



**UNIVERSITY  
OF OULU**

TIETO- JA SÄHKÖTEKNIIKAN TIEDEKUNTA

**Sara Nikula**

## **EDA-signaalin automaattinen virheensuodatus**

Kandidaatintyö  
Tietotekniikan tutkinto-ohjelma  
12/2019

**Nikula S. (2019) EDA-signaalin automaattinen virheensuodatus.** Oulun yliopisto, tietotekniikan tutkinto-ohjelma. Kandidaatintyö, 27 s.

## TIIVISTELMÄ

Tässä kandidaatintyössä tutkin ihmiskehosta mitatun sähkönjohtavuussignaalin eli EDA-signaalin automaattista virheensuodatusta. EDA-signaali on kytköksissä ihmisen parasympaattiseen hermostoon, joten sen avulla voidaan tulkita ihmisen tunnetiloja ja vireystilaa.

Tavallinen EDA-signaali etenee muutaman sekunnin mittaisissa nousevissa ja laskevissa aalloissa. Tässä tutkielmassa käyttämäni data on kerätty Empatica E4 -rannekkeella, joka mittaa EDA-signaalin lisäksi myös kiihtyvyydataa. Koska mittalaite asetetaan koehenkilön käteen, se on altis henkilön liikkeiden aiheuttamille virheille. Nämä virheet saattavat näkyä signaalissa sekä nollaan pudonneena signaalina että epätavallisen jyrkkinä piikkeinä.

EDA-signaalin keräämistä, sen ongelmia ja virheensuodatusta on käsitelty kirjallisuudessa ja siihen on esitetty erilaisia ratkaisuja. Tässä tutkielmassa käyn läpi aiheeseen liittyviä artikkeleita ja esittelen muutamia ratkaisuehdotuksia virhepiikkien eli niin sanottujen artefaktien automaattiseksi tunnistamiseksi. Esittelen myös oman ratkaisuni, joka perustuu signaalin pätkien luokitteluun normaaleiksi tai todennäköisesti virheellisiksi muodon ja amplitudin perusteella. Leimoja on neljää eri tyyppiä: 0 – normaalia signaalia; 1 – signaali on pudonnut alle minimiamplitudin; 2 – signaalissa on liikkumisesta johtuva jyrkkä piikki; ja 3 – signaalissa on tuntemattomasta syystä johtuva jyrkkä piikki. Luokittelussa käytetään apuna EDA-signaalin kanssa samanaikaisesti kerättyä kiihtyvyydataa. Analysoitu ja luokiteltu signaali esitetään graafisesti väriä leiman mukaan vaihtavalla viivalla.

Esittelen lyhyesti ohjelmani tuottamia tuloksia ja arvioin niiden oikeellisuutta. Lopuksi esitän ehdotuksia siitä, kuinka työtä voisi jatkaa, ja arvioin ratkaisematta jääneitä ongelmia.

**Avainsanat: EDA, signaalin suodatus, artefakti**

**Nikula, S. (2019) Automatic error detection in EDA signal.** University of Oulu, Degree Programme in Computer Science and Engineering. Bachelor's Thesis, 27 p.

## **ABSTRACT**

This is a bachelor's thesis on automatic detection of artefacts in EDA signal (*Electrodermal Activity*). EDA measures electrical characteristics of human skin. These characteristics are connected to the parasympathetic nerve system and thus they reflect emotions and arousal of the person being monitored.

My data is collected using *Empatica E4 wristband* and it consists of several hours of EDA and acceleration data from three different days. Wristbands are liable to errors due to rapid movement of the test person, and these errors can be seen in the signal as steep peaks or very low amplitude level.

Filtering these errors from EDA signal has been discussed in several articles, many of which also provide solutions to this problem. In this thesis I present some of these articles. I also suggest my own solution, which is a Python program that measures amplitude and derivative of the signal. Concurrently gathered acceleration data is used in determining whether the signal is erroneous due to rapid movements of the test person. Every sample of the signal is being labeled based on these properties. There are four different labels: 0 – normal signal; 1 – amplitude of the signal is too low; 2 – fast change in signal level due to movement of the test person; and 3 – fast change in signal level due to unknown reason. The program creates metadata, which contains information about proportions of the different labels. Lastly, labeled signal is presented as a multicoloured line.

