

Karoliina Talvitie-Lamberg, Minna Silvennoinen, Hannu Moilanen ja
Jari Korpela

Tekoäly sosiaalisen syrjäytymisen ennakoivassa tunnistamisessa - Case THL kouluterveyskysely



Informaatioteknologian tiedekunnan julkaisuja
No. 62/2018

Editor: Pekka Neittaanmäki

Covers: Petri Vähäkainu ja Matti Savonen

Copyright © 2018

Petri Vähäkainu ja Jyväskylän yliopisto

ISBN 978-951-39-7574-6 (verkkokj.)

ISSN 2323-5004

Jyväskylä 2018

Tekoäly sosiaalisen syrjäytymisen ennakoivassa tunnistamisessa - Case THL kouluterveyskysely

Karoliina Talvitie-Lamberg

Minna Silvennoinen

Hannu Moilanen

Jari Korpela

Tämä julkaisu on toteutettu osana WHC-hanketta, johon Jyväskylän yliopisto on saanut rahoituksen Business-Finlandilta.

Business Finland-hanke: WHC

KUVIOT

KUVIO 1. OTE KYSELYSTÄ JOHON MERKITTYY TUTKIJATRIANGULAATION AVULLA SAADUT PISTEMÄÄRÄT SEKÄ LUOKITUS	11
KUVIO 2. LUOKITUSTYÖSKENTELEN KUVAUS.....	12
KUVIO 3. NÄYTÖNKAAPPAUS MODELER -OHJELMASTA, JOSSA NÄKYVÄT KOULUTERVEYSKYSELYN TIETOJA KÄSITTELEVÄT SOLMUT JA STREAMI	14
KUVIO 4. VASTAUKSEN PISTEARVON MÄÄRITTELY SPSS MODELER –OHJELMAN NODESSA.....	14
KUVIO 5. SYRJÄYTYMISEN RISKILUOKITUKSEN MUODOSTUMINEN TUTKIJA- JA ASIANTUNTIJA- ARVIOIDEN KAUTTA	17
KUVIO 6. KARTTA TUTKIJOIDEN JA ASIANTUNTIJOIDEN PISTEITYKSEN KESKIARVON PERUSTEELLA, JOKA KUVAU RISKISSÄ OLEVIA OPPILAINEN JAKAUTUMISEN KESKI-SUOMEN KUNTIIN (MONTAKO PROSENTTIA KUNNISSA ON RISKISSÄ OLEVIA OPPILAITA).....	18
KUVIO 7. PLOTTAUS TUTKIJOIDEN JA ASIANTUNTIJOIDEN PISTEITYKSEN KESKIARVON PERUSTEELLA. PLOTTAUS KUVAU KUNTIEN KOKONAISRISKIPISTEITÄ (Y) SEKÄ RISKISSÄ OLEVIA OPPILAINEN MÄÄRÄÄ (PLOTIN KOKO).....	19
KUVIO 8. KUNTAKOHTAISET SYRJÄYTYMISRISKIPISTEIDEN KESKIARVOT TUTKIJOIDEN JA ASIANTUNTIJOIDEN PISTEITYKSEN KESKIARVON PERUSTEELLA. NUMERO KUNNAN KESKELLÄ KERTOO KYSELYYN VASTANNEIDEN OPPILAINEN LUKUMÄÄRÄN.....	20
KUVIO 9. TAULUKKO, KUNNAN OPPILAINEN SAAMASTA KESKIMÄÄRÄISESTÄ PISTEMÄÄRÄSTÄ JA ALALUOKKIEN VERTAILUSTA KESKENÄÄN MAKSIMIIN SUHTEELLISILLA ARVOILLA	21
KUVIO 10. HISTOGRAMMI KUVAU ÄÄNEKOSKEN EI RISKISSÄ OLEVIA OPPILAINEN LUKUMÄÄRIÄ (Y) SUHTEELLISTEN SOSIAALISIA SUHTEITA PERUSTEELLA	22
KUVIO 11. RANDOM TREES - LUOKITTELEVA MALLI	36
KUVIO 12. TÄRKEIMMÄT KYSYMYKSET, ELÄMÄÄN TYYTYVÄISYYS	37
KUVIO 13. PÄÄTTELYPUU MALLIN TULOKSET KAIKILLA KYSYMYKSILLÄ.....	39
KUVIO 14. PÄÄTTELYPUUSSA JOPA 10 KYSYMYKSELLÄ SAADAAN TARKKUUDEKSI 85 %	39
KUVIO 15. ENNUSTEISIIN KÄYTETTÄVÄT INPUT-TIETOJEN SEKÄ ENNUSTETTAVAN TARGET-TIEDON MÄÄRITTELY	40
KUVIO 16. PÄÄTTELYPUU, JOSSA 14 'TOP DECISION RULES' – KYSYMYSTÄ. (96 % TARKKUUS)	40
KUVIO 17. MODELER –OHJELMAN LÖYTÄMÄT MERKITTÄVIMMÄT KYMMENEN KYSYMYSTÄ, JOTKA ENNUSTAVAT SYRJÄYTYMISEN RISKIÄ THL:N KYSELYAINEISTOSSA.....	41

SISÄLLYSLUETTELO

1	Hankkeen lähtökohtia - Tekoälypohjaisen teknologian mahdollisuudet, syrjäytymisen riskit ja ennakoiva tunnistaminen	1
2	Sosiaalisen syrjäytymisen määrittely	3
2.1	Nuorten syrjäytymisen riskitekijät	4
2.1.1	Terveys, terveystottumukset ja elämänhallinta	4
2.1.2	Koulutus, koulunkäynnin haasteet ja tulevaisuus	5
2.1.3	Ulkopuolisuus, yksinäisyys ja sosiaaliset ongelmat	6
2.1.4	Elämänasenne ja elämänhallinta	7
2.1.5	Perhe, asuminen ja ylisukupolvisuus	8
2.1.6	Tuen ja avun puute	9
2.2	Tutkimuskysymys	9
3	Menetelmät ja työprosessi	10
3.1	Menetelmän valinta	10
3.2	Tutkijatriangulaatio, aineiston luokittelu ja pisteytys	10
3.3	Data-analyysi SPSS Modelerilla	13
4	Tulokset	16
4.1	Syrjäytymisriskin luokitus ja pisteytys	16
4.2	SPSS Modeler syrjäytymisennusteen työkaluna	18
4.2.1	Suurimmat syrjäytymisriskipisteet	18
4.2.2	Suurimman syrjäytymisriskin oppilaiden sijoittuminen Keski-Suomen alueelle	19
4.2.3	Alaluokkien keskinäinen vertailtavuus	20
4.3	Merkittävien kysymysten löytyminen	21
5	THL aineiston soveltuvuus ja SPSS Modeler tunnistettujen syrjäytymisen indikaattoreiden/riskitekijöiden kartoittamisessa - Keskeiset tulokset ja niiden johtopäätökset	23
6	Jatkotutkimukset ja tulevaisuus	26
7	Lähteet	29
	Liitteet	36

1 HANKKEEN LÄHTÖKOHTIA - TEKOÄLYPOHJAISEN TEKNOLOGIAN MAHDOLLISUUDET, SYRJÄYTYMISEN RISKIT JA ENNAKOIVA TUNNISTAMINEN

Olemassa olevan big datan eettinen hyödyntäminen yhteiskunnallisesti merkittävien ongelmien ratkaisussa erilaisten sosiaalisten innovaatioiden avulla on tulevaisuuden yksi kehityssuunta. Tässä raportissa esitellään tekoälyratkaisu, joka kehitettiin sosiaalisen syrjäytymisen ennakoivaan tunnistamiseen. Ratkaisun pohjana käytettiin Terveyden ja hyvinvoinnin laitoksen (THL) kansallista 8 - 9 luokkaisille tehtyä kouluterveyskyselyaineistoa. Aineistosta etsittiin aiempaan tutkimuskirjallisuuteen perustuen syrjäytymiseen ja osattomuuteen liittyviä riskejä kartoittavat kysymykset ja näiden pohjalta luotiin ei yksilöitävissä oleva oppilaskohtainen riskiarvio, joka on identifioitavissa koulutasolle. Mallin avulla voidaan laskea riskiarvio syrjäytymisuhan alla olevien oppilaiden määrästä sekä yksittäiselle koululle, että alueelle, esimerkiksi kunnalle. Syrjäytymisen ennakoivaan tunnistamiseen kehitetty ratkaisu hyödyntää IBM:n SPSS modeler -ohjelmaa, joka on tiedon analysointiin tarkoitettu ohjelma graafisella käyttöliittymällä.

Hankkeen lähtökohdat

Sosiaalinen syrjäytyminen on tällä hetkellä merkittävä yhteiskunnallisen keskustelun ja mielenkiinnon kohde. Syrjäytymiseen liittyvässä tutkimuskirjallisuudessa sekä julkisessa keskustelussa syrjäytymiskehitystä on usein tarkasteltu ns. riskifaktoriparadigman kautta (Crow, 2004). Syrjäytymiskehitykseen liittyviä tunnistettuja riskitekijöitä ovat tällöin muun muassa *työttömyys, heikko koulutus, yksinäisyys, huono terveys sekä syrjäytymisen ylisukupolvisuus* (Coles, 2000; MacDonald, 2006; Terveyden ja hyvinvoinnin laitos, 2016). Syrjäytyminen on paitsi suuri yhteiskunnallinen kuluerä, myös laajempi osattomuuteen liittyvä ongelma ja sitä kautta esimerkiksi radikalisoitumiskehitykseen vaikuttava tekijä. Näin syrjäytyminen on merkittävä ongelma niin yhteiskunnallisen vakauden kuin yksilön hyvinvoinninkin näkökulmasta. Jyväskylän Yliopiston IT-tiedekunnan hankkeessa haluttiin tutkia suomalaisen sosiaali- ja terveydenhuollon sekä hyvinvointiin liittyvän datan tekoälypohjaista hyödyntämistä ja teknologioiden kyvykkyyttä yhteiskunnallisesti merkittävässä ratkaisussa. Hankkeen ensimmäisessä vaiheessa vuoden 2017 aikana innovoitiin laajan noin 250 sosiaali- terveys- hyvinvointi- ja it-alan ammattilaisen voimin ratkaisuja yhdessä tunnistettuihin merkittäviin terveyden ja hyvinvoinnin alueen ongelmiin. Ongelmalähtöiseen työskentelyyn pohjaten hankkeessa tuotettiin yhteensä 35 tekoälypohjaisia teknologioita hyödyntävää käytätapauskuvausta tunnistettujen ongelmien ratkaisuksi. Yksi keskeinen tunnistettu ongelma oli lasten ja nuorten syrjäytymiskehitys sekä yksinäisyys. Tästä ongelmasta muodostettiin hankkeessa tekoälyteknologiaa hyödyntävä käytätapauskuvaus: ”lasten- ja nuorten sosiaalisen syrjäytymisen ennaltaehkäisy”, jonka päätarkoituksena on syrjäytymiskehityksen ennakoiva tunnistaminen sivistys-, sosiaali-, ja terveysdatan avulla ja sen pohjalta toteutettava varhaisen puuttumisen malli lähiverkoston avulla (esimerkiksi koulu- ja varhaiskasvatushenkilökunta). Toteutuessaan käytätapauksen ratkaisu mahdollistaa ennakoivan ja

ennaltaehkäisevän syrjäytymisen ja huono-osaisuuden tekoälypohjaisen tunnistamisen, osana lasten, nuorten ja perheiden palveluintegraatiota.

Hankkeen ensimmäisen vaiheen käyttötapausmäärittelyn jälkeen (ongelman tunnistaminen ja ratkaisu) haluttiin kakkosvaiheessa käytännössä tarkastella, *kuinka tekoälypohjaiset teknologiat, etenkin ennakoivan analytiikan ratkaisut, voisivat todellisuudessa toimia syrjäytymiskehityksen ennaltaehkäisevässä tunnistamisessa*. Koulu- ja opiskeluympäristössä kerätään jatkuvasti keskeistä yksilön hyvinvointiin ja osallisuuteen liittyvää dataa, mutta sitä hyödynnetään vielä riittämättömästi. Mikäli ennakoivaa syrjäytymiskehitystä ja varhaisen puuttumisen mallia halutaan tosiasiaassa toteuttaa, aineiston käyttömahdollisuuksia on tärkeää tutkia. *Tekoälyteknologioiden kyvykkyyttä lähdettiin testaamaan olemassa olevan data-aineiston avulla, jota kerätään laajalti, mutta jota ei sellaisenaan vielä hyödynnetä syrjäytymisen riskitekijöiden tunnistamisessa*. Tällainen aineisto saatiin käyttöön Terveyden ja Hyvinvoinnin (THL) laitokselta. THL:n 2017 kouluterveyskysely tehdään joka toinen vuosi 8 - 9 luokkalaisille. Kyselyssä kartoitetaan yksilön omaa kokemusta muun muassa koulunkäynnistä, perheestä, terveyden tilasta, terveystottumuksista, sosiaalisista suhteista sekä päihteiden käytöstä.

2 SOSIAALISEN SYRJÄYTYMISEN MÄÄRITTELY

Syrjäytymisen käsite on lähellä huono-osaisuuden käsitettä ja yksilötasolla siinä on kyse useimmiten siihen liittyvien tekijöiden kasautumisesta, mutta kasautuvat tekijät voivat kuitenkin olla toisistaan riippumattomia (Rintanen, 2000). Sosiaalisen syrjäytymisen ja huono-osaisuuden vastakäsitteeksi määritellään usein yhteiskunnallinen osallisuus ja sen kokemus (Nivala & Rynänen, 2014). Yksilön osallisuus on näin osallisuutta yhteiskunnan tarjoamista mahdollisuuksista, jotka turvaavat yksilön elintason ja elämänlaadun. Näin ymmärrettynä yhteiskunnallinen osallisuus on elinolosuhteisiin liittyvä kysymys ja osallisuutta esimerkiksi niin koulutuksesta, työstä kuin yhteiskunnan palveluista. Sosiaalinen syrjäytyminen määrittyy puolestaan osattomuutena näistä.

Sosiaalisesta syrjäytymisestä puhuttaessa etusijalle nostetaan monesti sen yhteiskunnalle aiheuttamat kulut, jolloin syrjäytyminen määritellään elinolosuhteiden kautta. Sosiaalinen syrjäytyminen johtaa yhteiskunnallisella tasolla työttömyyden ja kouluttamattomuuden kierteeseen ja menetettyyn työpanokseen. Esimerkiksi vuonna 2010 syrjäytyneitä 15–29-vuotiaita nuoria oli yhteensä noin 51 300 (Palola, Hannikainen-Ingman & Karjalainen, 2012), noin 5 prosenttia kaikista tämän ikäisistä nuorista. Tämä aiheuttaa noin 75 000 € kustannukset vuodessa nuorta kohden (Palola ym., 2012). Syrjäytyneistä nuorista työttömiä työnhakijoita oli 18 800 ja muita työvoiman ulkopuolisia 32 500, lisäksi osa on ”kadoksissa”, ts. työttömät nuoret jotka eivät ole työnhakijoina (Palola ym., 2012). Osa väestöstä näyttääkin syrjäytyneen pysyvästi (Sosiaali- ja terveystieteiden tutkimuskeskus, 2015). Eurostatin arvion mukaan jopa 168 000 suomalaista lasta on syrjäytymisen ja/tai köyhyyden vaarassa. Tämä tarkoittaa jopa 15 % nuorista. (Pelastakaa lapset ry, 2016).

Suomessa syrjäytymistä ja syrjäytymiseen vaikuttavia tekijöitä on tutkittu viimeisen kolmen vuosikymmenen aikana laajasti. Syrjäytyminen on laaja käsite ja tutkimuksen näkökulmasta riippuen syrjäytyminen on määritelty hieman eri tavoin. Syrjäytyminen tarkoittaa joutumista yhteiskunnan ”yhteisyyden” (Helne, 2002) tai valtakulttuurin määrittämän elämäntavan ulkopuolelle (Siljander, 1996, 3) ja yksilön palauttaminen takaisin sen piiriin voidaan katsoa olevan lähtökohtana syrjäytymisen vastaiselle työlle (Helne, 2002). Usein syrjäytyminen kuvataan pitkänä prosessina, johon liittyy useiden syrjäytymisen riskitekijöiden kasaantumista nuoren elämässä pidemmän ajan kuluessa. Nuoren elämässä yksittäinen tunnistettu syrjäytymisriski tai elämään liittyvät vaikeudet eivät välttämättä horjuta elämänhallintaa, mikäli muut hyvinvoinnin osa-alueet ovat kunnossa (Berg, Huurre, Kiviruusu & Aro, 2011). Syrjäytymisen ehkäisemiseksi tulisikin keskittyä kasautuvan huono-osaisuuden tunnistamiseen, erityisesti sosiaalisten suhteiden, elämäntyyliin, koulutuksen ja sosioekonomisen aseman näkökulmasta (Berg ym., 2011). Tällä hetkellä syrjäytymisen tunnistaminen tapahtuu liian myöhäisessä vaiheessa, kun useimmat syrjäytymisen riskitekijät ovat jo ehtineet kasautua. Syrjäytymisriskin tunnistaminen edellyttää varhaista moniammatillista yhteistyötä ja saumatonta palveluketjua toteutuakseen (Halonen, Aaltonen, Hämäläinen, Karppi, Kaukinen, Kervilä, Lehtinen, Pere, Puukka, Siitonen, Silvennoinen & Talo, 2007).

2.1 Nuorten syrjäytymisen riskitekijät

Syrjäytyneet ovat moniongelmaisia ihmisiä, joiden elämänhallinta on heikentynyt ja ongelmat kasaantuneet tai pitkittyneet, jolloin ne eivät ole enää yksilön hallittavissa. (Lämsä, 1998; Siljander, 1996). Syrjäytymisvaarassa olevien tunnistaminen on haastavaa, sillä lasten ja nuorten syrjäytymisessä erilaiset huono-osaisuuden osa-alueet ketjuuntuvat ja kasautuvat (Terveyden ja hyvinvoinnin laitos, 2018). Tässä luvussa esittelemme keskeisiä tutkimuskirjallisuudessa tunnistettuja syrjäytymiskehitykseen johtavia osa-alueita.

