

Joonas Karhumaa

**Koneoppimismallit lukemisen erityisvaikeuden
tunnistamisessa**

Tietotekniikan kandidaatintutkielma

16. toukokuuta 2024

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Tekijä: Joonas Karhumaa

Yhteystiedot: joonas.a.karhumaa@student.jyu.fi

Työn nimi: Koneoppimismallit lukemisen erityisvaikeuden tunnistamisessa

Title in English: Machine learning models in detecting dyslexia

Työ: Kandidaatintutkielma

Sivumäärä: 21+0

Tiivistelmä: Lukemisen erityisvaikeus on yksi yleisimmistä oppimisvaikeuksista. Lukihäiriöllä on merkittävä vaikutus yksilön itsetuntoon ja koulumenestykseen, mutta aikaisella diagnosoinnilla ja kuntouttavalla interventiolla sen vaikutuksia voidaan minimoida. Tutkielman tarkoituksena on selvittää, että voidaanko lukihäiriö tunnistaa tehokkaasti ja luotettavasti silmänliikkeistä koneoppimismallien avulla. Tämän lisäksi tutkitaan, mitkä käytössä olevista koneoppimismalleista soveltuvat lukihäiriön tunnistamiseen parhaiten ja onko eri mallien välillä merkittävää eroa.

Avainsanat: lukihäiriö, dysleksia, koneoppiminen, koneoppimismallit

Abstract: Reading disability is one of the most common learning disabilities. Dyslexia has a significant impact on an individual's self-esteem and academic performance, but with early diagnosis and rehabilitative intervention, its effects can be minimised. The aim of this thesis is to investigate whether dyslexia can be effectively and reliably identified from eye movements using machine learning models. It will also investigate which of the existing machine learning models are best suited for identifying dyslexia and whether there are significant differences between the different models.

Keywords: dyslexia, machine learning, machine learning models

Kuviot

Kuvio 1. Silmänliikkeisiin liittyviä käsitteitä.....	4
Kuvio 2. Esimerkki, miltä silmänliikkeet voivat näyttää lukemisen aikana. Muokattu alkuperäisestä (Sekhar, Chandrashekar ym. 2023).	5
Kuvio 3. Lukihäiriöisen koehenkilön silmänliikkeet lukemisen aikana. Muokattu alkuperäisestä (Sekhar, Chandrashekar ym. 2023).	6
Kuvio 4. Pöydälle asetettava EyeLink 1000 Plus -järjestelmä. Koehenkilön pää pidetään paikoillaan leuka- ja otsatuen avulla. (Moreira ym. 2022)	7
Kuvio 5. Esimerkki kaksiulotteista luokitteluongelmasta. Ympyröillä merkityt tukivektorit pyrkivät jakamaan datapisteet klustereihin siten, että erottavilla datapisteillä on suurin mahdollinen marginaali hypertasoon. Muokattu alkuperäisestä (Cortes ja Vapnik 1995).....	11
Kuvio 6. Monikerroksisen perseptroniverkon eteenpäin suuntautuvat neuronikerrokset. Muokattu alkuperäisestä (Popescu ym. 2009).....	12

Sisällys

1	JOHDANTO	1
2	LUKIHÄIRIÖ JA SILMÄNLIIKKEET	3
2.1	Lukihäiriö	3
2.2	Lukihäiriöön liitetyt silmänliikkeet.....	4
2.3	Silmänliikkeiden mittaaminen	6
3	KONEOPPIMISESSA KÄYTETTY DATA	8
4	KONEOPPIMISMALLIT LUKEMISEN ERITYISVAIKEUDEN TUNNISTA- MISESSA	10
4.1	Koneoppimismallit	10
4.2	Koneoppimismallien ongelmat	12
5	JOHTOPÄÄTÖKSET.....	14
	LÄHTEET	15

1 Johdanto

Lukihäiriö on yksi yleisimmistä oppimisvaikeuksista ja sillä on merkittävä vaikutus yksilön itsetuntoon ja oppimismenestykseen. Lukihäiriön aikaisella tunnistamisella ja nopealla kuntouttavalla interventiolla voidaan sen negatiivisia vaikutuksia vähentää huomattavasti. Lisäksi pelkästään diagnoosin merkitys yksilölle voi olla merkittävä.

Koneoppimismallit oppivat niille syötetystä opetusdatasta ja soveltavat opetusdatasta saatua tietoa ennennäkemättömään dataan. Tämän vuoksi koneoppimismallit soveltuvat erinomaisesti esimerkiksi erilaisten luokitteluongelmien ratkaisuun. Tässä tutkielmassa tutkitaan onko koneoppimismallien mahdollista oppia luokittelemaan koehenkilöt silmänliikkeistä saatavasta datasta lukihäiriöisiin ja kontrolliryhmään.

Tutkielma pyrkii selvittämään, että voidaanko lukihäiriö tunnistaa luotettavasti silmänliikkeistä erilaisten koneoppimismallien avulla. Miten eri koneoppimismallit soveltuvat lukihäiriön tunnistamiseen ja onko niiden välillä merkittäviä eroja. Entä, liittyykö koneoppimismallien käyttöön mahdollisesti ongelmia?

Tutkielma toteutetaan kirjallisuuskatsauksena. Tutkielman aiheesta löytyy artikkeleita usean vuosikymmenen ajalta, joten materiaalia kirjallisuuskatsauksen tekemiseen on hyvin saatavilla. Lähteiden hakemisessa hyödynnetään pääasiassa sekä Scopus- että Google Scholar-tietokantoja. Lisäksi artikkeleita on valittu aikaisemmin löydettyjen artikkeleiden viittauksista. Hakukielenä tutkimusprosessissa toimii englanti ja tärkeimpiä hakusanoja ovat muun muassa "eye movement", "dyslexia", "reading", "eye tracking", "machine learning", "AI" ja "visual attention span".

Tutkielmassa keskitytään lukihäiriön tunnistamiseen lapsilla. Silmänliikkeissä tutkielma keskittyy kielellisten tehtävien aikana esiintyviin tyypillisiin silmänliikkeisiin. Tutkielmassa ei ole tarkoitus ottaa kantaa nykyiseen diagnosointiin ja sen ongelmiin, eikä toisaalta ehdottaa vaihtoehtoisia diagnosointimenetelmiä.

