumäärän verran, koska verkolla tehtävän mallintamisen tilastollinen tarkkuus riippuu tämän vaiheen kestosta. [Kohonen, 2001]

Optimaaliset opetusparametrit joudutaan etsimään melko pitkälle kokeellisesti, mutta niiden vaihtelu ei kuitenkaan herkästi vaikuta tuloksiin [Juhola, 2004]. On kuitenkin muutamia naapuruston ja oppimismopeuden valintaan liittyviä ohjeita, jotka kannattaa huomioida. Kartanlaajuuden järjestymisen onnistumiseksi naapuruston on oltava alussa riittävän suuri, esimerkiksi yli puolet verkon läpimitasta [Kohonen, 2001] tai jopa kaikki verkon solmut [Juhola, 2004]. Naapurustoa pienennetään opetuksen jatkuessa lineaarisesti niin, että se järjestyksenvaiheen loppuessa käsittää esimerkiksi enää voittajasolmun viereiset naapurit. Hienosäätövaiheen ajan naapurusto yleensä pidetään koko ajan samankokoisena. Näin menetellen voidaan välttää toisistaan hyvin erillisten alueiden muodostuminen, joiden välillä järjestyssuunnat muuttuvat epäjatkuvasti. [Kohonen, 2001]

Oppimismopeuden $\alpha(t)$ suhteen menetellään vastaavasti. Sen arvo pidetään huipussaan opetuksen alussa, josta sitä pienennetään monotonisesti. Verkon alustustavasta riippuu, kuinka suurella arvolla täytyy aloittaa. Jos neuroneiden painovektorit on alustettu satunnaisilla arvoilla, aloitetaan korkeista arvoista, jopa läheltä arvoa yksi. Jos taas kartta on opetusta aloitettaessa jo alkuun järjestetty syötedatan todennäköisyysfunktio mukaiseksi, voidaan aloittaa matalalla oppimismopeudella ja tällöin myös naapurusto voidaan pitää jo alussa pienenä. Hienosäätövaiheen aikana oppimismopeudelle käytetään pieniä arvoja, esimerkiksi arvoja 0,02 tai tämän alle, pienentäen sitä joko lineaarisesti tai eksponentiaalisesti. [Kohonen, 2001]

Opetusparametreja käsitellään lisää kohdassa 4.4 eritellessä tässä tutkimuksessa tehtyjä valintoja niiden suhteen.

2.4. Kalibrointi

SOM on opetuksen jälkeen kalibroitava, jotta sitä voidaan käyttää luokittelutehtävissä [Bach *et al.*, 2003; Kohonen, 2001]. Kalibroinnin tarkoituksena on määritellä, mitä tutkittavia asioita neuronit edustavat. Kalibrointia kutsutaankin myös solmujen

nimeämiseksi tai merkitsemiseksi (labeling). Kalibrointi tehdään käyttämällä tunnettua opetusdataa, ts. opetusdatan osajoukkoa, joka on etukäteen analysoitu ja tulkittu ja josta tiedetään, mitä tutkittavia asioita yksittäiset syötteet edustavat. Esimerkiksi luokitteluongelmissa tulee tietää, mitä luokkia syötteet edustavat. Tämä kalibrointidata syötetään opetetulle verkolle ja tutkitaan, mihin neuroneihin syötteet kuvautuvat eli mitkä ovat syötteiden voittajasolmut ts. eniten syötteitä muistuttavat painovektorit. Osa neuroneista ei välttämättä edusta mitään syötteitä. [Kohonen, 2001]

Kun verkko on kalibroitu, voidaan sen avulla selvittää tuntemattoman syötedatan ominaisuuksia ja luokitella syötevektoreita. Verkolle syötetään tällaista dataa ja syötevektorit kuvautuvat parhaiten niitä vastaaviin mallivektoreihin. Kun kalibroinnin ansiosta tiedetään, mitä luokkia mallivektorit edustavat, saadaan näin myös selville tuntemattoman datan ominaisuuksia. [Kohonen, 2001] Vastaavalla tavalla voidaan myös testata verkon onnistuneisuutta. Opetetulle ja kalibroidulle verkolle syötetään opetuksessa käyttämätöntä tunnettua dataa ja tarkistetaan, kuvautuuko data oikeita asioita edustaviin neuroneihin, ts. osaako verkko tunnistaa datan oikein. Näin menetellään myös tässä tutkimuksessa. Diagnooseiltaan tunnettuja verkolle tuntemattomia testitapauksia käytetään verkkorakenteiden luokittelukyvyyn arvioimiseen.

SOM:n kalibrointiin käytetään monenlaisia menetelmiä, siihen on kehitetty mm. tilastollisiin oletuksiin klustereiden muodosta perustuvia algoritmeja [Bach *et al.*, 2003]. Tässä tutkimuksessa käytettyä verkkojen kalibrointimenetelmää käsitellään kohdassa 4.3.

3. Aineistosta ja tutkimusongelmasta

Tässä tutkimuksessa luokiteltiin huimauspotilaiden dataa korvalääketieteen alalta Kohosen itseorganisovilla verkoilla. Tavoitteena oli rakentaa neuroverkkomalleja, jotka pystyisivät mahdollisimman hyvin luokittelemaan vaikeasti toisistaan erotettavia tauteja, joiden diagnosointi on hyvin haastavaa parhaille erikoislääkäreillekin.

3.1. Aineisto ja tutkimusongelma

Tutkimus on jatkoa aikaisemmalle [Juhola *et al.*, 2001], jossa tämän saman aineiston varhaisempaa versiota luokiteltiin käyttäen monikerroksisia perceptron-verkkoja ja Kohosen itseorganisovivia verkkoja. Aineisto oli kerätty korva-, nenä- ja kurkkutautien klinikalla Helsingin yliopistollisessa keskussairaalassa ja käsitti ensin 564 tapausta [Auramo *et al.*, 1993; Kentala *et al.*, 1995]. Kyseistä tutkimusta varten aineisto laajennettiin 883 tapaukseen, jolloin siinä olevat tautitapaukset sisälsivät tämän tutkimuksen aineiston pääaineiston eli ensimmäisen osajoukon kokonaan ja lisäksi pienemmät määrät tautiluokkia 7, 8 ja 9: 16, 39 ja 13 tapausta. Aineistossa oli kaikkiaan 170 muuttujaa, joista kuitenkin vain osassa oli tieto riippuen potilaan oireista ja tehdyistä tutkimuksista. Näistä muuttujista oli päätöspuita käyttäen löydetty 38 tärkeintä [Viikki *et al.*, 1999], joista edelleen tärkeimmiksi havaittiin 9 ja lopulta 5 muuttujaa.

Juhola *et al.* [2001] luokittelivat 5:tä, 6:ta ja 8:aa tautia MLP- ja SOM-verkoilla käyttäen edellä mainittuja 5:tä, 9:ää ja 38:aa tärkeintä muuttujaa. Pienimmässä viiden tautiluokan luokittelussa mukaan otettiin viisi suurinta tautiluokkaa. Kuutta tautiluokkaa luokiteltaessa yhdistettiin edelliseen loput pienimmät yhdeksi luokaksi. Suurimpaan testattuun tautimäärään sisällytettiin kaikki aineiston taudit kaksi pienintä yhdistäen.

Tehtäessä luokittelua suurimmalle määrälle tautiluokkia käyttämällä 5:tä ja 9:ää muuttujaa onnistuttiin neljän suurimman tautiluokan kanssa hyvin, mutta loppuja ei pystytty luokittelemaan onnistuneesti. Viittä tautiluokkaa luokiteltaessa samoja muuttujia käyttäen tulokset parantuivat vain tautiluokkien 2 ja 5 osalta. SOM ja MLP pärjäsivät suurimpien tautiluokkien kanssa yhtä hyvin, mutta 38:aa tärkeintä muuttujaa käytettäessä MLP tunnisti vain suurimman tautiluokan, kun SOM onnistui kuitenkin hiukan paremmin, tosin senkin luokittelukyky huononi merkittävästi. Tutkimuksessa

todettiin, että jotta olisi pystytty hyödyntämään 38 muuttujan käytön myötä lisääntynyt tietomäärä, olisi tautiluokkien tapausten määrää pitänyt lisätä verkkojen topologioiden kasvaessa. MLP:lle todettiin kyseistä dataa käytettäessä opetusdatan kooksi riittävän viisi kertaa verkon yhteyksien määrä, mutta SOM:lle opetusdatan koko ei ollut näin tärkeä. Sen sijaan tautiluokkien mahdollisimman tasainen jakauma olisi ollut tarpeen molemmille verkkotyypeille, jotta kaikki taudit olisi ollut mahdollista luokitella onnistuneesti.

Tässä tutkimuksessa jatkettiin saman aineiston uudemman version luokittelua käyttäen Kohosen itseorganisoituvia verkkoja, joiden soveltuvuudesta tämän aineiston luokitteluun haluttiin lisätietoa. Luokittelussa toivottiin pystyttävän käyttämään kaikkia 38:aa tärkeintä muuttujaa. Jos luokittelu ei onnistuisi näin suurella muuttujamäärällä suhteessa pieneen tapausten määrään, asetettiin vaihtoehdoiksi 13 ja 9 muuttujaa. Tällöin 13 muuttujan joukko käsitti 9 tärkeimmän muuttujan lisäksi 4 muuta tärkeää muuttujaa. Tärkeimpiä 5:tä muuttujaa ei enää oltu kiinnostuneita käyttämään muuttujamäärän suppeuden vuoksi, ja koska sen tulokset tiedettiin jo riittävän hyvin. Tämän lisäksi oltiin kiinnostuneita saamaan selville, pystyttäisiinkö myös pienempiä tauteja [7, 10] luokittelemaan onnistuneesti. Taulukossa 1 esitetään aineiston tautiluokat.

Selostetun aikaisemman tutkimuksen [Juhola *et al.*, 2001] jälkeen aineistoa on luokiteltu useissa muissakin tutkimuksissa [Autio *et al.*, 2007; Juhola and Laurikkala, 2003; Juhola and Siermala, 2005; Siermala and Juhola, 2006; Siermala *et al.*, 2007]. Soveltamalla pääkomponenttianalyysiä [Juhola and Laurikkala, 2003] pystyttiin luokittelussa käyttämään kaikkia 38:aa tärkeintä muuttujaa ja parantamaan sensitiivisyyttä ja positiivista ennustearvoa. Negatiivista tässä lähestymistavassa kuitenkin oli se, että alkuperäisten muuttujien sijaan käytettiin niiden muunnoksia. Tässä suhteessa parempi menetelmä tautiluokkien epätasaisen jakauman aiheuttamien ongelmien kiertämiseksi oli tehdä oma MLP-verkko jokaiselle kuudelle luokiteltavalle tautiluokalle [Juhola and Siermala, 2005; Siermala and Juhola, 2006; Siermala *et al.*, 2007]. Kunkin verkon opetukseen käytettiin vain yhden tautiluokan tapaukset ja tämän komplementtina yhtä suuri määrä keinotekoisesti luotuja tapauksia, jotka olivat muuttuja-avaruudessa tasaisesti jakautuneita tautiluokan tapausten ulkopuolisessa arvoalueessa. Näin saatiin käytettyä kaikki 38 tärkeintä muuttujaa muuntamattomana alkuperäisessä muodossaan. Aikaisemman tutkimuksen [Juhola *et al.*, 2001] luokittelutuloksiin saatiin näin parannusta vähintään 10 % ja esimerkiksi vaikeasti luokiteltava pieni tautiluokka 4 tunnistettiin parhaimmillaan 100 %:n sensitiivisyydellä.

Taulukko 1. Diagnoosien luokat.

Luokka	Diagnoosi
1	akustikusneurinooma, akustikus neurinoma, kuulotasapainohermon neurilemmooma, vestibulaari schwannoma (VS) (acoustic neurinoma, vestibular schwannoma)
2	hyvänlaatuinen kohtauksittainen asentohuimaus (benign paroxysmal positional vertigo)
3	Ménièreen tauti (Ménière's disease)
4	äkillinen tai puoliäkillinen kuulon voimakas heikkeneminen tai kuurous (sudden deafness)
5	traumaattinen huimaus (traumatic vertigo)
6	vestibulaarineuriitti (vestibular neuritis)
7	hyvänlaatuinen rekurrentti asentohuimaus (benign recurrent vertigo)
8	vestibulopatia
9	sentraalinen leesio (central lesion)
10	migreeninen heitehuimaus (migraine associated dizziness)
-1	ei diagnoosia

Tautiluokkien epätasaisen jakauman negatiivista vaikutusta luokitteluun yritettiin vähentää myös aluksi selostetussa tutkimuksessa [Juhola *et al.*, 2001] tasoittamalla jakaumaa sekä suuria tautiluokkia pienentäen että tapauksia lisäten [Swingler, 1996] pieniin tautiluokkiin, joista ensimmäisellä tavalla pienen tautiluokan luokittelua onnistuttiin parantamaan verkon kokonaissuoriutumisen kuitenkin huonontuessa tiedon vähenemisen seurauksena. Tapauksen keinotekoinen lisääminen ei tuolloin vaikuttanut tuloksiin, mutta kokeiltaessa tätä myöhemmin edeltävää kehittyneemmällä tavalla [Autio *et al.*, 2007] samalle aineistolle, onnistui pienten luokkien luokittelu paremmin. Kokonaistarkkuuteen sillä ei kuitenkaan ollut merkittävää vaikutusta, koska suurimman luokan tulokset vastaavasti huononivat. Tässä tutkimuksessa tautiluokkien jakaumaa ei tasoitettu, koska luokittelun onnistumista haluttiin selvittää alkuperäisiä tautitapausten määriä käyttäen.

3.2. Aineiston osajoukot

Aineistossa oli yhteensä 1159 tapausta ja se jaettiin tutkimuksessa neljään osajoukkoon. Ensimmäinen ja tämän tutkimuksen kannalta tärkein osajoukko eli pääjoukko oli diagnoosiluokkia [1, 6] koskevat 815 tapausta, joissa diagnoosit olivat aineiston varmimpia ja niiden luokittelun odotettiin onnistuvan parhaiten. Näillä tapauksilla tehtiin neuroverkoille sekä opetus että testaus.

Toisen osajoukon tapaukset olivat samojen tautiluokkien tapauksia kuin pääjoukon mutta diagnooseiltaan epävarmempia. Tautien diagnosoinnissa oli ollut epävarmuutta, koska jopa parhaiden erikoislääkäreiden on vaikea erottaa näitä tauteja toisistaan ja osassa tapauksista tauti saattaa myös ajan kuluessa muuttua toiseksi taudiksi. Nämä tapaukset käytettiin vain ensimmäisen osajoukon parhaan verkon testaukseen, koska niille ei huonomman laadun takia kannattanut opettaa omia verkkoja. Luokittelutuloksistakaan ei odotettu yhtä hyviä kuin ensimmäiselle osajoukolle.

Kolmas osajoukko oli aineiston tapaukset, joissa diagnoosi oli [7, 10]. Tämä osajoukko yhdistettiin kaikkien muiden diagnoosillisten tapausten kanssa ja tällä datalla sekä opetettiin että testattiin verkkoja. Tautiluokka 10 jouduttiin kuitenkin jättämään kokonaan pois, koska sen tapausten määrä ei riittänyt kymmeneen erilliseen testijoukkoon ja niiden opetusjoukkovastineisiin jakamiseen. Tässä osajoukossa oli mukana niin pieniä luokkia, että niiden luokittelun onnistumista pidettiin hyvin epätodennäköisenä, mutta verkkojen suoriutuminen haluttiin testata.

Neljäs osajoukko käsitti tapaukset, joiden diagnooseja ei tiedetty (-1). Tämä data vain syötettiin pääjoukolla opetetulle verkolle, koska oli mielenkiintoista nähdä, kuinka verkko datan luokittelisi ja mikä olisi luokkien välinen jakauma. Suurimman osan tapauksista odotettiin luokittuvan Ménièreen taudiksi, joka oli enemmistönä tässä aineistossa ja joksi myös epävarma lääkäri saattaisi diagnosoida tämän aineiston tauteja.

Taulukoissa 2 – 5 esitetään tautien määrät aineiston osajoukoissa.

Taulukko 2. Luokkajakauma aineiston ensimmäisessä osajoukossa.

Tautiluokka	Lkm	%
1	130	16,0
2	146	17,9
3	313	38,4
4	41	5,0
5	65	8,0
6	120	14,7
Yhteensä	815	100,0

Taulukko 3. Luokkajakauma aineiston toisessa osajoukossa.

Tautiluokka	Lkm	%
1	1	0,9
2	27	23,3
3	37	31,8
4	6	5,2
5	8	6,9
6	37	31,9
Yhteensä	116	100,0

Taulukko 4. Luokkajakauma aineiston kolmannessa osajoukossa yhdistettynä kaikkien muiden diagnoosillisten tapausten kanssa. Tauti 10 jätettiin pois sen tapausten vähäisen määrän takia, seitsemää tapausta ei voitu jakaa kymmeneen erilliseen testijoukkoon ja niiden opetusjoukkovastineisiin.

Tautiluokka	Lkm	%
1	131	12,7
2	173	16,8
3	350	34,1
4	47	4,6
5	73	7,1
6	157	15,2
7	20	1,9
8	55	5,3
9	24	2,3
Yhteensä	1030	100,0

Taulukko 5. Aineiston neljännessä osajoukossa oli vain diagnoosittomat tapaukset.

Tautiluokka	Lkm	%
-1	122	100,0
Yhteensä	122	100,0

3.3. Aineiston esikäsittely

Neuroverkkojen opetukseen tulee käyttää riittävästi ongelman hyvin kattavaa korkealaatuista dataa. Laadun varmistamiseksi data on esikäsiteltävä ennen käyttöä opetukseen.

3.3.1. Puuttuvat arvot

Neuroverkoille annettavan aineiston täytyy olla täydellistä, ts. siinä ei saa olla puuttuvia arvoja. Jos parametrilla on korkeintaan 20 – 30 % puuttuvia arvoja, ne voidaan paikata dataan ja sovellukseen soveltuvalla arviointimenetelmällä. [Juhola, 2004]

Tämän tutkimuksen aineistossakin oli puuttuvia arvoja, mutta ne oli jo edeltävän tutkimuksen [Juhola *et al.*, 2001] yhteydessä korvattu luokkakohtaisilla estimaateilla eli oli arvioitu, mitä arvoja muuttujassa on keskimäärin kyseisen taudin tapauksessa. Binääri- ja luokitteluasteikon muuttujille oli laskettu luokkamoodit sekä järjestysasteikon muuttujille ja jatkuville muuttujille luokkamediaanit.