2.1.1 Terveys, terveystottumukset ja elämänhallinta

Terveyteen liittyvät syrjäytymisen ulottuvuudet voivat olla käänteisessä yhteydessä toisiinsa, esimerkiksi heikko terveys voi johtaa syrjäytymisen eskaloitumiseen ja toisaalta jokin hyvinvoinnin ongelma, kuten taloudelliset seikat, voi johtaa terveydentilan heikkenemiseen (Rintanen, 2000). Syrjäytymisen riskiä lisäävinä tekijöinä voidaan pitää muun muassa kroonisia sairauksia, jotka heikentävät toiminnallisuutta sekä psyykkisiä sairauksia, jotka aiheuttavat häiriökäyttäytymistä - sairauksiin liittyvät olosuhteet ovat ratkaisevia siinä miten ne vaikuttavat yksilön elämään (Rintanen, 2000). Rintasen (2000) mukaan pitkäaikaissairaudet lisäävät nuoren syrjäytymisriskiä. Syrjäytymisvaarassa olevilla nuorilla esiintyy suhteellisen runsaasti lieviä fyysisiä sairauksia (Halonen ym., 2004). Erityisesti psyykkiset sairaudet ja mielenterveyden ongelmat ovat merkittävä syrjäytymisen riskitekijä (Halonen ym., 2004; Ek, Saari, Viinamäki & Järvelin, 2004). Ekin ym. (2004) mukaan huonoksi koettu terveydentila vaikuttaa enemmän syrjäytymisriskiin kuin varsinainen lääkärin antama diagnoosi. Itse koettu hyvä fyysinen terveys on yhteydessä myös vahvaan elämänhallintaan, joka puolestaan suojaa syrjäytymiseltä (Kinnunen, 2003). Kouluttamattomat ja syrjäytyneet nuoret ovat muita *sairaampia*, terveydenhuollon kulut ovat keskimäärin 7-kertaisia ei-syrjäytyneisiin verrattuna ja esimerkiksi mielenterveyslääkkeitä käyttää keskimäärin noin 60% pysyvästi syrjäytyneistä.

Syrjäytymiskehitys on yhteydessä myös puutteellisiin *elämänhallintataitoihin, terveystietoihin ja koulunkäyntitaitoihin*. Huonot terveystottumukset ja elämänhallinnan vaikeudet, kuten häiriintynyt vuorokausirytm, runsas päihteiden käyttö sekä vähäinen liikunta ja terveystietämyksen puute ovat tunnettuja syrjäytymisen riskitekijöitä (Rintanen, 2000; Halonen ym., 2007; Luopa, Lommi, Kinnunen & Jokela, 2010; Ek ym., 2007). Kinnusen (2003) mukaan esimerkiksi tupakointi korreloi usein heikompien elämänhallintataitojen kanssa. Aiempien kouluterveyskyselyiden mukaan ammattikoululaisilla on keskimäärin huonommat terveystottumukset kuin lukiolaisilla, ja nuoren terveyskäyttäytymisellä on havaittu olevan vaikutusta syrjäytymiseen (Halonen ym., 2007).

2.1.2 Koulutus, koulunkäynnin haasteet ja tulevaisuus

Nuoren koulutuksen puute on yksi merkittävin syrjäytymisriskiin vaikuttava tekijä (Myrskylä, 2012). Ongelmat koulunkäynnissä, heikko koulumenestys- ja motivaatio, huono kouluviihtyvyys sekä oppimisvaikeudet vaikuttavat nuorten syrjäytymisriskiin (Savolainen, 2001; Nurmi, 2011; Onatsu-Arvilommi & Nurmi, 2000; Nurmi, Salmela-Aro & Ruotsalainen 1994). Heikko menestys peruskoulussa vaikuttaa toisen asteen koulutukseen pääsyyn ja kouluttamattomuus johtaa epävarmaan ja heikkoon asemaan työmarkkinoilla. Tilastoihin perustuvissa tutkimuksissa (muun muassa Myrskylä, 2012) syrjäytynyt nuori on 15 - 29-vuotias, jolla *ei ole peruskoulun jälkeistä koulutusta, joka ei opiskele*, ja joka ei myöskään käy töissä tai hoida lasta kotona. On kuitenkin huomattava, että jotkut nuoret ovat itse valinneet koulutuksen ja työn ulkopuolelle jättäytymisen ja tällöin heidän kohdalla ei ole mielekästä puhua perinteisestä syrjäytymisestä (Järvinen & Jahnuainen, 2001). Matalan koulutustason on havaittu liittyvän myös terveysongelmiin (Rintanen, 2000; Kestilä 2008) liialliseen alkoholin käyttöön, puutteellisiin elämäntaitoihin (Martelin, Koskinen, Kestilä & Aromaa, 2005; Kestilä, 2008) ja rikollisuuteen (Aaltonen, 2013). Jatkokoulutuksen on nähty korreloivan positiivisesti mielenterveyteen (Behrman, 2015) ja esimerkiksi toisen asteen jatkokoulutuksen on todettu suojaavan masennusoireilta (McFarland, 2018).

Oppimisvaikeuksia omaavilla nuorilla syntyy vähemmän onnistumisen kokemuksia, mikä voi vaikuttaa nuoren henkiseen hyvinvointiin ja minä-kuvan kehittymiseen (Halonen ym., 2004). Häiriökäyttäytymistä koulussa voidaan pitää yhtenä syrjäytymisen riskitekijöistä, sillä usein häiriökäyttäytyminen liittyy nuoren henkiseen pahoinvointiin (Rintanen, 2000). Koulukiusaaminen on myös osoittautunut mielenterveydelliseksi ja kehitykselliseksi riskitekijäksi, joka lisää nuoren syrjäytymisriskiä (Kuronen, 2010). Terveiden ja hyvinvointi laitoksen toteuttamissa kouluterveyskyselyissä 2000 - 2009 näkyy eroja terveystottumuksissa kouluasteen ja sukupuolten välillä siten, että pojilla syrjäytymisriskit olivat suuremmat, samoin ammattikoululaisilla verrattuna lukiolaisiin (Luopa ym., 2010).

Perusasteen jälkeisen matalan koulutustason on havaittu useassa tutkimuksessa olevan merkittävä syrjäytymisriski ja työelämään kiinnittymistä heikentävä tekijä (Myrskylä, 2012; Ilmakunnas, Kauppinen & Kestilä, 2015). Erityisesti, jos nuori ilmoittaa että ei aio jatkaa opintojaan peruskoulun jälkeen, kasvaa nuoren syrjäytymisriski tilastojen mukaan huomattavasti. Tulevaisuuden suunnitelmien puute on myös yhdistetty huono-osaisuuteen ja siihen liittyvään näköalattomuuteen (Reiter, 2003) ja masennukseen (Lehtinen, 1994). Nuorisobarometri 2014 mukaan noin kolme neljästä nuoresta oli sitä mieltä, että syrjäytyminen johtuu ainakin jonkin verran tulevaisuudenuskon puutteesta (Myllyniemi, 2014). Kouluterveyskyselyssä 2017 kartoitetaan 8 - 9-luokkalaisten tulevaisuuden suunnitelmia kysymällä, millaisia opiskelusuunnitelmia nuorella on peruskoulun jälkeen.

2.1.3 Ulkopuolisuus, yksinäisyys ja sosiaaliset ongelmat

Karvosen ja Kestilän (2014) mukaan nuorten merkittävimmät syrjäytymisriskit liittyvät toimeentuloon, työllisyyteen ja huonoon terveydentilaan, mutta uutena ilmiönä syrjäytymisvaaraan kytkeytyy myös yksinäisyys. Vuoden 2014 nuorisobarometrin mukaan nuoret itse arvioivat ystävien puutteen keskeisimmäksi syrjäytymisen syyksi. Nuoret itse määrittelevätkin syrjäytymisen yhteisöllisenä ilmiönä, jonka tärkein syy on yksinäisyys (Aaltonen, Berg & Ikäheimo, 2015; Törrönen ja Vornanen, 2002) ja sosiaalisten verkostojen ulkopuolelle jääminen (Myllyniemi, 2009; Gretschel & Myllyniemi, 2017). Sosiaalisella ympäristöllä eritoten sosiaalisten suhteiden ja lähisuhteiden merkityksessä on tärkeä rooli nuoren syrjäytymiskehityksessä (Lehtonen ja Kallunki, 2013). Usean tutkimuksen mukaan koettu yksinäisyys ja vertaisryhmien ulkopuolelle jääminen sekä lisäävät syrjäytymisriskiä ja ovat merkittäviä jo alkaneen syrjäytymisprosessin etenemisessä (Witvliet, Brendgen, Van Lier, Koot & Vitaro, 2010; Lehtonen ja Kallunki, 2013). Riskitekijöitä, joita myös kouluterveyskyselyssä selvitettiin ovat esimerkiksi vanhemmuus tai sen puute sekä irrallisuus koulun kulttuurista ja toiminnoista ja elämänhallinnan vaikeudet. Näiden tekijöiden merkitystä ovat käsitelleet muun muassa Aaltonen, Berg & Ikäheimo (2015), heidän mukaansa elämänvalintojen kaventuminen johtaa ulkopuolisuuteen paitsi ihmissuhteista, myös esimerkiksi koulusta ja sitä kautta yhteiskunnasta.

Yksinäisyyden kytkös syrjäytymiseen ja masennukseen on selkeä (Terveystieteiden tutkimuskeskus, 2018). Sosiaalisen syrjäytymisen on osoitettu myös olevan keskeinen yhteisötason selittävä tekijä esimerkiksi nuorten itsemurhissa. Itsemurhat ovatkin Suomessa keskeinen kuolemansyy nuorilla; 20 - 29-vuotiaissa joka kolmas kuolemantapaus oli itsemurhan aiheuttama ja 35 - 44-vuotiaissa joka kuudes. Alle 25-vuotiaita oli joka kymmenes. Suomessa itsemurhien lukumäärät ovat selvästi korkeammat verrattuna muihin pohjoismaihin ja länsimaihin; EU-maihin verrattuna määrät ovat puolitoistakertaisia (Findikaattori). Huono-osaisuus ja syrjäytyminen, eritoten sosiaalisten suhteiden osalta ennustavaa myös voimakkaasti ylipäänsä nuoren ennen aikaista kuolemaa (Berg, Lundborg & Vikström, 2017, 1794).

Useissa tilastollisissa tutkimuksissa syrjäytyminen määritellään elinolosuhteisiin liittyvien indikaattorien kautta vaikka syrjäytyminen voi olla laajempaa osattomuutta ja passiivisuutta esimerkiksi yhteiskunnallisessa osallistumisessa ja *sosiaaliseen ympäristöön kiinnittymisessä*. *Sosiaaliseen yhteisöön ja yhteisöihin kuulumisen* on todettu suojaavan syrjäytymiseltä, sillä se lisää hyvinvointia ja edistää sekä fyysistä että psyykkistä terveyttä (Hyypä, 2002). Nuori elää vuorovaikutuksessa erilaisten sosiaalisten ympäristöjen, kuten perheen, suvun, koulun, kavereiden, harrastusyhteisöjen ja myös sosiaalisen median kautta muodostuvien yhteisöjen kanssa, jotka kaikki voivat osaltaan vaikuttaa nuoren syrjäytymiskehitykseen epäsuotuisasti. Nuori omaksuu näissä erilaisissa sosiaalisissa ympäristöissä sosiaalisaation kautta kulttuurisia malleja ja tapoja, jotka voivat olla myös haitallisia nuoren kehityksen näkökulmasta. Vuoden 2014 nuorisobarometrin mukaan toiseksi tärkein syrjäytymistä selittävä tekijä on vastanneiden nuorten kokemusten mukaan

joutuminen väärään seuraan eli nuoret itse korostavat sosiaalisen ympäristön ja eritoten kavereiden merkitystä syrjäytymisprosessissa (Myllyniemi, 2014).

Kun syrjäytymistä tarkastellaan osallisuuden perspektiivistä, voidaankin se nähdä sosiaaliseen yhteisöön kiinnittymisen näkökulmasta (Anttiroiko, 2003; Sassi, 2002). Kiinnittyminen voi toteutua esimerkiksi vertaisyhteisöllisyyden kautta harrastusyhteisöihin kuulumisena ja yhteiskunnallisen osallisuuden näkökulmasta ilmetä esimerkiksi osallistumisena mielenilmauksiin, protesteihin tai kuluttajaboikotteihin (Loader, Vromen & Xenos, 2014). Eritoten digitaaliset nettipohjaiset alustat mahdollistavat tämän niin sanotun ”networked individualism” -liikehdinnän (Rainie & Wellman, 2012; Quinn & Papacharissi, 2017) ja sille ominaisen huomiotalouden (van Zoonen, Vis & Mihelj, 2010; Livingstone, 2008; Lange, 2007). Digitaalisten alustojen algoritmit muokkaavat keskeisesti nykypäivän sosiaalisia todellisuuksia (Just & Latzer, 2017) kun sosiaaliseen yhteisöön kiinnittyminen tapahtuu enenevässä määrin digitaalisilla alustoilla ja niiden kautta syntyvissä todellisuuksissa. Osallisuus ja osallisuuden kokemus liittyy näin keskeisesti yksilön kokemukseen omasta merkityksellisyydestä ja on osa identiteettityötä (Hjarvard, 2013). Sosiaaliseen yhteisöön kiinnittymisen näkökulmasta syrjäytyminen on näin osattomuutta (vertais)yhteisöön kuulumisesta ja siihen liittyvää ulkopuolisuuden kokemusta. On esitetty, että tällä voi olla yhteys muun muassa radikalisoitumiskehitykseen, kun lähtökohtaisesti itsensä ulkopuoliseksi kokeva yksilö hakee tunnustusta ja kuulumisen tunnetta johonkin tarjolla olevaan yhteisöön. Osallisuus onkin laajasti ottaen kiinnittymistä hyvinvoinnin lähteille (kuten ilmeisimmät toimeentulo, asuminen, koulutus ja harrastukset) mutta myös aineettomiin asioihin, kuten luottamus, turva ja luovuus (Isola, Kaartinen, Leemann, Lääperi, Schneider, Valtari & Keto-Toko, 2017).

Syrjäytymisellä osattomuuden näkökulmasta taas viitataan yleisimmin prosessiin, johon vaikuttavia riskitekijöitä ovat vanhemmuuden puute, irrallisuus koulukuluttuurista, putoaminen koulutuksen ja/tai työmarkkinoiden ulkopuolelle ja elämönhallinnan ongelmat (Aaltonen ym., 2015). Se ilmenee elämänvalintoja kaventavana ulkopuolisuutena yhteiskunnan valtavirrasta, instituutioista ja ihmissuhteista (Törrönen ja Vornanen, 2002).

2.1.4 Elämänasenne ja elämönhallinta

Positiivinen elämänasenne on yksi tärkeä syrjäytymiseltä suojaava tekijä. Positiivinen elämänasenne on yhdistetty niin sanottu yksilön psykologiseen joustavuuteen ja yksilön kykyyn selviytyä negatiivisista ja stressaavista elämäntilanteista (Tugade, Fredrickson & Feldman Barrett, 2004; Fletcher & Sarkar, 2013; Ong, 2006). Vaikka lapsuudessa ja nuoruudessa olisi syrjäytymisriskiä lisääviä tapahtumia ja tekijöitä, optimistinen elämänasenne voi auttaa suojaamaan niiden haitallisilta vaikutuksilta (Kangas, 2010). Myönteisen elämänasenteen ja positiivisen ajattelun on havaittu vaikuttavan ihmisen kokonaisvaltaiseen hyvinvointiin, sosiaalisiin suhteisiin ja oppimiseen (Ojanen, 2007). Vastaavasti negatiivisella elämänasenteella on havaittu yhteys myös yksilön terveydentilaan ja syrjäytymisriskiä lisäävien pitkäaikaissairauksien kehittymiseen (Neuvonen, Rusanen, Solomon, Ngandu, Laatikainen, Soininen & Tolppanen, 2014, 2205; Tindle, Chang, Kuller,

Manson, Robinson, Rosal & Matthews, 2009). Vanhempien elämänsenteella on myös vaikutusta nuoren syrjäytymiskehitykseen. Erityisesti äidin passiivinen elämänsenne, joka korostaa yhteiskunnan vastuuta perheen elinoloista vaikuttaa nuoren työttömyysriskiin ja sitä kautta syrjäytymisriskiin (Ek ym., 2004).

Elämänsenne ja elämänhallinta ovat puolestaan toisiinsa kytköksissä. Elämänhallinnan tunne on tärkeää nuoren psyykkisen ja henkisen hyvinvoinnin kannalta ja niin ikään syrjäytymiseltä suojaava tekijä. Nuoren tunne-elämän tasapaino, elämään tyytyväisyys, osallisuuden kokemus ja vaikutusmahdollisuudet omaan elämään ovat yhteydessä hyvään elämänhallintaan, ja jos näissä tekijöissä ilmenee puutteita, syrjäytymisriski kasvaa (Kinnunen, 2003; Lämsä, 2010; Nivala, 2013). Elämänhallinta rinnastetaan usein nuoren toiminnalliseen kapasiteettiin, joka sisältää hänen fyysiset, psyykkiset ja sosiaaliset voimavaransa (Rauhala, 1998, 3). Syrjäytymisprosessissa keskeisiä vaikuttavia riskitekijöitä on elämänhallinnan ongelmat (Aaltonen ym., 2015) ja Rauhalan (1998) mukaan elämänhallinnan ongelmat johtavat nuorella syrjäytymisprosessien alkamiseen. Syrjäytyminen ja elämänhallinta voidaan ymmärtää toistensa vastakohtina, joten kaikki syrjäytymisen riskitekijät olisivat samalla myös elämänhallinnan vaikeuksia (Lämsä, 2010).

2.1.5 Perhe, asuminen ja ylisukupolvisuus

Syrjäytymiskehityksen yhtenä merkittävimpänä riskitekijänä pidetään ylisukupolvisuutta (Coles, 2000; MacDonald, 2006). Ylisukupolvisuudella tarkoitetaan huono-osaisuuden periytymistä joko yhdellä tai useammalla elämän alueella. Useat tutkimukset korostavatkin perheen vaikutusta nuoren syrjäytymisprosessissa (Myrskylä, 2012; Paananen, Ristikari, Merikukka, Rämö & Gissler, 2012; Hilli, Ståhl, Merikukka & Ristikari, 2017), sillä nuori oppii perheen kautta erilaiset normit, arvot ja käyttäytymismallit, jotka tukevat selviytymistä kodin ulkopuolella muissa sosiaalisissa ympäristöissä. Terveiden ja hyvinvoinnin laitoksen Kansallinen syntymäkohortti 1987 – tutkimuksessa on seurattu eri rekisteritietojen pohjalta kaikkia noin 60 000 Suomessa vuonna 1987 syntyneitä henkilöä sikiöajalta kahteenkymmeneenviiteen ikävuoteen saakka. Tutkimus (Hilli ym., 2017; Paananen ym., 2012) nostaa merkittäviksi syrjäytymisriskeiksi neljä perheeseen liittyvää tekijää: vanhempien matala koulutustaso, vanhempien päihde- ja mielenterveysongelmat, perheen köyhyys ja perheen rakenteessa tapahtuneet suuret muutokset, kuten vanhemman kuolema tai avioero (Hilli ym., 2017; Paananen ym., 2012). Huono-osaisuuden on havaittu periytyvän sukupolvelta toiselle ja vanhempien alhainen sosioekonominen asema lisää syrjäytymisriskiä (Myrskylä, 2012; Paananen ym., 2012). Erityisesti äidin matala koulutustaso ja ikä lisäävät syrjäytymisriskiä (Witting & Keski-Petäjä, 2016; Haapasalo, 2006; Savioja, 2007; Alatupa, Karppinen, Keltikangas-Järvinen & Savioja, 2007). Perheen etninen tausta myös vaikuttaa syrjäytymisriskiin; maahanmuuttajataustaisista nuorista syrjäytyy viisi kertaa enemmän kantaväestöön verrattuna (Myrskylä, 2012). Rintasen (2000) mukaan myös vanhempien sisarusten lukumäärää vaikuttaa syrjäytymisriskiin. Vuonna 1987 Suomessa syntyneiden kohorttitutkimuksen perusteella kodin ulkopuolelle sijoittaminen lisää syrjäytymisriskiä, sillä sijoitetuille lapsille ja nuorille on kasautunut useita perhe- ja elinoloihin liittyviä epäsuotuisia tekijöitä (Kestilä, Paananen, Väisänen, Muuri,

Merikukka, Heino & Gissler, 2012). Aaltosen mukaan myös vanhemmuuden puute on syrjäytymisprosessin riskitekijä (Aaltonen ym., 2015).