Ensiksi, tutkielmassa perehdytään lukihäiriöön sekä siihen liitettyihin tyypillisiin silmänliikkeisiin. Mitkä ovat lukihäiriölle tyypillisiä silmänliikkeitä ja miten niitä mitataan? Toiseksi,

tutkielmassa selvitetään minkälaista dataa silmänliikkeistä saadaan kerättyä ja miten sitä käsitellään ennen koneoppimismalleille syöttämistä. Kolmanneksi, perehdytään tarkemmin eri koneoppimismalleihin. Voidaanko niiden avulla tunnistaa lukihäiriö silmänliikkeistä ja onko eri koneoppimismallien välillä eroja. Luvussa pohditaan myös koneoppimismalleihin mahdollisesti liittyviä ongelmia. Tutkielman viimeisessä luvussa kerrataan tutkimuksessa saatuja tuloksia.

2 Lukihäiriö ja silmänliikkeet

Lukihäiriöllä on merkittävä vaikutus yksilön elämään muun muassa negatiivisten tunteiden ja koulumenestyksen myötä. Mahdollisimman aikaisella diagnosoinnilla ja kuntouttavalla interventiolla lukihäiriön negatiivisia vaikutuksia voidaan minimoida. Silmänliiketutkimuksesta voidaan saada apua lukihäiriön diagnosointiin ja seulontaan. Tässä luvussa perehdytään lukihäiriöön ja siihen liittyviin silmänliikkeisiin.

2.1 Lukihäiriö

Lukihäiriö on yksi yleisimmistä oppimishäiriöistä. Käypä hoito -suosituksen (2019) mukaan kehityksellinen kielihäiriön esiintyvyys on 1–7 % väestöstä, riippuen käytetyistä diagnostisista kriteereistä. Bishop (2014) kertoo, ettei lasten kielihäiriöiden tutkimuksessa ole olemassa yhtenäistä termistöä tai selkeää konsensusta diagnostisista kriteereistä. Tässä tutkielmassa käytetään termiä lukihäiriö, kun puhutaan lukemisen erityisvaikeudesta ja kehityksellisestä kielihäiriöstä.

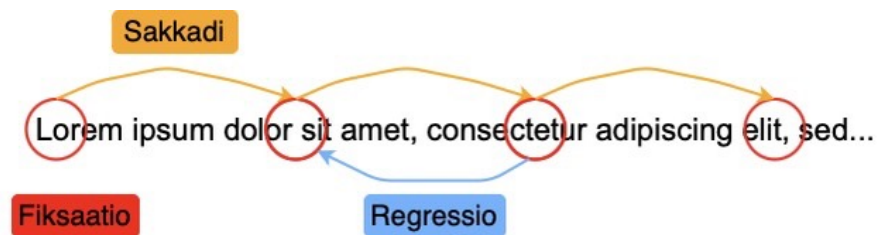
Lukihäiriöllä on merkittävä vaikutus yksilön elämään. Humphreyn (2002) mukaan lukihäiriöiset oppilaat ilmentävät epävarmaa käytöstä, välttävät stressaavia tilanteita sekä kysyvät apua ja vakuutteluja enemmän kuin vertaisensa. Humphrey (2002) ja Glazzard (2010) ovat yhtä mieltä siitä, että lukihäiriöllä on merkittävä vaikutus yksilön itsevarmuuteen ja omanarvontunteeseen. Glazzard (2010) lisäksi havainnoi, että lukihäiriöiset oppilaat tunsivat itsensä tyhmiksi, olivat pettyneitä itseensä ja tunsivat olevansa erilaisia useammin kuin verrokkiryhmä. Toisaalta Glazzard (2010) myös huomasi, että diagnoosilla oli merkittävä vaikutus edellä mainittuihin ongelmiin. Diagnoosiin saatuaan lukihäiriöiset oppilaat eivät enää pitäneet itseään tyhminä ja he kokivat myös ymmärtävänsä itseään paremmin. Diagnoosin jälkeen he eivät myöskään enää olleet yksin ongelmiensa kanssa.

Kehityksellisen kielihäiriön diagnosoinnissa tulee ottaa huomioon useita tekijöitä. On erityisen tärkeää erottaa onko kyse puhtaasta kielihäiriöstä vai johonkin muuhun tekijään, kuten esimerkiksi autismikirjoon, neurologiseen sairauteen tai kuulovammaan, liittyvästä kielihäiriöstä. Täytyy kuitenkin muistaa, etteivät edellä mainitut tekijät poissulje kuntoutuksen

tarvetta. (Bishop 2014; Mainela-Arnold 2019).

2.2 Lukihäiriöön liitetyt silmänliikkeet

Vuosien saatossa useat tutkimukset ovat liittäneet lukihäiriöön sille tyypillisiä silmänliikkeitä. Pavlidis (1981) väittää, että lukihäiriölle tyypillisiä silmänliikkeitä esiintyy sekä kielellisten että ei-kielellisten tehtävien aikana. De Luca ym. (1999) jakavat Pavlidiksen (1981) havainnon tyypillisistä silmänliikkeistä kielellisten tehtävien aikana. He eivät kuitenkaan huomaneet poikkeavia silmänliikkeitä ei-kielellisten tehtävien aikana. Hutzler ja Wimmer (2004), Rello ja Ballesteros (2015), Clifton ym. (2016), Raatikainen ym. (2021) ja Shalileh ym. (2023) ovat kaikki yhtä mieltä kielellisten tehtävien aikana esiintyvistä lukihäiriölle tyypillisistä silmänliikkeistä, mutta silmänliikkeiden alkuperästä ei ole päästy yksimielisyyteen. Pavlidis (1981) ehdottaa, että silmänliikkeet itsessään johtavat ongelmiin ja näin ollen aiheuttavat lukihäiriön. Prado, Dubois ja Valdois (2007) ovat kuitenkin sitä mieltä, että lukihäiriö on silmänliikkeiden aiheuttaja.



Kuvio 1. Silmänliikkeisiin liittyviä käsitteitä.