3.3.2. Tuloluokkien jakauma

Tulosluokkien tulisi olla mahdollisimman samankokoisia, jotta neuroverkko oppisi luokat toisistaan erottavat ominaisuudet luokkien todennäköisyyden sijaan [Juhola, 2004]. Tämä ei toteudu tässä aineistossa. Luokat ovat hyvin erikokoisia, ja osa niistä on erittäin pieniä. Kuten kohdassa 3.1 kerrottiin, aineiston jakaumaa onkin yritetty aikaisemmissa tutkimuksissa tasoittaa suuria tautiluokkia pienentämällä [Juhola *et al.*, 2001] ja pieniä tautiluokkia suurentamalla [Autio *et al.*, 2007; Juhola *et al.*, 2001]. Näistä ensimmäistä menetelmää käyttäen kokonaissuoriutumisen huononi ja jälkimmäinen ei tuonut siihen merkittävää parannusta. Tässä tutkimuksessa luokittelun

onnistumista haluttiin selvittää alkuperäisiä tautitapausten määriä käyttäen, joten tautiluokkien jakaumaa ei tasattu.

3.3.3. Muuttujien normalisointi

Syötemuuttujat kannattaa normalisoida, jotta ne tulisivat neuroverkossa tasapuolisesti käytetyiksi. Jos joillakin syötemuuttujilla olisi toisia suurempia arvoja, niillä olisi pienempiarvoisia muuttujia suurempi vaikutus neuroverkon vasteisiin. Normalisointitekniikoita on useita, joista yksinkertaisin on arvojen skaalaus samalle vaihteluvälille, esimerkiksi välille [0, 1]. Vaihteluvälille normalisoinnissa kuitenkin esiintyy tiedon vääristymisen ongelmaa. Parempi normalisointimenetelmä on standardoida muuttujat. Tässä menetelmässä arvoista vähennetään keskiarvo ja sen jälkeen jaetaan keskihajonnalla, jolloin keskihajonta saadaan arvoon yksi ja keskiarvo arvoon nolla. [Juhola, 2004] Tässä tutkimuksessa aineistolle tehtiin muuttujien standardointi.

3.4. Aineiston jakaminen opetukseen ja testaukseen

Neuroverkot opetettiin ja testattiin käyttäen kymmenkertaista *ristiinvalidointia* (cross-validation). Kymmenen opetus- ja testijoukkoparia muodostettiin sisällyttämällä kuhunkin testijoukkoon 10 % aineistosta ja opetusjoukkoon vastaavasti 90 % aineistosta. Edeltävässä tutkimuksessa [Juhola *et al.*, 2001] tämä tehtiin 60 %:n opetusjoukon osuudella aineistosta. Opetusjoukon osuutta kasvattamalla verkkojen opetukseen saatiin enemmän tapauksia, minkä toivottiin edesauttavan niiden luokittelukykyyn parantumista.

Kaikki kymmenen testijoukkoa muodostettiin toisistaan täysin erillisiksi. Tapaukset näihin valittiin satunnaisesti, mutta koko datan luokkajakaumaa noudattaen. Koska neuroverkon testaus tulee suorittaa opetusjoukkoon kuulumattomilla tapauksilla, kunkin testijoukon opetusjoukkovastine koostui testijoukkoon sisällyttämättömistä tapauksista.

Itseorganisoituminen on stokastinen prosessi [Kohonen, 2001], kuten neuroverkkojen oppiminen yleensäkin [Haykin, 1994], ja siten sen luokittelutulokset vaihtelevat. Tämän vuoksi jokaisella opetus- ja testijoukkoparilla opetettiin ja testattiin verkko kymmenen kertaa parhaille verkkorakenteille, jolloin verkkorakennetta kohti verkon opetus ja testaus tehtiin 100 kertaa. Parhaita opetusparametreja etsittäessä opetus ja testaus tehtiin

kuitenkin vain kerran kullakin opetus- ja testijoukkoparilla, koska verkkojen opetus oli melko hidasta varsinkin suurten verkkojen tapauksessa. Näin ehdittiin testata kattavammin opetusparametreja.

Luokittelun onnistumista arvioitiin erilaisilla tunnusluvuilla, joita käsitellään kohdassa 4.2. Nämä tunnusluvut laskettiin jokaisesta testitilanteesta, joita oli 10 tai 100 riippuen siitä, oliko kyseessä opetusparametrien optimointitilanne vai parhaan verkkorakenteen testaus. Näistä laskettiin keskiarvot ja -hajonnat sekä mediaanit, jotka olivat verkkorakenteen kokonaistulokset.

4. Testit ja tulokset

Tässä luvussa kuvataan työn empiirinen osuus eli käsitellään tehtyjä neuroverkkoja ja niillä saatuja tuloksia.

4.1. Käytetyt laitteistot ja ohjelmistot

Neuroverkot tehtiin käyttäen Matlab-ohjelmistoa [Matlab, 2007], jonka Neural Network Toolbox -lisäosassa (V5.1) on funktioita neuroverkkojen luontia ja käsittelyä varten. Koska neuroverkkojen kehitysprosessi on monivaiheinen ja sitä jouduttiin toistamaan useita kertoja, tehtiin skriptejä ja funktioita aineiston jakamiseen osajoukkoihin, osajoukkojen jakamiseen opetus- ja testijoukkoihin, neuroverkkojen luontiin, opetukseen ja testaukseen, opetuksen jälkeiseen verkon kalibrointiin sekä luokittelutulosten laskentaan ja niiden analysointia tukemaan.

SOM-algoritmin kompleksisuus on lineaarinen suhteessa syötevektoreiden määrään ja neliöllinen suhteessa neuroneiden määrään. Siten se soveltuu käytettäväksi suurtenkin datamäärien kanssa, mutta suurten verkkojen opetus vie paljon aikaa. [Vesanto and Alhoniemi, 2000] Suurten verkkojen opetus olikin Matlabissa hyvin hidasta, etenkin kun se tehtiin 10x1-ristiinvalidointia käytettäessä 10 kertaa jokaista testattavaa verkkorakennetta kohti ja lopullisissa testeissä 10x10-ristiinvalidoinissa 100 kertaa. Aluksi käytössä olleesta tietokoneesta (Mobile AMD Athlon 64 3000+ 2,0 GHz, 1,5 GB muistia ja käyttöjärjestelmä Windows XP) opetus vei lisäksi kaikki tehot, eikä koneella voinut tehdä mitään muuta samanaikaisesti. Ajan käytön tehostamiseksi kone vaihdettiin tehokkaampaan: Intel Core 2 Duo 2,2 GHz, 2 GB muistia ja käyttöjärjestelmä Windows Vista.

Skriptien ajon nopeutumista testattiin kahdella testiajolla, joissa molemmissa suoritettiin 10x1-ristiinvalidointi eli käsiteltiin 10 SOM:ia. Uudella koneella ajot sujuivat lähes kolmanneksen nopeammin ja veivät koneesta tehoja vain vaihdellen 50 – 75 %, joka oli tervetullut parannus edellisen koneen 100 %:iin, sillä se mahdollisti muun samanaikaisen käytön.

Ajojen nopeutumiseen saattoi oman lisänsä tuoda myös uudempi versio Matlabista [2007], johon myös jouduttiin päivittämään tietokoneen vaihdon myötä, koska edellisellä koneella käytetyn Matlabin version [Matlab, 2006] ei luvattu toimivan uuden

koneen Windows Vista -käyttöjärjestelmässä. Uudemmassa versiossa oli tuki moniytimisyydelle, mikä saattoi joiltain osin nopeuttaa skriptien suoritusta. Lisäksi sen suorituskykyä oli parannettu edellisessä koneessa käytetystä versiosta joillakin alustoilla, mm. kyseisessä koneessa olleessa XP:ssä. Siten uudemman Matlabin asentaminen vanhallekin koneelle olisi saattanut hiukan pienentää suoritusajojen eroa.

Näiden testien jälkeen skriptien ajoa pyrittiin edelleen nopeuttamaan lisäämällä uuteen koneeseen muistia, jolloin muistin kokonaisuudeksi tuli 3 GB. Koska verkkojen opetus vei edelleen paljon aikaa, alettiin verkkoja tehdä myös vanhalla koneella ja tätä varten siihenkin asennettiin uudempi versio Matlabista. Tämän jälkeen skriptien suoritusajojen ei enää vertailtu koneiden välillä.

Tarkastellaan vielä verkkorakenteiden opetukseen kulunutta aikaa. Esimerkkisuoritusajojen mittaamiseksi ajettiin testejä, joissa opetettiin opetusparametreiltaan eroavia verkkoja. Kun varsinaisissa verkkorakenteita testaavissa testeissä yksittäisiä verkkoja opetettiin joko 10 tai 100 riippuen siitä, oliko kyse opetusparametrien optimoinnista vai viimeisistä perusteellisemmista testeistä, näissä suoritusajatesteissä opetettiin ajan säästämiseksi vain yksi verkko jokaista testattavaa verkkorakennetta kohti. Suoritusajatestit suoritettiin vain vanhemmalla koneella.

Ensin opetettiin aineiston kolmannelle osajoukolle testatuista verkkorakenteista pienin ja suurin eli verkkorakenteet kokoa 3×3 ja 30×30 ja mitattiin yhden verkon opetukseen kulunut aika. Opetusaineistona käytettiin yhtä kymmenestä verkkorakennetesteistäkin käytetystä opetusjoukosta. Pienen 3×3 -kokoisen verkon opetukseen kului aikaa vain alle 11 sekuntia, kun taas suuren 30×30 -kokoisen verkon opetus kesti peräti yli 3,5 tuntia. Koon lisäksi näillä verkoilla oli eroa myös opetuksen kestossa, koska niissä käytettiin opetuksen hienosäätövaiheelle Kohosen [2001] esittämää verkon koosta riippuvaa minimikestoa, jolloin pienempää verkkoa opetettiin 6 kokonaisepookkia ja suurempaa 489 kokonaisepookkia. Järjestymisvaiheen kesto oli molemmilla 1000 askelta eli 1,08 epookkia. Opetusparametrijohdistelmät olivat samoja, joita käytettiin verkkorakennetesteissä etsittäessä parasta kokoa ja joista kerrotaan alakohdassa 4.5.2.1.

Näiden lisäksi haluttiin havainnollistaa, kuinka pelkkä opetusajan lisääminen vaikuttaa suoritusajoihin ja opetettiin 30×30 -kokoista verkkoa äskeiseen verrattuna kaksinkertaisella kokonaisepookkimäärällä käyttäen järjestymisvaiheeseen ensin saman ajan kuin äsken ja sen jälkeen puolet koko opetuksen kestosta. Näin meneteltiin, koska järjestymisvaiheen uskottiin saattavan olla hienosäätövaihetta hitaampi suorittaa päivitettävän naapuruston ollessa tällöin suurempi. Suoritusajoin kaksinkertaistui opetusajan kaksinkertaistamisen myötä, mutta järjestymisvaiheen osuuden huomattava

lisääminen pidensi suoritukseen kulunutta aikaa enää vain vartin verran. Näitä viimeiseksi testattuja opetusparametriyhdistelmiä ei käytetty verkkorakennetesteissä, mutta aineiston ensimmäiselle osajoukolle opetettiin melko suuriakin verkkoja vielä pitkäkestoisemmin. Kyseisiä opetuksen kestoa optimoivia testejä käsitellään alakohdassa 4.5.1.2. Suoritusaikatesteissä saadut tulokset esitetään yhteenvetona taulukossa 6.

Taulukko 6. Esimerkkejä itseorganisoituvien verkkojen opetukseen kuluneista ajoista. Opetukseen käytetty syötedatajoukko on kaikissa näissä sama. Esitettävät suoritusajat ovat yhden verkon opetukseen kuluneita aikoja. Varsinaisissa verkkorakennetesteissä yksittäisiä verkkoja tehtiin joko 10 tai 100 jokaiselle tutkimuksessa testatulle verkkorakenteelle.

Verkon koko	Kokonaisepookit	Järjestymisvaiheen osuus opetuksesta epookkeina	Suoritus aika
3x3	6	1,08	11 sek
30x30	489	1,08	3 h 36 min
30x30	978	1,08	7 h 14 min
30x30	978	489	7 h 30 min

4.2. Luokittelun onnistumisen arviointimenetelmät

Verkkojen luokittelukykyä mitattiin tunnusluvuilla *sensitiivisyys* (sensitivity), *spesifisyys* (specificity), *positiivinen ennustearvo* (positive predictive value, PPV) ja *kokonaistarkkuus* (total accuracy). Näistä sensitiivisyys kertoo, kuinka hyvin tunnistetaan oikein positiiviset tapaukset ja spesifisyys, kuinka hyvin tunnistetaan oikein negatiiviset tapaukset [Altman and Bland, 1994a]. Positiivinen ennustearvo on testin luotettavuuden mittari ilmaistessaan, kuinka todennäköisesti positiivinen tulos todella kertoo olemassa olevasta taudista [Altman and Bland, 1994b]. Kokonaistarkkuus kertoo oikein tunnistettujen, sekä oikeiden positiivisten että oikeiden negatiivisten tapausten, suhteen kaikkiin tapauksiin.

Esitetään seuraavaksi näiden tunnuslukujen kaavat. Kaavoissa käytettävät lyhenteet tp, fn, fp ja tn viittaavat verkkojen vasteesta jokaiselle tautiluokalle erikseen laskettuihin tapauksiin:

- tp oikeiden positiivisten (true positive) lukumäärä
- fn väärin negatiivisten (false negative) lukumäärä
- fp väärin positiivisten (false positive) lukumäärä

tn oikeiden negatiivisten (true negative) lukumäärä.

Nämä arvot saatiin syöttämällä testidatan jokainen syötevektori verkolle ja vertaamalla verkolta saatua tautiluokitustulosta syötteen todelliseen tiedettyyn tautiin. Oikea ja väärä viittaavat siihen, menikö luokittelu oikein vai väärin ja negatiivinen ja positiivinen, oliko tautia vai ei. Oikea positiivinen tarkoittaa siis, että tauti tunnistettiin oikein, ja väärä negatiivinen, että vaikka testivektori edusti tautia, sitä ei tunnistettu tähän tautiluokkaan kuuluvaksi. Tunnusluvut laskettiin näistä seuraavasti:

$$\text{Sensitiivisyys} \quad r = \frac{tp}{tp + fn} 100\%$$

$$\text{Spesifisyys} \quad q = \frac{tn}{tn + fp} 100\%$$

$$\text{Positiivinen ennustearvo} \quad p = \frac{tp}{tp + fp} 100\%$$

$$\text{Kokonaistarkkuus} \quad t = \frac{tp + tn}{tp + tn + fp + fn} 100\%$$

Nämä tunnusluvut laskettiin jokaiselle tautiluokalle erikseen ja niistä laskettiin 10x10- tai 10x1-ristiinvalidointiajojen keskiarvo, keskihajonta ja mediaani tautiluokittain. Parhaita opetusparametreja etsittäessä käytettiin 10x1-ristiinvalidointia eli kymmenellä opetus- ja testijoukkoyhdistelmällä tehtiin vain yksi ajo eli yksi verkon opetus ja testaus kullakin, jotta ehdittiin testata opetusparametreja kattavammin, koska opetusprosessi oli hyvin aikaavievä. Vasta viimeiset parhaiden verkkorakenteiden testit suoritettiin perusteellisemmalla ja siten luotettavammalla 10x10-ristiinvalidoinnilla, jossa jokaisella opetus- ja testijoukkoyhdistelmällä tehtiin 10 verkon opetusta ja testausta.

Tautiluokkien tunnuslukujen keskiarvoista laskettiin lisäksi verkkorakenteelle keskiarvo sekä tautiluokkien tapausten määrillä painotettuna että painottamattomana. Parhaaksi verkkorakenteeksi valittiin kokonaisuutena parhaat luokittelutulokset tuottava verkkorakenne, jolloin tautiluokkien erikokoisuuden vuoksi tarkkailtiin tautiluokkien tapausten määrillä painotettuja keskiarvoja. Tunnusluvuista seurattiin erityisesti sensitiivisyyttä.

4.3. Verkkojen kalibrointi

Kuten kohdassa 2.4 kerrottiin, SOM on kalibroitava opetuksen jälkeen, jotta sitä voidaan käyttää testidatan luokitteluun. Seuraavassa selostetaan tutkimuksessa käytetty kalibroitimenetelmä eli kerrotaan, kuinka määritettiin, mitä tauteja neuronit opetuksen tuloksena edustivat.

Opetuksen jälkeen opetetulle verkolle syötettiin koko opetusdata, jonka kaikkien syötevektoreiden todelliset tautiluokat olivat tiedossa, ja tarkistettiin syötevektoreiden neuroneihin kuvautuminen. Näistä laskettiin kuhunkin neuroniin sijoittuneiden tautitapausten määrät, joista määritettiin enemmistöperiaatteella, mitä tautia kukin solmu edusti. Osa neuroneista ei välttämättä opetuksen tuloksena edusta mitään syötteitä.

Mikäli eri tauteja edustavia syötevektoreita kuvautui samaan solmuun yhtä paljon, ei enemmistöperiaate enää riittänyt ratkaisemaan solmun edustusta. Kohonen [2001] kehottaa tällaisessa tilanteessa joko valitsemaan voittajan näistä satunnaisesti tai käyttämään jotain muuta menetelmää. Satunnaisen valinnan sijaan haluttiin käyttää jotain perustellumpaa ratkaisutapaa ja harkittiin kahta vaihtoehtoa: valintaa todennäköisyyden perusteella ja valintaa naapurisolmujen tautiedustusten mukaan.

Naapurisolmujen tautiedustusten apuna käyttämisen perusteena oli se, että SOM:ssa neuronit järjestyvät oppimisen aikana kuvaamaan syötedataa siten, että toisiaan muistuttavat syötevektorit kuvautuvat kartalle neuroneihin lähelle toisiaan [Kohonen, 2001]. Näin ollen saman taudin edustajien odotettiin sijoittuvan kartalle neuroneihin lähelle toisiaan, ja jos siis naapurit edustaisivat samoja tauteja kuin tasapelisolmun voittajaehdokkaat, voitaisiin voittajaksi valita näistä se, jota naapureilla on eniten. Tätä menetelmää pidettiin mielenkiintoisena, koska se ei todennäköisyyden perusteella valitsemisen tavoin suosi suurinta tautiluokkaa Ménièreen tautia (tautiluokka 3), johon asentohuimaus (tautiluokka 2) ja vestibulaari schwannoma eli akustikusneurinooma (tautiluokka 1) sekoittuvat joissakin tapauksissa niin luokiteltaessa kuin lääkäreiden diagnooseissakin, ja luokittelun toivottiin siten onnistuvan paremmin.

Todennäköisyyden perusteella valittaessa voittajaksi valitaan tasapeliin yltäneistä taudeista se, joka on todennäköisin voittaja eli jonka määrä opetusdatassa on tautivaihtoehtojen suurin. Todennäköisyysperiaatetta voidaan pitää melko perinteisenä tapana ratkaista tällainen tasapelitilanne, ja vaikka naapurustomallin toimivuus oli mielenkiintoisempi kokeilla, pidettiin todennäköisyysperiaate mukana menetelmien vertailun mahdollistamiseksi.

Esitetään seuraavaksi yhteenvedona ja tarkentaen tasapelitilanteiden ratkaisumallit. Voittajan valitsemista tasapelitaudeista naapurisolmujen tautiedustusten mukaan kutsutaan tästä eteenpäin lyhyemmin naapurustomalliksi. Naapurustomallia kokeiltiin kahdella toisistaan hiukan eroavalla tavalla, joten solmujen tautiedustukset määritettiin yhteensä kolmella tavalla ja näin saatavia tuloksia oli mahdollista vertailla keskenään.