2.1.6 Tuen ja avun puute

Yhteiskunnan ja perheen tuella on vaikutusta nuoren syrjäytymiskehitykseen. Vuoden 2009 nuorisobarometrin mukaan yli puolet nuorista oli sitä mieltä, että avun saamisen vaikeus selittäisi syrjäytymistä paljon tai jonkin verran (Myllyniemi, 2009). Vanhempien emotionaalinen tuki ja ohjaus ovat nuorille erittäin tärkeää. (Paunonen, Vehniäinen & Julkunen, 1999). Varhaislapsuudessa lapsi oppii myös sosioemotionaalisia taitoja, joiden puute korreloi myöhemmin lisääntyneen riskiin päihteidenkäytön, mielenterveysongelmien, matalan koulutustason ja työttömyyden osalta (Jones, 2015).

Koululla on tärkeä syrjäytymistä ehkäisevä rooli yhteiskunnassa. Koulun tehtävä on tunnistaa lapsen syrjäytymisriskiä lisääviä tekijöitä mahdollisimman varhaisessa vaiheessa ja tarjota riittävä apua ja tukea, jotta syrjäytymiskehitys pysähtyisi (Alatupa ym., 2007). Opettajat voivat vaikuttaa syrjäytymiseen muun muassa puuttamalla kiusaamiseen, lisäämällä yhteisöllisyyttä luokassa ja opettamalla erilaisuuden hyväksymistä, tunneälyä sekä ihmissuhdetaitoja (Salmivalli, 2010; Veijola, 2003). Syrjäytymisen ehkäisyn näkökulmasta ryhmäkokojen pitäminen riittävän pieninä helpottaa opettajaa tunnistamaan syrjäytymiseen viittaavia signaaleja oppilaista ja tarjoamaan apua. Kouluterveydenhoitajilla, koulupsykologeilla ja koulukuraattoreilla on myös tärkeä tehtävä syrjäytymistä ennaltaehkäisevässä työssä. Henkilöstöresurssit ovat kuitenkin yhä liian vähäisiä etenkin kuraattorien ja psykologien osalta, ja palvelut jakautuvat maantieteellisesti epätasaisesti (Wiss, Ståhl, Saaristo, Kivimäki, Frantsi-Lankia & Rimpelä, 2017). Puutteellisten tukitoimien vuoksi koulu voi omilla toimintatavoillaan myös lisätä syrjäytymistä ja pahimmillaan leimata nuoria omilla mekanismeillaan (Kuula, 2000).

2.2 Tutkimuskysymys

Kuten edellä esitellystä tutkimuskirjallisuudesta ilmenee, syrjäytyminen on monitahoinen ongelma. Syrjäytymiskehitys on pitkä ajallinen prosessi, jossa erilaiset osattomuuden ja passiivisuuden osa-alueet johtavat kasaantuneeseen huono-osaisuuteen (Berg ym., 2017). Syrjäytymiskehityksen ennaltaehkäisy, tai ainakin ennakoiva tunnistaminen, voi siis olla mahdollista vain näiden monialaisten osattomuuden muotojen tunnistamisen avulla. Tässä tutkimuksessa esitetään yksi mahdollinen malli, jonka avulla monialaista ja kasautuvaa osattomuutta on mahdollista tunnistaa ennakoivasti. Tutkimuskysymyksiämme on:

- 1) kuinka tekoälypohjaiset teknologiat voivat toimia ennakoivasti syrjäytymisen tunnistamisessa ja**
- 2) tarkastella tekoälyteknologioiden kyvykkyyttä olemassa olevan data-aineiston avulla, tässä tapauksessa THL:n kouluterveyskyselyn avulla.**

3 MENETELMÄT JA TYÖPROSESSI

3.1 Menetelmän valinta

Tutkimuksessa ei ollut mahdollista käyttää ohjatun koneoppimisen työkaluja, sillä käytössä ollut aineisto ei ollut siihen soveltuvaa. Ohjattuun koneoppimisen käyttämiseksi tarvittaisiin esimerkiksi vanhoista kouluterveyskyselyistä syrjäytyneiden nuorten vastauksista koostuva opetusdata, jolla ohjelma opetettaisiin tunnistamaan syrjäytymisriskissä olevien nuorten vastukset. Koska kouluterveyskyselyt ovat nimettömiä, tällaista opetusdataa ei löytynyt, joten ohjattu koneoppiminen suljettiin ulos. Tutkimuksen menetelmäksi valittiin ohjaamaton tutkijatriangulaation ja SPSS -Modeler-työkalulla tapahtuva data-analyysi syrjäytymisriskin tunnistamisessa.

3.2 Tutkijatriangulaatio, aineiston luokittelu ja pisteytys

Tutkimusta varten kartoitettiin aiemmin kuvattu laaja joukko nuorten syrjäytymisen riskitekijöihin liittyviä tutkimuksia, joiden perusteella luotiin luokituspohja nuoren syrjäytymisriskiin ja huono-osaisuuteen vaikuttavista tekijöistä. Taustamateriaalin tuella THL:n kyselyaineistoa tarkasteltiin suhteessa *tutkimuskirjallisuudesta esiin nouseviin nuorten syrjäytymisen tunnistettuihin riskitekijöihin*. THL:n kysymyspatteriston soveltuvuutta tarkasteltiin tutkijatriangulaation avulla kolmessa eri vaiheessa. Tarkoituksena oli luoda aineistosta kysymyskohtainen painoarvoluokitus SPSS Modeler analyysia varten.

Ensimmäisessä vaiheessa hankkeen moniammatillinen tutkijatiimi kävi läpi kyselyn. Tutkijat arvioivat itsenäisesti toisistaan riippumatta kysymysten soveltuvuuden syrjäytymisen ennustamisessa yksitellen, aiheen tutkimuskirjallisuuteen ja omaan substanssiosaamiseensa perustuen. Kysymykset, joilla ei katsottu olevan merkitystä tai suoraa osoitettua yhteyttä tutkimuskirjallisuudessa esitettyyn syrjäytymisriskin arviointiin, jätettiin arvioinnin ulkopuolelle (esimerkiksi taustatiedot, seurustelusuhte, työskentely koulunkäynnin ohella). Työskentelyssä arvioitiin kysymyskohtaisesti sen merkittävyys painoarvolla 0 - 5 syrjäytymisriskiin nähden. Samalla määriteltiin mihin aiemmassa tutkimuksessa tunnistettuihin syrjäytymisen riskitekijöiden luokkaan kukin yksittäinen kysymys voidaan liittää. Tätä kautta luotiin **syrjäytymisriskiluokitus**, jonka avulla voitaisiin tarkastella erilaisten syrjäytymiseen liittyvien riskitekijöiden kasautuvuutta ja keskinäisiä suhteita. Syrjäytymisriskin mittariston pohjana käytettiin niin sanottu painoarvokertoimia, minkä perusteella jokainen tutkija loi oman pisteytyksensä. Tämän jälkeen kysely käytiin tutkijatiimin kanssa yhdessä läpi keskustellen. Jos tutkijoiden pisteytyksissä oli eroja, päädyttiin sen pistemäärän kannalle, johon enemmistö oli päätenyt tai jos kaikilla oli eriävä arvio, laskettiin keskiarvo. Jokaiselle kysymykselle saatiin lopulta painoarvo. Lisäksi keskusteltiin kysymyksen luokituksesta ja kysymyksen luokka päätettiin samaan tapaan kuin pisteytys, jos luokitus erosi tutkijatiimin sisällä, päätös tehtiin joko enemmistön päätöksellä tai keskusteltiin luokituksesta niin että löydettiin yhteisymmärrys. Kuvassa 1 on ote kyselystä, johon on merkitty tutkijatriangulaation avulla saadut

pistemäärät sekä luokka 3.3 Kokemukset koulusta, kouluviihtyvyys ja koulussa jaksaminen(LKA:3.3) johon kysymys on sijoitettu.



8. Mitä mieltä olet opettajistasi?

	Täysin samaa mieltä	Samaa mieltä	Eri mieltä	Täysin eri mieltä
Opettajat rohkaisevat minua ilmaisemaan oman mielipiteeni oppitunneilla	<input type="checkbox"/>	<input type="checkbox"/>	0,5	1
Opettajat ovat kiinnostuneita siitä, mitä minulle kuuluu	<input type="checkbox"/>	<input type="checkbox"/>	1,5	3
Opettajani odottavat minulta liikaa	2	1	<input type="checkbox"/>	<input type="checkbox"/>
Opettajat kohtelevat meitä oppilaita oikeudenmukaisesti	<input type="checkbox"/>	<input type="checkbox"/>	1	2

9. Miten helppoa koulussasi on olla oma itsensä?

- erittäin helppoa
 melko helppoa
 ei helppoa eikä vaikeaa
 melko vaikeaa
 erittäin vaikeaa

Mean: (a-d)

Max: $8/4 = 2$

LKA: 3.3

KUVIO 1. Ote kyselystä johon merkitty tutkijatriangulaation avulla saadut pistemäärät sekä luokitus

Toisessa vaiheessa kysymysten painoarvot ja luokittelu validoitiin ulkopuolisen asiantuntija-arvioinnin avulla. Asiantuntijoita rekrytoitiin eri ammatti- ja tutkimusaloilta siten, että saatiin kattavasti mukaan alat, joilla toimitaan lasten- ja nuorten syrjäytymistekijöihin liittyen. Asiantuntija/ammattialat olivat; psykologia, sosiaalityö, erityisopetus ja koulupsykologia. Asiantuntijat ohjeistettiin pisteyttämään kysymykset samaan tapaan kuin tutkijat olivat tehneet. Lisäksi tutkijoiden aiemmin toteuttama luokittelu annettiin heille tiedoksi ja pyydettiin tarkastelemaan sitä kriittisesti oman käytännön asiantuntemuksen ja ammattikirjallisuuden valossa. Asiantuntijat ohjeistettiin luokittelemaan jokainen kysymys - olemassa olevaan luokitukseen pohjaten ja tarvittaessa määrittämään kysymykselle uusi luokka. Asiantuntijoista neljä (4) sosiaalityön tutkija, sosiaalityöntekijä, kouluterveydenhoitaja ja koulupsykologi toteuttivat pisteytyksen ja luokittelun itsenäisesti, psykologian tutkijat muodostivat kolmen hengen tiimin, joka teki luokittelun yhdessä.

Kolmannessa vaiheessa tutkijat kokoontuivat uudelleen tarkastelemaan pisteytystä ja luokittelua. Heidän aiemmin toteuttamaansa pisteytystä ja luokittelua verrattiin asiantuntijoiden tekemään pisteytykseen ja luokitteluun. Asiantuntijoiden pisteytyksiä verrattiin samaan tapaan kuin aiemmin tutkijoiden erillisiä pisteytyksiä, jos eriäviä arvioita ilmeni, päädyttiin enemmistön päätökseen tai keskiarvoon. Lisäksi luokittelua tarkasteltiin sisällöllisesti asiantuntijoiden ehdotusten ja kommenttien perusteella ja siihen tehtiin muutoksia, jos useampi kuin yksi asiantuntija oli ehdottanut muutosta ja ehdotus oli samansuuntainen. Muutoksia tehtiin yhteensä 24, joista luokkien sisäisiä eli pienempiä muutoksia tuli 12 ja luokkien välisiä suurempia muutoksia 12.



KUVIO 2. Luokitustyöskentelyn kuvaus

Neljännessä vaiheessa kunkin luokan keskimääräinen painokerroin otettiin tavoitteeksi luokkien välisissä painokertoimien suhteissa. Tämän painotuksen jälkeen kustakin luokasta saadaan maksimissaan pisteitä samoissa suhteissa kuin ovat keskimääräiset painotukset tehtävissä. Näin luokat pysyvät keskenään oikeissa suhteissa, vaikka koululaiskyselyssä kysymysten määrä luokittain vaihteleeekin. Ilman tämän vaiheen painotusta jokin luokka voisi saada eniten pisteitä, vaikka tehtävissä olisikin pienimmät painokertoimet. Vastaavasti tärkeäksi koettu luokka ei vaikuttaisi kokonaispisteiden kertymään, jos koko luokassa on vähän kysymyksiä.

Luokittain saavutettavat maksimipisteet vaiheen kolme jälkeen ovat:

Luokka 1:	51,15	pistettä
Luokka 2:	223,575	pistettä
Luokka 3:	40,264	pistettä
Luokka 4:	82,95	pistettä
Luokka 5:	11,167	pistettä
Luokka 6:	21,0	pistettä

Kun nämä pisteet jaettiin luokissa olevien tehtävien lukumäärillä, saatiin luokkien keskimääräiset painokertoimet.

Luokka 1:	14 kpl =	3,654 pistettä/tehtävä
Luokka 2:	36 kpl =	6,210 pistettä/tehtävä
Luokka 3:	13 kpl =	3,097 pistettä/tehtävä
Luokka 4:	22 kpl =	3,770 pistettä/tehtävä
Luokka 5:	3 kpl =	3,722 pistettä/tehtävä
Luokka 6:	6 kpl =	3,500 pistettä/tehtävä

Luokkien maksimipisteet tulevat jakautumaan näin suhteessa ylläoleviin painokertoimiin. Oikea jakauma saadaan jakamalla luokassa saatavat pisteet kyseisen luokan tehtävien lukumäärällä. Eli luokan 1 tehtävästä myös oppilaan saama pistemäärä jaetaan 14:llä. Luokan 2 tehtävissä pisteet jaetaan 36:lla jne. Pisteiden jako tehtiin kuhunkin tehtävään, jolloin maksimipisteet vähenivät, mutta luokkien suhde saatiin vastaamaan luokkien keskimääräistä painokerrointa.

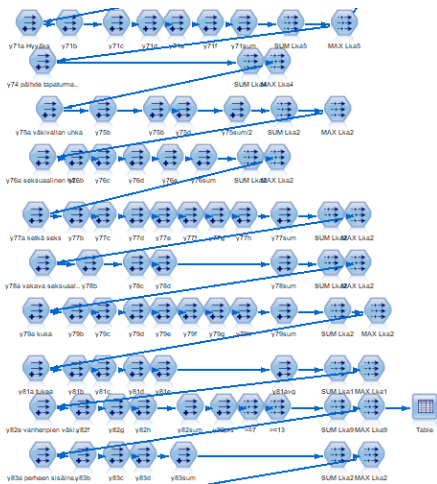
Edellä esitellyn neljän vaiheen tuloksena luotiin tehtäväkohtainen pisteytys ja luokitustyöskentelyn pohjalta koko kyselyaineisto käytiin läpi SPSS Modelerin avulla. Tämän avulla muodostettiin syrjäytymisriskiarvio, joka esitetään seuraavaksi.

3.3 Data-analyysi SPSS Modelerilla

SPSS -pohjaisen aineiston analyysin avulla laskettiin kasautuvaa osattomuutta ja ennustettiin sen aiheuttamaa potentiaalista syrjäytymisriskiä. Data-analyysissä päädyttiin käyttämään SPSS Modeleria, koska esimerkiksi ei-ohjattu koneoppiminen olisi painottanut kaikkia kysymyksiä samantarvoisesti. Tutkimuksessa käytettiin siksi menetelmänä *SPSS Modelerin avulla luotua pisteytysmallia*, joka perustui tutkija-asiantuntija-arvion pisteytykseen kysymysten vaikutuksesta syrjäytymisriskiin. Palolan ym. (2012) mukaan noin 5 prosenttia ikäluokan nuorista syrjäytyy ja Eurostatin mukaan 20–24-vuotiaista suomalaisnuorista 15,7 prosenttia oli vuonna 2015 koulutuksen ja työelämän ulkopuolella (Alatalo, Mähönen & Räisänen, 2017). Koska syrjäytymisen todellinen tulevaisuuden toteuma tässä aineistossa ei ollut tiedossa, eikä nyt vastanneista voi tehdä johtopäätöksiä tulevasta toteutuneesta syrjäytymisestä, päädyimme tutkimusten (Palola ym., 2012; Alatalo ym., 2017) perusteella arvioimaan, että suurimmassa syrjäytymisvaarassa voisi olla noin 10 prosenttia oppilaista. Tässä tutkimuksessa SPSS Modeler erotteli oppilaiden pisteiden perusteella oppilaista osajoukon, johon kuului 10 prosenttia oppilaista. He saivat koko joukosta korkeimmat kokonaispisteet ja heidän katsottiin olevan vastausten perusteella suurimmassa syrjäytymisriskissä. SPSS Modeleria käytettiin tutkimuksessa kolmella eri tavalla, jotka on kuvattu tässä luvussa ja havainnollistettu lisäksi kuvina.