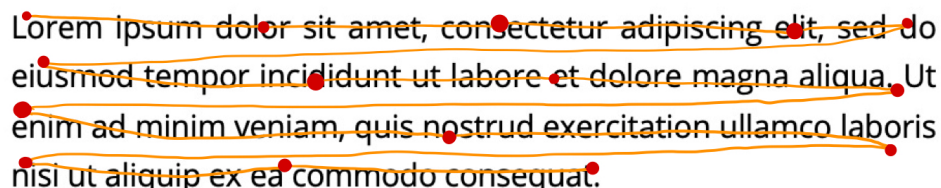
Kuviossa 1 on esitetty silmänliiketutkimukseen liittyviä käsitteitä. Fiksaatiot eli katseen kohdennukset ovat esitettyinä punaisina ympyröinä. Sakkadit eli silmänliikkeet fiksaatiosta toiseen puolestaan ovat esitettyinä oransseilla nuolilla ja regressiot eli palautumiset aiempaan fiksaatioon sinisellä nuolella.

De Luca ym. (1999) havainnoivat, että yksittäisten grafeemien lukeminen oli virheetöntä sekä lukihäiriöisillä koehenkilöillä että verrokkiryhmällä. Grafeemit ovat kirjoitusjärjestelmän pienimpiä yksittäisiä yksikköjä. Näitä ovat muun muassa yksittäiset kirjaimet sekä tavumerkit. Pitkissä sanoissa, jotka sisälsivät kuudesta kahdeksaan kirjainta, lukihäiriöisten koehenkilöiden fiksaatiot olivat noin kaksi kertaa pidempiä kuin verrokkiryhmällä. Lisäksi lukihäi-

riöisten koehenkilöiden kokonaisluku-aika oli 2–6 kertaa pidempi kuin verrokkiryhmällä. De Luca ym. (1999) huomasivat, että sanojen esiintyvyyksiheys sekä pituus vaikuttivat fiksaatioiden keston. Cliftonin ym. (2016) mukaan edellä mainittujen tekijöiden lisäksi sanojen ennustettavuudella kontekstissa on vaikutus fiksaatioiden keston.

Seuraavassa tutkimuksessaan De Luca ym. (2002) vertailivat silmänliikkeitä tavallisten sanojen ja pseudosanojen lukemisen välillä. He huomasivat, että lukihäiriöisellä tutkimusryhmällä silmänliikkeet olivat päinvastaiset verrokkiryhmään nähden. Lukihäiriöisillä koehenkilöillä oli enemmän sakkadeja lyhyiden pseudosanojen lukemisessa kuin lyhyiden sanojen lukemisessa. Pitkien pseudosanojen ja tavallisten sanojen välillä ei kuitenkaan ollut eroa sakkadien määrässä. Kontrolliryhmä puolestaan luki lyhyet pseudosanat ja lyhyet sanat samalla määrällä sakkadeja. Pitkissä pseudosanoissa puolestaan heillä tuli enemmän sakkadeja kuin tavallisissa pitkissä sanoissa.

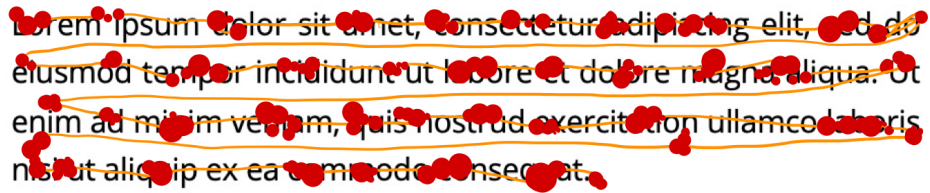
Rello ja Ballesteros (2015) ja Sekhar, Chandrashekar ym. (2023) ovat yhtä mieltä siitä, että lukihäiriöisillä koehenkilöillä fiksaatiot ovat verrokkiryhmää pidempiä. Molempien tutkimuksissa syntyi havainto myös lyhyemmistä sakkadeista, joihin liittyy lisäksi erittäin satunnaisia silmänliikkeitä. Rello ja Ballesteros (2015) huomasivat myös, että lukihäiriöisillä koehenkilöillä oli lukiessa enemmän regressioita kuin verrokkiryhmällä.



>Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Kuvio 2. Esimerkki, miltä silmänliikkeet voivat näyttää lukemisen aikana. Muokattu alkuperäisestä (Sekhar, Chandrashekar ym. 2023).

Kuvio 2 esittää, miten ei-lukihäiriöisen koehenkilön silmänliikkeet voivat näyttäytyä tekstikatkelmaan sijoitettuna ja kuvio 3 puolestaan esittää lukihäiriöisen koehenkilön silmänliikkeitä lukemisen aikana. Kuvioissa punaiset ympyrät symboloivat fiksaatioita ja niiden kestoa. Oranssit viivat symboloivat oikealla suuntautuvia sakkadeja. Kuvioista ilmenee lukihäiriössä esiintyvät pidemmät fiksaatiot sekä lyhyemmät sakkadit.



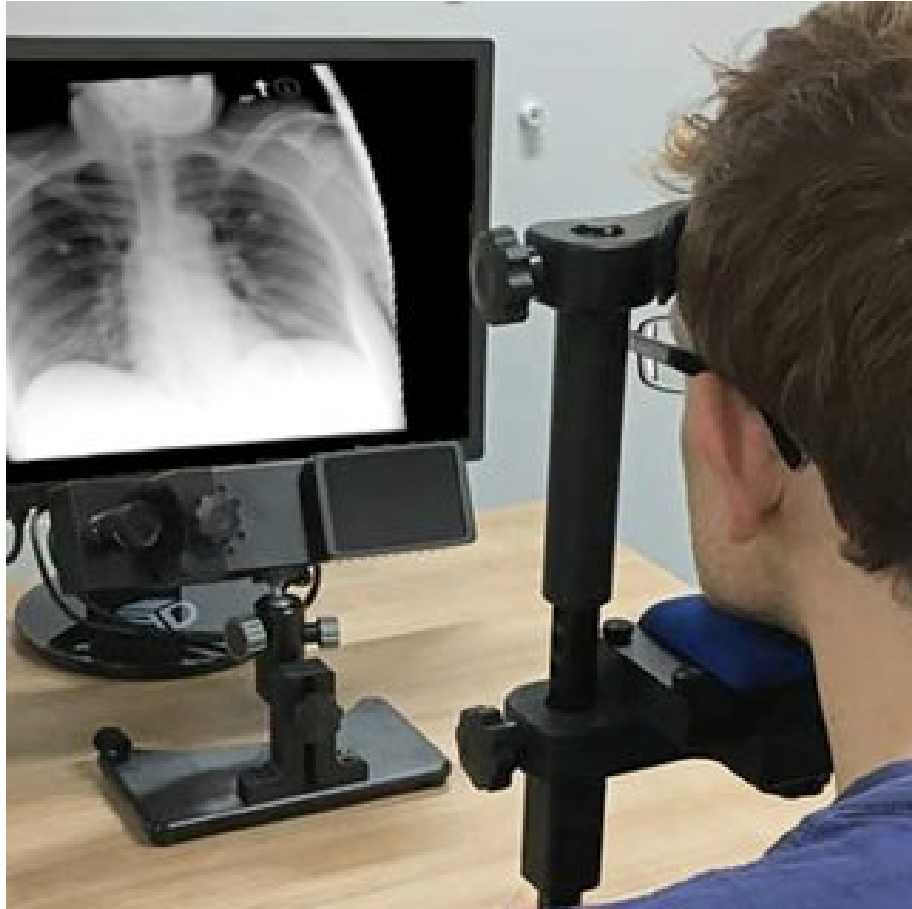
Kuvio 3. Lukihäiriöisen koehenkilön silmänliikkeet lukemisen aikana. Muokattu alkuperäisestä (Sekhar, Chandrashekar ym. 2023).