1) **Naapurustomalli 1:** Lähinaapurusto tautien etsintäalueena.

Tässä mallissa otettiin huomioon vain lähinaapurusto eli tasapelineuronin viereiset neuronit. Jos lähinaapuruston neuronit edustivat samoja tauteja kuin keskusneuronin tasapeliin yltäneet taudit, valittiin keskusneuronin voittajaksi tauti, jota oli naapurustossa eniten. Tässä määrällä tarkoitetaan naapureiden tautiedustuksia eikä naapureihin kuvautuneiden tauteja edustavien syötevektoreiden määriä, eli voittajataudiksi keskusneuronille valittiin tauti, jota edusti suurempi osa naapureita kuin muita keskusneuronin tasapelitauteja. Jos lähinaapurustosta ei löytynyt keskusneuronin tasapelitauteja edustavia neuroneita, valittiin voittaja tasapelitaudeista todennäköisyysperiaatteella. Mikäli naapureilla oli joitakin keskusneuronin tasapelitauteja yhtä paljon, valittiin voittaja kyseisistä taudeista todennäköisyysperiaatteella, ts. vaikka keskusneuronilla olisi ollut useampia tasapelitauteja kuin naapurustossa yhtä suuria määriä olevia kiinnostavia tauteja, olisi todennäköisyysperiaatetta sovellettu vain näihin naapurustosta löytyneisiin tasapelitauteihin.

2) **Naapurustomalli 2:** Laajennettava naapurusto tautien etsintäalueena.

Kun ensimmäisessä naapurustomallissa otettiin huomioon vain lähinaapurusto, oli mahdollista joutua käyttämään todennäköisyysperiaatetta, vaikka hiukan kauempana olisi ollut tilanteen ratkaisun kannalta kiinnostavia tauteja. Tässä mallissa tautien etsintäaluetta laajennettiin, kunnes löydettiin samoja tauteja kuin keskusneuronin tasapelitaudit sellaisina määrinä, että jotain näistä taudeista oli enemmän kuin muita etsittäviä tauteja. Jos siis löydettiin tasapelitauteja edustava naapurusto, jolla oli yhtä suuret määrät etsittäviä tauteja eikä tilannetta siten saatu ratkaistua, laajennettiin etsintöjä edelleen seuraaviin tasapelitauteja edustaviin naapureihin, mutta oltiin kiinnostuneita enää taudeista, joita oli yhtä paljon tällä lähemmällä naapurustolla. Näin meneteltiin, kunnes saatiin jollain naapurustolla tilanne ratkaistua tai kunnes saavutettiin kartan reuna, jolloin voittaja valittiin vielä tässä vaiheessa kiinnostuksen kohteena olevista tasapelitaudeista todennäköisyysperiaatteella.

Esimerkki. Jos keskusneuronilla on tasapelitauteina taudit x , y ja z ja ensimmäisellä löydetyllä tasapelitauteja edustavalla naapurustolla on yhtä paljon tauteja x ja y , jatketaan seuraaviin tasapelitauteja edustaviin naapureihin, mutta etsitään enää tauteja x ja y .

3) **Todennäköisyysperiaate**

Tasapelitaudeista valittiin todennäköisin voittaja eli tauti, jota oli eniten opetusdatassa. Jos näitäkin määriä olisi ollut yhtä paljon, olisi voittaja voitu valita näistä yhtä suuria määriä olevista taudeista satunnaisesti, mutta tämän aineiston tapauksessa määrät vaihtelivat.

Esiteltyjä tasapelitilanteiden ratkaisumalleja käytettiin vasta, kun oli ensin määritetty tautiedustukset kaikille sellaisille neuroneille, joissa oli yksi selkeä voittaja eli joiden tautiedustukset olivat määritettävissä enemmistöperiaatteella.

Näistä kolmesta tasapelimallista käytetään tulostarkasteluissa lyhenteitä N1-, N2- ja TN-malli.

4.4. Opetusparametrit

Verkkojen mahdollisimman hyvän luokittelukyvyyn saavuttamiseksi etsittiin parhaita aineistoon ja luokitteluongelmaan sopivia opetusparametreja. Optimaalisen parametriyhdistelmän löytämisen haasteellisuutta lisää se, että parametrivalinnat vaikuttavat toisiinsa ja koska erilaisia yhdistelmiä on siten hyvin suuri määrä, pystyttiin vain osa näistä testaamaan.

Tässä kohdassa käsitellään työssä käytettyjä opetusparametreja ja perusteluja niiden valinnalle. Käsiteltäviä opetusparametreja ovat verkon koko, alustustapa, hilan muoto, naapurusto, opetuksen kesto ja oppimisnopeus. Näitä parametreja käytettiin verkkojen opetuksessa, josta kerrotaan kohdassa 4.5.

4.4.1. Verkon koko

Verkon koon eli neuroneiden lukumäärän on oltava vähintään opetusdatassa olevien tautiluokkien määrä, jotta jokaista tautia edustamaan olisi mahdollista saada neuroni. Neuroneiden määrää kasvatettaessa verkon tarkkuus kasvaa, koska pystytään edustamaan tarkemmin yksittäisiä syötevektoreita, mutta toisaalta sen yleistyskyky alkaa jossain vaiheessa verkkoa suurennettaessa kärsiä neuroneiden oppiessa liian tarkasti yksittäisten syötteiden ominaisuudet, jolloin verkko ei enää suoriudu hyvin uusien tapausten luokittelusta. Tällainen ylisovittuminen on vaarana ainakin, jos neuroneita on paljon enemmän kuin syötteitä [Leflaive *et al.*, 2005]. Ylisovittuminen haluttiin välttää, joten kummallakin opetukseen käytettävällä datan osajoukolla verkon optimaalisen koon etsintä päätettiin viimeiseen $N \times N$ -arvoon, jossa neuroneiden määrä ei vielä ylittänyt opetukseen käytettyjen syötteiden määrää. Etsintä aloitettiin aineiston tautiluokkien määrästä, joka aineiston pääjoukossa oli kuusi ja kolmannessa osajoukossa yhdeksän, ja kokoa kasvatettiin tasaisesti päätearvoonsa. Kolmannessa osajoukossa oli alun perin kymmenen tautiluokkaa, mutta tautiluokka 10 jouduttiin jättämään pois sen tapausten vähäisen määrän vuoksi.

4.4.2. Verkon alustaminen

Verkkojen alustus tehtiin asettamalla kaikkien neuroneiden painovektorit samaan pisteeseen syötteiden arvoalueiden keskelle. Tämä menetelmä on hyvä asettaessaan solmut tasa-arvoiseen lähtöasetelmaan, kuten alakohdassa 2.3.3 todettiin. Vaikka tämän alustuksen jäljiltä kartta ei ole lineaarisen alustuksen tavoin alustavasti organisoitu, sillä ei ole merkitystä organisoitumisen ollessa ennen kaikkea verkon tehtävä.

Saman alustusmenetelmän käyttäminen kaikissa testeissä myös yksinkertaisti muiden opetusparametrien käyttöä, koska vaatimukset naapurustofunktiolle ja oppimiskertoimelle riippuvat käytetystä alustustavasta. Esimerkiksi lineaarista alustusta käytettäessä riittää opetuksessa pienempi oppimisnopeus ja kapeampi naapurustofunktio. [Kohonen, 2001; Kohonen, 2005]

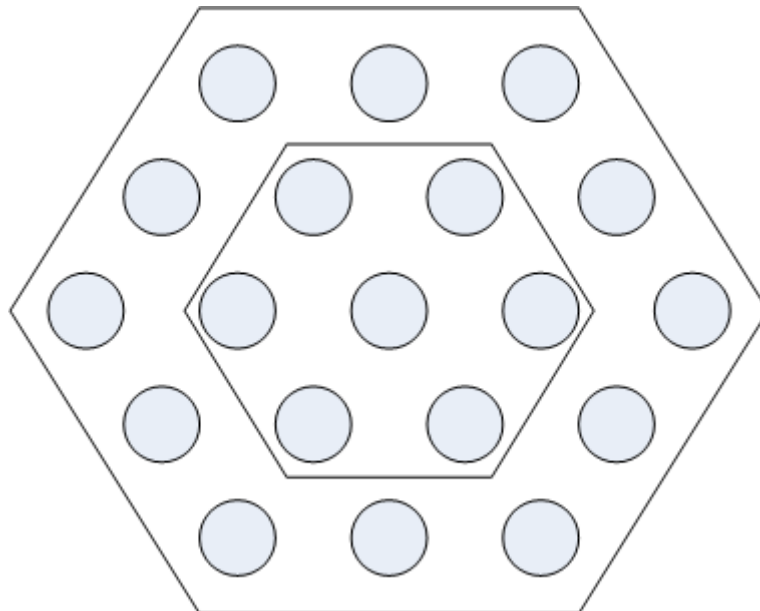
4.4.3. Hilan muoto

Kuten alakohdassa 2.1.3.2 todettiin, hila voi olla yksi- tai useampiulotteinen ja muodoltaan suorakulmainen, kuusikulmainen eli heksagonaalinen tai epäsäännöllinen. Useimmiten, kuten myös tässä tutkimuksessa, verkoille käytetään kaksiulotteista hilaa. Hilan muotona kokeiltiin molempia säännöllisiä hilarakenteita eli sekä suorakulmaista että heksagonaalihilaa.

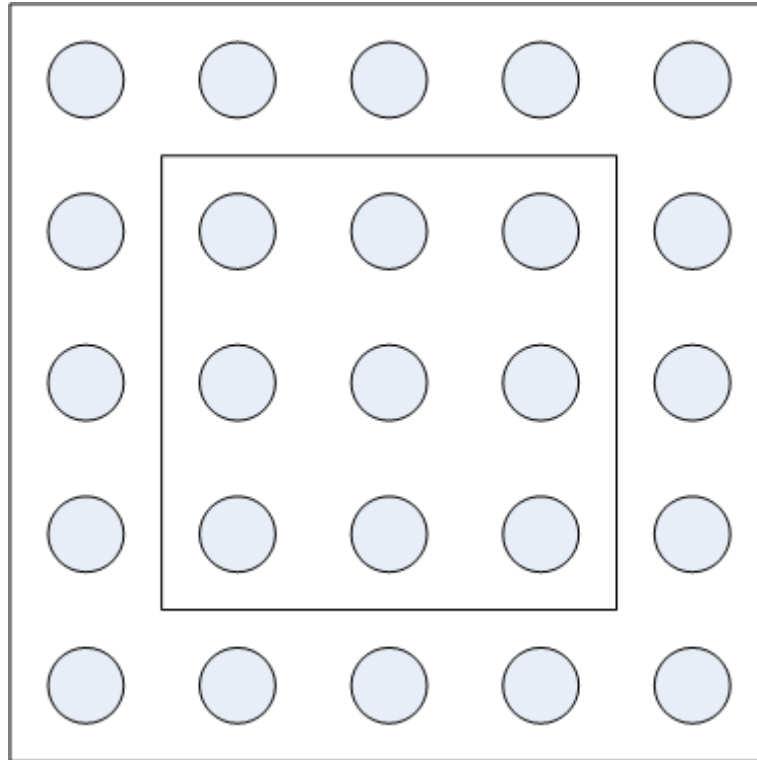
4.4.4. Naapurusto

Naapuruston suhteen tuli tehdä valinnat koskien sen muotoa ja kokoa.

Naapurustona käytettiin naapurijoukkoa eli hilassa voittajasolmun vieressä tietyllä säteellä sijaitsevia solmuja ja tämä joukko määritettiin verkon hilan muotoa mukaillen käyttäen heksagonaalihilalle kuusikulmion muotoista naapurustoa (kuva 4) ja suorakulmaiselle hilalle neliön muotoista naapurustoa (kuva 5). Naapuruston muoto ei kuitenkaan ole verkon suorituskyvyn kannalta olennainen [Matlab, 2007].



Kuva 4. Heksagonaalihila ja kuusikulmion muotoiset naapurustot, joista sisempi on naapurusto säteellä 1 ja ulompi on naapurusto säteellä 2. Kuva on piirretty Kohosta [2001] mukaillen.



Kuva 5. Suorakulmainen hila ja neliön muotoiset naapurustot, joista sisempi on naapurusto säteellä 1 ja ulompi on naapurusto säteellä 2. Kuva on piirretty Kohosta [2001] mukaillen.

Opetus aloitettiin käyttäen naapurustona koko kartan neuroneita ja sen kokoa pienennettiin opetuksen ensimmäisen vaiheen eli järjestysvaiheen aikana hienosäätövaiheelle asetettuun arvoon, jossa sitä pidettiin muuttamattomana koko hienosäätövaiheen ajan [Matlab, 2007].

Loppuvaiheessa käytetään usein pelkän keskusneuronin päivitystä ilman naapurustoa pyrittäessä sovittamaan neuronit mahdollisimman tarkasti opetusaineistoon. SOM:n käyttötarkoituksesta ja käytettävästä aineistosta riippuu, onko tämä järkevää. Koko perusjoukkoa analysoitaessa ja visualisoitaessa näin voidaan tehdä, mutta muuten vaarana on aineistoon ylisovittuminen, jota tosin ei voida estää pelkästään huolellisella naapuruston leveyden valinnalla. [Lampinen and Kostianen, 1999]

Tässä tutkimuksessa käytettäessä SOM:ia luokitteluun oli tärkeää säilyttää yleistyskyky. Tämän takia ei harkittu naapuruston päivittämisen pois jättämistä opetuksen loppuvaiheessa. Toisaalta naapurusto ei kuitenkaan saa olla hienosäätövaiheessa kovin suuri, jotta järjestysvaiheessa muodostettu järjestys pystytään säilyttämään ja hienosäätö kohdistuu tarkasti vain syötettä eniten vastaaviin painovektoreihin [Matlab, 2007]. Hienosäätövaiheessa naapurustona käytettiin siten

tähän vaiheeseen hyvin soveltuvaa [Kohonen, 2001; Matlab, 2007] keskussolmun viereistä lähinaapurustoa.

4.4.5. Opetuksen kesto

Luvussa 2.3.5 kerrottiin opetuksen jakautuvan järjestyms- ja hienosäätövaiheeseen, joista Kohosen [2001] mukaan järjestymsvaihe voi kestää esimerkiksi 1000 askelta, mutta hienosäätövaiheeseen tulee käyttää paljon enemmän aikaa, vähintään 500 kertaa verkon neuroneiden lukumäärän verran. Tässä tutkimuksessa tehtiin ensin erikokoisia verkkoja käyttäen näitä arvoja ja tämän jälkeen tutkittiin, pystytäänkö näin saatuja tuloksia parantamaan opetuksen kestoa kasvattamalla.

Koska syötevektoreita oli kummassakin opetukseen käytettävässä aineiston osajoukossa vähemmän kuin opetusaskeleita oli tehtävä, koko opetusdataa toistettiin esittäen syötevektorit aina erilaisessa satunnaisessa järjestyksessä [Matlab, 2007]. Tämä onkin Kohosen [2001] mukaan satunnaisen perusjoukosta valinnan ohella matemaattiselta kannalta suositeltava tapa syötevektoreiden kierrättämiseen, mutta myös opetusdatan toistaminen pitäen syötevektorit samassa järjestyksessä on käytännössä toiminut lähes yhtä hyvin.

Opetus tehtiin inkrementaaliseen tapaan eli verkolle esitettiin yksi syötevektori kerrallaan ja painovektoreita päivitettiin jokaisen syötevektorin esittämisen jälkeen [Matlab, 2007].

4.4.6. Oppimismisnopeus

Oppimismisnopeus määrittää, kuinka paljon voittajasolmun ja sen naapuruston mallivektoreita muutetaan vielä enemmän syötevektorin kaltaiseksi. Matlabissa [2007] voittajasolmua päivitetään suoraan oppimismisnopeuden mukaisesti ja sen naapurustoa puolella oppimismisnopeudella. Oppimismisnopeus vaihtelee opetuksen aikana siten, että järjestymsvaihe aloitetaan tälle vaiheelle asetetusta arvosta ja sitä lasketaan hienosäätövaiheelle asetettuun arvoon, josta sitä edelleen lasketaan hienosäätövaiheen aikana, mutta tällöin hyvin hitaasti [Matlab, 2007].

Tässä tutkimuksessa verkoille käytettiin oppimisnopeuksina yleisesti käytettyjä [Matlab, 2007] arvoja: järjestymisvaiheen alussa 0,9 ja hienosäätövaiheen alussa 0,02, joita myös Kohonen [2001] esimerkinomaisesti suosittelee.

4.5. Opetus ja testaus aineiston osajoukoilla

Kuten kohdassa 3.2 kerrottiin, verkkoja opetettiin aineiston ensimmäisellä ja kolmannella osajoukolla. Tässä luvussa selostetaan näille osajoukoille parhaiten soveltuvien verkkorakenteiden löytämiseksi tehdyt testit. Kohdassa 4.4 käsiteltiin opetusparametreja ja eriteltiin, mitä niistä päätettiin näissä testeissä optimoida ja mitkä kiinnittää käyttäen kaikille verkoille samoja arvoja ja toimintamenetelmiä. Verkon kokoa, opetuksen kestoa sekä hilan ja naapuruston muotoa päätettiin optimoida ja loput opetusparametrit kiinnitettiin muuttumattomiksi. Testiasetelmia raportoitaessa näitä muuttumattomia opetusparametreja ei enää erikseen mainita, joten kerrataan ne vielä lyhyesti.

Verkon alustaminen

Kaikki painovektorit alustettiin samaan pisteeseen syötteiden arvoalueiden keskelle.

Oppimisnopeus

Järjestymisvaiheen alussa käytettiin oppimisnopeutta 0,9 ja hienosäätövaiheen alussa 0,02.

Naapuruston koko

Opetuksen järjestymisvaiheen alussa naapurusto käsitti kaikki kartan solmut ja hienosäätövaiheen ajan vain lähinaapuruston eli solmut säteellä yksi.

Näiden lisäksi mainitaan vielä käytetyt syötemuuttujat. Kohdassa 3.1 kerrottiin, että tässä tutkimuksessa toivottiin pystyttävän käyttämään kaikkia 38:aa tärkeintä [Viikki *et al.*, 1999] muuttujaa, mutta tämän epäonnistumisen varalle vaihtoehdoiksi oli sovittu 13 tai 9 tärkeintä muuttujaa. Suurinta muuttujajoukkoa käyttäen kuitenkin saatiin hyviä tuloksia kummankin opetukseen käytettävän aineiston osajoukon tapauksessa, joten pienempiä muuttujajoukkoja ei ollut tarvetta kokeilla.

Seuraavissa alakohdissa selostetaan osajoukoilla tehdyt testit. Molemmille osajoukoille etsittiin ensin niiden luokitteluun parhaiten sopivaa verkon kokoa ja sen jälkeen tutkittiin opetuksen keston vaikutusta tuloksiin. Nämä testit tehtiin käyttäen verkoille

heksagonaalihilaa ja naapurustona kuusikulmion muotoista naapurijoukkoa. Lopuksi selvitettiin vielä, oliko suorakulmaiseen hilarakenteeseen ja neliön muotoiseen naapurustoon vaihtamisella vaikutusta edeltävissä testeissä löydettyjen parhaiden verkkorakenteiden tuloksiin.