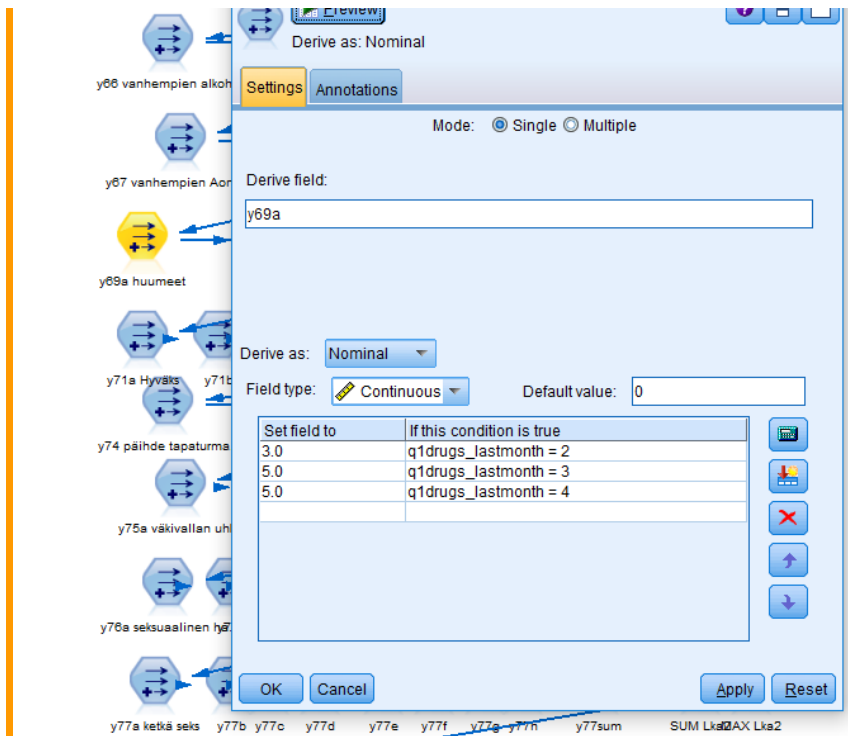
SPSS Modelerilla tietoa voidaan muokata ja yhdistellä tehtäväkohtaisesti niin kutsutuissa solmukohdissa. Näissä solmuissa on valmiita valikkoja tehtävään liittyen, tiedon esikäsittelystä aina ryhmittelyyn ja ennusteisiin asti. Tieto kulkee virtana solmukohdista toisiin käyttäjän määrittelemien ehdoin. Tekoälyyn liittyviä solmuja ovat nimen mukaiset modeler-solmut. Kuvio 3 on näytönkaappaus SPSS Modelerin näytöstä, joka kuvaa, miten kouluterveyskyselyn raakadata virtaa solmukohtien läpi. Jokainen solmukohta kerryttää oppilaan syrjäytymisriskipisteitä, mikäli oppilas on vastannut syrjäytymisriskin kannalta merkitykselliseen kohtaan. Oppilaan vastauksen perusteella hänen riskipisteensä muuttuvat tehtävä tehtävältä ennalta määrätyn tehtäväkohtaisen pisteytyksen mukaan. Jokaisen oppilaan vastaukset käyvät kerran läpi streamin, minkä jälkeen hänelle on kertynyt riskipisteet luokittain sovituilla pisteytyksillä. Pisteytyksen jälkeen streami ohjataan haluttuihin esitysgrafiikka-solmuihin, jossa kertynyt tieto esitetään solmussa määrättyllä tavalla. Grafiikan lisäksi streami voidaan ohjata esim. päätöspuu-solmulle, joka määrittelee säännöt, joilla

vastauksista saadaan päättelymalli kyseisen lopputuloksen saamiseksi. SPSS modelerissa on useita valmiita solmuja eri tehtäviä varten.



KUVIO 3. Näytönkaappaus Modeler -ohjelmasta, jossa näkyy kouluterveyskyselyn tietoja käsittelevät solmut ja streami

Ensimmäisenä SPSS Modelerin avulla määriteltiin kullekin kouluterveyskyselyn kysymykselle yksittäisessä solmukohdassa vastauksen pistearvo (0 - 5) syrjäytymisriskin näkökulmasta sekä syrjäytymisen riskiluokka (kuvio 4 alla). Esimerkiksi kysymys 69 käsittelee oppilaan marihuanan ja kannabiksen käyttöä (luokka 4.3) viimeisen kuluneen 30 päivän aikana. Mikäli oppilas on käyttänyt aineita kerran, hänelle vastauksesta tulee pisteitä 3, useammista käyttökertoista (2 - 4 ja 5 kertaa tai useammin) kertyy 5 pistettä.



KUVIO 4. Vastauksen pistearvon määrittely SPSS Modeler –ohjelman nodessa

Kun datavirta etenee solmukohtien eli yksittäisten kysymysten läpi SPSS Modeler laskee oppilaiden pistekertymät luokittain, kunkin oppilaan luokkakohtaisen kokonaispistemäärän sekä suhteellisen pistemäärän suhteessa luokan maksimipisteisiin. Kun datavirta on kulkenut kaikkien pisteitä kerryttävien solmukohtien läpi, tuloksia voidaan analysoida, visualisoida ja luokitella liittämällä siihen uusia esimerkiksi tekoäly- tai muihin tilastollisiin menetelmiin liittyviä modeler-solmuja. Kuvio 3 on näytönkaappaus SPSS Modelerin näytöstä, josta näkyy, miten datan visualisointi ja luokittelu etenee solmukohdissa.

Ensimmäinen tapa arvioida Keski-Suomen kuntien tilannetta nuorten syrjäytymisriskin näkökulmasta oli laskea kyselyyn vastanneiden oppilaiden keskimääräiset syrjäytymispisteet. Kokonaispisteiden jakaumaa voidaan vertailla myös eri riskiluokissa oppilaittain, kouluittain tai kunnittain.

Toinen tapa, jolla SPSS modeleria hyödynnettiin, oli arvioida syrjäytymisriskiä kokonaispisteiden lisäksi tutkimalla, paljonko kunta tai koulu poikkeaa normaalista 10 % määrästä syrjäytymisriskissä olijoista.

Kolmantena arviointitapana käytettiin niin sanottua päätöspuu -mallia koko Keski-Suomen aineistoon hakemalla merkittäviä kysymyssarjoja. Päätöspuu-mallilla haettiin syrjäytymisriskin näkökulmasta merkittäviä kysymyksiä, joilla voitaisiin päästä samaan lopputulokseen kuin koko kyselyn kokonaismäärältään 530 vastauksella.

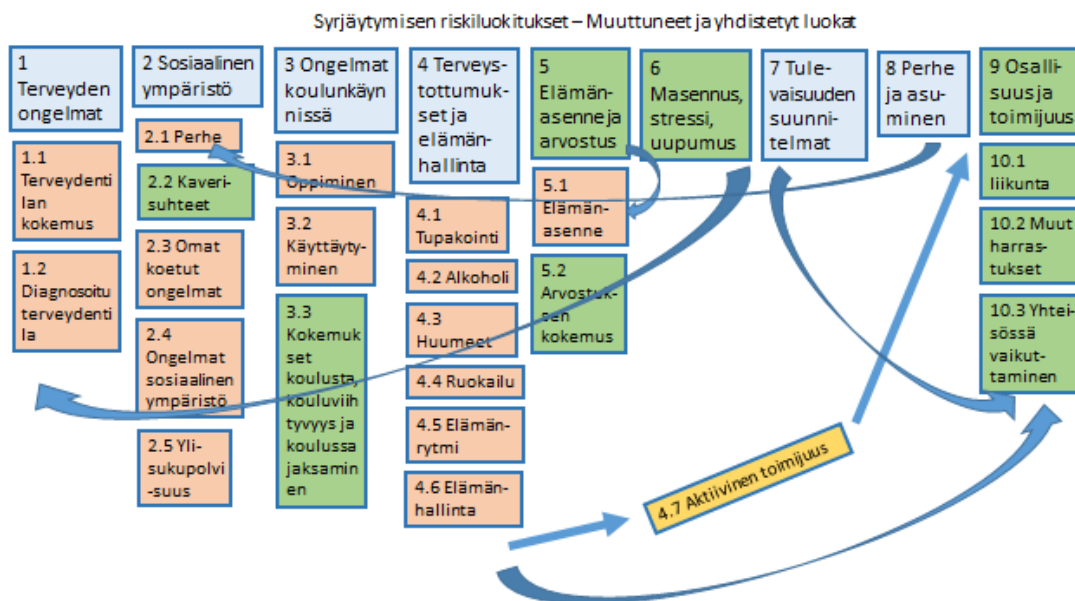
4 TULOKSET

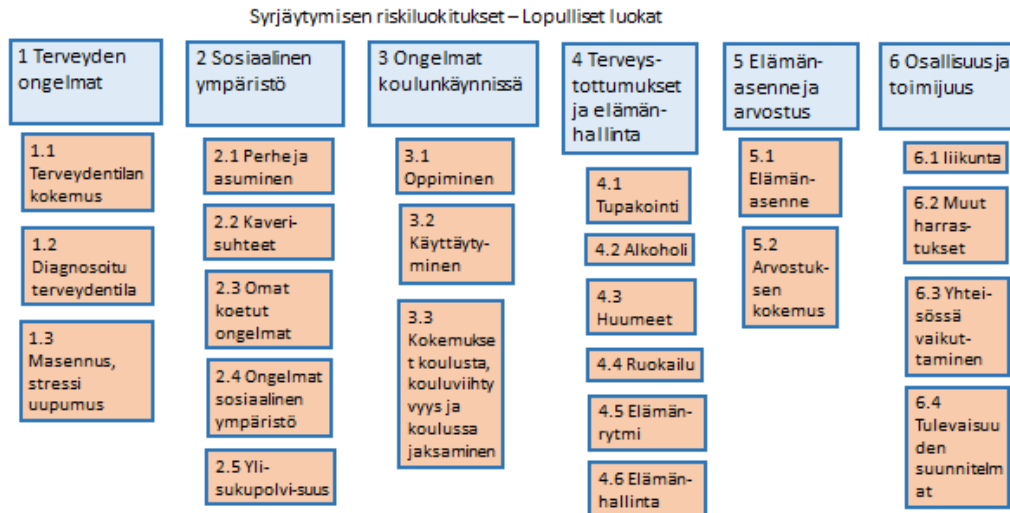
Tässä luvussa kuvataan tutkija-asiantuntija arvioon pohjaavan SPSS Modeler analyysin tulokset.

4.1 Syrjäytymisriskin luokitus ja pisteytys

Tutkijoiden ja asiantuntijoiden työskentelyn tuloksena luotiin ensimmäisen vaiheen jälkeen 28 kohtainen lista kyselystä tunnistetuista muuttujista syrjäytymisriskin laskemiseksi sekä niiden keskinäiset painoarvot.

Kyselyaineisto luokiteltiin tutkimuskirjallisuudesta esiin nousseiden teemojen avulla pääluokkiin. Pääluokkia oli alunperin yhteensä yhdeksän ja asiantuntijoiden arvion jälkeen päädyttiin kuuteen luokkaan. Pääosin asiantuntijat olivat samaa mieltä tutkijoiden kanssa luokituksista, mutta osaan luokista tuli pieniä muutosehdotuksia ja myös uusia luokkia luotiin. Lopullinen luokitus muutoksineen on esitetty seuraavassa kuviossa.





KUVIO 5. Syrjäytymisen riskiluokituksen muodostuminen tutkija- ja asiantuntija-arvioiden kautta

Muuttuneet ja kokonaan uudet luokat on merkitty vihreällä. Lisäksi keltaisella on merkitty luokka 4.7 joka on vaihtanut paikkaa alaluokasta yläluokaksi ja samalla nimi on hieman muuttunut. Luokka 2.6 puolestaan on muuttunut yläluokasta alaluokaksi. Kyselystä oli valittu 94 tehtävää, joissa oli yhteensä 266 kysymystä, joiden vastausta analysoitiin. Sisällöllisesti tarkasteltuna luokat ovat hyvin eri suuria, tarkasteltiinpa niitä kysymysten määrällä tai kysymysten alakohtien määrällä. Ylivoimaisesti suurin määrä kysymyksiä liittyi luokkaan 2. Sosiaalinen ympäristö (111 kysymystä/36 tehtävää), kun taas luokissa 5 ja 6 oli yhteensä 27 kysymystä / 9 tehtävää. Yhteensä 275 kysymyksestä yli kolmasosa käsitteli sosiaalista ympäristöä ja reilu neljäsosa käsitteli terveysasioita ja noin neljäsosa käsitteli koulunkäyntiä. Lisäksi noin kolmekymmentä kysymystä käsitteli terveystottumuksia ja elämäntapaa, johon sisältyivät esimerkiksi ruoka, lepo ja päihteiden käyttö.

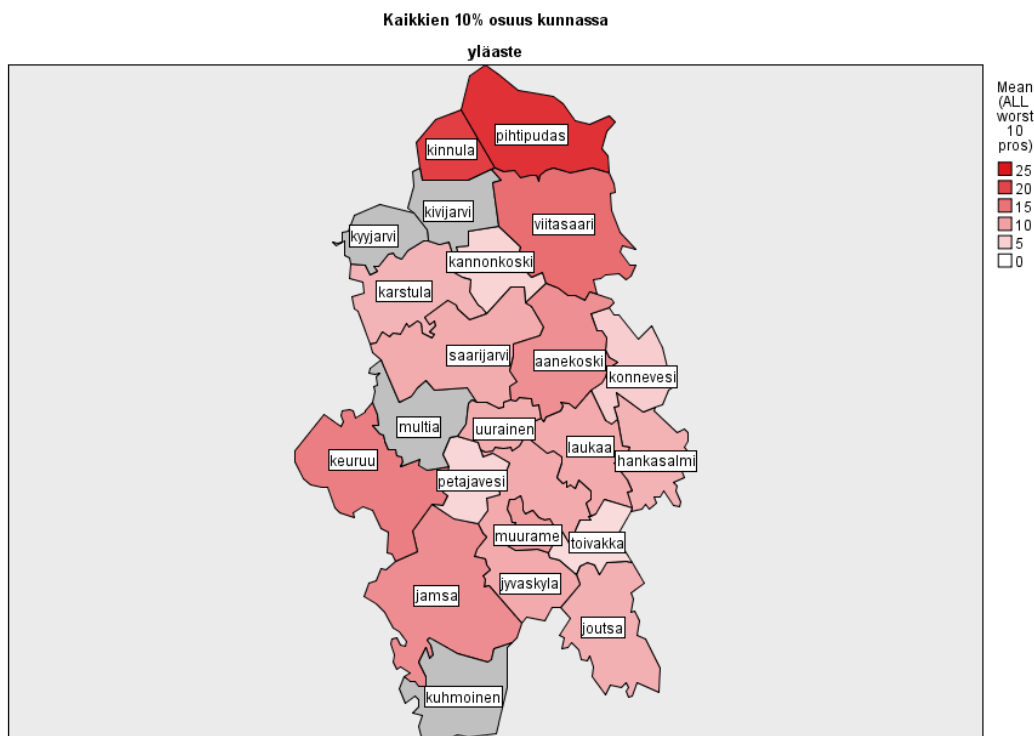
Kysymysten painokertoimet erosivat jonkin verran tutkijoiden alkuperäisen luokittelun ja asiantuntijoiden tekemän luokittelun kesken. Suurin osa kysymyksistä painotettiin joko täysin samalla tavalla tai lähes samalla tavalla. Suurin ero kysymysten merkittävässä painotuseroissa (ero yli 2 painokerroin yksikköä) oli, että asiantuntijat arvioivat syrjäytymisen riskit hieman tutkijoiden arviota korkeammiksi, kun vain satunnaisesti tutkijat olivat painottaneet yli kaksi yksikköä korkeammalle syrjäytymisen riskiä. Asiantuntijat painottivat tutkijoita enemmän syrjäytymisen riskitekijöinä sellaisia asioita, kuin koulukiusaaminen ja kouluun suhtautuminen, päihteiden käyttö ja niihin suhtautuminen sekä häirintään ja väkivaltaan liittyvät kysymykset. Jonkin verran tutkijoita enemmän asiantuntijat painottivat lisäksi yksinäisyyteen, yhteisöllisyyteen, mielialaan sekä saadun tuen ja avun saamiseen liittyviä kysymyksiä. Lisäksi tutkijoita enemmän asiantuntijoiden mielestä syrjäytymisriskin näkökulmasta tärkeitä olivat sosiaalisen median, internetin sekä pelaamiseen käytetty aika sekä etninen tausta.

4.2 SPSS Modeler syrjäytymisennusteen työkaluna

Pisteytyksen perusteella laskettiin syrjäytymisriskiarviot Keski-Suomen kouluista. Alueellisten erojen havainnollistamiseksi nämä riskit on kuvattu värikartoilla. Sekä tutkijoiden pisteytyksen että asiantuntijoiden pisteytysten avulla luotiin erilliset kartat. Kartat noudattivat pääosin toisiaan, joten voidaan arvoida, että ulkopuolisilla asiantuntijoilla ja hankkeen tutkijoilla oli suhteellisen yhtenäinen näkemys nuoren syrjäytymisriskiin liittyvistä tekijöistä. Seuraavaksi laskimme kysymyskohtaisesti SPSS Modelerilla tutkijoiden ja asiantuntijoiden pisteytysten keskiarvot ja toteutimme SPSS-modelerilla yhden kartan, joka yhdistää tutkijoiden ja asiantuntijoiden arviot syrjäytymisriskistä.

4.2.1 Suurimmat syrjäytymisriskipisteet

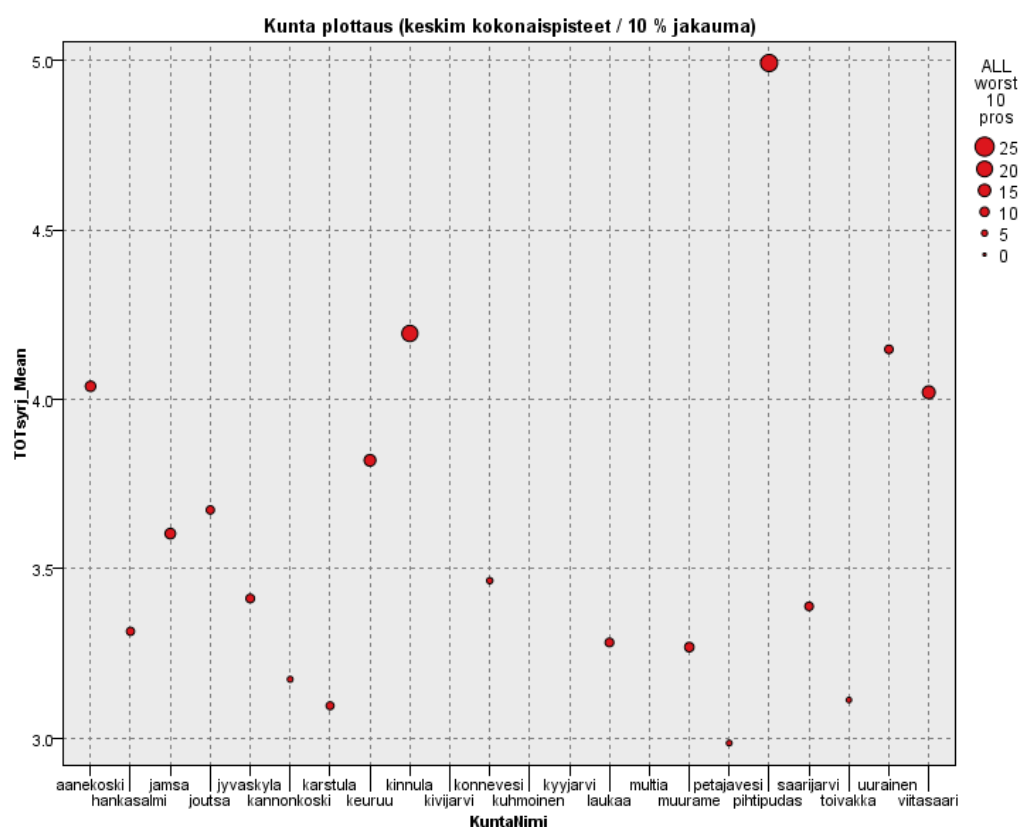
Kartalla kuvataan suurimmassa syrjäytymisriskissä olevien oppilaiden jakautumista kunnittain (kuvio 6). Kuviossa 6 värikartta esittää suurimpaan riskiluokkaan kuuluvien oppilaiden prosenttiosuuden kaikista vastanneista oppilaista (korkeimmat pisteet kerännyt 10 %). Esimerkiksi Muuramessa kyselyyn vastasi 240 oppilasta ja syrjäytymisriskiin kuului 23 oppilasta, eli 9,6 prosenttia paikkakunnan 8 - 9-luokkalaisista on suurimmassa syrjäytymisriskissä kouluterveyskyselyn vastausten perusteella. Neljällä paikkakunnalla kyselyyn vastanneiden oppilaiden lukumäärä oli alle 20, joten ne jätettiin vastaajajoukon pienuuden vuoksi kokonaan tutkimuksen ulkopuolelle ja merkittiin värikarttoihin harmaalla.