2.3 Silmänliikkeiden mittaaminen

Silmänliikkeet on eniten käytetty datalähde lukihäiriön tunnistamisessa. Shalileh ym. (2023) löysivät seitsemän uniikkia datajoukkoa silmänliikkeistä, kun taas EEG eli aivosähkökäyrällä tuotettuja datajoukkoja he löysivät kuusi kappaletta ja MRI eli magneettikuvantamisella tuotettuja datajoukkoja viisi kappaletta. Edellä mainittujen lisäksi lukihäiriön tunnistamisessa on käytetty fMRI-kuvantamista eli toiminnallista magneettikuvausta, kasvokuvaa tai -videota, lukitestejä sekä muita testituloksia (Shalileh ym. 2023).

Silmänliikkeiden mittaamiseen on käytössä eri valmistajien ratkaisuja. De Luca ym. (2002) hyödynsivät silmänliikkeiden mittaamiseen infrapunajärjestelmää (AMTech ET4 Eye Tracking System), jonka näytenopeus oli 500 Hz. Koehenkilöiden päät pidettiin paikoillaan kokeen aikana pää- ja leukatuen avulla. Järjestelmä kalibroitiin jokaisen testitilanteen aluksi. Raatikainen ym. (2021) käyttivät puolestaan silmänliikedataa, joka oli tuotettu käyttäen Eye-Link 1000 -järjestelmää 1000 Hz:n näytenopeudella. Laitteen kalibrointi suoritettiin ennen koetta ja se toistettiin mikäli huomattiin näkyviä päänliikkeitä. Shalileh ym. (2023) käyttivät datajoukkonsa muodostamiseen sekä Eyalink 1000 Plus- että Eyalink Portable Duo -järjestelmiä (ks. kuvio 4). Molempien järjestelmien näytenopeus oli 1000 Hz. Koehenkilöiden päät pidettiin paikoillaan leukatuen avulla.

Kuten edellä kerrotusta huomataan, voidaan silmänliikkeitä mitata erilaisten järjestelmien avulla. Yhteistä kuitenkin on järjestelmien tarpeeksi suuri näytenopeus, jotta pienimmätkin silmänliikkeet saadaan mitattua riittävällä tarkkuudella. Päänliikkeet tulee eliminoida joko erilaisten tukien avulla tai poistamalla niiden merkitys datasta jälkikäteen, sillä päänliikkeet saattavat aiheuttaa virheellisiä mittaustuloksia.



Kuvio 4. Pöydälle asetettava EyeLink 1000 Plus -järjestelmä. Koehenkilön pää pidetään paikoillaan leuka- ja otsatuen avulla. (Moreira ym. 2022)

3 Koneoppimisessa käytetty data

Valtaosa koneoppimismalleihin liittyvästä tutkimuksesta keskittyy mallien tehokkuuden ja tarkkuuden parantamiseen. Koska huomio on ollut pääasiassa tehokkuudessa, ei käytetyn datan laatuun ole kiinnitetty tarpeeksi huomiota. Datan merkitys koneoppimismalleissa on vähintään yhtä tärkeä kuin mallien tehokkuus ja tarkkuus, sillä virheet datasyötteessä voivat pahimmassa tapauksessa tehdä mallista käyttökeltottoman. Tässä kappaleessa tutkitaan minikälaista dataa lukihäiriöön ja silmänliikkeisiin liittyvässä tutkimuksessa on käytetty. (Polyzotis ym. 2019).

Shalileh ym. (2023) analysoivat tutkimuksessaan muun muassa fiksaatioiden kestoa millisekunneissa, fiksaatioita x- ja y-akseleilla, sanojen ohittamisen todennäköisyyttä, regressioita sekä kokonaisluku-aikaa. He havainnoivat, että fiksaatiot y-akselilla olivat merkittävämpiä kuin muunlaiset fiksaatiot. Rello ja Ballesteros (2015) puolestaan havainnoivat, että kokonaisluku-aika ja fiksaatioiden keskiarvo olivat hyödyllisimpiä piirteitä. Raatikainen ym. (2021) kokeilivat erilaisia piirrejoukkoja yksinkertaisesta aina laajentuvaan kompleksisuuteen. Yksinkertaisin piirrejoukko sisälsi neljä piirrettä, joita olivat fiksaatioiden kesto ja määrä sekä sakkadien kesto ja voimakkuus. Kompleksisemmat piirrejoukot sisälsivät muun muassa tilasiirtymämatriisien, sakkadien voimakkuuksien sekä fiksaatioiden kestojen keskiarvoja. Keskiarvojen käytöllä Raatikainen ym. (2021) pyrkivät vähentämään datan dimensioita ja hälyä.