After discussing my methodology I present and assess the results yielded by my program. In the last section I discuss unsolved problems and propose possible themes for future work.

**Key words: EDA, signal filtering, artifacts**

## **ALKUSANAT**

Tämä on Oulun yliopiston tietotekniikan tutkinto-ohjelmaan kuuluva kandidaatintyö. Kiitän avusta kandidaatintyöni ohjaajaa Pekka Siirtolaa. Kiitokset myös aineistoon liittyvissä kysymyksissä auttaneille Henna Mönttiselle ja Satu Tammiselle.

## SISÄLLYSLUETTELO

TIIVISTELMÄ.....	2
ABSTRACT.....	3
ALKUSANAT.....	4
SISÄLLYSLUETTELO.....	5
LYHENTEIDEN JA MERKKIEN SELITYKSET.....	6
1. JOHDANTO.....	7
2. EDA-SIGNAALI JA SEN MITTAAMINEN.....	8
2.1. Fysiologinen perusta ja käyttötarkoitukset.....	8
2.2. Aiempia EDA-signaalia hyödyntäviä tutkimuksia.....	9
2.3. EDA-signaalin mittauksessa ilmenevät virheet.....	10
2.3.1. Erilaisia tapoja tunnistaa EDA-signaalin virheitä.....	10
2.4. Empatica-mittarin esittely.....	11
2.4.1. Empatica E4 wristbandin teknisiä tietoja.....	12
3. DATA.....	14
4. METODOLOGIA.....	15
5. TULOKSET.....	19
6. YHTEENVETO JA POHDINTA.....	25
LÄHTEET.....	26

## LYHENTEIDEN JA MERKKIEN SELITYKSET

EDA	ihon sähkönjohtavuutta mittaava signaali. Lyhenne tulee sanoista <i>Electrodermal Activity</i> . [1]
g	putoamiskiihtyvyyden yksikkö, joka tarkoittaa vapaasti putoavan kappaleen kiihtyvyyttä. Putoamiskiihtyvyys maapallon pinnalla on noin 9,81 metriä neliösekunnissa. [2]
Hertsi (Hz)	taajuuden yksikkö, joka tarkoittaa kertoja sekunnissa. [3]
SCL	ihon sähkönjohtavuussignaalin yleinen pidemmän aikavälin taso. Lyhenne tulee sanoista <i>Skin Conductance Level</i> . [4]
SCR	ihon sähkönjohtavuussignaalin nopeasti muuttuva osa. Lyhenne tulee sanoista <i>Skin Conductance Response</i> . [4]
Siemens	sähkönjohtavuutta kuvaava johdannaisyksikkö ja resistanssia kuvaavan ohmin käänteisyksikkö. <i>Mikrosiemens</i> on siemensin miljoonasosa. [3]

## 1. JOHDANTO

Tämä kandidaatintyö käsittelee ihon sähkönjohtavuustasoa kuvaavan EDA-signaalin automaattista käsittelyä. EDA tulee sanoista *Electrodermal Activity*. Käyttämäni data on kerätty ranteessa pidettävän mittalaitteen avulla. Kyseinen ranneke on tarkoitettu pidettäväksi kädessä normaalin elämän aikana, jolloin esimerkiksi rannekkeen pitäjän liikkuminen saattaa aiheuttaa signaaliin erilaisia virheitä. Signaali saattaa joko pudota hetkellisesti noltaan tai siihen voi tulla joko ylös- tai alaspäin suuntautuva jyrkkä häiriöpiikki. Joskus tällaiset virheelliset signaalin osat on mahdollista poistaa silmämääräisesti jälkikäteen. Aina se ei kuitenkaan ole mahdollista. Esimerkiksi jatkuvasti sensoridatasta oppivat mallit voivat häiriintyä huonolaatuisesta datasta.

Tässä kandidaatintyössä tutkin aiheeseen liittyvää kirjallisuutta ja kehittelen itse Python-ohjelmointikielellä eräänlaisen ratkaisun EDA-signaalin virheellisten osuuksien automaattiseksi tunnistamiseksi. Työssä käyttämäni data on kolmelta eri päivältä peräisin oleva testihenkilön EDA-signaali sekä samanaikaisesti kerätty kiihtyvyydata. Molemmat on kerätty Empatica E4 -rannekkeella testihenkilön ranteesta. Data sisältää sekä hyvälaatuista signaalia että virheellisiä osuuksia.