4.5.1. Pääjoukko

4.5.1.1. Verkon koko

Sopivimman verkon koon etsintä toteutettiin tekemällä erikokoisia verkkoja, joiden muut opetusparametrit pidettiin mahdollisimman muuttumattomina, jotta voitiin nähdä verkon koon vaikutus tuloksiin. Testattavat verkkokoot määritettiin alakohdassa 4.4.1 kerrotulla tavalla. Testaus aloitettiin aineiston osajoukossa olevien tautiluokkien lukumäärää vastaavalla neuronimäärällä, josta sitä kasvatettiin ja lopuksi testattiin viimeinen neuronimäärä, joka ei vielä ylittänyt opetukseen käytettyjen syötteiden lukumäärää. Verkon kokoa muutettaessa verkon leveys ja korkeus pidettiin samansuuruisina lukuun ottamatta ensimmäistä verkkoa, jonka koko 2×3 määräytyi pääjoukon tautien lukumäärän mukaan. Testatut verkon koot olivat siten 2×3 , 3×3 , 4×4 , ..., 27×27 .

Muut opetusparametrit pidettiin verkon kokoa muutettaessa mahdollisimman muuttumattomina. Neuronit järjestettiin heksagonaalihilaan ja naapurustona käytettiin kuusikulmion muotoista naapurijoukkoa. Ainoastaan opetuksen kestoa muutettiin, koska minimivaatimus tälle riippuu verkon koosta. Tässäkin järjestymisvaiheen askelten lukumäärä pidettiin vakiona arvossa 1000 ja vain hienosäätövaiheen kestoa muutettiin asettaen se verkon koosta riippuvaan minimiin, joka on 500 kertaa verkon neuroneiden lukumäärä [Kohonen, 2001]. Mikäli näiden muodostama opetuksen kokonaiskesto epookkeina ei ollut kokonaisluku, se pyöristettiin seuraavaan kokonaislukuun. Käytetyt opetusparametrit esitetään liitteessä 1 A.

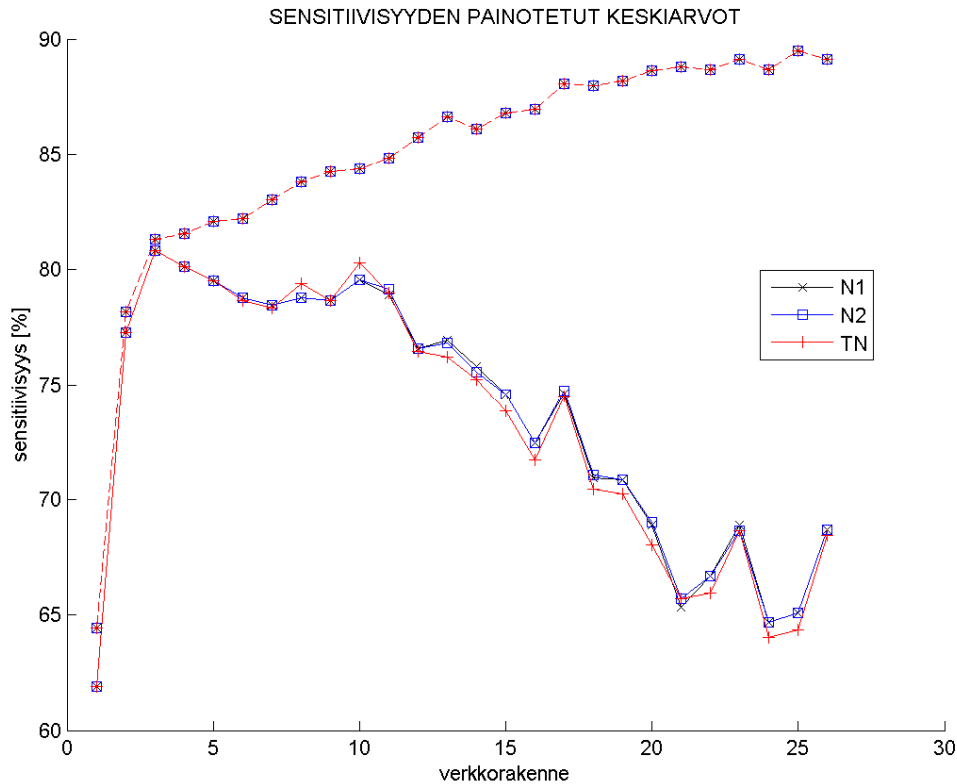
Esitellään seuraavaksi saatuja tuloksia, joista tarkkaillaan erityisesti testidatan tuloksia, koska se kertoo verkolle tuntemattomana datana verkon todellisen luokittelukyvyyn. Opetusdatan tuloksia verrataan testidatan tuloksiin datan ylisovittumisen havaitsemiseksi, koska tehtäessä SOM:lla luokittelua, jossa tarkoituksena on pystyä tekemään tuntemattomille tapauksille oikea diagnoosi eikä esimerkiksi koko perusjoukon kattavaa visualisointia tai data-analyysia, verkon yleistyskyky on tärkeää säilyttää hyvänä. Opetusdataan ylisovittuminen voidaan havaita testidatan tulosten

heikentymisestä opetusdatan tulosten edelleen parantuessa. Jos opetusdatan tuloksia ei tarkkailtaisi testidatan tulosten rinnalla opetusparametrin arvoa muutettaessa, voitaisiin tulosten huononemisen tulkita myös johtuvan joidenkin muiden opetusparametrien muutostarpeesta suhteessa muuttuneeseen opetusparametriin, esimerkiksi kokoa suurennettaessa järjestymisvaiheen kestoa saattaisi olla tarvetta kasvattaa. Opetusdatan tulosten paraneminen ilmaisee verkon opetusparametrien olevan oppimiselle suotuisat, jolloin testidatan tulosten samanaikainen huononeminen kertoo yleistyskyvyn heikentymisestä liiallisen tarkkuuden seurauksena.

Yleensä neuroverkon validointi suoritetaan käyttäen opetus- ja testidatasta erillistä validointidataa [Juhola, 2004], mutta myös edellä kuvattu testidatan käyttäminen validointidatana soveltuu ylisovittumisen valvomiseen, koska olennaisinta on validointidatan erillisuus opetusdatasta. Erillisiä validointijoukkoja ei muodostettu datan vähyiden vuoksi. Kaikki käytettävissä oleva aineisto haluttiin hyödyntää opetuksessa ja testauksessa.

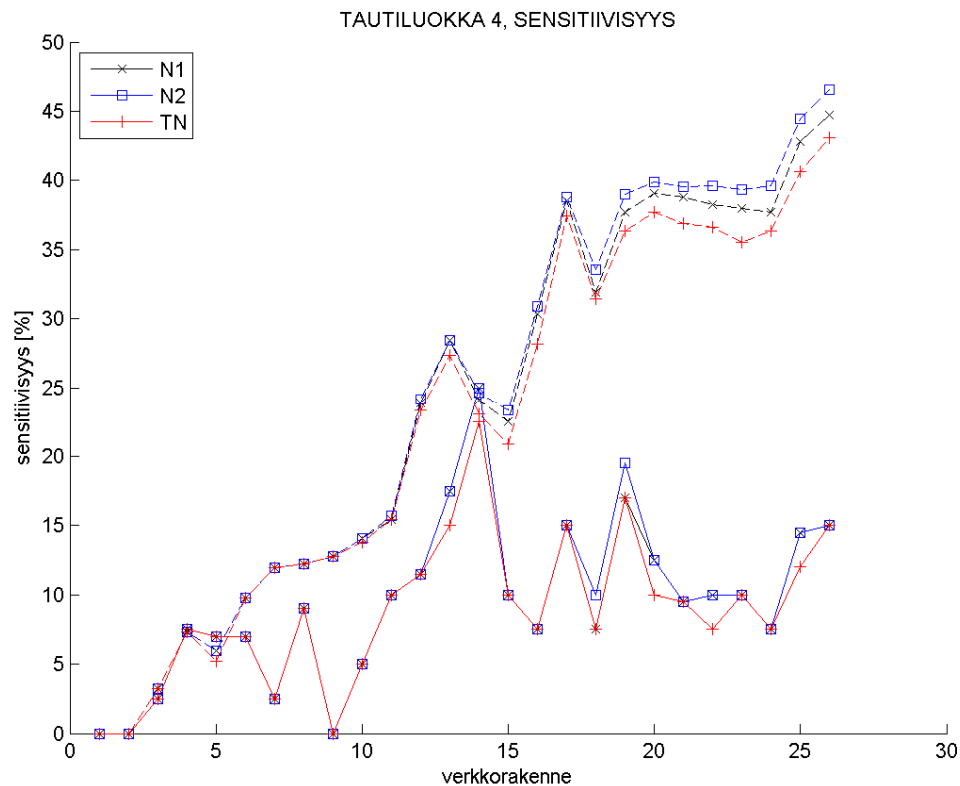
Tulokset

Verkon koon muuttamisen vaikutus opetus- ja testidatan tuloksiin esitetään kuvassa 6, jossa x-akselin numerot 1, ..., 26 vastaavat verkon kokoja 2×3 , ..., 27×27 ja y-akselilla esitetään sensitiivisyyden painotetut keskiarvot. Testidatan tulosten havaitaan verkon koon suurentuessa ensin parantuneen opetusdatan tulosten tavoin kolmanteen testattuun kokoon (4×4) asti, mutta tämän jälkeen alkaneen ensin hitaasti ja myöhemmin merkittävästi huonontua opetusdatan tulosten edelleen parantuessa loppuun asti solmumäärää kasvatettaessa. Tästä voidaan tulkita kyseisen 4×4 -koon olevan riittävä tämän datan luokitteluksi ja myös yleistyskyvyltään paras opetus- ja testidatan tulosten ollessa tämän kokoisella verkolla lähes yhtä hyvät. Solmumäärää kasvatettaessa mallin tarkkuus on lisääntynyt yleistyskyvyn kärsiessä. Vaikka testidatan tulosten huononeminen on ollut alussa hidasta ja esimerkiksi kymmenes testattu koko (11×11) on vielä lähes yhtä hyvä kuin paras koko, on neuroverkoista yleensä järkevintä valita yksinkertaisin ongelman ratkaisemaan pystyvä malli [Swingler, 1996] juuri tämän lisääntyvän tarkkuuden yleistyskykyä heikentävän vaikutuksen vuoksi.



Kuva 6. Verkon koon muuttamisen vaikutus opetus- ja testidatan sensitiivisyyden painotettuihin keskiarvoihin. Yhtenäisellä viivalla kuvataan testidatan tuloksia ja katkoviivalla opetusdatan tuloksia.

Kuitenkin koska itseorganisoituminen on stokastinen prosessi [Kohonen, 2001] ja sen luokittelutulokset vaihtelevat, saattaisi luokittelun hyvin lievä huononeminen koosta 4x4 kokoon 11x11 selittyä osin myös satunnaisella vaihtelulla. Tämän vuoksi päätettiin seuraaviin opetuksen keston vaikutusta selvittäviin testeihin parhaan 4x4-koon lisäksi ottaa mukaan lähes yhtä hyvän sensitiivisyyden tuottanut 11x11-kokoinen verkko. Lisäksi haluttiin tutkia, kuinka opetuksen keston lisääminen vaikuttaisi heikompia luokittelutuloksia saaneiden suurempien verkkojen suoriutumiseen. Tätä varten jatkotesteihin valittiin myös testattujen kokojen loppupäästä verkko kokoa 24x24 sekä keskiväliltä koko 15x15. Näistä jälkimmäinen oli myös mielenkiintoinen pystyessään näiden testien verkkorakenteista parhaiten luokittamaan vaikeasti tunnistettavan tautiluokan 4. Kyseisellä verkkorakenteella sensitiivisyyden painotettu keskiarvo saavutti arvon 25 %, kun paras 4x4-kokoinen verkkorakenne epäonnistui tässä saaden sensitiivisyydelle vain arvon 2,5 %. Sisällyttämällä parasta kokoa suuremmat verkkorakenteet jatkotesteihin toivottiin saatavan lisätietoa pienten tautiluokkien luokittelun onnistumisesta, mikä oli yksi tutkimustavoitteista. Kuvassa 7 esitetään verkon koon vaikutus tautiluokan 4 luokittelutuloksiin, joista esitetään sekä opetus- että testidatan tulokset.

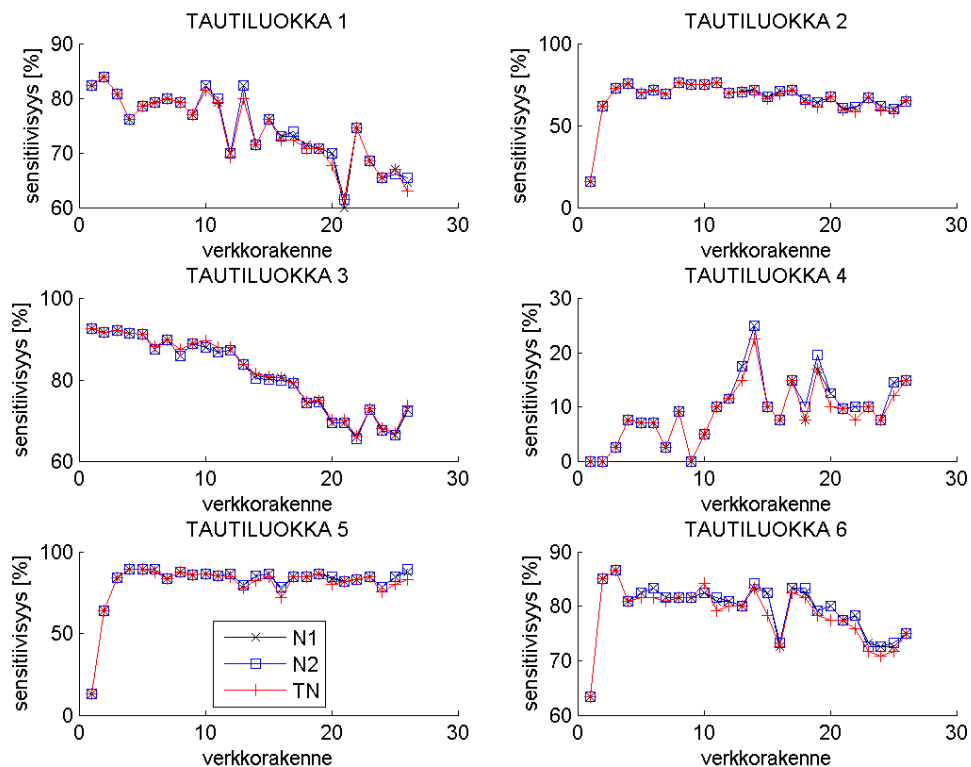


Kuva 7. Verkkorakenteiden sensitiivisyyden painotetut keskiarvot tautiluokalle 4. Kokonaisuutena paras verkkorakenne numero 3 ei tunnistanut tautia 4 kovinkaan hyvin ja sen sijaan isommat verkot suoriutuivat tästä paremmin. Yhtenäisellä viivalla kuvataan testidatan tuloksia ja katkoviivalla opetusdatan tuloksia.

Sekä kuvassa 6 että kuvassa 7 esitetään kolmella erilaisella verkon kalibroinnissa käytetyllä tasapelitilanteiden ratkaisumallilla saadut opetus- ja testidatan sensitiivisyyden painotetut keskiarvot. Näitä N1-, N2- ja TN-mallia käsiteltiin kohdassa 4.3. Kuvista huomataan, että näiden mallien tuottamissa tuloksissa ei ole paljoa eroa, mutta erot ovat kuitenkin kasvaneet solmumäärää lisättäessä. Tämä oli odotettavissa, koska neuronien järjestyessä approksimoimaan syötedatan todennäköisyysfunktioita suuret tautiluokat saavat pieniä tautiluokkia enemmän neuroneita käyttöönsä, jolloin pientä verkkoa käytettäessä neuroneihin kuvautumisessa ei helposti tule tasapelitilanteita, vaan paremminkin selviä voittajia, jolloin tasapelimalleja ei edes sovelleta.

Kuvassa 8 näytetään vielä verkon koon muuttamisen vaikutus kaikkien tautiluokkien sensitiivisyyteen testiaineiston osalta. Kuvasta voidaan havaita, että suurimman

tautiluokan 3 tulokset ovat olleet huipussaan heti alussa ja tämän jälkeen heikentyneet. Pienimmän tautiluokan 4 tunnistaminen sen sijaan on ollut parhaimmillaan testattujen kokojen keskivälillä, mutta lopussakin parempi kuin alussa. Tautiluokan 5 sensitiivisyys on ensin parantunut, mutta sen jälkeen pysynyt suurinpiirtein samana kokoa kasvatettaessa loppuun saakka. Lopuilla tautiluokilla sensitiivisyys on verkkokokoa kasvatettaessa alun nousun jälkeen laskenut vaihtelevalla voimakkuudella.



Kuva 8. Verkon koon kasvatuksen vaikutus tautiluokkien sensitiivisyyksiin.

4.5.1.2. Opetuksen kesto

Näissä testeissä tutkittiin, oliko kokotesteissä saatuja luokitustuloksia mahdollista parantaa opetuksen kesto kasvattamalla. Kokotesteissä järjestymsvaiheessa käytettiin 1000 askelta ja hienosäätövaiheessa verkon koosta riippuvaa minimikesto. Nyt kokeiltiin, pystyttäisiinkö jommankumman vaiheen tai molempien kesto kasvattamalla vaikuttamaan tuloksiin myönteisesti.

Testit toteutettiin kokotestien perusteella valituilla erikokoisilla verkoilla (4x4, 11x11, 15x15 ja 24x24), joiden valintaa jatkotesteihin perusteltiin kyseisten testien tulosten

analysoinnin yhteydessä. Verkot opetettiin vaihdellen opetuksen kestoja ja pitäen muut opetusparametrit samoina kuin edellisissä testeissä. Opetuksen keston testaaminen jaettiin kolmeen vaiheeseen.

Ensin testattiin järjestymisvaiheen keston vaikutusta tuloksiin, jolloin hienosäätövaihe pidettiin samassa arvossa eli verkon koon perusteella määräytyvässä minimiarvossa. Järjestymisvaiheen keston alkuarvoksi otettiin edellisissä testeissä käytetty 1000 askelta. Tätä arvoa kasvatettiin kullekin verkolle käytettyyn hienosäätövaiheen keston asti, jonka ainakin tulisi riittää ylärajaksi ko. keston riittäessä ainakin jossain määrin hienosäätövaiheeseenkin, joka kuitenkin on järjestymisvaihetta paljon enemmän aikaa vievä prosessi eli sopivan järjestymisvaiheen keston odotettiin jäävän paljon alle testatun ylärajan. Järjestymisen tapahtuu Kohosen [2001] mukaan yleensä suhteellisen nopeasti.