Demograafisesta datasta Shalileh ym. (2023) nostivat merkittävimiksi koehenkilöiden älykkyydosamäärän, sukupuolen ja iän. Rello ja Ballesteros (2015) myös korostavat iän merkitystä omassa artikkelissaan. Syyksi iän merkitykseen he mainitsevat, että lukutaito karttuu iän myötä, jonka vuoksi lukihäiriöiset saattavat selvittää lukemiseen liittyvät ongelmansa vanhetessaan. Rello ja Ballesteros (2015) ovat kuitenkin sitä mieltä, että nämäkin henkilöt olisi hyödyllistä tunnistaa lukihäiriöisiksi, vaikka suurimmat lukemiseen liittyvät ongelmat olisi selätetty. Toisaalta Raatikainen ym. (2021) eivät hyödyntäneet omassa mallissaan lainkaan demograafista dataa ja tästä huolimatta heidän mallinsa tuotti kiitettäviä tuloksia. Demograafisen datan hyödyistä ei siis ole selkeää yksimielisyyttä.

Koneoppimismalleissa käytetty data useimmiten vaatii esikäsittelyä, ennen kuin sitä voidaan hyödyntää täysimittaisesti. Shalileh ym. (2023) aloittivat oman datansa esikäsittelyn muuttamalla luokkamuuttujat numeerisiksi muuttujiksi. Muutoksen jälkeen datajoukot sekä niitä vastaavat yksittäiset muuttujat standartisoitiin min-max -menetelmällä, jossa jokaisen datapisteen piirre vähennetään vastaavasta minimiarvosta ja jaetaan vaihteluvälillä. Rello ja Ballesteros (2015) myös muuttivat luokkamuuttujat numeerisiksi muuttujiksi, kun he olivat ensin poistaneet epäonnistuneesti mitatut datapisteet datajoukosta.

Datan esikäsittelyn lisäksi koneoppimismallit usein hyötyvät myös hyperparametrien optimoinnista. Hyperparametrit ovat parametrejä, joiden avulla voidaan hallita koneoppimismallien oppimisprosessia. Feurer ja Hutter (2019) kertovat, että jokaisella koneoppimismallilla on hyperparametrejä, mutta kiinnostus niiden optimointiin on kasvanut monimutkaisempien ja laskennallisesti vaativampien mallien myötä.

Raatikainen ym. (2021) päättivät käyttää oman mallinsa hyperparametrien optimointiin kahden metodia, jossa käytettiin ruudukkoetsintää (engl. grid search) kahteen merkittävimpään parametriin ja loput parametrit jätettiin SciKit-learnin oletusasetuksille (ks. (Pedregosa ym. 2011)). Shalileh ym. (2023) puolestaan päättivät Yangin ja Shamin (2020) tekemän tutkimuksen pohjalta hyödyntää bayesilaista optimointia. Bayesilainen optimointi on iteratiivinen prosessi, jonka pääkomponentit ovat sijaismalli sekä hankintafunktio, joka päättää mitä kohtaa arvioidaan seuraavaksi (Feurer ja Hutter 2019).

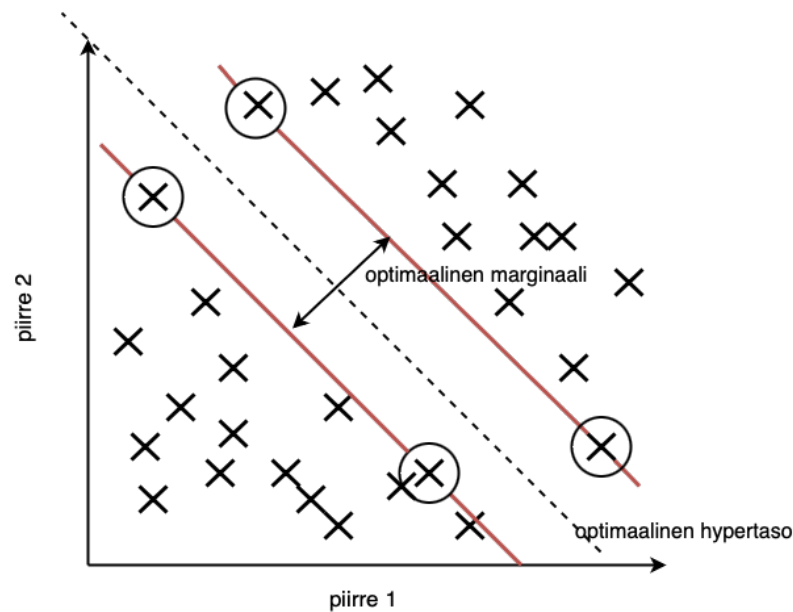
4 Koneoppimismallit lukemisen erityisvaikeuden tunnistamisessa

Koneoppimismallien avulla kyetään ratkaisemaan haastaviakin ongelmia. Voidaanko niillä kuitenkin luotettavasti todeta tai seuloa lukihäiriötä? Tässä luvussa tutkitaan, mitä eri koneoppimismalleja on käytetty lukihäiriön tunnistamiseen ja minkälaisia tuloksia eri malleilla on saatu. Tämän lisäksi luvun lopussa selvitetään koneoppimismallien käyttöön liittyviä ongelmia.

4.1 Koneoppimismallit

Tukivektorikone on lineaarinen luokitinmalli, joka jakaa datapisteet datajoukkoihin eli klustereihin siten, ettei yksikään datapiste jää marginaalitason väliin ja erottavien datapisteiden välimatka on mahdollisimman suuri. Kuviossa 5 on havainnollistettu kaksiulotteista luokitteluongelmaa, johon on käytetty tukivektorikonetta. Datapisteitä erottavaa vektoria kutsutaan hypertasoksi. Tukivektorikoneet soveltuvat erityisen hyvin numeeriseen ennustamiseen sekä luokitteluun, ja sitä onkin käytetty paljon lukihäiriön tunnistamisessa.