Ratkaisuni tunnistaa sekä signaalin putoamisen kynnsarvon alapuolelle että epäilyttävän jyrkät nousut ja laskut signaalin tasossa. Huonolaatuisiksi havaittu signaalin osa merkitään virheen tyyppiä vastaavalla leimalla. Samaan aikaan luodaan metadatan, josta käy ilmi, paljonko näytteitä on jouduttu poistamaan virheellisinä ja minkä laatuista virheestä oli kyse. Analysoinnin apuna käytetään EDA-datan kanssa samaan aikaan kerättyä kiihtyvyydetietoa, joka kertoo, onko rannekkeen pitäjä liikkunut voimakkaasti virhepiikin aikana. Koehenkilön rivakka liikehdintä antaa EDA-signaaliin ilmestyneille jyrkille piikeille luonnollisen selityksen.

Dataa käsittelevä ohjelma visualisoi analyysin lopputuloksen. Eri tavoin luokitellut signaalin osat esitetään eri värisellä viivalla.



OULUN YLIOPISTO  
UNIVERSITY of OULU

Kuva 1. yliopiston logo.

## 2. EDA-SIGNAALI JA SEN MITTAAMINEN

### 2.1. Fysiologinen perusta ja käyttötarkoitukset

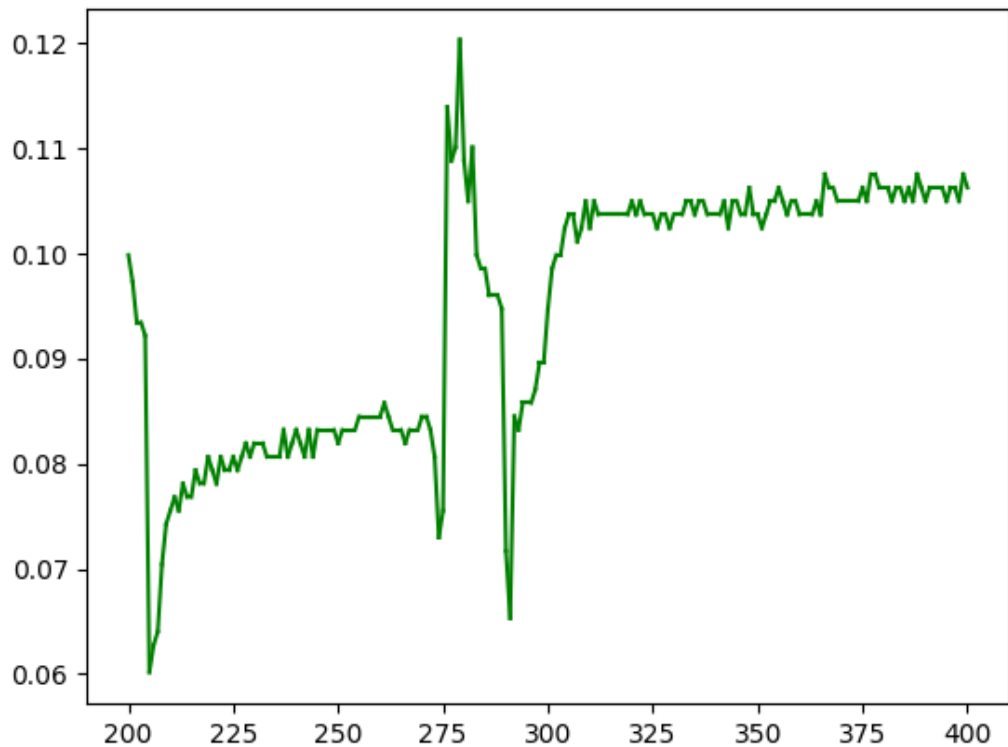
Ihmiskehossa on 2–3 miljoonaa hikirauhasta, eniten jalkapohjissa ja kämmenissä. Hikoilulla on tärkeä rooli ihmisen lämmönsäätelyssä. Hikoilua aiheuttavia tekijöitä on useita. Lämmöstä johtuva hikoilu aiheuttaa hien erityistä joka puolella kehoa. Pelosta ja ahdistuksesta johtuva hikoilu keskittyy pääasiassa kämmeniin, jalkapohjiin ja kainaloihin. Tulisten ruokien syömisestä aiheutuva hikoilu taas keskittyy kasvoille. [5 s. 167]