Tämän jälkeen testattiin, saataisiinko tuloksia edelleen parannettua hienosäätövaiheen kestoja kasvattamalla, minkä aikana järjestymisvaiheen kesto pidettiin äskeisessä vaiheessa parhaaseen sensitiivisyyden painotettuun keskiarvoon johtaneessa arvossa. Testattava arvoalue määritettiin kaksinkertaistamalla kyseisen järjestymisvaiheen keston yhteydessä käytetyt kokonaisepookit ja tämän jälkeen kokeilemalla vielä tätä selvästi suurempia kokonaisepookkeja. Arvojen testaamistiheys näin määritellyltä alueelta valittiin kokokohtaisesti, koska suurten verkkojen opetus kesti pienempien verkkojen opetusta huomattavasti kauemmin eikä arvoja siten suurten verkkojen suurilta arvoväleiltä voitu testata yhtä tiheästi kuin pienempien verkkojen suppeammilta arvoväleiltä. Kokonaisepookkimäärän kasvattaminen oli suoraa lisäystä hienosäätövaiheen keston, koska järjestymisvaiheen kesto pidettiin vakiona.

Lopuksi testattiin vielä pelkkä hienosäätövaiheen keston kasvattaminen pitäen järjestymisvaiheen kesto alkuperäisessä 1000 askeleessa. Tämä testattiin kasvattamalla kokonaisepookkeja yhtä suuriin arvoihin samoin periaattein kuin äsken, mutta aloittaen kasvatus lähtöarvosta, jota käytettiin kokotesteissä kyseistä kokoa olevalle verkkorakenteelle.

Opetusparametrit listataan liitteessä 1 kohdissa B – E. Jokaiselle verkkokoolle esitetään eri testivaiheissa käytetyt järjestymisaskleet ja kokonaisepookit. Järjestymisaskleet esitetään myös epookkeina, jotta järjestymisvaiheen osuutta opetuksen kokonaiskestosta on helpompi tarkastella.

Tulokset

Tarkastellaan, kuinka erikokoisten verkkorakenteiden tuloksiin pystyttiin vaikuttamaan opetuksen kestoa kasvattamalla. Taulukossa 7 esitetään parhaat opetuksen kestoa optimoimalla saadut sensitiivisyyden painotetut keskiarvot ja verrataan niitä kokotestien parhaisiin tuloksiin. Käsitellään seuraavaksi tuloksia kokokohtaisesti.

Taulukko 7. Verkkokokotestien ja opetuksen keston optimointitestien parhaat sensitiivisyyden painotetut keskiarvot.

Koko	Kokotestien paras sensitiivisyyden painotettu keskiarvo [%]	Kestotestien paras sensitiivisyyden painotettu keskiarvo [%]	Erotus [%]
4x4	80,84	81,96	1,12
11x11	80,29	82,47	2,18
15x15	75,80	79,69	3,89
24x24	68,89	68,45	-0,44

Kokotestien parhaalla 4x4-kokoisella verkkorakenteella pieni parannus tuloksiin oli selitettävissä satunnaisella vaihtelulla. Opetuksen keston lisäämisellä ei ollut säännönmukaista vaikutusta tuloksiin, vaikka kestot testattiin arvoväleiltään hyvin tiheästi. Paras sensitiivisyys 81,96 % saatiin 26 kokonaisepookilla kolmannessa vaiheessa, jossa hienosäätövaiheen kestoa kasvatettiin ja järjestymisaskalet pidettiin kokotesteissäkin käytetyssä 1000 askeleessa (1,37 epookkia).

Verkkorakenteelle kokoa 11x11 tuloksia pystyttiin parantamaan jonkin verran kasvattamalla sekä järjestymisvaiheen että hienosäätövaiheen kestoa. Testien paras sensitiivisyys 82,47 % saatiin testien toisessa vaiheessa, kun kasvatettiin hienosäätövaiheen kestoa käyttäen järjestymisvaiheen keston ensimmäisen vaiheen parhaan sensitiivisyyden tuottanutta 45625 askelta eli 62,41 epookkia. Hienosäätövaiheen keston kasvattaminen oli kuitenkin kannattavaa vain aluksi, jonka jälkeen sensitiivisyys alkoi laskea. Parhaan sensitiivisyyden tuotti 219 kokonaisepookkia, joka oli ensimmäisen vaiheen kokonaisepookkien ja tämän kaksinkertaistetun arvon puolestavälistä. Pelkällä hienosäätövaiheen keston kasvatuksella sensitiivisyyttä ei saatu kokotestien tuloksista parannettua.

15x15-kokoinen verkkorakenne hyötyi testatuista erikokoisista verkkorakenteista eniten opetuksen keston optimoinnista. Parannusta sensitiivisyyden painotettuun keskiarvoon saatiin parhaimmillaan lähes 4 % saavutettaessa arvo 79,69 %. Tämä arvo saatiin testien ensimmäisessä vaiheessa kasvatettaessa järjestymisvaihetta arvovälillä [1000, 112500]

testaten se ääripäistään ja kolmesta niiden välillä olevasta arvosta tasaisin välein, joista parhaaksi osoittautui toinen testattu arvo 28875 järjestymisaskelta (epookkeina lähes 40), jolloin kokonaisepookkimäärä oli 194. Testien toisessa vaiheessa käytettäessä tätä järjestymisvaiheen kestoa ja kasvatettaessa hienosäätövaiheen kestoa ei sensitiivisyyttä enää saatu parannettua.

Koska tämä verkon koko valittiin mukaan näihin testeihin osittain sen takia, että se oli tautiluokan 4 parhaiten diagnosoineena mielenkiintoinen, on paikallaan käsitellä opetuksen keston vaikutusta tämän taudin luokitteluun. Sensitiivisyys huononi selvästi opetusta lisättäessä. Kokotestien parasta tulosta 25 % ei pystytty parantamaan mutta se saatiin toistettua juuri samoja opetusvaiheiden kestoja käyttäen. Kokonaisuutena parhaaseen sensitiivisyyteen johtaneella edellä eriteltyllä opetuksen kestolla tautiluokan 4 luokittelussa päästiin enää vain 10 %:n sensitiivisyyteen. Kaikki testatut keston lisäämiset, eli järjestymisvaiheen ja hienosäätövaiheen keston sekä näiden molempien kasvattaminen, vaikuttivat tuloksiin negatiivisesti.

Verkkorakenteella kokoa 24x24 kokonaissensitiivisyyttä ei saatu parannettua opetuksen kestoa lisäämällä. Huomionarvoista tuloksissa kuitenkin oli, että tautiluokan 4 sensitiivisyyttä saatiin näissä testeissä parannettua jopa 27,5 %:iin, kun se kokotesteissä oli tälle koolle vain 10 %. Jos SOM:ia olisi käytetty esimerkiksi data-analyysiin, jossa yleistyskykyä ei tarvitsisi säilyttää, vaan sen sijaan pyrittäisiin mahdollisimman suureen tarkkuuteen, järjestymisvaiheen kestoa olisi kannattanut lisätä, koska opetusdatan tulokset tällöin edelleen paranivat.

Yhteenveto

Opetuksen kesto kasvattamalla kokotesteissä saatu verkkorakenteiden paremmuusjärjestys ei juuri muuttunut, mikä osoitti kokotesteissä käytettyjen opetuksen kestojen olleen jo hyvin valittuja. Suurten verkkojen tuloksia ei yleisesti ottaen saatu pienten verkkojen tuloksia paremmiksi. Ainoastaan 11x11-kokoisen verkkorakenteen sensitiivisyys nousi noin puoli prosenttia 4x4-kokoisen verkkorakenteen sensitiivisyyttä korkeammaksi, kun ero kokotesteissä oli niillä samaa luokkaa toisinpäin. Nämä erot ovat kuitenkin niin pieniä, että käytännössä voidaan sanoa näiden verkkorakenteiden olevan yhtä hyvät. Tämän takia jatkotesteihin valittiin näistä kumpikin, koska haluttiin vielä selvittää, kuinka hilan ja naapuruston muoto vaikuttaisi näiden verkkorakenteiden tuloksiin. Lisäksi 4x4-kokoisesta verkkorakenteesta jatkotesteihin sisällytettiin opetuksen kesto kasvatettaessa parhaan sensitiivisyyden tuottaneen 26 kokonaisepookin lisäksi jo kokotesteissä käytetty 13 kokonaisepookin verran opettaminen, koska parhaan arvon tulkittiin olevan lähinnä satunnaisuudesta johtuva.

4.5.1.3. Hilan ja naapuruston muoto

Tähän mennessä tehdyissä testeissä kaikki verkot oli tehty käyttäen heksagonaalihilaa ja kuusikulmion muotoista naapurustoa. Näissä testeissä kokeiltiin, pystyttäisiinkö tuloksia parantamaan muuttamalla verkon hila suorakulmaiseksi ja opetuksessa käytettävän naapuruston muoto neliön muotoiseksi.

Hilan ja naapuruston muodon vaikutusta testattiin edellisten testien parhailla 4x4- ja 11x11-kokoisilla verkkorakenteilla, joista ensimmäistä opetettiin sekä 13 että 26 kokonaisepookkia käyttäen kummassakin 1000 järjestymissäaskelta, ja jälkimmäistä opetettiin 219 kokonaisepookkia 62,41 epookin järjestymissävaiheen osuudella. Jokaiselle verkkorakenteelle tehtiin testit muuttaen opetusparametreista vain hilan ja naapuruston muotoa ja kutakin yhdistelmää testattiin kolme kertaa tulosten vaihtelun havaitsemiseksi.

Tulokset

Jokaisella testatulla verkkorakenteella heksagonaalihilaa ja kuusikulmion muotoista naapurustoa käyttäen sensitiivisyyden painotetut keskiarvot vaihtelivat kolmen eri testikerran välillä enemmän kuin suorakulmaisella hilalla ja neliön muotoisella naapurustolla. Ensimmäisellä yhdistelmällä saatiin kaikkien verkkorakenteiden testien matalin ja korkein arvo. Keskimääräisen hilan ja naapuruston sopivuuden arvioimiseksi näistä painotetuista keskiarvoista laskettiin keskiarvo ja korkeamman keskiarvon saaneen hilan ja naapuruston yhdistelmän tulkittiin sopivan kyseiselle verkkorakenteelle paremmin.

Opetettaessa 4x4-kokoista verkkorakennetta 13 epookkia tuottivat suorakulmainen hila ja neliönmuotoinen naapurusto paremman keskiarvon 80,34 % toisen yhdistelmän saadessa keskiarvon 79,50 %. Nämä olivat yhtä hyvät tulokset antaneiden N1- ja N2-mallien tulokset, mutta TN-malli suoriutui lähes yhtä hyvin. Kun samankokoista verkkorakennetta opetettiin 26 epookkia, saatiin heksagonaalihilalla ja kuusikulmion muotoisella naapurustolla hiukan parempi keskiarvo 79,67 % toisen yhdistelmän arvolla ollessa 79,48 %. Tässä tasapelimallien tulokset eivät poikenneet toisistaan.

Suuremmalle 11x11-kokoiselle verkkorakenteelle soveltui paremmin suorakulmainen hila ja neliönmuotoinen naapurusto, joilla keskiarvoksi saatiin 80,02 %, kun toisella yhdistelmällä arvo oli 79,61 %. Nämä arvot saatiin käyttäen kalibroinnissa N2-mallia, joka toimi tasapelimalleista parhaiten. N1-mallin tulokset olivat lähes yhtä hyviä kuin

N2:n poiketen sen tuloksista vain jälkimmäisen yhdistelmän arvolla 79,53 %. Huonoiten suoriutui TN-malli, jonka vastaavat arvot olivat 79,90 % ja 79,01 %.

Kaikista verkkorakenteista keskimääräisesti parhaiten luokittelu onnistui 4x4-kokoisella 13 epookkia opetetulla verkkorakenteella, mutta koska erot olivat hyvin pieniä, päätettiin vielä tehdä seuraavaksi raportoitavat perusteellisemmat ja siten luotettavammat 10x10-ristiinvalidointitestit näille kaikille. Kyseisten testien tulosten perusteella aiottiin valita tämän aineiston osajoukon luokitteluun parhaiten sopiva verkkorakenne. Tässä vaiheessa oltiin 11x11-kokoista verkkorakennetta kiinnostuneempia valitsemaan parhaaksi jompikumpi 4x4-kokoisista verkkorakenteista, koska ei kannata käyttää ongelman ratkaisuun 121 solmua, jos sen saa yhtä hyvin ratkaistua 16 solmulla. Testeihin valittiin verkkorakenteille käytettäviksi näissä testeissä parhaiksi osoittautuneet hilan ja naapuruston muotoyhdistelmät, vaikka tulokset eivät yhdistelmien välillä paljoa vaihdelleetkaan.

4.5.1.4. Viimeiset 10x10-ristiinvalidointitestit

Tähän mennessä kaikissa testeissä käytetyn 10x1-ristiinvalidoinnin sijaan suoritettiin nämä viimeiset testit käyttäen perusteellisempaa 10x10-ristiinvalidointia, joka tehtiin kolmelle edellä eritellylle verkkorakenteelle. Jokaiselle verkkorakenteelle suoritettiin 10x10-ristiinvalidointitestit kolme kertaa tulosten vaihtelun havaitsemiseksi. Taulukossa 8 kerrataan näissä testeissä käytettyjen verkkorakenteiden tiedot ja esitetään niiden testijärjestys, johon viitataan tulostarkasteluissa.

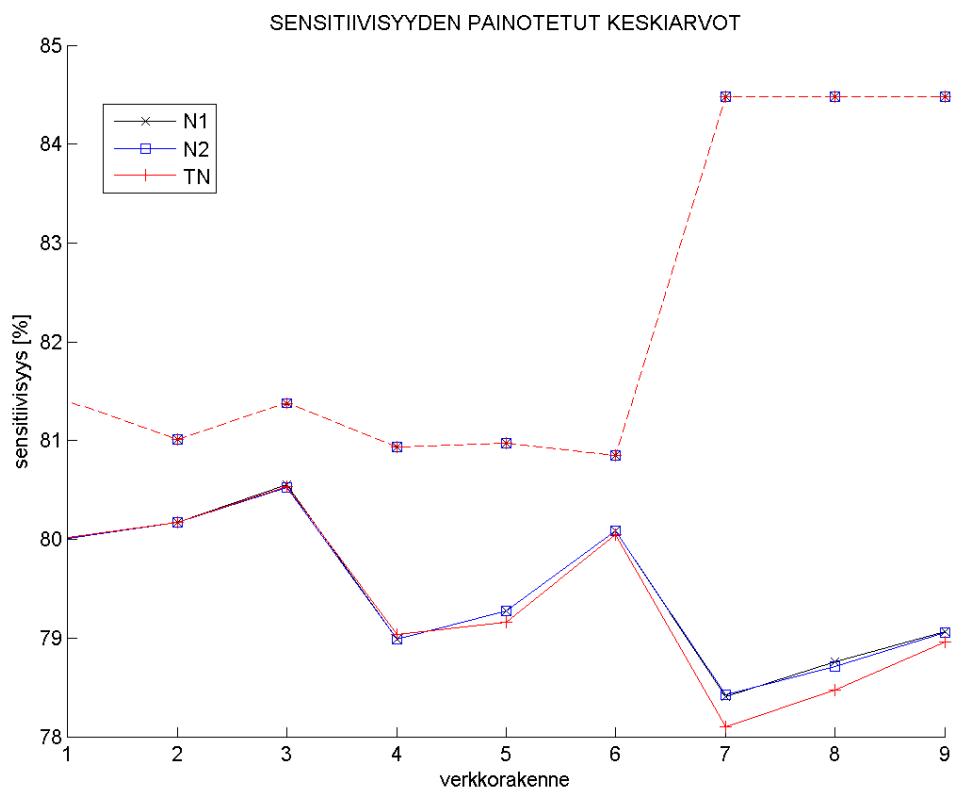
Taulukko 8. Viimeisissä 10x10-ristiinvalidointitesteissä testatut verkkorakenteet.

10x10-ristiin-validointi-testien numerot	Verkon koko	Opetuksen kokonais-epookit	Järjestymis-vaiheen osuus kokonais-epookeista	Hila	Naapuruston muoto
[1, 3]	4x4	13	1,37	suorakulmahila	neliö
[4, 6]	4x4	26	1,37	heksagonaalihila	kuusikulmio
[7, 9]	11x11	219	62,41	suorakulmahila	neliö

Tulokset

Toistettaessa 10x10-testit kolme kertaa voitiin parhaaksi verkkorakenteeksi todeta 13 epookkia opetettu 4x4-kokoinen verkkorakenne. Tällä saatiin sekä korkein

sensitiivisyyden painotettu keskiarvo 80,55 % että myös paras kolmesta testistä laskettu sensitiivisyyden painotettujen keskiarvojen keskiarvo 80,25 %. Vastaavat kolmen testin keskiarvot olivat 79,46 % toiselle 4x4-kokoiselle verkkorakenteelle, jota opetettiin 26 kokonaisepookkia, ja 78,75 % verkkorakenteelle kokoa 11x11. Näiden tulosten myötä oli vielä selvempää, että pienet verkot eivät ole ainoastaan yhtä hyviä luokittelijoita kuin suurempi 11x11, vaan jopa parempia, vaikka 11x11-kokoisen verkkorakenteen tulokset opetuksen kestoa optimoivissa testeissä ylittivätkin pienempien tulokset. Kuvassa 9 näytetään testattujen verkkorakenteiden kolmen testin sensitiivisyyden painotetut keskiarvot esittäen x-akselilla 10x10-ristiinvalidointitestien numerot, joihin viitattiin taulukossa 8 verkkorakenteet listattaessa.



Kuva 9. Viimeisten 10x10-ristiinvalidointina suoritettujen testien tulokset. X-akselilla esitetään 10x10-ristiinvalidointitestien numerot, joihin viitattiin taulukossa 8 verkkorakenteet listattaessa. Y-akselilla esitetään sensitiivisyyden painotetut keskiarvot [%]. Testidatan tulokset esitetään yhtenäisellä ja opetusdatan tulokset katkoviivalla.

Taulukossa 9 esitetään jokaisesta 10x10-ristiinvalidointitestistä verkkorakenteen tunnuslukujen painotetut keskiarvot. Esitettävät tulokset on saatu käyttäen N1-mallia, joka tuotti parhaat tulokset, vaikka eri tasapelimallien tuloksissa ei suuria eroja

ollutkaan. Huonoimmat tulokset tuottaneella TN-mallilla oli tunnusluvuissa eroa parhaaseen vain korkeintaan 0,37 %.

Taulukko 9. Viimeisten 10x10-ristiinvalidointitestien verkkorakenteiden tunnuslukujen painotetut keskiarvot, jotka saatiin N1-mallia käyttäen. Ristiinvalidointitesteissä testatut verkkorakenteet kerrottiin taulukossa 8.