KUVIO 6. Kartta tutkijoiden ja asiantuntijoiden pisteytyksen keskiarvon perusteella, joka kuvaa riskissä olevien oppilaiden jakautumisen Keski-Suomen kuntiin (montako prosenttia kunnissa on riskissä olevia oppilaita)

4.2.2 Suurimman syrjäytymisriskin oppilaiden sijoittuminen Keski-Suomen alueelle

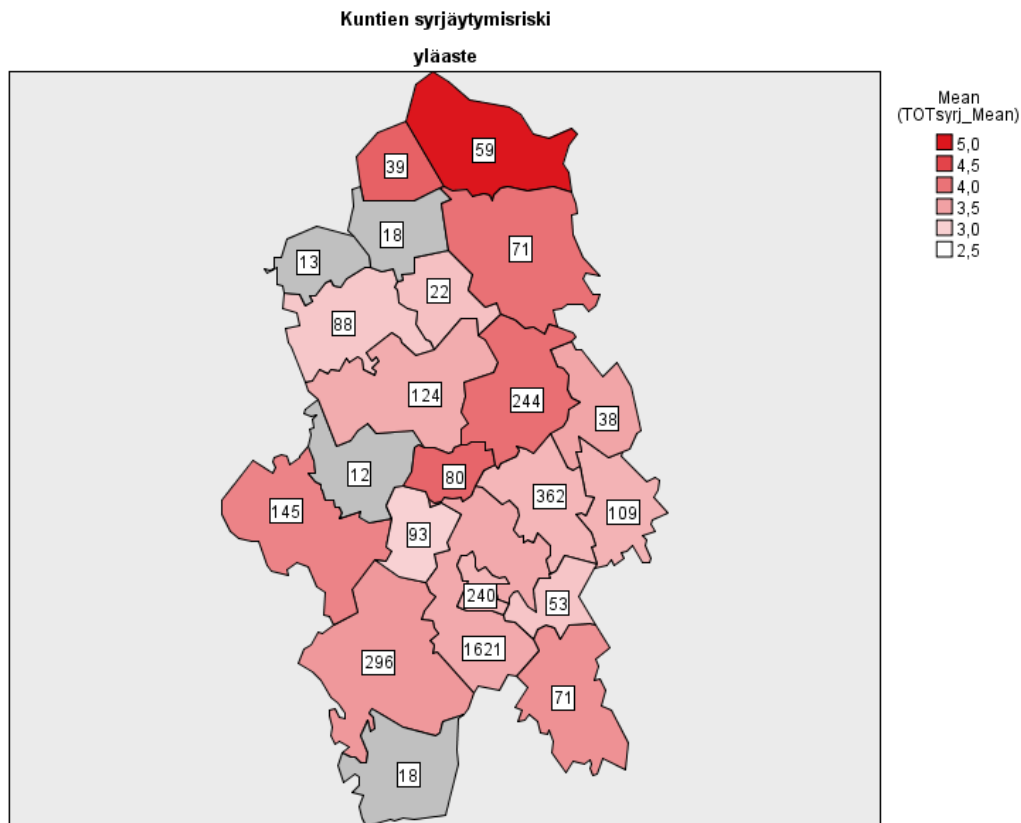
Tämän jälkeen arvioitiin syrjäytymisriskiä kokonaispisteiden lisäksi tutkimalla, paljonko kunta tai koulu poikkesi 10% määrästä syrjäytymisriskissä olijojen osalta kokonaispisteistä ja syrjäytymisriskissä olevien nuorten määrästä, jossa y-akselilla (kuviossa 7) näkyvät kunnan keskimääräiset kokonaispisteet ja plotin koko kuvastaa paljonko oppilaista kuuluu syrjäytymisriskissä oleviin. Esimerkiksi tutkijoiden pisteytyksen mukaan Muuramessa oppilaiden syrjäytymisriskipisteet olivat 13,7 % maksimista ja alle 10 prosenttia oppilaista kuuluu riskiryhmään (Kuviossa 7). Vastaavat plottaukset tehtiin myös asiantuntijoiden pisteille (Kuviossa 7) sekä tutkijoiden ja asiantuntijoiden pisteytysten keskiarvolle (Kuviossa 7). Seuraavat kartat ja plottaukset tuovat tietoa siitä, miten suurimmassa syrjäytymisriskissä (10%) olevat 8 - 9-luokkalaiset ovat sijoittuneet Keski-Suomen alueelle.



KUVIO 7. Plottaus tutkijoiden ja asiantuntijoiden pisteytyksen keskiarvon perusteella. Plottaus kuvaa kuntien kokonaisriskipisteitä (y) sekä riskissä olevien oppilaiden määrää (plotin koko)

Kolmas tapa arvioida Keski-Suomen kuntien tilannetta nuorten syrjäytymisriskin näkökulmasta oli laskea kyselyyn vastanneiden oppilaiden keskimääräiset syrjäytymispisteet, sekä miten ne ovat maantieteellisesti jakautuneet. Alla olevassa kuvassa näemme syrjäytymisriskiarvion Keski-Suomen alueelta, kyselyyn vastanneiden oppilaiden keskimääräisten syrjäytymispisteiden osalta. Korkeimmat syrjäytymisriskipisteet ovat korkeimmat tummanpunaisella merkityissä kunnissa.

Seuraavassa karttakuvauksessa esitetään syrjäytymisriskipisteiden jakautumisesta.



KUVIO 8. Kuntakohtaiset syrjäytymisriskipisteiden keskiarvot tutkijoiden ja asiantuntijoiden pisteytyksen keskiarvon perusteella. Numero kunnan keskellä kertoo kyselyyn vastanneiden oppilaiden lukumäärän

4.2.3 Alaluokkien keskinäinen vertailtavuus

Oppilaan saamien pisteiden lisäksi SPSS Modeleriin tallennettiin tieto alaluokittain mahdollisesta maksimipistemäärästä. Näin voitiin vertailla oppilaan pisteitä suhteessa maksimipistemäärään (alaluokittain). Samaa alaluokittelua voidaan käyttää vertailtaessa eroja syrjäytymiseen vaikuttavien riskitekijöiden alaluokissa eri kunnissa ja kouluissa. Alla olevalla taulukolla havainnollistamme eri koulujen saamat keskimääräiset kokonaispistemäärät sekä alaluokkien pistemäärien osuudet mahdollisesta maksimipistemäärästä kyseisessä alaluokassa. Koska alaluokkien mahdolliset maksimipistemäärät vaihtelevat, alaluokkien vertailu keskenään on helpompaa näillä maksimiin suhteellisilla arvoilla, kun kaikkien luokkien arvoväli on 0 % - 100 %. Taulukossa näkyy, että esim. Pihtiputaan kunnan oppilaat ovat saaneet keskimäärin 26 % kaikkien luokkien yhteispistemäärän maksimipisteistä. Kunnan oppilaiden saama keskimääräinen pistemäärä terveydentila-luokassa on ollut 11,2 % terveydentila-luokan maksimipisteistä.

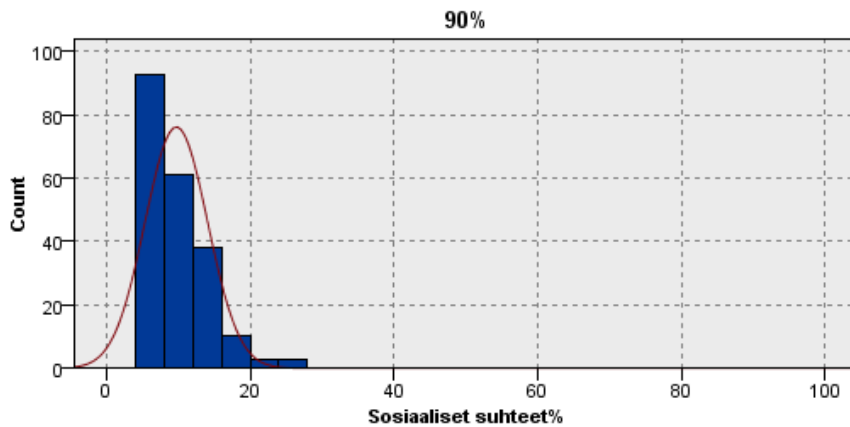
KuntaNimi	TOTsyrj_Mean	Kokonaisuus%...	Terveystentia%...	Sosiaaliset suhteet%...	Koulunkäynt%...	Terveystottumukset%...	Elämänasenne%...	Osallisuus%...
multia	6.330	26.472	15.169	36.780	39.965	20.891	17.786	23.413
pihtipudas	4.993	20.879	11.238	19.768	37.265	19.183	14.799	26.695
kinnula	4.194	17.540	10.333	18.707	35.651	14.587	10.046	18.132
uurainen	4.147	17.344	9.800	13.234	35.178	16.675	12.509	22.545
aanekoski	4.038	16.888	10.173	10.621	36.285	14.631	14.060	23.219
viitasaari	4.021	16.814	11.380	9.048	36.431	11.810	12.865	28.404
kivijärvi	3.924	16.412	5.337	11.540	36.185	11.957	10.862	29.762
keuruu	3.820	15.975	10.253	10.268	35.646	13.155	11.590	22.299
kuhmoinen	3.770	15.768	10.717	6.609	35.894	9.431	14.718	27.315
joutsa	3.673	15.362	9.755	8.669	35.043	12.028	13.475	21.194
jamsa	3.604	15.071	9.592	9.287	35.296	11.859	10.773	21.115
konnevesi	3.465	14.489	12.039	5.741	36.966	14.303	9.918	17.638
kyjyjarvi	3.443	14.399	12.025	4.551	35.214	11.759	9.989	23.352
jyväskylä	3.412	14.269	9.590	8.512	35.507	11.007	10.362	18.177
saarijärvi	3.389	14.173	9.600	10.250	35.317	10.039	10.303	15.716
hankasal...	3.315	13.864	9.892	8.010	35.992	9.567	10.105	17.377
laukaa	3.282	13.726	9.133	7.398	34.815	9.153	9.920	19.988
muurame	3.269	13.669	9.529	7.836	34.610	9.459	11.178	16.925
kannonko...	3.173	13.271	9.355	4.786	35.470	11.914	6.716	21.104
toivakka	3.113	13.016	9.872	5.538	34.756	9.201	9.279	18.329
karstula	3.096	12.946	8.196	6.325	34.081	7.814	9.184	20.400
petajavesi	2.985	12.485	8.211	5.022	35.942	9.429	9.092	16.244

KUVIO 9. Taulukko, kunnan oppilaiden saamasta keskimääräisestä pistemäärästä ja alaluokkien vertailusta keskenään maksimiin suhteellisilla arvoilla

4.3 Merkittävien kysymysten löytäminen

Aineistosta haluttiin selvittää myös keskeisiä merkittäviä kysymyksiä, jotka kuvastavat parhaiten riskiarviota. Aineiston analysointiin päätettiin soveltaa päätöspuu -mallia. Tällöin viidellä merkittävällä kysymyksellä päästiin 77 % todennäköisyydellä samaan tulokseen syrjäytymisriskin tunnistamisessa, kuin koko aineistolla. Kun merkittävien kysymysten määrää nostettiin kymmeneen, se antoi luotettavuudeltaan jo 85 % todennäköisyyden syrjäytymisriskin tunnistamisessa. Tärkeimmät kymmenen kysymystä, joilla voitiin tarkimmin ennustaa syrjäytymisriskipisteitä verrattuna koko kyselyyn, olivat sukupuoli, oma ja vanhempien suhtautuminen koulunkäyntiin, kokemus luokan työrauhasta ja ilmapiirin sallivuudesta, terveyden ja hyvinvoinnin ammatillinen tuki kouluympäristössä sekä perhetausta. Erilaisissa ryhmissä merkittävien kysymysten sarjassa löydettiin eroja. Merkittävimpiä kysymys-vastausyhdistelmiä korkeimman syrjäytymisriskiennusteen saaneissa olivat heikko elämäntyytyväisyys ja huono kokemus elämässä pärjäämisestä, vaikka päihteitä ei käytettäisi ollenkaan, toisena merkittävänä syrjäytymisriskiä ennustavana yhdistelmänä nousi esiin vaikeudet ongelmien hallinnassa ja siihen liittyvät alakuloisuuden, masentuneisuuden ja toivottomuuden kokemukset sekä pelko tulevista vastoinkäymisistä. Alhaisimman syrjäytymisriskin yksilöiden merkittävimpiä kysymys-vastausarvoja olivat hyvä ongelmien käsittelytaito, elämän tarkoituksellisuuden kokemus, mielihyvän tuntemukset sekä kyky hallita tunteita.

Yksi visuaalinen tapa tarkastella koulun vastauksia toteutettiin vertaamalla syrjäytymisriskissä olevien (10 %) pistejakaumaa muihin (90 %), katso kuvio 10. Kullekin syrjäytymisen alaluokalle haettiin jakaumat(histogrammi), jossa x-akselilla on alaluokasta saatu pisteväli ja y-akselilla, montako oppilasta kyseisen pistevälin tuloksen (0 - 100%) on saanut. Alla oleva histogrammi kuvaa Äänekosken oppilaiden sosiaaliset suhteet - pistejakaumaa. Oppilaista on mukana Keski-Suomen 90 % osuus, joka ei kuulu 10 %:n ryhmään eniten pisteitä saaneista. Kuvioista näkee, että kenelläkään oppilaalla, joka ei ole syrjäytymisriskissä, ei ole ongelmia sosiaalisissa suhteissa.



KUVIO 10. Histogrammi kuvaa Äänekosken ei riskissä olevien oppilaiden lukumääriä (y) suhteellisten sosiaaliset suhteet perusteella

Kootussa kuvassa alla on erään koulun alaluokittain eroteltuna histogrammit. Ylärivissä ovat syrjäytymisriskissä olevat (10 %) ja alarivissä 90 % oppilaista, jotka eivät ole syrjäytymisriskissä. Histogrammi kuvaa miten syrjäytymisriskissä olevien oppilaiden pistejakauma eroaa muista. Mitä enemmän riskissä olevien pisteet kasaantuvat oikealle, sitä isompi riski kyseinen luokka on.

Esimerkiksi yllä olevassa koulun kuvassa syrjäytymisriskissä olevista oppilaista löytyy monta oppilasta, joilla ei ole sosiaalisissa suhteissa ongelmia. Myös oppilailla, joilla ei ole syrjäytymisriskiä, lähes kaikilla sosiaaliset suhteet ovat hyvässä kunnossa. Käytännössä tällainen tarkastelutapa auttaa hahmottamaan esimerkiksi koulutasolla sellaisia keskeisiä riskitekijäluokkia, joissa ennaltaehkäisevää tukea olisi hyödyllistä tarjota. Ylläolevan kuvan esimerkissä kyseisen koulun oppilaiden sosiaalisissa suhteissa ei näytä olevan ongelmia syrjäytymiskehityksen näkökulmasta.

5 THL AINEISTON SOVELTUVUUS JA SPSS MODELER TUNNISTETTUIEN SYRJÄYTYMISEN INDIKAATTOREIDEN/RISKITEKIJÖIDEN KARTOITTAMISESSA - KESKEISET TULOKSET JA NIIDEN JOHTOPÄÄTÖKSET

Tällä hankkeella oli useita tavoitteita. Keskeisimpinä tavoitteina oli

1) selvittää, kuinka tekoälypohjaiset teknologiat voisivat toimia ennakoivasti syrjäytymisen tunnistamisessa ja

2) tarkastella tekoälyteknologioiden kyvykkyyttä olemassa olevan data-aineiston avulla, tässä tapauksessa THL:n kouluterveyskyselyn avulla.

Aineisto eli THL:n kouluterveyskysely valikoitui, sillä sitä kerätään laajasti, mutta ei sellaisenaan vielä hyödynnetä ennustamiseen, kuten syrjäytymisen riskitekijöiden ennakoivaan tunnistamiseen. Tulokset voidaan ryhmitellä neljään osaan.

Ensimmäiseksi kouluterveyskyselystä tuotettiin tutkijatriangulaation avulla **syrjäytymisriskiluokitus** ja pisteytys, joka syötettiin SPSS Modeleriin. Tämän avulla laskettiin niin sanottu **syrjäytymisriskikartta**. Kartan avulla voidaan tarkastella maakunnittain, alueittain ja koulukohtaisesti syrjäytymisriskissä olevien oppilaiden määriä. Kartan avulla on mahdollista erotella syrjäytymisriskin kasautumista tehdyn luokituksen pohjalta (esimerkiksi koetun terveyden, elämänhallinnan, sosiaalisen ympäristön ongelmien, perhetilanteen ja kouluviihtyvyyden tiedot yhdistettyinä.) Tämän avulla esimerkiksi yksittäisen koulun näkökulmasta voidaan tarkastella kehityskohteita, joissa oppilaat tarvitsevat eniten tukea syrjäytymisen ehkäisyyn näkökulmasta sekä eri luokkien välisiä kytköksiä.

Kartat antavat mahdollisuuden tarkastella tilannetta kunnassa koulukohtaisella jaottelulla ja esimerkiksi tunnistaa koulukohtaisia painopisteitä syrjäytymiskehityksen torjumiseksi. Syrjäytymisriskiluokitusta voidaan käyttää myös laajemmasta yhteiskunnallisesta näkökulmasta kokonaistilanteen kuvaamiseen sekä ennaltaehkäisevien resurssien suuntaamiseen poliittisessa päätöksenteossa.

Tutkimuksen uutuusarvona voidaan todeta, että aiemmin THL:n kouluterveyskyselyä ei ole hyödynnetty syrjäytymisen riskitekijöiden tunnistamiseen, vaikkakin useat kysymykset kartoittavat syrjäytymisen kannalta keskeistä yksilön osallisuutta ja osallisuuden kokemusta niin yhteiskunnallisen osallisuuden (koulutus, työ, hyvinvointipalvelut, päätöksenteko) (Nivala & Rynnänen, 2014) kuin sosiaaliseen yhteisöön kiinnittymisenkin näkökulmasta (Anttiroiko, 2003; Hanhivaara, 2006; Sassi, 2002). Tutkimus osoitti, että suuri osa kyselyn kysymyksistä pystyttiin liittämään johonkin syrjäytymisen riskitekijäluokkaan. Kouluterveyskyselyaineisto on anonymia ja siksi verrattaen helposti käytettävissä tutkimukseen. Anonymiteetilla on kuitenkin haittapuolena se, ettei yksittäisiä oppilaita pystytä tunnistamaan, eikä siten kohdentamaan tukitoimia yksilötasolle.