Rello ja Ballesteros (2015) hyödynsivät tukivektorikonetta omassa tutkimuksessaan, jossa he jakoivat datan kymmeneen noin yhtä suureen alijoukkoon ja käyttivät 90 % datasta tilastollisen mallin kouluttamiseen. Heidän mallinsa pystyi tunnistamaan lukihäiriön oikein 910 kertaa 1135:stä lukemasta. Malli tunnisti siis lukihäiriön oikein 80,18 % todennäköisyydellä loppujen lukemien jäädessä vääriksi negatiiviksi. Raatikainen ym. (2021) käyttivät myös tukivektorikonetta ja pääsivät hieman parempaan tulokseen. Heidän mallinsa pystyi tunnistamaan lukihäiriön 89,7 % tarkkuudella ja 84,8 % herkkyydellä. Parikhin ym. (2008) mukaan tarkkuus tarkoittaa saatujen oikeiden negatiivisten tulosten osuutta kaikista "terveistä" tutkittavista. Herkkyys puolestaan tarkoittaa testissä saatujen oikeiden positiivisten tulosten osuutta kaikista lukihäiriöisistä tutkittavista. Tukivektorikoneen lisäksi heidän koneoppimismallinsa hyödynsi satunnaisia metsiä, joita käytettiin datan piirrevalinnassa. Piirrevalinnasta on kerrottu tarkemmin luvussa 3. Edellä mainitun hybridimallin lisäksi Raatikainen ym. (2021) kokeilivat myös pelkkiä satunnaisia metsiä lukihäiriön tunnistamisessa, jonka

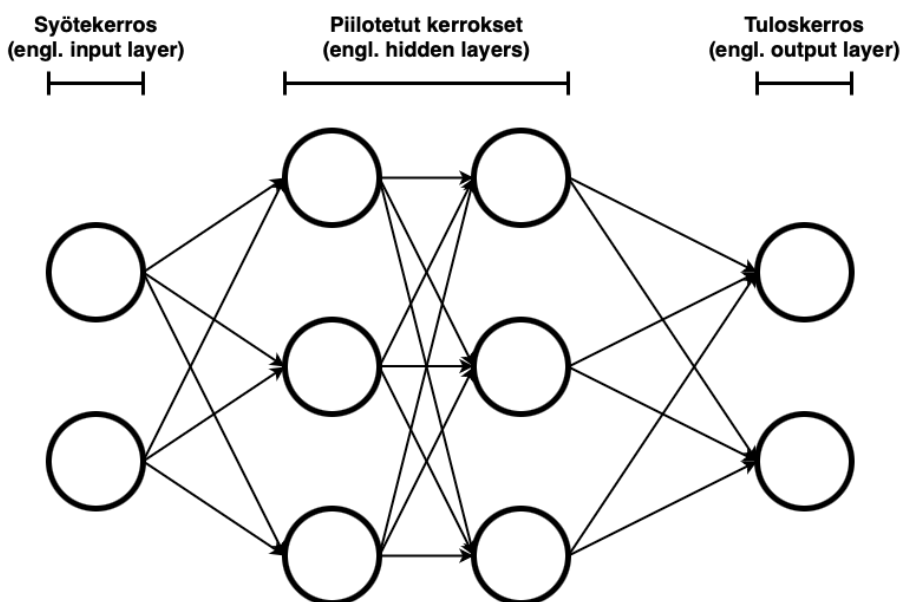


Kuvio 5. Esimerkki kaksiulotteista luokitteluongelmasta. Ympyröillä merkityt tukivektorit pyrkivät jakamaan datapisteet klustereihin siten, että erottavilla datapisteillä on suurin mahdollinen marginaali hypertasoon. Muokattu alkuperäisestä (Cortes ja Vapnik 1995).

avulla he pääsivät myös tyydyttävään tulokseen. Satunnaisia metsiä hyödyntävällä mallilla saavutettiin käytetystä datasetistä riippuen noin 80 % tarkkuus herkkyyden vaihdellessa 30 - 50 % välillä.

Shalileh ym. (2023) päättivät omassa tutkimuksessaan vertailla eri koneoppimismalleja sekä datalähteitä. He huomasivat, ettei mikään datalähde yksinään tuottanut tyydyttäviä tuloksia milläkään koneoppimismallilla. Kuitenkin datalähteitä yhdistelemällä päästiin hyviin tuloksiin, jolloin monikerroksista perseptroniverkkoa (engl. multilayer perceptron, MLP) hyödyntävä malli saavutti 91,3 % tarkkuuden ja 91,1 % herkkyyden. Popescun ym. (2009) mukaan monikerroksinen perseptroniverkko on neuroverkko, jossa signaaleja kuljetetaan eteenpäin verkon sisällä (ks. kuvio 6). Monikerroksinen perseptroniverkko ei sisällä silmukoita, eikä yksittäisen neuronin ulostulo vaikuta neuronin itseensä. Monikerroksista perseptroniverkkoa käytettäessä Shalileh ym. (2023) yhdistivät datalähteeksi fiksaatoraportteja sekä demograafista dataa. Samaa datajoukkoa kokeiltiin myös muilla koneoppimismalleilla ja tulokset olivat hyvin vaihtelevia. Epätarkin malli heidän tutkimuksensa perusteella oli logistinen regressioanalyysi, joka pystyi tunnistamaan lukihäiriön 57,3 % tarkkuudella. Kyseisen mal-

lin herkkyys oli 65,8 %. Huomattavaa on, että tukivektorikonetta hyödyntämällä Shalileh ym. (2023) pääsivät lähes samaan tarkkuuteen kuin Rello ja Ballesteros (2015) vuosia aiemmin. Molempien tukivektorikonemallin tarkkuus oli noin 80 %.



Kuvio 6. Monikerroksisen perseptroniverkon eteenpäin suuntautuvat neuronikerrokset. Muokattu alkuperäisestä (Popescu ym. 2009).

Koneoppimismalleja hyödyntämällä on päästy tyydyttäviin tuloksiin lukihäiriön tunnistamisessa silmänliikedatasta. Joissakin tapauksissa koneoppimismallit ovat hyötyneet, kun silmänliikedataan on yhdistetty demograafista dataa, kun taas toisissa tapauksissa pelkkä silmänliikedata on ollut riittävä hyvien tulosten saavuttamiseksi.

4.2 Koneoppimismallien ongelmat

Koneoppimismallien käyttäminen ei koskaan ole täysin ongelmatonta. Wagstaff (2012) mainitsee koneoppimismallien haasteiksi muun muassa varsinaisen ongelman määrittämisen, relevantin datan keräämisen, pätevän koneoppimismallin valitsemisen käsillä olevaan ongelmaan sekä tulosten tulkinnan. Selkeää rajaa onnistuneelle koneoppimismallille on hankala asettaa ja se riippuukin huomattavasti käsiteltävästä ongelmasta. Wagstaff (2012) mainitsee, että 80 % tarkkuus iiriksen tunnistuksessa voi olla riittävä kasvitieteilijöille, mutta toisaalta he voivat vaatia 99 % (tai jopa korkeampaa) tarkkuutta sienien luokitteluun syötäväksi

kelpaaviksi ja myrkyllisiksi.