10x10-ristiinvalidointi-testin numero	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonaistarkkuus
1	80,02	91,60	78,05	91,19
2	80,17	91,65	78,28	91,25
3	80,55	91,76	78,49	91,41
4	78,98	91,79	77,37	90,85
5	79,27	91,68	77,45	90,93
6	80,09	92,10	78,12	91,28
7	78,40	92,25	77,76	90,73
8	78,76	92,35	77,73	90,91
9	79,06	92,49	78,17	91,05

4.5.1.5. Paras verkkorakenne

Edellä käsitellyissä testeissä aineiston pääjoukon luokitteluun parhaiten soveltuvaksi osoittautui 4x4-kokoinen suorakulmahilainen verkkorakenne, jota opetettiin 13 kokonaisepookkia 1,37 epookin järjestymisvaiheen osuudella neliönmuotoista naapurustoa käyttäen. Tarkastellaan seuraavaksi kyseisen verkkorakenteen tuloksia tarkemmin. Taulukossa 10 esitetään verkkorakenteen edellisessä alakohdassa selostettujen kolme kertaa toistettujen 10x10-ristiinvalidointitestien sensitiivisyyksien painotettuja keskiarvoja, joista verkkorakenteen sensitiivisyys on 100 yksittäisen verkon keskiarvo painotettuna koko datan osajoukon tautiluokkien tapausten määrillä ja yksittäisten verkkojen keskiarvot on painotettu kyseisten verkkojen testauksessa käytetyillä tarkoilla tautitapausten määrillä.

Taulukko 10. Kolmen 10x10-ristiinvalidointitestin sensitiivisyyksien vertailua. Sensitiivisyydet ovat tautiluokkien painotettuja keskiarvoja. Verkkorakenteen sensitiivisyys on 100 yksittäisen verkon keskiarvo painotettuna koko datan osajoukon tautiluokkien tapausten määrillä ja yksittäisten verkkojen keskiarvot on painotettu kyseisten verkkojen testauksessa käytetyillä tarkoilla tautitapausten määrillä.

10x10-ristiinvalidointi-testin numero	Verkkorakenteen sensitiivisyys	Yksittäisten verkkojen korkein sensitiivisyys	Yksittäisten verkkojen sensitiivisyyksien mediaani
1	80,02 %	91,57 %	80,37 %
2	80,17 %	93,98 %	79,01 %
3	80,55 %	92,77 %	80,37 %

Verkkorakenteen kolmesta 10x10-ristiinvalidointitestistä korkeimman sensitiivisyyden painotetun keskiarvon 80,55 % tuotti kolmas testi, jonka 100 yksittäisen verkon korkein sensitiivisyys oli 92,77 % ja sensitiivisyyksien mediaani 80,37 %. Korkein yksittäisen verkon sensitiivisyys 93,98 % saatiin toisessa testissä, jossa verkkojen sensitiivisyyksien mediaani oli 79,01 %. Ensimmäisen testin mediaani oli sama kuin verkkorakenteen parhaan sensitiivisyyden tuottaneen kolmannen testin, mutta sen korkein arvo 91,57 % jäi näiden testien matalimmaksi. Nämä luetellut arvot olivat käytännössä samoja kaikilla tasapelimalleilla poiketen ainoastaan kolmannen testin verkkorakenteen sensitiivisyyden painotetun keskiarvon osalta N1-mallin ollessa vain prosentin sadasosien verran muita parempi.

Tarkastellaan vielä tuloksia tautiluokittain. Koska verkkorakenteen kolmen 10x10-testin tuloksissa ei ollut suurta vaihtelua, riittää diagnoosien luokittelun onnistumista tutkia yhden testin osalta. Tähän valittiin parhaan sensitiivisyyden painotetun keskiarvon tuottanut testi, koska myös sen yksittäisten verkkojen mediaani 80,37 % esiintyi kahdessa näistä kolmesta testistä, jolloin sen voitiin katsoa edustavan hyvin näitä kaikkia testejä. Taulukossa 11 esitetään parhaan verkkorakenteen yksittäisten tautiluokkien luokittelun onnistumista kuvaavien tunnuslukujen keskiarvot ja niiden keskiarvot tautiluokkien tapausten määrillä painotettuina ja painottamattomina, taulukossa 12 mediaanit ja taulukossa 13 keskihajonnat. Esitettävät tulokset ovat N1-mallilla saatuja.

Taulukko 11. Aineiston pääjoukon luokitteluun parhaiten soveltuvan verkkorakenteen luokittelun tunnuslukujen keskiarvot.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonais-tarkkuus
1	80,92	95,78	80,00	93,43
2	70,74	96,54	84,10	91,95
3	92,87	82,72	77,94	86,62
4	1,60	99,79	3,03	94,86
5	87,33	98,93	88,74	98,00
6	83,25	98,57	91,70	96,28
Keskiarvo	69,45	95,39	70,92	93,53
Painotettu keskiarvo	80,55	91,76	78,49	91,41

Taulukko 12. Aineiston pääjoukon luokitteluun parhaiten soveltuvan verkkorakenteen luokittelun tunnuslukujen mediaanit.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonais-tarkkuus
1	84,62	97,01	81,82	93,90
2	71,43	98,48	88,89	92,50
3	93,55	81,99	75,72	86,90
4	0,00	100,00	0,00	95,00
5	100,00	98,68	87,50	97,62
6	87,50	98,59	92,31	96,25

Taulukko 13. Aineiston pääjoukon luokitteluun parhaiten soveltuvan verkkorakenteen luokittelun tunnuslukujen keskihajonnat.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonais-tarkkuus
1	12,49	2,57	9,40	2,24
2	26,20	3,94	15,43	5,27
3	5,23	8,51	8,88	5,85
4	7,68	0,80	15,27	0,72
5	18,40	1,35	13,17	1,84
6	16,73	1,54	7,54	2,59

Tuloksista voidaan huomata kaikkien tautiluokkien luokittelun onnistuneen hyvin lukuun ottamatta pienintä tautiluokkaa 4, joka tunnistettiin parhaalla verkkorakenteella vain keskimääräisellä sensitiivisyydellä 1,60 %. Sen vaikutus näkyikin

verkkorakenteen sensitiivisyyden ja positiivisen ennustearvon painottamattomissa keskiarvoissa, jotka jäävät painotettuja keskiarvoja selvästi huonommiksi pienimmän tautiluokan vaikuttaessa laskentaan yhtä paljon kuin suuremmat tautiluokat. Aikaisemmissa testeissä havaittiin tämän pienimmän luokan luokittelun onnistuvan parhaiten suuremmilla verkkorakenteilla, jotka olivat kokonaistulosten eli kaikkien tautien luokittelun kannalta liian tarkkoja. Tämä on selitettävissä SOM:n syötedatan todennäköisyystiheysfunktion approksimointiominaisuudella [Kohonen, 2001]. Syötedatan jakauma vaikuttaa siihen, kuinka paljon neuroneita järjestyy edustamaan eri luokkia. Näin ollen neuronimäärää kasvatettaessa pienet tautiluokat kyllä saavat suuremmalla todennäköisyydellä solmuja käyttöönsä, mutta niin saavat myös suuremmat tautiluokat, jolloin mallista tulee liian tarkka suurten tautiluokkien osalta ja luokittelussa tarpeellinen yleistyskyky heikkenee. Näissä aineiston pääjoukolla tehdyissä testeissä todettiin, että tälle näin epätasaisesti jakautuneelle aineistolle ei voitu rakentaa itseorganisoituvaa verkkoa, jolla sekä suurten että pienten tautiluokkien luokittelu olisi onnistunut optimaalisesti. Vaatimukset mallin tarkkuudelle olivat suurten ja pienten tautiluokkien osalta liian erilaiset.

Pääjoukon parasta verkkorakennetta testattiin myös aineiston toisella ja neljännellä osajoukolla, mitä varten verkkorakennetta edustamaan valittiin yksittäisiä tämän verkkorakenteen verkkoja. Tässä haluttiin huomioida sekä se, mihin tämä verkkorakenne parhaimmillaan pystyy että mikä on sen keskimääräinen suoriutumisen, jolloin käyttöön valittiin parhaan sensitiivisyyden 93,98 % tuottanut verkko ja lisäksi kaikkien kolmen testin 300 verkon sensitiivisyyksistä laskettua mediaania 80,12 % lähimpänä oleva verkko, jonka sensitiivisyys oli 80,25 %. Tasapelimallit tuottivat tässä saman mediaanin. Verkkojen testausta mainituilla osajoukoilla käsitellään kohdassa 4.6.

4.5.2. Kolmas osajoukko

Kuten kohdassa 3.2 kerrottiin, aineiston kolmannen osajoukon tapaukset käytettiin yhdistäen ne kaikkien muiden diagnoosillisten tapausten kanssa eli ensimmäisen ja toisen osajoukon kanssa, jolloin neuroverkkojen opetukseen ja testaukseen käytettävä datajoukko muodostui lähes 80-prosenttisesti aineiston ensimmäisestä osajoukosta. Näin ollen tämän datajoukon oletettiin käyttäytyvän SOM:lla luokiteltaessa hyvin samantapaisesti kuin kyseinen ensimmäinen osajoukko ja sillä saatua tietämystä voitiin soveltaa tähän datajoukkoon. Tällä datajoukolla ei myöskään ollut tarpeen suorittaa yhtä perusteellisia testejä kuin ensimmäisellä osajoukolla, koska ensimmäisellä osajoukolla saatiin jo selvitettyä monia tutkittavia asioita, kuten pienten tautiluokkien

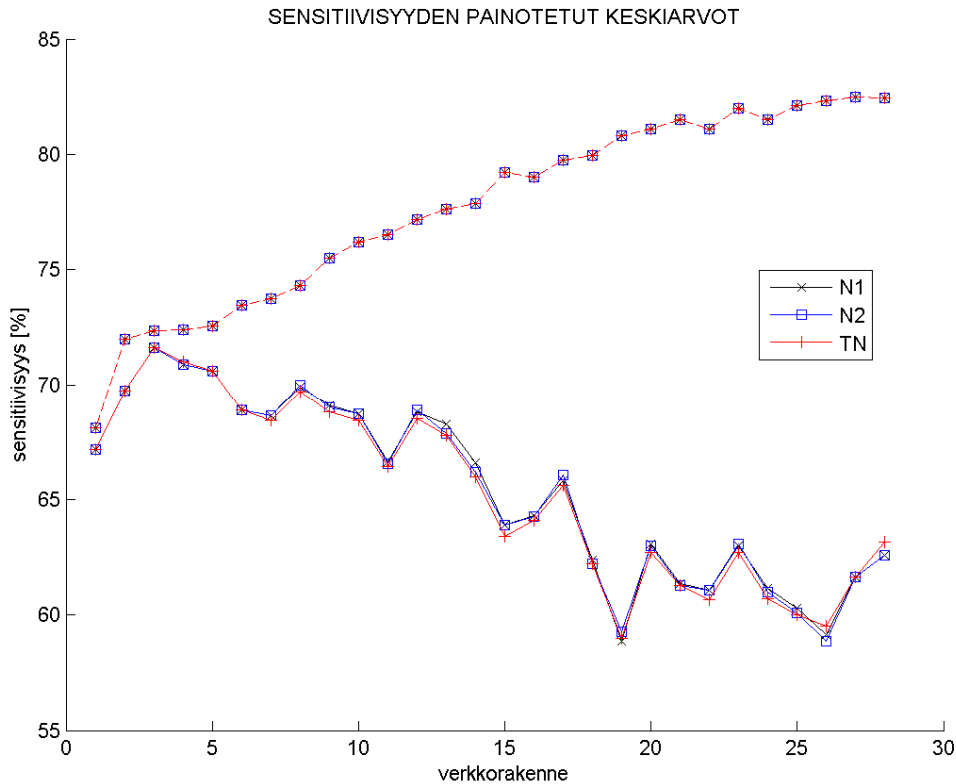
luokittelun onnistuminen. Pieniä tautiluokkia ei oletettu pystyttävän luokittelemaan tällä aineistolla ainakaan ensimmäistä osajoukkoa paremmin tämän datan sisältäessä myös epävarmempia diagnooseja, jolloin tämän datan luokittelun ei yleensä uskottu onnistuvan yhtä hyvin kuin varmemmista diagnooseista koostuvan pääjoukon.

4.5.2.1. Verkon koko

Kuten pääjoukolle, myös tälle aineiston osajoukolle aloitettiin parhaiden opetusparametrien etsiminen optimoimalla verkon kokoa. Testattavat verkon koot määritettiin samoin periaattein kuin pääjoukolle ja kuten alakohdassa 4.4.1 kerrottiin, joten tälle osajoukolle testattiin koot 3x3, 4x4,..., 30x30. Muut opetusparametrit asetettiin myös kuten pääjoukolle. Opetusparametritaulukko esitetään liitteessä 2 A.

Tulokset

Luokittelu erikokoisilla verkoilla onnistui odotetusti hyvin samaan tapaan kuin aineiston ensimmäisellä osajoukolla. Verkon kokoa kasvatettaessa sensitiivisyydet ensin paranivat testidatan tulosten alkaen tämän jälkeen huonontua opetusdatan tulosten edelleen parantuessa. Kyse oli jälleen ylisovittumisesta dataan eli yleistyskyvyn heikentymisestä tarkkuuden lisääntymisen seurauksena. Paras sensitiivisyyden painotettu keskiarvo 71,61 % saavutettiin kolmannella kokeillulla koolla 5x5, joka valittiin ainoana verkkorakenteena jatkotesteihin. Vaikka tämänkin jälkeen tulokset olivat hyviä, tiedettiin jo ensimmäisen osajoukon testien perusteella, että opetuksen kestoa optimoimalla ei kokotesteissä saatuun verkkorakenteiden paremmuusjärjestykseen juuri pystytä vaikuttamaan. Kokoa muutettaessa saadut sensitiivisyyden painotetut keskiarvot esitetään kuvassa 10.

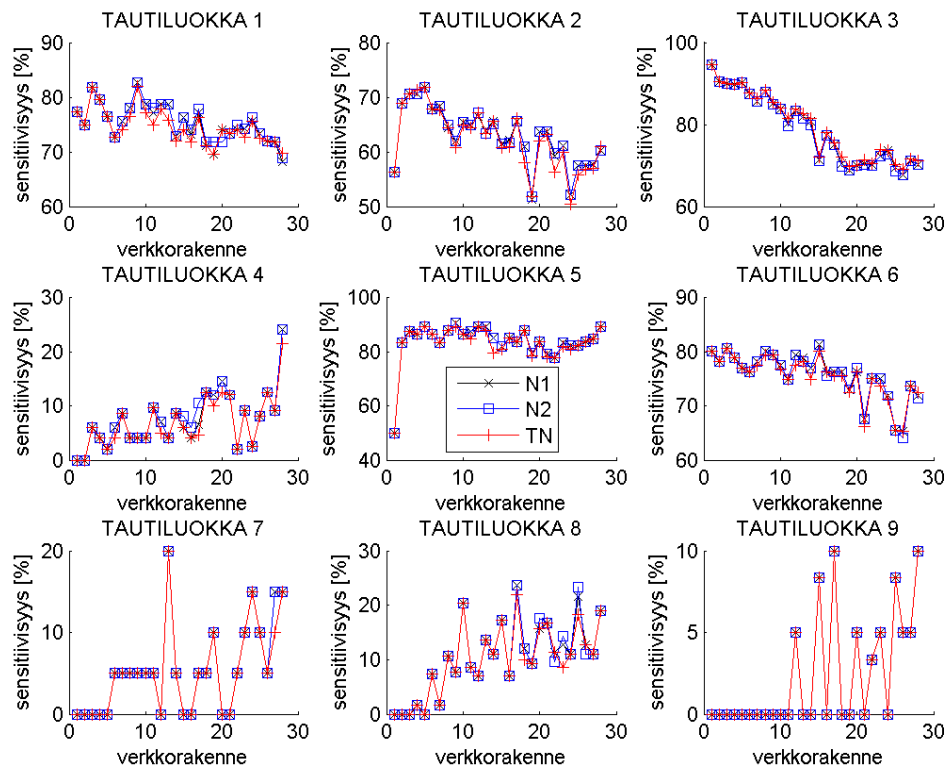


Kuva 10. Verkon koon muuttamisen vaikutus opetus- ja testidatan sensitiivisyyden painotettuihin keskiarvoihin. Yhtenäisellä viivalla kuvataan testidatan tuloksia ja katkoviivalla opetusdatan tuloksia.

Todetaan vielä solmumäärän lisäämisen vaikutuksen tautiluokkien sensitiivisyyksiin olevan hyvin samantapainen kuin pääjoukolla. Pieniä tautiluokkia pystyttiin tunnistamaan paremmin, mutta samalla suurten tautiluokkien sensitiivisyys huononi lisääntyneen tarkkuuden aiheuttaman yleistyskyvyn heikentymisen vuoksi (kuva 11). Tämä johtuu SOM:n syötedatan todennäköisyystiheysfunktion approksimointiominaisuudesta, mikä aiheuttaa neuroneiden keskittymisen alueille, joissa syötevektoreita on eniten. Suuret tautiluokat valtaavat pienten karttojen solmut. Solmumäärää lisättäessä pienet tautiluokat kyllä saavat enemmän sijaa kartalla, mutta niin saavat suuretkin tautiluokat, jolloin niiden osalta liiallinen tarkkuus heikentää yleistyskykyä ja siten luokittelusta suoriutumista.

Kuvasta 11 voidaan kuitenkin havaita tautiluokan 4 tuloksissa pieni eroavuus pääjoukon kokotestien tuloksiin. Pääjoukon kokotesteissä pienimmän tautiluokan 4 tulokset olivat parhaimmillaan testattujen kokojen puolivälissä. Nyt kuitenkin huippu saavutettiin vasta suurimmalla testatulla verkkorakenteella. Tämä johtune satunnaisesta vaihtelusta ja mahdollisesti osin myös siitä, että nyt kun mukana oli enemmän neuroneista kilpailevia tautiluokkia, tautiluokka 4 sai vasta myöhemmin riittävimpien määrän neuroneita tapauksiaan edustamaan. Toisaalta kun pääjoukolla

optimoitiin verkkorakenteiden opetuksen kestoa, saavutettiin kaikkien testien korkein sensitiivisyys tautiluokalle 4 samoin testattujen kokojen loppupäästä olevalla verkkorakenteella.



Kuva 11. Verkon koon kasvatuksen vaikutus tautiluokkien sensitiivisyyksiin.

4.5.2.2. Opetuksen kesto

Opetuksen kestoa optimoitiin samoin periaattein kuin pääjoukon tapauksessa eli kuten alakohdassa 4.5.1.2 määriteltiin. Opetusparametrit esitetään liitteessä 2 B.

Tulokset

Testeissä tuloksia ei juuri pystytty parantamaan. Paras saavutettu sensitiivisyyden keskiarvo oli 71,86 % eli vain hiukan parempi kuin kokotesteissä saatu 71,61 %. Kyseinen arvo saatiin 39 kokonaisepookilla kolmannessa vaiheessa pelkkää hienosäätövaihetta kasvatettaessa ja pitäen järjestyminen vaiheen kesto samana kuin kokotesteissä eli 1000 järjestyminenaskelissa, joka tämän osajoukon tapauksessa vastasi 1,08 epookkia. Tässäkään ei kuitenkaan pääjoukon pienen 4x4-kokoisen verkkorakenteen tavoin esiintynyt säännönmukaista vaikutusta tuloksiin opetuksen

kestoja kasvatettaessa. Näin ollen tämän hyvin pienen parannuksen tuloksiin tulkittiin jälleen johtuneen satunnaisesta vaihtelusta, joten jatkotesteissä päätettiin testata kyseinen 5x5-kokoinen verkkorakenne sekä optimoidulla että alkuperäisellä kokotesteissä käytetyllä opetuksen kestolla.