Kyselyssä ei ole henkilötietoja, joilla yhdistää riskitieto todelliseen syrjäytymistilanteeseen. Näin ollen myöskään syrjäytymisriskien ennustettavuutta ei tässä tutkimuksessa pystytty arvioimaan henkilökohtaisella tasolla. Tähän tarvittaisiin tunnistellista dataa esimerkiksi kouluterveydenhuollosta, jonka käyttäminen vaatii tarkat tutkimuslupaprosessit. Toisaalta se, että mallin avulla ei tunnisteta syrjäytymiskehitystä yksilötasolla, on yksi mahdollinen ratkaisu käynnissä olevaan kriittiseen yhteiskunnalliseen keskusteluun ylipäänsä tekoälyteknologioiden hyödyntämisestä riskiyrityksien tunnistamisessa ja siihen liittyvissä eettisissä kysymyksissä. Sen sijaan että leimataan yksilötason riskiyritykset, kiinnitetään mallin avulla huomiota laajempiin esimerkiksi koulukohtaisiin ongelmiin, joiden tunnistamisesta hyötyvät niin matalan kuin korkeinkin syrjäytymisriskin oppilaat.

Toiseksi tekoälypohjaisten teknologioiden testaamisen näkökulmasta havaittiin, että testattu SPSS Modeler toimi hyvin mallin luomiseksi asiantuntija-arvioista. **Teknologia on käyttökelpoinen ja hyödynnettävissä skaalautuvasti.** Samaa mallia on helppo käyttää muille maakunnille ja koko Suomen alueelle. Ennusteiden tekoa kuitenkin vaikeuttaa huomattavasti 'supervised'-menetelmien vaatima tieto lopputuloksesta (henkilötieto puuttuu). Sama ongelma koskee täysin koneoppimisperusteista arviota, koska todellista tietoa syrjäytymisriskissä olijoista ei ole. Kaikki tulokset perustuvat näin ollen asiantuntijoiden perusteella tehtyyn pisteytykseen. Olisikin perusteltua validoida nämä arviot vertaamalla tuloksia toteutuneisiin tilanteisiin.

Kolmanneksi tuloksena oli **työkalu ja uudenlainen malli hyödyntää olemassa olevaa kouluympäristössä kerättävää dataa.** Alueen subtanssiosajaat (esimerkiksi sivistystoimen, sosiaalityön ja terveydenhuollon ammattilaiset) voivat itse käyttää työkalua ja se on helposti muunnettavissa myös jatkokäyttöön. Ratkaisua voidaan jatkossa hyödyntää mm. yksittäisen koulun kohdalla toiminnan ohjaukseen liittyvien päätösten tukena ja kuntatasolla resurssointipäätösten apuna. Työkalun luokitusperusteet ja painoarvokertoimet ovat tarkennettavissa todellisen käytön myötä. Toisin sanoen työkalua voi validoida sen käyttöönoton myötä.

Neljäs päätulos on käytetyn datan sisältöön liittyvä. Nyt analysoitu kyselyaineisto soveltui hyvin esitellyn ennustemallin luomiseen. Kiinnostavaa oli, että kun kyselystä etsittiin **merkittävimpiä kysymyksiä jotka korreloivat syrjäytymisriskin merkittävyyden kanssa,** huomattiin että huomattavasti pienemmällä kysymyspatteristolla olisi päästy lähes yhtä tarkkaan ennustemalliin. Merkittävien kysymysten sisältö sinänsä kuvastaa asiantuntijoiden ja tutkijoiden luomaa pisteytystä, mutta kiinnostavaa kuitenkin on tarkastella näitä merkittävimpiä kysymyksiä myös sisällöllisesti. Tarkimmin syrjäytymisriskipisteitä ennustavaan kymmenen kysymyksen joukkoon kuuluivat muun muassa oma ja vanhempien suhtautuminen koulunkäyntiin, kokemus luokan työrauhasta ja ilmapiirin sallivuudesta sekä saatavissa oleva terveyden ja hyvinvoinnin ammatillinen tuki kouluympäristössä, kaikki kokonaisuuksia joihin ainakin osaltaan voitaisiin vaikuttaa esimerkiksi koulukohtaisella tarkemmalla resurssoinnilla ja koulussa toimivien kasvatuksen ammattilaisten ennakoivalla puuttumisella. Merkittävimpiä kysymys-vastausyhdistelmiä korkeimman syrjäytymisriskiennusteen saaneissa olivat heikko elämään tyytyväisyys ja huono kokemus elämässä

pärjäämisestä, vaikka päihteitä ei käytettäisi ollenkaan, toisena merkittävänä kysymys-vastausyhdistelmänä nousi esiin vaikeudet ongelmien hallinnassa ja siihen liittyvät alakuloisuuden, masentuneisuuden ja toivottomuuden kokemukset sekä pelko tulevista vastoinkäymisistä. Alhaisimman syrjäytymisriskin yksilöiden kysymys-vastauksissa korostui puolestaan hyvä ongelmien käsittelytaito, elämän tarkoituksellisuuden kokemus, mielihyvän tuntemukset sekä kyky hallita tunteita. Kiinnostavaa on, että kaikissa näissä ryhmissä nousee esille yksilön kokemus omasta tilanteestaan ja pärjäämisestään. Tämä korostaa yksilön kokemuksen huomioon ottamisen keskeisyyttä syrjäytymisprosessin ymmärtämisessä.

Aineiston jatkohyödyntäminen olisi järkevää. Tutkimuksessa käytettyä THL aineistoa olisi mahdollista hyödyntää tässä tutkimuksessa kuvatulla menetelmällä myös muissa nuorten terveyteen ja hyvinvointiin liittyvissä teemoissa. Aineistosta luontevasti esiin nousevia teemoja voivat olla esim. koulunkäyntiin, terveyskäyttäytymiseen tai osallisuuteen liittyvät alueet.

Tässä **tutkimuksessa kehitetty menetelmä on yleistettävissä** myös muihin laajoihin kyselyaineistoihin. Kuvatun menetelmän avulla nähtiin, että asiantuntijatiетoon pohjautuva tekoälyteknologian hyödyntäminen on suhteellisen helppoa.

6 JATKOTUTKIMUKSET JA TULEVAISUUS

Tutkimuksen pohjalta löydettiin tärkeitä jatkotutkimusaiheita ja tarpeita tekoälyteknologioiden yleisen kehittämisen, hyödynnettävän datan, tekoälyteknologian tämänhetkisen kyvykkyyden ja syrjäytymisilmiön näkökulmasta. Alla tarkastelemme näitä lyhyesti:

Tekoälyteknologian kehityksen näkökulmasta tutkimus osoittaa teknologioiden sovellettavuutta niin sanottu sosiaalisten innovaatioiden osalta. Etenkin eurooppalaisessa julkisen vallan agendasetting linjauksissa tekoälykehityksen tärkeäksi painopisteeksi on nostettu esiin niin sanottu sosiaaliset innovaatiot ja sosiaaliin kysymyksiin liittyvät tekoälypohjaiset ratkaisut (katso esimerkiksi Ranskan tekoälykehityksen poliittiset linjaukset, Villani, 2018). Tämä voidaan nähdä eurooppalaisena vastavetona esimerkiksi Kiinassa jo käytössä olevaan sosiaalisen luototuksen sovellukseen (Social Credit System (SCS), joka mahdollistaa tarkan yksilökontrollin ja valvonnan olemassa olevan käyttäytymistiedon perusteella. Tässä esitetty malli voidaan ymmärtää yhtenä tulevaisuuden tekoälypohjaisena sosiaalisen innovaation ratkaisuna, joka pohjaa suomalaiselle yhteiskunnalle ominaiseen ajatukseen teknologioiden tasa-arvoistavasta vaikutuksesta.

Hyödynnettävän datan näkökulmasta tutkimus avaa uusia mahdollisuuksia. Se osoittaa, että laajoja kansallisia hyvinvointiin liittyviä aineistoja on hyödyllistä tarkastella myös spesifimmän ongelman näkökulmasta. Nyt kehitettyä Keski-Suomen syrjäytymisriskimallia voidaan jatkossa hyödyntää kuvaamaan koko Suomea. Tämä edellyttää koko Suomen THL-kouluterveyskyselydatan käsittelyä vastaavalla tavalla kuin Keski-Suomen aineistolle tehtiin. Lisäksi työkaluun on jatkossa mahdollista yhdistää muuta syrjäytymisen liittyvää ennustavaa dataa koulumaailmasta (esimerkiksi käyttäytymisdataa wilma -järjestelmästä, dataa koulupsykologin ja kouluterveyspalvelujen käytöstä), kuntakohtaisia tietoja esim. terveys- ja sosiaalitoimen palvelujen käytöstä sekä työllisyystietoja. Kouluissa tai kunnissa tapahtuneista muutoksista voidaan saada lisätietoa useamman vuoden kouluterveyskyselyaineistojen yhdistämisellä. Syrjäytymisen ennustettavuutta voidaan myös arvioida mallintamalla riski 10 vuotta sitten tehdyllä kyselyllä ja vertaamalla sen värikarttaa toteutuneeseen tilanteeseen 25-vuotiailla nuorilla. Näin olisi mahdollista tarkentaa pisteytystä vastaamaan paremmin toteumaa. Nykyisellään kehitetty malli mahdollistaa kuitenkin nopean tilannekuvan saamisen niin kansallisella tasolla kuin yksittäisen koulun tai maakunnankin näkökulmasta.

Tekoälypohjaisen teknologian kyvykkyyden testaamisen ja teknologian hyödynnettävyyden näkökulmasta tutkimus avaa uusia mahdollisuuksia niin tutkimuksellisesti kuin konkreettisen sovellettavuudenkin tasolla. SPSS Modeleria ja vastaavia menetelmiä on mahdollista hyödyntää kyseisen tyyppisen aineiston analyysissä jatkossakin. Kiinnostava jatkotutkimuksen aihe olisi ohjatun koneoppimisen hyödyntäminen vastaavanlaisessa aineistossa. Tämä edellyttäisi kouluterveyskyselyaineistoa todellisista syrjäytyneistä toisin sanoen tietoa minkälaiset vastaukset ovat tyyppisiä todellisille syrjäytyneille. Näin mallia voitaisiin opettaa tunnistamaan yksilöllinen syrjäytymisriski. Luonnollisesti syrjäytymisriskin koneoppimiseen perustuvan mallin kehittäminen

vaatii tarkan eettisen arvion mallin konkreettisesta soveltamisesta; toisin sanoen missä ja kenen toimesta mallia hyödynnettäisiin ja minkälaisiin kohteisiin sen tuottamaa tietoa sovellettaisiin.

Konkreettisen sovellettavuuden näkökulmasta tämänhetkisen mallin jatkohyödyntäminen mahdollistaa koulu- ja kuntakohtaisen ongelmien tarkemman tunnistamisen. Esimerkiksi tulokset -osiossa kuvatun päättelypuu -mallin avulla voidaan tarkentaa yksittäiselle koululle tai kunnalle olennaisimmat kysymysten yhdistelmät ja sitä kautta ominaisimmat ongelmakokonaisuudet. Tulevaisuudessa tämä voi auttaa tunnistamaan esim. yksittäisen koulun toimintakulttuurisia ongelmia, joihin on mahdollista vaikuttaa muun muassa uudella resurssoinnilla. Kehitetty malli voi tulevaisuudessa toimia niin yksittäisen opettajan kuin kuntapäätäjienkin päätöksenteon tukena.

Syrjäytymisen moninaisen ilmiön tunnistamisessa yhä edelleen useat niin sanottu heikot signaalit jäävät tunnistamatta. Nyt tehty tutkimus osoittaa, että jo olemassa olevia data-aineistoja on mahdollista hyödyntää uudelleen datan käsittelyn avulla, niin, että myös tällaiset niin sanottu heikot signaalit ja niiden varhainen kasaantuminen on mahdollista tunnistaa. On esitetty, että syrjäytymiskehityksen katkaisemisessa tulevaisuudessa pääpainon tulisi olla peruspalveluissa ja ennaltaehkäisevässä toiminnassa (Sosiaali- ja terveysministeriö, 2015) eikä nykyisissä korjaavissa sote-palveluissa. Lisäksi lasten ja perheiden palvelut ovat irtautuneet perheiden arjesta eivätkä vastaa yksilöllisiä tarpeita (Torkki, Leskelä, Maksimainen, Niemelä, Koukkula, Torvinen, Mulari, Välimaa & Rimpelä, 2016). Tässä esitetty malli auttaa tunnistamaan orastavaa kasautuvaa huono-osaisuutta ja siitä potentiaalisesti aiheutuva syrjäytymistä. Näin tutkimus tarjoaa konkreettisen työkalun syrjäytymisen ennaltaehkäisyyn sellaisille kouluympäristön toimijoille, jotka aidosti ovat läsnä lasten ja nuorten päivittäisessä arjessa.

Vertailuaineisto ja tulosten luotettavuuden arvio

Kyselyn soveltuvuutta syrjäytymisriskin ennustamiseen voidaan tulevaisuudessa tarkentaa vertailuaineiston avulla. THL:ltä olemme saaneet käyttöön vuoden 1987 syntymäkohortti aineiston (muun muassa työllisyys- ja koulutustiedot) sekä tuolle ikäluokalle vuonna 2003 tehdyn THL:n kouluterveyskyselyn. Kysely on käyty läpi hankkeen tutkijatiimin voimin, tarkoituksena selvittää kyselyjen keskinäinen vastaavuus. Kysymyspatteristot näyttävät heijastelevan kullekin aikakaudelle ominaisia (julkisen) keskustelun aiheita. Näin esimerkiksi aiemmassa kyselyssä kartoitettiin laajalti syömishäiriöihin liittyvää käyttäytymistä, mitä ei kysytty lainkaan vuoden 2017 tutkimuksessa. Vuonna 2017 taas esimerkiksi huumeiden käyttöön liittyvien kysymysten määrä ja tarkkuus oli selvästi kasvanut, heijastellen päihteiden käytössä tapahtunutta (todellista tai odotettavissa olevaa) muutosta. Kyselyt vastasivat kuitenkin toisiaan verrattain hyvin, ja kaikkiin vuoden 2017 aineiston luokituksiin löydettiin kysymykset myös vuoden 2003 aineistosta. Syrjäytymisriskin ennusteen osuvuus on mahdollista testata käytössämme olevalla syntymäkohortti aineistolla ja verrata todellisten syrjäytyneiden määrää suhteessa syrjäytymisriskiennusteen antamaan tulokseen. Huomioitavaa toki on, että tilastollisesti niin sanottu syrjäytyneiden määrä (elinolosuhteisiin liittyvät indikaattorit kuten esim. työttömyys, kouluttamattomuus, toimeentulotuki) ja itsensä syrjäytyneiksi

kokevien määrä eivät täysin korreloi. Huomioitavaa myös on, että aineisto pohjaa yksilölliseen kokemustietoon. Kysely kerätään oppilailta koulukohtaisesti oppitunneilla. Näin esimerkiksi kyselyn tekemiseen käytetty aika vaihtelee yksilötasolla suuresti joten on mahdollista että osaan kysymyksistä on vastattu niin sanottu "läpällä". Vastausten luotettavuudessa on siten suuria yksilökohtaisia eroja.

7 LÄHTEET

Aaltonen, M. 2013. Socioeconomic Differences in Crime and Victimization: A Register-Based Study. Helsinki: Oikeuspoliittisen tutkimuslaitos.

Aaltonen, S., Berg, P. & Ikäheimo, S. 2015. Nuoret luukulla -Kolme näkökulmaa syrjäytymiseen ja nuorten asemaan palvelujärjestelmässä. Nuorisotutkimusverkosto/ Nuorisotutkimusseura Verkkojulkaisuja 84. Saatavilla: 2.10.2018
<http://www.nuorisotutkimusseura.fi/images/julkaisuja/nuoretluukulla.pdf>

Alatalo, J., Mähönen, E. & Räisänen, H. 2017. Nuorten ja nuorten aikuisten työelämä ja sen ulkopuolisuus. Helsinki: Työ- ja elinkeinoministeriö.

Alatupa, S., Karppinen, K., Keltikangas-Järvinen, L. & Savioja, H. 2007. Koulu, syrjäytyminen ja sosiaalinen pääoma – Löytyykö huono-osaisuuden syy koulusta vai oppilaasta? Sitran raportteja 75. Helsinki: Edita Prima Oy.

Anttiroiko, A-V. 2003. Kansalaisten osallistuminen, osallisuus ja vaikuttaminen tietoyhteiskunnassa. Teoksessa P. Bäcklund (toim.) Tietoyhteiskunnan osallistuva kansalainen. Tapaus Nettimaunula. Helsinki: Helsingin kaupungin tietokeskus.

Behrman, J.A. 2015. The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment. *Social Science & Medicine* 127, 108 - 115.

van den Berg, G., Lundborg, P. & Vikström, J. 2018. The economics of grief. *The Economic Journal*, 127,1794 - 1832.

Berg, N., Huurre, T., Kiviruusu, O. & Aro, H. 2011. Nuoruusiän huono-osaisuus ja sen kasautumisen yhteys kuolleisuuteen. Seurantatutkimus 16-vuotiaista nuorista. Sosiaalilääketieteellinen Aikakauslehti v. 48, 3. numero. Saatavilla: 27.10.2017 <https://journal.fi/sla/article/view/4499>

Coles, B. 2000. Joined up youth research, policy and practice. United Kingdom: Youth Work Press.

Crow, I., France, A., Hacking, S. & Hart, M. 2004. Does Communities that Care work? An evaluation of a community-based risk prevention programme in three neighbourhoods. United Kingdom: Joseph Rowntree Foundation.

Ek, E., Saari, E., Viinamäki, L. & Järvelin, M.-R. 2004. Nuorten aikuisten työelämästä syrjäytyminen ja sosiaaliturvan käyttö. Sosiaali- ja terveysturvan tutkimuksia. Helsinki: Kela.

Fletcher, D. & Sarkar, M. 2013. Psychological resilience: A review and critique of definitions, concepts, and theory. *European Psychologist*, 18, 12 - 23.

Findikaattori. Itsemurhat. Findikaattorin internetsivusto. Saatavilla: 2.10.2018
<http://www.findikaattori.fi/fi/10>

Gretschel, A. & Myllyniemi, S. 2017. Työtä, koulutus- tai harjoittelupaikkaa ilman olevien nuorten käsityksiä tulevaisuudesta, demokratiasta ja julkisista palveluista – Nuorisobarometrin erillisnäyte/aineistonkeruu. Helsinki: Nuorisotutkimusseura ry.

Haapasalo, J. 2006. Kasvuympäristön varhaiset riskitekijät rikollisen käyttäytymisen kehityksessä. Teoksessa P. Honkatukia & J. Kivivuori (toim.) Nuorisorikollisuus. Määrä, syyt ja kontrolli. Oikeuspoliittisen tutkimuslaitoksen julkaisuja 221.