Tutkielman aiheen ympäriltä merkittävin ongelma koneoppimismallien käytössä liittyy käytettyjen opetusdatajoukkojen vääristymiin. Shalileh ym. (2023) vertailevat seitsemässä eri tutkimuksessa käytettyjen koeryhmien kokoa (kontrolliryhmä vs. korkea riski lukihäiriöön -ryhmä). Kahdessa tarkastelussa olleessa tutkimuksessa lukihäiriöjoukon koko oli suurempi kuin kontrolliryhmän koko. Esimerkiksi toisessa näissä tutkimuksista lukihäiriöisten osuus oli 53 % koko tutkimusjoukosta. Tämä eroaa huomattavasti Käypä hoidon (2019) arviosta lukihäiriön esiintyvyydestä, joka on 1–7 % väestöstä. Vääristyneen datan johdosta koneoppimismalleilla on siis todellisuudesta eroava käsitys lukihäiriön esiintyvyydestä, joka voi johtaa virheellisiin tuloksiin.

Dataan liittyvien ongelmien lisäksi ongelmia ilmaantuu myös tulosten tulkinnassa. Ensimmäinen, lukihäiriön diagnosoinnista ei ole täyttä yksimielisyyttä (Bishop 2014). Miten voidaan olla varmoja, että koneoppimismalli on tunnistanut lukihäiriön oikein, kun diagnostisista kriteereistä ei ole varmuutta? Toiseksi, mikä koneoppimismallin tarkkuus on tarpeeksi tarkka. Mikäli tietty malli tunnistaa lukihäiriön onnistuneesti kahdeksalta henkilöltä kymmenestä, mutta kaksi henkilöistä jää ilman diagnoosia ja apua, voidaanko mallia pitää tarpeeksi luotettavana?

Yhteenvetona voidaan todeta, että koneoppimismallien käyttäminen ei ole ongelmatonta, jonka vuoksi onkin tärkeää punnita niiden hyötyjä suhteessa mahdollisiin haittoihin. Luvussa 2.1 kerrottiin lukihäiriön aiheuttamista ongelmista yksilön elämään, jotka voivat olla hyvinkin merkittäviä. Haittojen lisäksi täytyy muistaa Glazzardin (2010) tekemät havainnot diagnoosin positiivisesta vaikutuksesta yksilön negatiivisiin tunteisiin. Haittoihin ja hyötyihin peilattaessa väärät positiiviset tulokset tuskin aiheuttavat ongelmia yksilölle, mutta toisaalta mikäli niiden määrä kasvaa merkittävästi, aiheuttaa se varmasti ongelmia resurssien riittävyydelle interventoiden ja kuntoutuksen järjestämisessä. Vääriä positiivisia haitallisempaa ovat kuitenkin väärät negatiiviset tulokset. Nämä aiheuttavat sen, että väärän negatiivisen tuloksen saaneet yksilöt jäävät ilman apua ja joutuvat elämään lukihäiriön aiheuttamien ongelmien ja negatiivisten tunteiden kanssa.

5 Johtopäätökset

Tutkielman ensisijaisena tavoitteena oli selvittää, että voidaanko lukihäiriö tunnistaa luotettavasti silmänliikkeistä koneoppimismalleja apuna käyttäen. Parhaimmillaan koneoppimismallien avulla lukihäiriö kyettiin tunnistamaan erinomaisilla tuloksilla. On tärkeää kuitenkin muistaa, ettei yksikään koneoppimismalli tuottanut tyydyttäviä tuloksia yksittäisestä datalähteestä. Tämän vuoksi malleihin käytettyyn dataan tulee kiinnittää erityisen tarkkaa huomiota, jotta saadut tulokset ovat luotettavia.

Toisena tutkimuksen tavoitteena oli selvittää, onko eri koneoppimismallien välillä huomattavia eroja. Tutkielmaan valituissa artikkeleissa useimmin käytetty koneoppimismalli oli tukivektorikone, jonka avulla on päästy noin 80–90 % tarkkuuteen. Parhaat tulokset on tuottanut monikerroksinen perseptroniverkko, jonka tarkkuus oli 91,3 % ja herkkyys 91,1 %. Samaa datajoukkoa käyttämällä eri koneoppimismallien tarkkuus ja herkkyys vaihtelivat merkittävästi. Testatuista malleista parhaat tulokset saavutti jo edellä mainittu monikerroksinen perseptroniverkko ja heikoimmat tulokset puolestaan logistinen regressioanalyysi. Eri mallien välillä voidaan havaita eroja, joten ei ole yhdentekevää, mitä mallia päätetään käyttää.

Viimeisenä tavoitteena tutkielmassa oli tarkoitus perehtyä koneoppimismallien ongelmiin. Vääristyneen opetusdatan käyttö voi johtaa myös vääristyneisiin tuloksiin. Tällöin osa henkilöistä voi jäädä ilman diagnoosia ja toisaalta osa taas voi saada diagnoosin virheellisesti. Voidaanko koneoppimismallien tuottamiin tuloksiin luottaa? Voidaanko vastuu diagnosoinnista antaa koneoppimismallien päätettäväksi vai tulisiko niitä käyttää vain lukihäiriön seurlontaan?

Tulevaisuudessa voisi olla kannattavaa tutkia, mikä silmänliikkeistä saatavasta ja demografisesta datasta on kaikkein hyödyllisintä koneoppimismallien kannalta. Koneoppimismalleja olisi mielekästä tutkia myös koeryhmässä, jossa lukihäiriön esiintyvyys olisi lähempänä todellista esiintyvyyttä väestössä. Tämän jälkeen voitaisiin toteuttaa luotettavampaa vertailua eri diagnosointimenetelmien välillä ja miettiä, onko koneoppimismalleissa todella tulevaisuus lukihäiriön tunnistamisessa silmänliikkeiden avulla.