EDA on lyhenne sanoista *Electrodermal Activity*, ja se viittaa ihmisen ihon sähkönjohtavuuteen. Ihon sähkönjohtavuutta voidaan mitata kahdella eri tavalla. Endosomaattisessa mittauksessa mittaava laite ei käytä ulkoista virtaa apuna, vaan mittaa vain ihon itsensä aiheuttamat sähkönjohtavuusvaihtelut. Eksosomaattisessa mittaustyyliässä iholle johdetaan pieni määrä joko tasa- tai vaihtovirtaa, jonka avulla mitataan ihon resistanssia sekä sähkönjohtavuutta. [1 s. 1–3]

EDA-signaali voidaan edelleen jakaa tekijöihin. Termi SCR tulee sanoista *Skin Conductance Response* ja SCL sanoista *Skin Conductance Level*. SCL kuvaa sähkönjohtavuuden yleistä tasoa, joka vaihtelee pidemmällä aikavälillä. SCR taas kuvaa sähkönjohtavuussignaalin nopeasti muuttuvaa komponenttia. [4]

SCR-taso vaihtelee vuoroin nousten ja vuoroin laskien yleensä 1–5 sekunnin mittaisissa sykleissä. Näistä muutaman sekunnin mittaisista SCR-tapahtumista voidaan tutkia muun muassa seuraavia ominaisuuksia: amplitudi eli nousun korkeus, nousuaika eli aika siitä, kun signaalin taso alkaa nousta, siihen, kun se saavuttaa maksiminsa, sekä laskuaika eli aika, joka kuluu, kun signaalin taso laskee huipun tasosta esimerkiksi puoleen. [1 s. 2–3]





Kuva 2. Tutkielmassa käyttämäni datan perusteella piirrettyä EDA-signaalia. Kuvassa näkyy, että signaalin hitaammin muuttuvan perustason lisäksi siinä on lyhyempiä, nopeammin nousevia ja laskevia aaltoja. Y-akselin yksikkönä on mikrosiemens ja x-akselin yksikkönä näytteen indeksi. Näytteistystaajuus on 4 Hz, eli kuvassa näkyvä 200 näytteen jakso on kerätty 50 sekunnin aikana.

## 2.2. Aiempia EDA-signaalia hyödyntäviä tutkimuksia

EDA-signaali mittaa sähkönjohtavuutta, joka on yhteydessä hikoiluun [6, kappale 3.12]. EDA-signaalin mittausta ja analysointia voidaan käyttää apuna monenlaisissa tutkimuksissa, jotka liittyvät ihmisen stressitasoon ja vireystilaan. Aiemmissa tutkimuksissa sen avulla on tutkittu mm. konferenssi- ja kokouksien osallistujien ja heidän yleisönsä innostuneisuutta [4], hätäkeskuspäivystäjien stressitasoja [7] sekä vanhusten mielentiloja [8]. Ihmisen mielentilaa voidaan mitata myös esimerkiksi kasvonilmeiden tai puheen perusteella. EDA-signaaliin perustuvan monitoroinnin etuna näihin tekniikoihin nähden on se, että autonomisesta hermostosta riippuva EDA-signaali ei vaihtele yksilöiden välillä yhtä voimakkaasti kuin puhetapa tai kasvonpiirteet. Lisäksi EDA-signaali ei ole ihmisen hallittavissa ja peiteltävissä yhtä hyvin kuin puhe ja ilmeet. Monet ihmiset myös pitävät videointia ja nauhoitusta epämiellyttävänä. [8] Ranteessa pidettävä mittalaite ei aiheuta tällaista ongelmaa.