Halonen, J-P., Aaltonen, T., Hämäläinen, A., Karppi, S-L., Kaukinen, J., Kervilä, A., Lehtinen, M., Pere, E., Puukka, P., Siitonen, V., Silvennoinen, S. & Talo, S. 2007. Syrjäytymisvaarassa olevien vajaakuntoisten nuorten kuntoutustarpeen arviointi. Sosiaali- ja terveysturvan katsauksia 73. Helsinki: Kela.

Hanhivaara, P. 2006. Maailmaa syleilevä osallisuus–osallisuuden suhde kouluun. *Nuorisotutkimus*, 24(3), 29 - 38.

Helne, T. 2002. Syrjäytymisen yhteiskunta. Helsinki: Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskus.

Hilli, P., Ståhl, T., Merikukka, M. & Ristikari, T. 2017. Syrjäytymisen hinta–case investoinnin kannattavuuslaskemasta. Julkarin internetsivusto. Saatavilla: 2.10.2018
<http://www.julkari.fi/handle/10024/135714>

Hjarvard, S. 2013. *The Mediatization of Culture and Society*. London: Routledge.

Hyyppä, M. 2002. *Elinvoimaa yhteisöstä. Sosiaalinen pääoma ja terveys*. Jyväskylä: PS-kustannus.

Ilmakunnas, I., Kauppinen, T. M. & Kestilä, L. 2015. Sosioekonomisten syrjäytymisriskien kasautuminen vuonna 1997 syntyneillä nuorilla aikuisilla. *Yhteiskuntapolitiikka*, 80(3), 247 - 262.

Isola, A-M., Kaartinen, H., Leemann, L., Lääperi, R., Schneider, T., Valtari, S. & Keto-Toko, A. 2017. Mitä osallisuus on? Osallisuuden viitekehystä rakentamassa. Helsinki: Suomen Yliopistopaino Oy.

Jones, C.I. 2015. The facts of economic growth. Working Paper 2114. Cambridge: National bureau of economic research.

Just, N. & Latzer, M. 2017. Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, 39(2), 238 - 258.

Jäppinen, A. & Veijola, E. 2003. Lasten ja nuorten syrjäytymisen ennaltaehkäisy koulutuksen alalla. Opetusministeriön työryhmämuistioita ja selvityksiä, 4.

Järvinen, T. & Jahnukainen, M. 2001. Kuka meistä onkaan syrjäytynyt? Marginalisaation ja syrjäytymisen käsitteellistä tarkastelua. Teoksessa M. Suutari (toim.) Vallattomat marginaalit. Yhteisöllisyyksiä nuoruudessa ja yhteiskunnan reunoilla. Helsinki: Nuorisotutkimusverkosto.

Karvonen, S. & Kestilä, L. 2014. Nuorten aikuisten syrjäytymisvaaraan liittyvä huono-osaisuus. Teoksessa M. Vaarama, S. Karvonen, L. Kestilä, P. Moisio & A. Muuri (toim.) Suomalaisten hyvinvointi. Tampere: Terveyden ja hyvinvoinnin laitos.

Kinnunen, P. 2003. Nuoren elämäntilannetta ja siihen yhteydessä olevia tekijöitä. Pro gradu – tutkielma. Kuopion yliopisto. Hoitotieteen laitos.

Kestilä, L., Paananen, R., Väisänen, A., Muuri, A., Merikukka, M., Heino, T. & Gissler, M. 2012. Kodin ulkopuolelle sijoittamisen riskitekijät. Rekisteripohjainen seurantatutkimus Suomessa vuonna 1987 syntyneistä. Julkari internetsivusto. Saatavilla: 2.10.2018
<http://www.julkari.fi/bitstream/handle/10024/102955/kestila.pdf?sequence=1&isAllowed=y>

Kestilä, L. 2008. Pathways to Health. Determinants of Health, Health Behaviour and Health Inequalities in Early Adulthood. Kansanterveyslaitoksen julkaisuja A23/2008. Helsinki: Terveyden ja hyvinvoinnin laitos.

Kangas, O. 2010. Perhepiirissä. Lapsuudentapahtumat, terveys ja elämäntyytyväisyys. K. Koskenvuo & U. Hämäläinen (toim.) Helsinki: Kela.

Kuronen, I. 2010. Peruskoulusta elämäkouluun - Ammatillisesta koulutuksesta syrjäytymisvaarassa olevien nuorten aikuisten tarinoita peruskoulusuhteesta ja elämäkulusta peruskoulun jälkeen. Väitöskirja. Koulutuksen tutkimuslaitos. Jyväskylä: Jyväskylän yliopisto.

Kuula, R. 2000. Syrjäytymisvaarassa oleva nuori koulun paineessa: Koulu ja nuorten syrjäytyminen. Joensuu: Joensuun yliopisto.

Lange, P. G. 2007. Publicly private and privately public: Social networking on YouTube. *Journal of computer Mediated communication*, 13(1), 361 - 380.

Lehtinen, V. 1994. Depression hoito terveyskeskuksessa. Stakes oppaita 22. Jyväskylä: Gummerus kirjapaino Oy.

Lehtonen, O. & Kallunki, V. 2013. Nuorten aikuisten syrjäytymiskierre. *Janus Sosiaalipolitiikan ja sosiaalityön tutkimuksen aikakauslehti*, 128 - 143.

Livingstone, S. 2008. Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media & Society*, 10 (3), 393 - 411.

Loader, B. D., Vromen, A. & Xenos, M. A. 2014. The networked young citizen: Social media, political participation and civic engagement. *Information, Communication & Society* 17:2, 143 - 150.

Luopa, P., Lommi, A., Kinnunen, T. & Jokela, J. 2010. Nuorten hyvinvointi Suomessa 2000-luvulla. Kouluterveyskysely 2000 - 2009. Helsinki: Terveystieteiden tutkimuskeskus. Saatavilla: 2.10.2018 <https://www.julkari.fi/bitstream/handle/10024/80115/91431fe2-cfa6-4909-9363-75eda1839dc3.pdf?sequence=1>

Lämsä, A.-L. 2010. Tuhat tarinaa lasten ja nuorten syrjäytymisestä. Lasten ja nuorten syrjäytyminen sosiaalihuollon asiakirjojen valossa. Väitöskirja. Kasvatustieteiden tiedekunta. Oulu: Oulun yliopisto.

Lämsä, A.-L. 1998. Ihminen tarvitsee sitä, mitä ilman on paha olla – Tuhat tarinaa lasten ja nuorten selviytymisestä. Teoksessa V-M. Ulvinen (toim.) Lasten ja nuorten syrjäytymistä koskevan tutkimuksen mahdollisuudet. Helsinki: Nykypaino Oy.

MacDonald, R. 2006. Social exclusion, youth transitions and criminal careers: Five critical reflections on 'risk'. *Australian & New Zealand Journal of Criminology* 39.3, 371 - 383.

Martelin, T., Koskinen, S., Kestilä, L. & Aromaa, A. 2005. Terveystieteiden ja toimintakyvyn vaihtelu asuinalueen, koulutuksen ja kotitaloustyyppin mukaan. Teoksessa S. Koskinen, L. Kestilä, T., Martelin & A. Aromaa (toim.) Suomalaisen hyvinvointi 2014 -tutkimuksen perustulokset 18–29-vuotiaiden terveydestä ja siihen liittyvistä tekijöistä. Helsinki: Kansanterveyslaitos.

McFarland, S.A. 2018. Interest Groups and the Policymaking Process: Sources of Countervailing Power in America. M.P. Petracca, (toim.) *The Politics Of Interests Interest Groups Transformed*. New York: Routledge.

Myllyniemi, S. 2009. Taidekohtia. Nuorisobarometri 2009. Nuorisotutkimusverkosto/Nuorisotutkimusseura, julkaisuja 97, Nuorisosiain neuvottelukunta, julkaisuja 41. Helsinki: Opetusministeriö.

Myllyniemi, S. 2014. Ihmisarvoinen nuoruus. Nuorisobarometri 2014. Saatavilla: 2.10.2018 https://tietoanuorista.fi/wp-content/uploads/2015/03/Nuorisobarometri_2014_web.pdf

Myrskylä, P. 2012. Hukassa–Keitä ovat syrjäytyneet nuoret. EVA analyysi, 19. Saatavilla: 2.10.2018 <https://www.eva.fi/wp-content/uploads/2012/02/Syrjailyminen.pdf>

Neuvonen, E., Rusanen, M., Solomon, A., Ngandu, T., Laatikainen, T., Soininen, H. & Tolppanen, A. M. 2014. Late-life cynical distrust, risk of incident dementia, and mortality in a population-based cohort. *Neurology*, 82(24), 2205- 2212.

Nivala, E. & Ryyänen, S. 2013. Kohti sosiaalipedagogista osallisuuden ideaalia. *Sosiaalipedagoginen aikakauskirja*, 14, 9 - 41.

Nurmi, J.-E., Salmela-Aro, K. & Ruotsalainen, H. 1994. Cognitive and attributional strategies among unemployed young adults – A case of the failure-trap strategy. *European Journal of Personality*, 8, 135–148.

Nurmi, J. E. 2011. Miksi nuori syrjäytyy. *NMI Bulletin*, 2, 28 - 35.

Ojanen, M. 2007. Positiivinen psykologia. Helsinki: Edita.

Onatsu-Arviolommi, T. & Nurmi, J.-E. 2000. The role of task-avoidant and task-focussed behaviors in the development of reading and mathematical skills during the first school year: A cross-lagged longitudinal study. *Journal of Educational Psychology*, 92, 478–491.

Ong, A. 2006. Mutations in Citizenship. *Theory, Culture & Society* 23(2 - 3), 499 - 531.

Paananen, R., Ristikari, T., Merikukka, M., Rämö, A. & Gissler, M. Lasten ja nuorten hyvinvointi Kansallinen syntymäkohortti 1987 -tutkimusaineiston valossa. Raportti 52/2012. Helsinki: Terveyden ja hyvinvoinnin laitos.

Palola, E., Hannikainen-Ingman, K. & Karjalainen, V. 2012. Nuorten syrjäytymistä on tutkittava pintaa syvemmin. *Yhteiskuntapolitiikka-YP 77 (2012): 3*.

Paunonen, M. & Vehviläinen-Julkunen, K. 1999. Perhe hoitotyössä: teoria, tutkimus ja käytäntö. Helsinki: WSOY.

Pelastakaa lapset ry. 2016. Köyhyys ja syrjäytyminen uhkaavat yli 26 miljoonaa lasta Euroopassa. Pelastakaa lapset ry:n internetsivusto. Saatavilla: 10.9.2018 <https://www.pelastakaa lapset.fi/uutiset/koyhyys-ja-syrjailyminen-uhkaavat-yli-26-miljoonaa-lastat-euroopassa/>

Quinn, K. & Papacharissi, Z. 2017. Our Networked Selves: Personal Connection and Relational Maintenance in Social Media Use. *The SAGE Handbook of Social Media*.

- Rainie, L. & Wellman, B. 2012. *Networked: The new social operating system*. *Mit Press*.
- Rauhala, P-L. 1998. Mistä ehkäisevässä sosiaalipolitiikassa on kysymys? Käsitteellistä ja historiallista tarkastelua. *Stakesin tutkimuksia 90*. Helsinki: Stakes.
- Reiter, H. 2003. Past, Present, Future. Biographical Time Structuring of Disadvantaged Young People. *Young. Nordic Journal of Youth Research 11 (3)*, 253 - 279.
- Rintanen, H. 2000. Terveys ja koulutuksellinen syrjäytyminen nuoren miehen elämäkulussa. Väitöskirja. Terveystieteiden laitos. Tampere: Tampereen yliopisto.
- Salmivalli, C. 2010. Koulukiusaamiseen puuttuminen. Kohti tehokkaita toiminta-malleja. Jyväskylä: PS-kustannus.
- Sassi, S. 2002. Kulttuurinen identiteetti ja osallisuus. Teoksessa P. Bäcklund, J. Häkli & H. Schulman (toim.) *Osalliset ja osaajat. Kansalaiset kaupungin suunnittelussa*. Helsinki: Gaudeamus.
- Savioja, H. 2007. Koulutekijät nuorten syrjäytymisriskiä selittämässä. Teoksessa S. Alatupa (toim.), *Koulu, syrjäytyminen ja sosiaalinen pääoma: Löytyykö huono-osaisuuden syy koulusta vai oppilaasta*. Sitran raportteja 75. Helsinki: Sitra.
- Savolainen, H. 2001. Explaining mechanisms of educational career choice. A follow-up study of the educational career choices of a group of youths that finished compulsory education in 1990. *Kasvatustieteellisiä julkaisuja 69*. Joensuu: Joensuun yliopisto.
- Siljander, P. 1996. Syrjäytyminen – aatteiden murroksen kriisi. Teoksessa P. Siljander & V.M. Ulvinen (toim.) *Syrjäytymisestä selviytymiseen – vaikeuksien kautta elämäntilanteen hallintaan*. Oulun yliopiston kasvatustieteiden tiedekunnan opetusmonisteita ja selosteita 66/1996. Oulu: Oulun yliopisto monistus- ja kuvakeskus.
- Sosiaali- ja terveysministeriö. 2015. Syrjäytymistä, köyhyyttä ja terveysongelmia vähentävän poikkialueellisen ohjelman (2011 - 2015) loppuarvio. Saatavilla: 10.9.2018 <https://stm.fi/documents/1271139/1448516/Syrj%C3%A4ytymisen+v%C3%A4hent%C3%A4minen+loppuarvio+16.4.2015+%284%29.pdf/2e6be4ae-72a5-4c94-a1ff-7ae2e389af04/Syrj%C3%A4ytymisen+v%C3%A4hent%C3%A4minen+loppuarvio+16.4.2015+%284%29.pdf.pdf>
- Terveyden ja hyvinvoinnin laitos. 2018. Syrjäytyminen ja syrjäytymisen riskitekijät. Terveyden ja hyvinvoinnin laitoksen internetsivusto. Saatavilla: 1.10.2018 https://thl.fi/fi/web/lapset-nuoret-ja-perheet/tyon_tueksi/nuorten-syrjailyminen/syrjailyminen-ja-syrjailytymisen-riskitekijät

Tindle, H. A., Chang, Y. F., Kuller, L. H., Manson, J. E., Robinson, J. G., Rosal, M. C. & Matthews, K. A. 2009. Optimism, cynical hostility, and incident coronary heart disease and mortality in the Women's Health Initiative. *Circulation*, 120(8), 656 - 662.

Torkki, P., Leskelä, R-L., Maksimainen, A., Niemelä, P., Koukkula, L., Torvinen, A., Mulari, M., Välimaa N. & Rimpelä, M. Sosiaali- ja terveystieteiden tutkimuskeskuksen tutkimusraportti ja säästöpotentiaalinen arviointi. Valtioneuvoston selvitys ja tutkimustoiminnan julkaisusarja 42/2016. Saatavilla: 1.10.2018 <http://tietokayttoon.fi/documents/10616/1266558/Sote-palveluiden+kehitt%C3%A4mis-+ja+s%C3%A4st%C3%A4st%C3%B6potentiaalinen+arviointi+%2B+liite/070196e5-69da-42b4-9401-5e8ea9a6f78b?version=1.0>

Tugade, M. M., Fredrickson, B. L. & Feldman Barrett, L. 2004. Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of personality*, 72(6), 1161 - 1190.

Törrönen, M. & Vornanen, R. 2002. Emotionaalinen huono-osaisuus peruskoululaisten korostamana syrjäytymisenä. *Nuorisotutkimus* 4 (20), 33 - 42.

Villani, C. 2018. For a meaningful artificial intelligence. towards a French and European strategy. Mission assigned by the Prime Minister Édouard Philippe.

Wiss, K., Ståhl, T., Saaristo, V., Kivimäki, H., Frantsi-Lankia, M. & Rimpelä, A. 2017. Terveystieteiden tutkimuskeskuksen tutkimusraportti ja säästöpotentiaalinen arviointi. Valtioneuvoston selvitys ja tutkimustoiminnan julkaisusarja 42/2016. Saatavilla: 1.10.2018 <http://tietokayttoon.fi/documents/10616/1266558/Sote-palveluiden+kehitt%C3%A4mis-+ja+s%C3%A4st%C3%A4st%C3%B6potentiaalinen+arviointi+%2B+liite/070196e5-69da-42b4-9401-5e8ea9a6f78b?version=1.0>

Witting, M., & Keski-Petäjä, M. 2016. Vanhempien koulutus vaikuttaa lasten valintoihin. *Tieto & Trendit*, 2.

Witvliet, M., Brendgen, M., Van Lier, P. A., Koot, H. M. & Vitaro, F. 2010. Early adolescent depressive symptoms: Prediction from clique isolation, loneliness, and perceived social acceptance. *Journal of abnormal child psychology*, 38(8), 1045 - 1056.

van Zoonen, L., Vis, F. & Mihelj, S. 2010. Performing citizenship on YouTube: activism, satire and online debate around the anti-Islam video Fitna. *Critical Discourse Studies*, 7(4), 249 - 262.

LIITTEET

LIITE 1: Merkittävimmät kysymykset syrjäytymisriskin kannalta

Tutkijoiden ja asiantuntijoiden pisteytykset. Lihavoitu fontti tarkoittaa, että sekä tutkijat, että asiantuntijat pitivät riskiä merkittävänä, Italic-fontti kuvaa asiantuntijoiden arvioita ja tavallinen fontti tutkijoiden näkemystä.

Kysymyksiä tutkittiin Random Trees luokittelevalla mallilla, jossa kohdemääre oli 'kuuluuko vastaaja 10 % ryhmään', jolla on syrjäytymisriski. Mallin rakentamisessa käytettiin 3816 oppilaan vastauksia. Määrävinä tekijöinä olivat 266 vastausvaihtoehtoa, joille annettiin luokitteleva-tyyppinen tieto.

Random Trees		
Model Information		
Target Field	yli10pros	
Model Building Method	Random Trees Classification	
Number of Predictors Input	266	
Model Accuracy	0,924	
Misclassification Rate	0,076	
Records Summary		
Records	Number	Percent
Included	3 816	100,00
Excluded	0	0,00
Total	3 816	100,00

KUVIO 11. Random Trees - luokitteleva malli

Oppilaat, jotka vastasivat alla olevan mukaisesti **eivät ole** syrjäytymisriskissä.

Valitse vaihtoehto, joka parhaiten kuvaa kokemuksiasi viimeisen kahden viikon aikana:

Olen käsitellyt ongelmia hyvin = 1,3,4,5 (ei 2, harvoin)

JA

Kuinka hyvin seuraavat väittämät kuvaavat tuntemuksiasi tällä hetkellä?

(1) *Tunnen, että elämälläni on päämäärä ja tarkoitus* = 1234 (ei 5, täysin eri mieltä)

JA

Kuinka usein seuraavat ongelmat ovat vaivanneet sinua lomakkeen täyttöö edeltäneen kahden viikon aikana?