Lähteet

- Bishop, D. V. M. 2014. “Ten questions about terminology for children with unexplained language problems”. *International Journal of Language & Communication Disorders* 49 (4): 381–415. <https://doi.org/10.1111/1460-6984.12101>.
- Clifton, Charles, Fernanda Ferreira, John M. Henderson, Albrecht W. Inhoff, Simon P. Livesedge, Erik D. Reichle ja Elizabeth R. Schotter. 2016. “Eye movements in reading and information processing: Keith Rayner’s 40year legacy”. *Journal of Memory and Language* 86:1–19. ISSN: 0749-596X. <https://doi.org/10.1016/j.jml.2015.07.004>.
- Cortes, Corinna ja Vladimir Vapnik. 1995. “Support-Vector Networks”. *Machine Learning* 20, numero 3 (syyskuu): 273–297. ISSN: 1573-0565. <https://doi.org/10.1023/A:1022627411411>.
- De Luca, M., M. Borrelli, A. Judica, D. Spinelli ja P. Zoccolotti. 2002. “Reading words and pseudowords: An eye movement study of developmental dyslexia”. *Brain and Language* 80 (3): 617–626. ISSN: 0093-934X. <https://doi.org/10.1006/brln.2001.2637>.
- De Luca, M., E. Di Pace, A. Judica, D. Spinelli ja P. Zoccolotti. 1999. “Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia”. *Neuropsychologia* 37 (12): 1407–1420. ISSN: 0028-3932. [https://doi.org/10.1016/S0028-3932\(99\)00038-X](https://doi.org/10.1016/S0028-3932(99)00038-X).
- Feurer, Matthias ja Frank Hutter. 2019. “Hyperparameter optimization”. *Automated machine learning: Methods, systems, challenges*, 3–33.
- Glazzard, Jonathan. 2010. “The impact of dyslexia on pupils’ self-esteem”. *Support for Learning* 25 (2): 63–69. <https://doi.org/10.1111/j.1467-9604.2010.01442.x>.
- Humphrey, Neil. 2002. “Teacher and pupil ratings of self-esteem in developmental dyslexia”. *British Journal of Special Education* 29 (1): 29–36. <https://doi.org/10.1111/1467-8527.00234>.

- Hutzler, F. ja H. Wimmer. 2004. “Eye movements of dyslexic children when reading in a regular orthography”. *Brain and Language* 89 (1): 235–242. ISSN: 0093-934X. [https://doi.org/10.1016/S0093-934X\(03\)00401-2](https://doi.org/10.1016/S0093-934X(03)00401-2).
- Käypä hoito. 2019. “Kehityksellinen kielihäiriö (kielellinen erityisvaikeus, lapset ja nuoret)”. Viitattu 6. maaliskuuta 2024. <https://www.kaypahoito.fi/hoi50085>.
- Mainela-Arnold, Elina. 2019. “Terminologia, kriteerit ja liitännäisoireet kehityksellisessä kielihäiriössä”. Teoksessa Loukusa, S., Tolonen, AK., Nylund, A. & Kautto, A.(toim.) *Kehityksellisen kielihäiriön monet kasvot. Puheen ja kielen tutkimuksen yhdistyksen julkaisuja* 51:9–18.
- Moreira, Catarina, Isabel Blanco Nobre, Sandra Costa Sousa, João Madeiras Pereira ja Joaquim Jorge. 2022. “Improving x-ray diagnostics through eye-tracking and XR”. Teoksessa *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 450–453. IEEE.
- Parikh, Rajul, Annie Mathai, Shefali Parikh, G Chandra Sekhar ja Ravi Thomas. 2008. “Understanding and using sensitivity, specificity and predictive values”. *Indian journal of ophthalmology* 56 (1): 45–50.
- Pavlidis, G.Th. 1981. “Do eye movements hold the key to dyslexia?” *Neuropsychologia* 19 (1): 57–64. ISSN: 0028-3932. [https://doi.org/10.1016/0028-3932\(81\)90044-0](https://doi.org/10.1016/0028-3932(81)90044-0).
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel ym. 2011. “Scikit-learn: Machine Learning in Python”. *Journal of Machine Learning Research* 12:2825–2830.
- Polyzotis, Neoklis, Martin Zinkevich, Sudip Roy, Eric Breck ja Steven Whang. 2019. “Data validation for machine learning”. *Proceedings of machine learning and systems* 1:334–347.
- Popescu, Marius-Constantin, Valentina E Balas, Liliana Perescu-Popescu ja Nikos Mastorakis. 2009. “Multilayer perceptron and neural networks”. *WSEAS Transactions on Circuits and Systems* 8 (7): 579–588.

Prado, Chloé, Matthieu Dubois ja Sylviane Valdois. 2007. “The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span”. *Vision Research* 47, numero 19 (syyskuu): 2521–2530. ISSN: 0042-6989. <https://doi.org/10.1016/j.visres.2007.06.001>.

Raatikainen, Peter, Jarkko Hautala, Otto Loberg, Tommi Kärkkäinen, Paavo Leppänen ja Paavo Nieminen. 2021. “Detection of developmental dyslexia with machine learning using eye movement data”. *Array* 12 (joulukuu): 100087. ISSN: 2590-0056. <https://doi.org/10.1016/j.array.2021.100087>.

Rello, Luz ja Miguel Ballesteros. 2015. “Detecting readers with dyslexia using machine learning with eye tracking measures”. Teoksessa *Proceedings of the 12th International Web for All Conference*. W4A '15. Florence, Italy: Association for Computing Machinery. ISBN: 9781450333429. <https://doi.org/10.1145/2745555.2746644>.

Sekhar, SR, Swathi Chandrashekar ym. 2023. “Eye Movement Interpretation for Detecting Dyslexia Using Machine Learning Techniques”. *International Journal of Computing and Digital Systems* 13 (1). ISSN: 2210-142X. <https://doi.org/10.12785/ijcnds/130162>.

Shalileh, S., D. Ignatov, A. Lopukhina ja O. Dragoy. 2023. “Identifying dyslexia in school pupils from eye movement and demographic data using artificial intelligence”. *PLoS ONE* 18 (11 November). ISSN: 1932-6203. <https://doi.org/10.1371/journal.pone.0292047>.

Wagstaff, Kiri L. 2012. “Machine learning that matters”. Teoksessa *Proceedings of the 29th International Conference on International Conference on Machine Learning*, 1851–1856.

Yang, Li ja Abdallah Shami. 2020. “On hyperparameter optimization of machine learning algorithms: Theory and practice”. *Neurocomputing* 415:295–316. ISSN: 0925-2312. <https://doi.org/10.1016/j.neucom.2020.07.061>.