4.5.2.3. Hilan ja naapuruston muoto

Testiasetelmat olivat samat kuin pääjoukolla eli kuten alakohdassa 4.5.1.3 kerrottiin. Edellä kerrottujen verkkorakenteiden hila muutettiin suorakulmaiseksi ja opetuksessa käytettävän naapuruston muoto neliön muotoiseksi. Tässä kuitenkin suoritettiin testit jo 10x10-ristiinvalidointina, jotta tuloksia voitiin käyttää parhaan verkkorakenteen suoriutumisen arviointiin. Kummallekin verkkorakenteelle testattiin molemmat muotoyhdistelmät toistaen 10x10-ristiinvalidointi kolme kertaa, jotta voitiin havaita tuloksissa mahdollisesti esiintyvä vaihtelu.

Tulokset

Verkkorakenteille laskettiin kolmen 10x10-ristiinvalidointitestin keskiarvo sensitiivisyyksien painotetuista keskiarvoista käyttäen tähän tasapelimallien tuloksista parhaita. Keskiarvot eivät juuri vaihdelleet, paras arvo oli 70,54 % ja huonoin 70,32 %. Kumpaakin opetuksen kestoja käytettäessä saatiin heksagonaalihilalla ja kuusikulmionaapurustolla hiukan paremmat tulokset, mutta tulosten vaihdellessa näin vähän ei ollut juuri merkitystä, kumpaa näistä muotoyhdistelmistä käytettiin. Tasapelimallitkaan eivät juuri tuoneet vaihtelua tuloksiin, mutta parhaan muotoyhdistelmän parhaat arvot saatiin käyttäen N1- ja N2-mallia erojen ollessa kuitenkin merkityksettömiä.

Näissä perusteellisemmissä 10x10-ristiinvalidointitesteissä voitiin varmistua myös siitä, että tällä pienellä verkolla ei ollut väliä, opetettiinkö sitä 15 vai 39 kokonaisepookkia. Näistä ensimmäinen oli vain 0,04 % jälkimmäistä parempi kolmen testin sensitiivisyyksien painotettujen keskiarvojen keskiarvoltaan, kun tarkasteltiin parhaiden muotoyhdistelmien tuottamia tuloksia. Näiden välillä ei siten voida sanoa olevan eroa, mutta analysoidaan kuitenkin seuraavaksi parhaana verkkorakenteena 15 epookkia opetettua heksagonaalihilaista verkkorakennetta, joka oli opetusparametreiltaan täysin sama kuin jo ensimmäisessä verkon kokoa optimoivissa testeissä. Koska tällä verkkorakenteella kolmen 10x10-ristiinvalidointitestin sensitiivisyyden painotetut keskiarvot vaihtelivat vain 70,45 % – 70,67 %, tarkastellaan niistä vain yhden tuloksia. Erojen ollessa merkityksettömät voitiin tarkasteltavaksi valita paras. Kyseisessä testissä

verkkorakenteen sensitiivisyyden painotettu keskiarvo oli 70,67 % N1- ja N2-malleilla saatuna. Näistä tarkasteltavaksi valittiin N1-mallin tulokset.

4.5.2.4. Paras verkkorakenne

Parasta verkkorakennetta käsiteltiin edellä, mutta kerrataan vielä sen opetusparametrit. Parhaaksi osoittautui 5x5-kokoinen heksagonaalihilainen verkkorakenne, jota opetettiin 15 epookkia 1000 askeleen eli 1,08 epookin järjestyisvaiheen osuudella käyttäen kuusikulmion muotoista naapurustoa. Kyseinen opetuksen kesto oli Kohosen [2001] suositusten mukainen, ja kesto kasvatamalla ei saatu merkityksellistä parannusta sensitiivisyyteen.

Esitetään parhaalla verkkorakenteella suoritetun luokittelun tuloksia tautiluokittain. Taulukossa 14 esitetään yksittäisten tautiluokkien luokittelun tunnuslukujen keskiarvot ja niiden keskiarvot tautiluokkien tapausten määrillä painotettuina ja painottamattomina, taulukossa 15 mediaanit ja taulukossa 16 keskihajonnat. Esitettävät arvot ovat N1-mallin tuloksia.

Taulukko 14. Aineiston kolmannen osajoukon luokitteluun parhaiten soveltuvan verkkorakenteen luokittelun tunnuslukujen keskiarvot.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonais-tarkkuus
1	78,12	95,88	75,57	93,62
2	70,20	90,66	60,37	87,25
3	89,86	78,93	69,71	82,67
4	3,10	99,70	6,45	95,30
5	87,43	98,93	87,09	98,11
6	78,92	97,17	84,84	94,37
7	1,50	99,81	1,30	97,90
8	0,17	99,80	0,33	94,48
9	0,00	100,00	0,00	97,68
Keskiarvo	45,48	95,65	42,85	93,49
Painotettu keskiarvo	70,67	90,21	62,88	89,56

Taulukko 15. Aineiston kolmannen osajoukon luokitteluun parhaiten soveltuvan verkkorakenteen luokittelun tunnuslukujen mediaanit.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonais-tarkkuus
1	76,92	96,55	75,74	93,46
2	70,59	91,67	59,55	86,96
3	91,43	78,46	68,97	82,00
4	0,00	100,00	0,00	95,28
5	86,61	98,97	87,50	98,04
6	75,00	97,66	84,21	94,00
7	0,00	100,00	0,00	98,04
8	0,00	100,00	0,00	94,39
9	0,00	100,00	0,00	98,00

Taulukko 16. Aineiston kolmannen osajoukon luokitteluun parhaiten soveltuvan verkkorakenteen luokittelun tunnuslukujen keskihajonnat.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonais-tarkkuus
1	13,79	2,91	11,83	2,52
2	20,93	3,76	11,16	4,19
3	5,95	7,95	7,75	5,32
4	8,46	0,87	19,00	0,76
5	14,64	1,11	12,07	1,57
6	15,07	2,33	10,72	2,73
7	8,57	1,00	10,22	0,87
8	1,67	0,67	3,33	0,73
9	0,00	0,00	0,00	0,42

Tuloksista huomataan suurimpien tautiluokkien luokittelun onnistuneen melko hyvin myös tälle epävarmempia tapauksia sisältävälle aineiston osajoukolle. Pienten tautiluokkien luokittelu on kuitenkin epäonnistunut, koska tämä kokonaisuutena kaikille tautiluokille parhaat tulokset antava verkkorakenne on liian pieni pienten tautiluokkien tunnistamiseen, sillä pienellä verkolla ei riitä solmuja edustamaan niitä suurten tautiluokkien vallatessa kartan solmut. Valittaessa paras verkkorakenne sensitiivisyyden painotettujen keskiarvojen perusteella pienet tautiluokat eivät vaikuttaneet valintaan yhtä paljon kuin suuret tautiluokat. Pienten tautiluokkien vaikutus on nähtävissä paremmin painottamattomissa keskiarvoissa, jotka ovat sensitiivisyyden ja positiivisen ennustearvon osalta huomattavasti painotettuja keskiarvoja matalampia.

4.6. Testaus aineiston osajoukoilla

Tässä kohdassa käsitellään aineiston toisen ja neljännen osajoukon käyttöä ja niillä saatuja tuloksia. Kuten kohdassa 3.2 kerrottiin, näillä osajoukoilla ei opetettu verkkoja, vaan niillä vain testattiin pääjoukon luokitteluun parhaiten soveltuvaa verkkorakennetta. Tätä varten kyseistä verkkorakennetta edustamaan valittiin sekä parhaiten että keskimääräisesti suoriutuvat verkot, joiden valintaa käsiteltiin alakohdassa 4.5.1.5.

4.6.1. Toinen osajoukko

Aineiston toista osajoukkoa käyttäen saatiin parhaan verkkorakenteen parhaalla verkolla taulukossa 17 esitetyt tulokset ja parhaan verkkorakenteen keskimääräisellä verkolla taulukossa 18 esitetyt tulokset.

Taulukko 17. Tautiluokkien luokittelun tulokset tunnusluvuittain ja niiden painotetut keskiarvot, jotka saatiin parhaan verkkorakenteen parhaalla verkolla.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonaistarkkuus
1	0,00	99,13	0,00	98,28
2	59,26	96,63	84,21	87,93
3	97,30	65,82	57,14	75,86
4	0,00	100,00	0,00	94,83
5	87,50	100,00	100,00	99,14
6	70,27	100,00	100,00	90,52
Painotettu keskiarvo	73,28	88,31	76,62	86,13

Taulukko 18. Tautiluokkien luokittelun tulokset tunnusluvuittain ja niiden painotetut keskiarvot, jotka saatiin parhaan verkkorakenteen keskimääräisellä verkolla.

Tautiluokka	Sensitiivisyys	Spesifisyys	Positiivinen ennustearvo	Kokonaistarkkuus
1	0,00	99,13	0,00	98,28
2	59,26	95,51	80,00	87,07
3	94,59	62,03	53,85	72,41
4	0,00	100,00	0,00	94,83
5	50,00	98,15	66,67	94,83
6	62,16	98,73	95,83	87,07
Painotettu keskiarvo	67,24	86,30	70,96	83,43

Verrataan saatuja tuloksia pääjoukon luokittelun onnistumiseen. Tautiluokkaa 1 ei tunnistettu tällä osajoukolla ollenkaan, vaikka se pääjoukolla tunnistettiin sensitiivisyydellä 80,92 %. Tämä ei kuitenkaan ollut kovinkaan merkittävä tulos, koska tässä osajoukossa oli vain yksi tapaus kyseistä tautia. Pääjoukkoa huomattavasti huonommin tunnistettiin lisäksi tautiluokat 2 ja 6, mutta toisaalta pääjoukkoa paremmin tautiluokka 3. Pienintä tautiluokkaa 4 ei voitu löytää, koska sitä ei edustanut yksikään solmu kummassakaan käytetyssä verkossa, kuten kuvissa 12 ja 13 näytetään. Kyseisistä kuvista nähdään myös, että tautiluokat ovat SOM:lle tyypillisesti ryhmittyneet kartalle muodostaen yhtenäisiä alueita.

6	2	2	5
6	3	3	3
1	3	3	3
1	3	3	3

Kuva 12. Solmujen tautiedustukset aineiston pääjoukolla opetetun parhaan verkkorakenteen parhaassa verkossa, jota käytettiin aineiston toisen ja neljännen osajoukon luokitteluun.

6	6	1	1
5	3	3	3
5	3	3	3
2	3	3	3

Kuva 13. Solmujen tautiedustukset aineiston pääjoukolla opetetun parhaan verkkorakenteen keskimääräisessä verkossa, jota käytettiin aineiston toisen osajoukon luokitteluun.

Verrattaessa tämän osajoukon luokittelun onnistumista parhaalla ja keskimääräisellä verkolla huomataan sensitiivisyyksien poikkeavan toisistaan eniten tautiluokan 5 osalta. Parhaalla verkolla se tunnistettiin sensitiivisyydellä 87,50 %, mutta keskimääräisellä verkolla huomattavasti huonommin, 50 %:n sensitiivisyydellä. Keskimääräisellä verkolla luokittelu onnistui myös selvästi huonommin tautiluokan 6 osalta ja jonkin verran huonommin tautiluokan 3 osalta.

4.6.2. Neljäs osajoukko

Koska neljännen osajoukon diagnooseja ei tunnettu, tämän osalta oltiin kiinnostuneita, kuinka pääjoukon paras verkkorakenne datajoukon luokittelisi. Tässä parasta verkkorakennetta edustavana verkkona käytettiin ainoastaan parhaan sensitiivisyyden painotetun keskiarvon tuottanutta verkkoa. Luokittelusta ei luonnollisestikaan voitu laskea sen onnistumista mittaavia tunnuslukuja diagnoosien ollessa tuntemattomia, mutta taulukossa 19 esitetään, mihin tautiluokkiin verkko tapaukset sijoitti. Odotetusti suurin osa tapauksista sijoittui näistä taudeista yleisimpään tautiluokkaan 3. Pienintä tautiluokkaa 4 ei voitu löytää aineistosta, koska kuten edellisessä alakohdassa todettiin, kyseinen tauti ei ollut tässä verkossa edustettuna.

Taulukko 19. Aineiston neljännen osajoukon tuntemattomien tapausten sijoittuminen tautiluokkiin pääjoukolla opetettua verkkoa käyttäen.

Tautiluokka	Tapausten lukumäärä	%
1	6	4,92
2	36	29,51
3	69	56,56
4	0	0,00
5	4	3,28
6	7	5,74
Yhteensä	122	100,00

5. Yhteenveto

Tutkimuksessa luokiteltiin huimauspotilaiden dataa korvalääketieteen alalta Kohosen itseorganisoiduilla verkoilla. Aineisto jaettiin neljään osaan, joista kahta käytettiin neuroverkkojen opetukseen ja testaukseen ja kahta vain testaukseen. Kummallekin verkkojen opetukseen käytettävälle aineiston osajoukolle etsittiin niiden tautiluokkien luokitteluun parhaiten soveltuva verkkorakenne. Luokittelun onnistumista arvioitiin erilaisilla tunnusluvuilla, joista tarkkailtiin erityisesti luokittelun sensitiivisyyttä sen kertoessa oikein tunnistettujen positiivisten tapausten onnistumisesta. Parhaiksi verkkorakenteiksi valittiin osajoukot kokonaisuutena parhaiten luokittelevat verkkorakenteet. Koska tautiluokat olivat hyvin erikokoisia, luokittelun tunnusluvuista laskettiin verkkorakenteille keskiarvot tautiluokkien tapausten määrillä painotettuina. Siten parhaiden verkkorakenteiden valinnassa korostettiin suurimpien tautiluokkien luokittelun onnistumista, mutta tämän lisäksi tutkimuksessa tehtiin testejä aineiston pienimpien tautiluokkien luokittelun mahdollisuuden selvittämiseksi.

Käytettäessä neuroverkkoja luokitteluun aineiston epätasainen jakauma on ongelmallinen. Tässäkin tutkimuksessa tautiluokkien erikokoisuus osoittautui luokittelun onnistumista vaikeuttavaksi tekijäksi. Itseorganisoiduessaan neuroneiden painovektorit järjestyvät approksimoimaan syötedatan todennäköisyysfunktioita, jolloin neuronit keskittyvät alueille, joissa syötteitä esiintyy enemmän. Suuret tautiluokat saavat siten pieniä tautiluokkia enemmän neuroneita tapauksiaan edustamaan. Pienillä verkoilla ei välttämättä riitä ollenkaan neuroneita pienten tautiluokkien edustamiseen, koska suuret tautiluokat valtaavat kartan solmut. Verkkojen kokoa kasvatettaessa pieniä tautiluokkia aletaan todennäköisemmin edustaa kartalla, mutta samalla myös suuria tautiluokkia edustavien neuroneiden määrä kasvaa. Tällöin mallista tulee suurten tautiluokkien osalta liian tarkka ja luokittelussa tarpeellinen yleistyskyky heikkenee. Tautiluokkien tasainen jakauma olisi siis tärkeää, jotta SOM:lla voitaisiin luokitella tauteja tasapuolisesti.

Tutkimuksessa keskityttiin erityisesti aineiston varimmista diagnooseista koostuvan ensimmäisen osajoukon luokitteluun, jota käyttäen selvitettiin huolellisimmin myös aineistossa olevan pienen tautiluokan 4 luokittelun onnistuminen. Kyseistä tautia pystyttiin tutkimuksessa suoritetuissa testeissä luokittelemaan parhaimmillaan lähes 30 %:n sensitiivisyydellä. Tähän vaadittiin kuitenkin melko suurikokoisia verkkoja, jotka olivat suurempien tautiluokkien optimaalisen luokittelun kannalta liian tarkkoja. Valittaessa paras verkkorakenne sensitiivisyyden tautiluokkien tapausten määrillä painotetun keskiarvon perusteella suuret tautiluokat painoivat valinnassa enemmän ja

parhaalla 4x4-kokoisella verkkorakenteella tautiluokkaa 4 ei pystytty luokittelemaan onnistuneesti. Kyseisen parhaan verkkorakenteen painotetut keskiarvot olivat sensitiivisyydelle 80,55 %, positiiviselle ennustearvolle 78,49 % ja kokonaistarkkuudelle 91,41 % eli luokittelu onnistui hyvin.

Aineiston kolmas epävarmempia diagnooseja sisältävä osajoukko käytettiin opetukseen ja testaukseen yhdistäen se ensimmäisen ja toisen osajoukon kanssa. Näin muodostetun datajoukon koostuessa lähes 80-prosenttisesti aineiston ensimmäisestä osajoukosta sen oletettiin käyttäytyvän luokittelussa hyvin samantapaisesti kuin ensimmäinen osajoukko. Tämä myös todettiin optimaalista verkon kokoa etsivissä testeissä, jotka suoritettiin yhtä perusteellisesti samoin periaattein kuin ensimmäiselle osajoukolle. Verkkokokotestit osoittautuivat ensimmäiselläkin osajoukolla tärkeimmiksi testeiksi parhaan verkkorakenteen valitsemiseksi. Suurten tautiluokkien sensitiivisyys huononi solmumäärää lisättäessä, kun taas pienempiä tautiluokkia pystyttiin tällöin tunnistamaan paremmin. Kokonaisuutena parhaaksi osoittautui jälleen testattujen kokojen alkupäästä kokoa 5x5 oleva verkkorakenne. Pienimpiä tautiluokkia 4, 7, 8 ja 9 pystyttiin testeissä tunnistamaan parhaimmillaan sensitiivisyyksillä 24 %, 20 %, 23,67 % ja 10 %. Jos kokonaisuutena parasta suurempien verkkorakenteiden opetuksen kestoa olisi kasvatettu, olisi pienten tautiluokkien tuloksia näistä arvoista todennäköisesti saatu vielä jonkin verran nostettua, kuten onnistuttiin tekemään myös ensimmäisellä osajoukolla sen pienelle tautiluokalle 4, tosin vain 2,5 %. Kuitenkin tällä datajoukolla keskityttiin kokotestien jälkeen kokonaisuutena parhaat tulokset tuottaneen verkkorakenteen optimointiin, koska ensimmäisen osajoukon testeissä todettiin vaatimusten mallin tarkkuudelle olevan liian erilaiset sekä suurten että pienten tautiluokkien luokitteluksi optimaalisesti. Parhaalla verkkorakenteella painotetuiksi keskiarvoiksi saatiin sensitiivisyydelle 70,67 %, positiiviselle ennustearvolle 62,88 % ja kokonaistarkkuudelle 89,56 %. Luokittelu onnistui siis melko hyvin tälle epävarmempiakin diagnooseja sisältävälle datajoukolle.