(6) *Taipumus harmistua tai ärsyyntyä helposti* = 1 tai 2 (ei lainkaan tai useana päivänä)

JA

(8) Vain vähäinen mielenkiinto tai mielihyvä erilaisten asioiden tekemisestä = 2 ,3, 4 (useammin kuin 'ei lainkaan')

JA

Mihin aiot ensisijaisesti hakea opiskelemaan peruskoulun jälkeen? Valitse yksi vaihtoehto.
= 1 tai 3. lukioon TAI ammatilliseen koulutukseen, jossa suoritetaan myös lukion

Top Decision Rules for 'yli10pros'				
Decision Rule	Most Frequent Category	Rule Accuracy	Ensemble Accuracy	Interestingness Index
(q4swemwbs = {1.0,3.0,4.0,5.0}) & (q1lifesatisfaction = {1.0,2.0,3.0,4.0}) & (q6gad7 = {1.0,2.0}) & (future_plans = {1.0,3.0}) & (q8gad7 = {2.0,3.0,4.0})	0	0.944	0.984	0.893
(lifecontentment = {1.0,3.0,4.0,5.0}) & (q1accept_drugs = {1.0}) & (q7lifesatisfaction = {3.0,4.0,5.0})	1	0.603	0.986	0.366
(lifecontentment = {4.0,5.0}) & (q6lifesatisfaction = {3.0,4.0,5.0}) & (q3drugs = {1.0}) & (q4ecigarette = {1.0})	1	0.557	0.957	0.320
(q9gad7 = {3.0,4.0}) & (q7gad7 = {3.0,4.0})	1	0.558	0.987	0.315
(q6lifesatisfaction = {1.0,3.0,4.0,5.0}) & (q2sitting = {1.0,4.0,7.0}) & (q5swemwbs = {1.0,2.0})	1	0.500	0.979	0.255

KUVIO 12. Tärkeimmät kysymykset, elämään tyytyväisyys

Tarkkuuden 0,557 syrjäytymisriskistä antaa esimerkiksi vastaussarja:

Oletko elämääsi tällä hetkellä?
(melko tyytymätön TAI erittäin tyytymätön)

JA

Kuinka hyvin seuraavat väittämät kuvaavat tuntemuksiasi tällä hetkellä?
Koen pärjääväni elämässäni: 3,4,5
(Ei samaa eikä eri mieltä TAI Eri mieltä TAI Täysin eri mieltä)

JA

Oletko koskaan kokeillut tai käyttänyt seuraavia aineita?
Lääkkeitä (rauhoittavia, uni- tai särkylääkkeitä) tai alkoholia ja lääkkeitä yhdessä päihtyäksesi: 1(En koskaan)

JA

Käytätkö sähkötupakkaa, joka sisältää seuraavia aineita?
jotain muuta: 1 (en lainkaan)

Kaksi alinta väitettä ei ole kuulunut pisteytettyihin tehtäviin.

Tarkkuuden 0,558 syrjäytymisriskistä antaa esimerkiksi vastaussarja:

Kuinka usein seuraavat ongelmat ovat vaivanneet sinua lomakkeen täyttöö edeltäneen kahden viikon aikana?

Alakuloisuus, *masentuneisuus,* *toivottomuus:*
(Suurimpana osana päivistä TAI Lähes joka päivä)

JA

Pelko siitä, että jotakin kauheaa saattaisi tapahtua:
(Suurimpana osana päivistä TAI Lähes joka päivä)

Mielenkiintoista oli, että ennustettaessa merkittäviä kysymyksiä, SPSS voi ottaa mukaan kaikki kyselyn vastaukset. Näin voi löytyä uusia merkittäviä tehtäviä, joita ei alunperin pisteytetty mukaan. Pisteytyksen perusteella tietyt oppilaat kuuluvat riskiryhmään. Ennusteen tuloksena pidetään oppilaita, jotka on merkitty riskiryhmäläisiksi. Ennustaviksi tekijöiksi voidaan laittaa kaikki kysymykset kyselystä. Silloin tulee merkittäviä kysymyksiä myös niistä tehtävistä, jotka ennustavat hyvin riskiluokassa olevia oppilaita. Ennustavina tekijöinä ei enää tällöin pidetä oppilaan tehtävästä saamia pisteitä vaan oppilaan vastauksia (rastin kohtaa alakysymyksessä).

Vaikka päästäänkin 94 % tarkkuuteen ennustettaessa, milloin oppilas *ei kuulu syrjäytymisriskiin, syrjäytyneiden* ennustaminen onnistuu vain 60 % tarkkuudella. On kuitenkin muistettava, että arvaamalla kyseiset prosentit ovat 90 % ja 10 %. Eli syrjäytymisriskissä olevan ennustaminen onnistuu 6 kertaa paremmin kuin arvaamalla. Mallista saadaan kymmenen parhaiten luokittelua ennustavaa vastausta.

Ne olivat:

- Mikä on sukupuolesi?
- Mitä pidät koulunkäynnistä tällä hetkellä?
- Miten tärkeänä vanhempasi pitävät koulunkäyntiäsi?
- Mitä mieltä olet luokastasi?
 - (1) Luokassani on hyvä työrauha
 - (2) Luokan ilmapiiri on sellainen, että uskallan vapaasti ilmaista mielipiteeni
- Oletko saanut tukea ja apua hyvinvointiisi seuraavilta koulun aikuisilta tämän lukuvuoden aikana?
 - (1) Terveystieteiltä
 - (3) Psykologilta
 - (5) Opettajalta
- Asutko molempien vanhempien kanssa yhteisessä kodissa?

Koska tulokset perustuvat itse tehtyyn pisteytykseen, mallin tarkkuudeksi saadaan odotetusti lähes 99 %.

The screenshot shows a hierarchical menu structure with three levels: 'Results for output field yli10pros', 'Overall Results', and 'Comparing \$R-yli10pros with yli10pros'. Below the third level is a table with the following data:

Correct	3 768	98,74%
Wrong	48	1,26%
Total	3 816	

KUVIO 13. Päättelypuu mallin tulokset kaikilla kysymyksillä

Päättelypuu mallin avulla voitaisiin helpottaa toistuvien kyselyjen tekemistä, jolloin vähemmällä aineistolla voitaisiin laskea syrjäytymisriskiin yhtä tarkasti.

Vähemmällä vaihtoehdoilla päättelyä kokeiltiin valitsemalla targetiksi flag-tyyppi: kuuluuko riskiryhmään. Tähän ryhmään kuulumisen on laskettu valitsemalla tietyt tehtävät ja pisteyttämällä niiden alikysymykset. Päättelypuu ei käytä pistetietoa vaan sille on annettu vain flag-tieto targetista sekä kohdat, joihin oppilas on laittanut rastin.

Tekemällä sama malli parhaalla 5 kysymyksellä saadaan tarkkuudeksi **77%**. Tekemällä sama malli parhaalla 10 kysymyksellä saadaan tarkkuudeksi **85 %**.

The screenshot shows a hierarchical menu structure with two levels: 'Results for output field yli10pros' and 'Comparing \$R-yli10pros with yli10pros'. Below the second level is a table with the following data:

Correct	3 251	85,19%
Wrong	565	14,81%
Total	3 816	

KUVIO 14. Päättelypuussa jopa 10 kysymyksellä saadaan tarkkuudeksi 85 %

Valitsemalla malliin vain 14 alikysymystä, jotka olivat kohdassa 'Top decision rules' saadaan tarkkuudeksi **96** %. Nämä vaihtoehdot eivät kuulu ryhmään: merkittävät kysymykset.

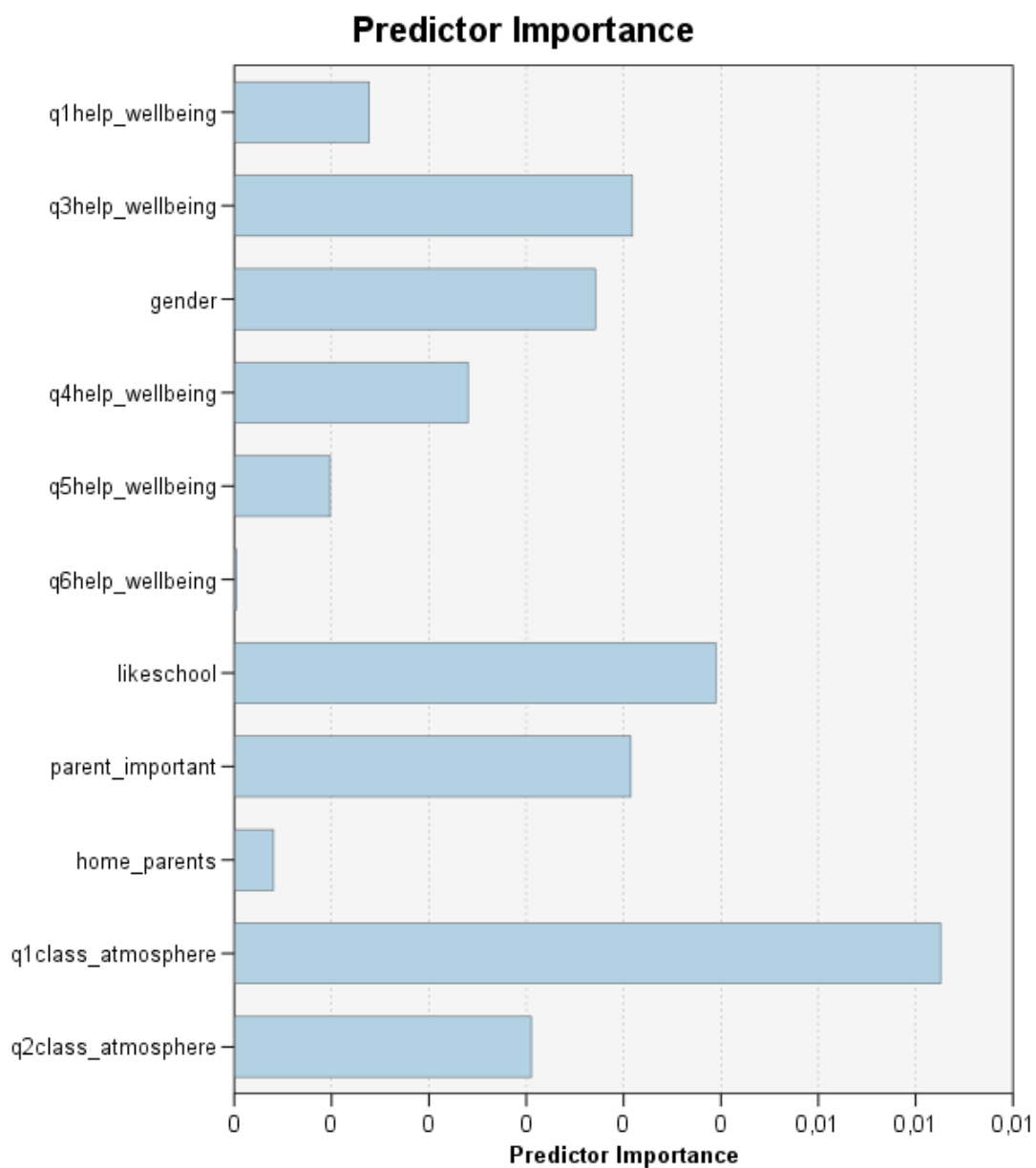
Field	Measurem...	Values	...	Check	
# q4swemwbs	Nominal	1.0,2.0,3.0,4....		None	Input
# q5swemwbs	Nominal	1.0,2.0,3.0,4....		None	Input
# q1accept_dr...	Nominal	1.0,2.0,3.0		None	Input
# q1lifesatisfa...	Nominal	1.0,2.0,3.0,4....		None	Input
# q6lifesatisfa...	Nominal	1.0,2.0,3.0,4....		None	Input
# q7lifesatisfa...	Nominal	1.0,2.0,3.0,4....		None	Input
# future_plans	Nominal	1.0,2.0,3.0,4....		None	Input
# q6gad7	Nominal	1.0,2.0,3.0,4.0		None	Input
# q7gad7	Nominal	1.0,2.0,3.0,4.0		None	Input
# q8gad7	Nominal	1.0,2.0,3.0,4.0		None	Input
# q9gad7	Nominal	1.0,2.0,3.0,4.0		None	Input
# q2sitting	Nominal	1.0,2.0,3.0,4....		None	Input
# q3drugs	Nominal	1.0,2.0,3.0,4.0		None	Input
# lifecontentm...	Nominal	1.0,2.0,3.0,4....		None	Input
# yli10pros	Flag	1/0		None	Target

KUVIO 15. Ennusteisiin käytettävät input-tietojen sekä ennustettavan target-tiedon määrittely

Results for output field yli10pros		
Comparing \$R-yli10pros with yli10pros		
Correct	3 676	96,33%
Wrong	140	3,67%
Total	3 816	

KUVIO 16. Päätelypuu, jossa 14 'top decision rules' – kysymystä. (96 % tarkkuus)

Näissä malleissa on huomioitava, että ne toimivat vain Keski-Suomen 2017 kyselyyn. Muiden läänien tietoja ei saatu kyseiseltä vuodelta. Niillä olisi voitu testata mallin toimivuutta toisessa läänissä.



The top 10 inputs are shown.

KUVIO 17. Modeler –ohjelman löytämät merkittävimmät kymmenen kysymystä, jotka ennustavat syrjäytymisen riskiä THL:n kyselyaineistossa

Liite 2: Kouluterveyskyselyn 2017 tulokset



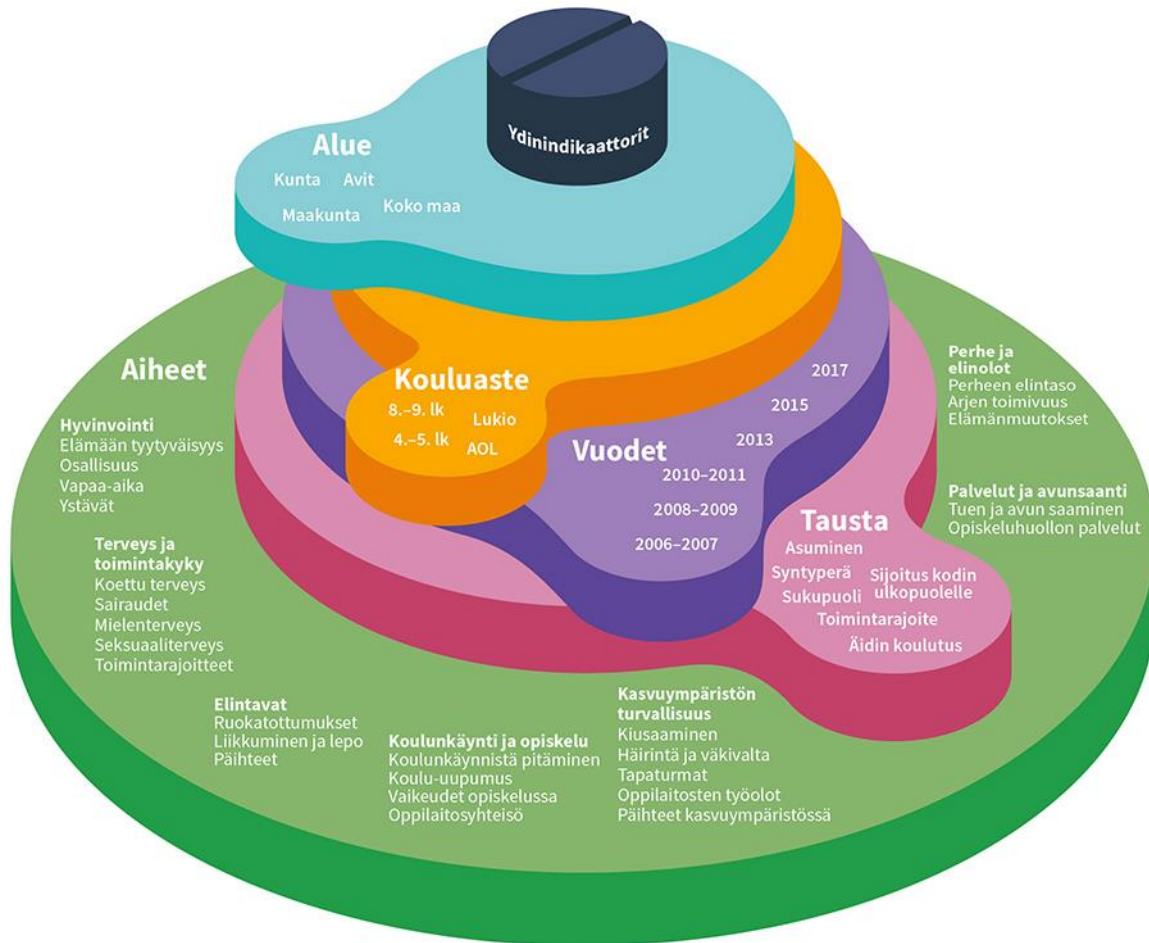
THL.FI AIHEET LAPSET, NUORET JA PERHEET TUTKIMUSTULOKSIA

Tutkimustuloksia

THL kerää tietoa lasten, nuorten ja perheiden hyvinvoinnista ja palveluista väestökyselyillä ja tilastotiedonkeruilla. Tietoa saadaan myös asiakastietojärjestelmistä ja rekistereistä.

Kouluterveyskyselyn 2017 tulokset

Kouluterveyskyselyn tulokset ovat saatavilla sähköisessä tulospalvelussa. Katso alla olevasta kuviosta tulosten tarkastelumahdollisuudet. **Pääset tuloksiin vasemmalla olevasta valikosta tai sivun lopusta.**



Kuvio: Kouluterveyskyselyn tulosten tarkastelumahdollisuudet sähköisessä tulospalvelussa

Kouluterveyskysely toteutetaan parittomina vuosina koko maassa. Oppilaat vastaavat kyselyyn koulupäivän aikana opettajan valvonnassa. Kysely on vapaaehtoinen ja nimetön.

Kyselyyn osallistujat

- perusopetuksen 4. ja 5. vuosiluokkien oppilaat ja heidän huoltajansa (ensimmäistä kertaa vuonna 2017)
- perusopetuksen 8. ja 9. vuosiluokkien oppilaat
- lukioiden 1. ja 2. vuoden opiskelijat
- ammatillisten oppilaitosten 1. ja 2. vuoden opiskelijat.

Kyselyn perustulokset ovat maksuttomia kunnille ja koulutuksen järjestäjille.

Tulosten tarkastelumahdollisuudet

- Perustulokset sisältävät noin 250 indikaattoria, jotka on jaoteltu seitsemään aihealueeseen. Osasta tuloksia on seurantatietoa vuodesta 2006 alkaen, jolloin voit tarkastella trendituloksia.
- Tulokset ovat saatavilla alueittain: koko maa, avi-alue, maakunta ja kunta. Voit verrata esimerkiksi oman maakuntasi tilannetta toisiin maakuntiin tai koko maan tilanteeseen.
- Vuoden 2017 tuloksia voidaan tarkastella väestöryhmittäin.
- Oppilaitoskohtaiset perustulokset on suojattu salasanalla.

Informaatioteknologian tiedekunnan julkaisuja
No. 62/2018

ISBN 978-951-39-7574-6 (verkkoj.)
ISSN 2323-5004