Aineiston toisen osajoukon tapausten luokittelun onnistuminen testattiin aineiston ensimmäisen osajoukon parhaan verkkorakenteen parhaalla ja keskimääräisellä eli mediaania lähimpänä olevalla yksittäisellä verkolla. Parhaalla verkolla sensitiivisyyden painotetuksi keskiarvoksi saatiin 73,28 % ja keskimääräisellä verkolla 67,24 %. Pienintä tautiluokkaa 4 ei voitu löytää käytetyillä verkoilla, koska sitä ei edustanut yksikään solmu kummassakaan kartassa. Suurin tautiluokka 3 sen sijaan tunnistettiin hyvin korkeilla sensitiivisyyksillä 97,30 % ja 94,59 %.

Aineiston neljännen osajoukon luokittelun onnistumista ei voitu arvioida tunnusluvuilla, koska tämän osajoukon tapausten diagnoosit olivat tuntemattomia.

Tapaukset syötettiin ensimmäisen osajoukon parhaan verkkorakenteen parhaalle yksittäiselle verkolle ja tarkasteltiin, mihin tautiluokkiin verkko tapaukset sijoitti. Odotetusti eniten tapauksia sijoittui aineiston suurimpaan tautiluokkaan 3.

Kun viimeksi Kohosen itseorganisoituvia verkkoja käytettiin tämän aineiston luokitteluun [Juhola *et al.*, 2001], niiden soveltuvuutta tehtävään lähinnä vain kokeiltiin alustavasti. Verkot osoittautuivat tuolloin lupaavaksi menetelmäksi ja tämä tutkimus tehtiinkin, koska verkkojen soveltuvuudesta aineiston luokitteluun kaivattiin lisätietoa. Nyt tavoitteena oli paneutua luokittelun optimointiin. Verkkojen opettamiseen käytettiin tällä kerralla ristiinvalidoinnissa suurempi osuus, minkä toivottiin edesauttavan verkkojen luokittelukyvyn parantumista. Kaikki 38 muuttujaa pystyttiinkin käyttämään nyt onnistuneemmin, erityisesti tautiluokat 3, 5 ja 6 tunnistettiin paremmin. Pienimpiä tautiluokkia ei viimeksi pystytty luokittelemaan onnistuneesti, kuten ei nytkään kokonaisuutena parhailla melko pienillä verkkorakenteilla, mutta suuremmilla verkkorakenteilla niitä tunnistettiin kuitenkin paremmin. Kummassakin tutkimuksessa todettiin tautiluokkien jakauman olevan tärkeä seikka luokiteltaessa aineistoa itseorganisoituvalla kartalla. Jakauman tulisi olla nykyistä tasaisempi, jotta voitaisiin muodostaa kaikki tautiluokat optimaalisesti luokitteleva malli.

Tarkastellaan aineiston pääjoukolla saatuja tuloksia vielä suhteessa muihin aikaisempiin sitä luokitelleisiin tutkimuksiin, joissa kaikki 38 tärkeintä muuttujaa on käytetty onnistuneesti. MLP-verkon kanssa pääkomponenttianalyysia soveltaneen tutkimuksen [Juhola and Laurikkala, 2003] sensitiivisyyden tautiluokkien tapauksilla painotettu keskiarvo oli yli 8 % parempi kuin tässä tutkimuksessa saavutettu. Jokaiselle tautiluokalle omaa MLP-verkkoa käytettäessä sensitiivisyys oli tutkimuksesta riippuen joko lähes 5 % korkeampi [Siermala and Juhola, 2006] kuin nyt tai vajaan prosentin matalampi [Juhola and Siermala, 2005; Siermala *et al.*, 2007]. Etuna pääkomponenttianalyysia soveltaneeseen tutkimukseen oli, että tässä tutkimuksessa muuttujat käytettiin muuntamattomina. Jokaiselle tautiluokalle oman MLP-verkon käyttämiseen verrattuna etuna oli, että kaikkien tautiluokkien luokitteluun riitti yksi verkko. Tässä tulee kuitenkin huomioida, että sekä pääkomponenttianalyysia sovellettaessa että jokaiselle tautiluokalle omaa verkkoa käytettäessä pieni tautiluokka 4 saatiin luokiteltua erittäin hyvin sensitiivisyyksillä 91 – 100 %, kun sen luokittelu ei tässä tutkimuksessa onnistunut parhaalla verkkorakenteella ja testeissä parhaimmillaankin vain alle 30 %. Toisaalta eroa luokittelun onnistumisessa oli myös toisinpäin. Eniten sitä ilmeni suurimman tautiluokan 3 luokittelussa, joka tässä tutkimuksessa onnistui sensitiivisyydellä 92,87 %, kun se omalla MLP-verkollaan luokiteltiin tutkimuksesta riippuen joko 73 %:n [Juhola and Siermala, 2005; Siermala *et*

al., 2007] tai 84 %:n sensitiivisyydellä [Siermala and Juhola, 2006]. Pääkomponenttianalyysiä soveltanut tutkimus onnistui tässäkin sensitiivisyydellä 92 %. Olisikin mielenkiintoista kokeilla, kuinka kävisi yhdistettäessä pääkomponenttianalyysin soveltaminen itseorganisoituvien verkkojen kanssa. Sillä voitaisiin saada pienten tautiluokkien tuloksia parannettua. Se olisikin mahdollisesti seuraava kokeilemisen arvoinen asia tämän aineiston luokittelussa.

Viiteluettelo

[Altman and Bland, 1994a] D. G. Altman and J. M. Bland, Statistics Notes: Diagnostic tests 1: sensitivity and specificity. *British Medical Journal* **308**, 6943 (June 1994), 1552.

[Altman and Bland, 1994b] D. G. Altman and J. M. Bland, Statistics Notes: Diagnostic tests 2: predictive values. *British Medical Journal* **309**, 6947 (July 1994), 102.

[Auramo *et al.*, 1993] Y. Auramo, M. Juhola, and I. Pyykkö, An expert system for the computer-aided diagnosis of dizziness and vertigo. *Medical informatics* **18**, 4 (1993), 293-305.

[Autio *et al.*, 2007] L. Autio, M. Juhola, and J. Laurikkala, On the neural network classification of medical data and an endeavour to balance non-uniform data sets with artificial data extension. *Computers in Biology and Medicine* **37**, 3 (March 2007), 388-397.

[Bach *et al.*, 2003] C. Bach, S. Bredl, W. Kossa, and B. Sick, Calibration of self-organizing maps for classification tasks. *IEEE International Conference on Systems, Man and Cybernetics, 2003* **3** (5-8 Oct. 2003), 2877-2882.

[Begg *et al.*, 2006] R. Begg, J. Kamruzzaman, and R. Sarker (Eds.), *Neural Networks in Healthcare: Potential and Challenges*. Idea Group Publishing, 2006.

[Cottrell and Verleysen, 2006] M. Cottrell and M. Verleysen, Introduction: Advances in self-organizing maps. In: *Neural Networks, Advances in Self Organising Maps - WSOM'05*, **19**, 6-7 (July-August 2006), 721-722.

[Haykin, 1994] S. Haykin, *Neural Networks. A Comprehensive Foundation*. Macmillan College Publishing Company Inc., 1994.

[Ilmoniemi, 2008] R. Ilmoniemi, Aivojen rakenne ja toiminta, <http://www.biomag.hus.fi/braincourse/L1.html#RTFTtoC7> (viitattu 4.3.2008).

[Joutsiniemi *et al.*, 1995] S.-L. Joutsiniemi, S. Kaski, and T.A. Larsen, Self-organizing map in recognition of topographic patterns of EEG spectra. *IEEE Transactions on Biomedical Engineering* **42**, 11 (Nov. 1995), 1062-1068.

[Juhola, 2004] M. Juhola, Neurolaskennan luentomateriaali, Tietojenkäsittelytieteiden laitos, Tampereen yliopisto, 2004.

[Juhola and Laurikkala, 2003] M. Juhola and J. Laurikkala, Increasing Accuracy of Neural Network Classification with Principal Component Analysis for Otoneurological and Female Urinary Incontinence Data, *XVIIIth International Congress of the European Federation for Medical Informatics (MIE 2003)*, Saint Malo, France: IOS Press, 2003: (CD-ROM).

[Juhola and Siermala, 2005] M. Juhola and M. Siermala, Assessment Methods of Neural Network Classification Applied to Otoneurological Data. *European Notes in Medical Informatics (Medical Informatics Europe 2005)*, 1 (1), Geneve, Switzerland, 28.8.-31.8.2005, 6 pp., European Federal for Medical Informatics.

[Juhola *et al.*, 1999] M. Juhola, J. Laurikkala, K. Viikki, Y. Auramo, E. Kentala, and I. Pyykkö, Neural network recognition of otoneurological vertigo diseases with comparison of some other classification methods. In: W. Horn, Y. Sharar, G. Lindberg, S. Andreassen, J. Wyatt (Eds.), *Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making (AIMDM'99)*, Aalborg, Denmark, 1999, pp. 217–226 (Lecture Notes in Artificial Intelligence 1620: Artificial Intelligence in Medicine, Springer, Berlin, 1999).

[Juhola *et al.*, 2001] M. Juhola, K. Viikki, J. Laurikkala, I. Pyykkö, and E. Kentala, On classification capability of neural networks: a case study with otoneurological data. In: V. Patel, R. Rogers, R. Haux (Eds.), *Proceedings of 10th World Congress on Health and Medical Informatics (medinfo2001)*, London, UK, 2001, 474–478 (IOS Press, Amsterdam, 2001).

[Kentala *et al.*, 1995] E. Kentala, I. Pyykkö, Y. Auramo, and M. Juhola, Database for vertigo. *Otolaryngology - Head and Neck Surgery* **112**, 3 (March 1995) 383-390.

[Kentala *et al.*, 1996] E. Kentala, I. Pyykkö, Y. Auramo, and M. Juhola, Otoneurological expert system. *The Annals of otology, rhinology, and laryngology* **105**, 8 (1996), 654-658.

- [Kohonen, 2001] T. Kohonen, *Self-Organizing Maps*. Springer, 2001.
- [Kohonen, 2005] T. Kohonen, The Self-Organizing Map (SOM), Introduction. <http://www.cis.hut.fi/projects/somtoolbox/theory/somalgorithm.shtml> (viitattu 28.2.2008), 2005.
- [Lampinen and Kostiainen, 1999] J. Lampinen and T. Kostiainen, Overtraining and model selection with the self-organizing map. *International Joint Conference on Neural Networks, 1999 (IJCNN'99)* **3** (1999), 1911-1915.
- [Leflaive *et al.*, 2005] J. Leflaive, R. Céréghino, M. Danger, G. Lacroix, and L. Ten-Hage, Assessment of self-organizing maps to analyze sole-carbon source utilization profiles. *Journal of Microbiological Methods* **62**, 1 (July 2005), 89-102.
- [Lyytinen ja Himberg, 1996] H. Lyytinen ja L. Himberg, *Psykologia 4: Ihmisen toiminnan neuropsykologia*, WSOY, 1996.
- [Matlab, 2006] Matlab R2006b, Version 7.3.0.267 (R2006b), The Mathworks Inc., 2006.
- [Matlab, 2007] Matlab R2007b, Version 7.5.0.342 (R2007b), The Mathworks Inc., 2007.
- [Oja and Kaski, 1999] E. Oja and S. Kaski (Eds.), *Kohonen Maps*. Elsevier Science B.V., 1999.
- [Oja ja Mannila, 2007] E. Oja ja H. Mannila, Datasta tietoon -kurssin luentomateriaali. Tietotekniikan osasto, Teknillinen korkeakoulu, 2007. Saatavissa <http://www.cis.hut.fi/Opinnot/T-61.2010/luennot.shtml>, viitattu 3.3.2008.
- [Rojas, 1996] R. Rojas, *Neural Networks. A Systematic Introduction*. Springer-Verlag, 1996.
- [Siermala and Juhola, 2006] M. Siermala and M. Juhola, Techniques for biased data distributions and variable classification with neural networks applied to otoneurological data. *Computer Methods and Programs in Biomedicine* **81** (2006), 128-136.
- [Siermala *et al.*, 2007] M. Siermala, M. Juhola, J. Laurikkala, K. Iltanen, E. Kentala, and I. Pyykkö, Evaluation and classification of otoneurological data with new data

analysis methods based on machine learning. *Information Sciences* **177**, 9 (1 May 2007), 1963-1976.

[Swingler, 1996] K. Swingler, *Applying Neural Networks. A Practical Guide*. Academic Press, 1996.

[Vesanto and Alhoniemi, 2000] J. Vesanto and E. Alhoniemi, Clustering of the self-organizing map. *IEEE Transactions on Neural Networks* **11**, 3 (May 2000), 586-600.

[Viikki *et al.*, 1999] K. Viikki, E. Kentala, M. Juhola, and I. Pyykkö, Decision tree induction in the diagnosis of otoneurological diseases. *Medical Informatics and the Internet in Medicine* **24**, 4 (October 1999), Taylor and Francis Ltd, 277-289.

Liite 1. Pääjoukon testien opetusparametritaulukoita

1 A. Verkkokokotestit

Verkko- rakenne	Leveys	Korkeus	Neuronien lukumäärä	Hienosäätö- askeleita vähintään	Järjestymis- askeleet	Kokonais- epookit
1	2	3	6	3000	1000	6
2	3	3	9	4500	1000	8
3	4	4	16	8000	1000	13
4	5	5	25	12500	1000	19
5	6	6	36	18000	1000	26
6	7	7	49	24500	1000	35
7	8	8	64	32000	1000	46
8	9	9	81	40500	1000	57
9	10	10	100	50000	1000	70
10	11	11	121	60500	1000	85
11	12	12	144	72000	1000	100
12	13	13	169	84500	1000	117
13	14	14	196	98000	1000	136
14	15	15	225	112500	1000	156
15	16	16	256	128000	1000	177
16	17	17	289	144500	1000	200
17	18	18	324	162000	1000	223
18	19	19	361	180500	1000	249
19	20	20	400	200000	1000	275
20	21	21	441	220500	1000	304
21	22	22	484	242000	1000	333
22	23	23	529	264500	1000	364
23	24	24	576	288000	1000	396
24	25	25	625	312500	1000	429
25	26	26	676	338000	1000	464
26	27	27	729	364500	1000	500

1 B. Opetuksen keston optimointitestit, verkkorakenteen koko 4x4

Verkkorakenne	Järjestymis- askeleet	Järjestymisvaiheen osuus epookkeina	Kokonaisepookit
Vaihe 1			
1	1000	1,37	13
2	2000	2,74	14
3	3000	4,10	16
4	4000	5,47	17
5	5000	6,84	18
6	6000	8,21	20
7	7000	9,58	21
8	8000	10,94	22
Vaihe 2			
9, 10, ..., 30	8000	10,94	23, 24, ..., 44
31	8000	10,94	100
32	8000	10,94	150
33	8000	10,94	200
34	8000	10,94	250
35	8000	10,94	300
Vaihe 3			
36, 37, ..., 66	1000	1,37	14, 15, ..., 44
67	1000	1,37	100
68	1000	1,37	150
69	1000	1,37	200
70	1000	1,37	250
71	1000	1,37	300

1 C. Opetuksen keston optimointitestit, verkkorakenteen koko 11x11

Verkkorakenne	Järjestymis- askeleet	Järjestymisvaiheen osuus epookkeina	Kokonaisepookit
Vaihe 1			
1	1000	1,37	85
2	15875	21,72	105
3	30750	42,07	125
4	45625	62,41	146
5	60500	82,76	166
Vaihe 2			
6	45625	62,41	219
7	45625	62,41	292
8	45625	62,41	450
9	45625	62,41	600
Vaihe 3			
10	1000	1,37	152
11	1000	1,37	219
12	1000	1,37	292
13	1000	1,37	450
14	1000	1,37	600

1 D. Opetuksen keston optimointitestit, verkkorakenteen koko 15x15

Verkkorakenne	Järjestymis- askeleet	Järjestymisvaiheen osuus epookkeina	Kokonaisepookit
Vaihe 1			
1	1000	1,37	156
2	28875	39,50	194
3	56750	77,63	232
4	84625	115,77	270
5	112500	153,90	308
Vaihe 2			
6	28875	39,50	291
7	28875	39,50	388
8	28875	39,50	500
9	28875	39,50	800
Vaihe 3			
10	1000	1,37	291
11	1000	1,37	388
12	1000	1,37	500
13	1000	1,37	800

1 E. Opetuksen keston optimointitestit, verkkorakenteen koko 24x24

Verkkorakenne	Järjestymis- askeleet	Järjestymisvaiheen osuus epookkeina	Kokonaisepookit
Vaihe 1			
1	1000	1,37	396
2	72750	99,52	494
3	144500	197,67	592
4	216250	295,83	690
5	288000	393,98	788
Vaihe 2			
6	216250	295,83	1035
7	216250	295,83	1380
8	216250	295,83	2000
Vaihe 3			
9	1000	1,37	690
10	1000	1,37	1035
11	1000	1,37	1380
12	1000	1,37	2000

Liite 2. Kolmannen osajoukon testien opetusparametritaulukoita

2 A. Verkkokokotestit

Verkko- rakenne	Leveys	Korkeus	Neuronien lukumäärä	Hienosäätö- askeleita vähintään	Järjestymis- askeleet	Kokonais- epookit
1	3	3	9	4500	1000	6
2	4	4	16	8000	1000	10
3	5	5	25	12500	1000	15
4	6	6	36	18000	1000	21
5	7	7	49	24500	1000	28
6	8	8	64	32000	1000	36
7	9	9	81	40500	1000	45
8	10	10	100	50000	1000	56
9	11	11	121	60500	1000	67
10	12	12	144	72000	1000	80
11	13	13	169	84500	1000	93
12	14	14	196	98000	1000	108
13	15	15	225	112500	1000	123
14	16	16	256	128000	1000	140
15	17	17	289	144500	1000	158
16	18	18	324	162000	1000	177
17	19	19	361	180500	1000	197
18	20	20	400	200000	1000	218
19	21	21	441	220500	1000	240
20	22	22	484	242000	1000	264
21	23	23	529	264500	1000	288
22	24	24	576	288000	1000	314
23	25	25	625	312500	1000	340
24	26	26	676	338000	1000	368
25	27	27	729	364500	1000	396
26	28	28	784	392000	1000	426
27	29	29	841	420500	1000	457
28	30	30	900	450000	1000	489

2 B. Opetuksen keston optimointitestit

Verkkorakenne	Järjestymis- askeleet	Järjestymisvaiheen osuus epookkeina	Kokonaisepookit
Vaihe 1			
1	1000	1,08	15
2	2000	2,17	16
3	3000	3,25	17
4	4000	4,33	18
5	5000	5,42	19
6	6000	6,50	21
7	7000	7,58	22
8	8000	8,67	23
9	9000	9,75	24
10	10000	10,83	25
11	11000	11,92	26
12	12000	13,00	27
13	12500	13,54	28
Vaihe 2			
14, 15, ..., 38	10000	10,83	26, 27, ..., 50
39	10000	10,83	100
40	10000	10,83	150
41	10000	10,83	200
42	10000	10,83	250
43	10000	10,83	300
Vaihe 3			
44, 45, ..., 78	1000	1,08	16, 17, ..., 50
79	1000	1,08	100
80	1000	1,08	150
81	1000	1,08	200
82	1000	1,08	250
83	1000	1,08	300