EDA-signaali on käyttökelpoista dataa myös lääketieteellisissä yhteyksissä, ja sen analyysin avulla voidaan saada tietoa esimerkiksi migreeni- ja epilepsia-kohtausten

luonteesta. Eräs tutkimus osoittaa rajuja muutoksia EDA-signaalissa epilepsiakohtausten aikana. Seitsemän epilepsiapotilaan EDA-signaalia tarkkailtiin päivin ja öin rannekkeen avulla. Datan analysointi osoitti signaalin amplitudin nousevan huomattavan korkealle henkilön kärsiessä kohtauksista. [9] Migreenikohtauksen lähestyminen taas voidaan mahdollisesti tunnistaa etukäteen henkilön nukkuessa kerätyn EDA-signaalin avulla. Useat EDA-signaalin vaihtelusta lasketut ominaisuudet yhdessä muodostavat dataa, jonka avulla migreenikohtauksen ennustaminen näyttää mahdolliselta. [10]

EDA-signaalin analysointi tarjoaa siis mielenkiintoisia näkymiä monenlaisiin tutkimuskohteisiin. Kuten edellä esitellyssä migreeniä käsittelevässä tutkimuksessa todetaan, henkilökohtaiset erot EDA-signaalin käyttäytymisessä voivat kuitenkin olla suuria, joten henkilökohtaisella tasolla käyttökelpoisten laskentakaavojen yleistäminen ei ole aivan yksinkertaista [10]. Vaikuttaakin siltä, että EDA-signaaliin pohjautuvat mallit toimivat varmimmin henkilökohtaisesti räätälöityinä.

### 2.3. EDA-signaalin mittauksessa ilmenevät virheet

Tavallinen EDA-signaali noudattelee tiettyä kaavaa. Se kestää yleensä ainakin yhden sekunnin ja on tietyn muotoinen: sillä on tietyt nousu- ja laskukulmat sekä tietty amplitudi. [11] Tästä muodosta selvästi eroavat signaalit kielivät virheistä datankeruuvaiheessa: signaali saattaa pudota noltaan tai se saattaa heittelehtiä arvaamattomasti ylös alas.

Varsin yleinen selittävä syy signaalissa oleville virheille on se, että ranteessa pidettävä laite ei pysy täydellisesti paikallaan henkilön liikkuaessa. Esimerkiksi kävelyn aikana mitatussa signaalissa ilmenevät äkilliset tason muutokset johtuvat siitä, että laitteen kontaktipinta heittelehtii edestakaisin, eivätkä suinkaan siitä, että henkilön sähkönjohtavuustaso tosiasiaassa muuttuisi näin äkillisesti. Asiaa koeolosuhteissa testanneet tutkijat ehdottavatkin, että kovan fyysisen ponnistelun jaksot pitäisi sulkea datasta kokonaan pois. Kevyempi liikuskelu ei kuitenkaan vielä tarvele signaalia. [12]

#### 2.3.1. Erilaisia tapoja tunnistaa EDA-signaalin virheitä

EDA-signaalin virhepiikit voidaan poimia ja poistaa joko silmämääräisesti tai koneellisesti. Myös yleisiä signaalin suodattimia voidaan käyttää. Hernandezin, Morrisin ja Picardin artikkelissa *Call Center Stress Recognition with Person-Specific Models* pyrittiin luomaan malli, jolla voitaisiin EDA-signaalin avulla tulkita henkilön stressitasoa. Tässä tutkimuksessa EDA-signaalin kohinaa sekä liikkeestä aiheutuneita artefakteja poistettiin eksponentiaalisella tasoituksella. [7] Tämän kandidaatintyön aiheeseen kyseinen tekniikka ei kuitenkaan sovi, koska kerätty data halutaan virheettömiltä osin säilyttää mahdollisimman alkuperäisenä. Tämä johtuu siitä, ettei ole vielä tiedossa, millaiset EDA-signaalissa esiintyvät ominaisuudet saattavat olla kiinnostavia. Virhepiikit on kuitenkin joka tapauksessa suodatettava pois, joten muita menetelmiä tarvitaan.

Toisessa tutkimuksessa [11] EDA-datan kanssa tekemisissä olleet asiantuntijat luokittelivat datanäytteitä silmämääräisesti joko puhtaisiin tai todennäköisesti virheellisiin näytteisiin. Asiantuntijoiden luokittelu perustui siihen, että tyypillinen SCR-tapahtumaa kuvaava signaali on tietyn pituinen ja muotoinen. Normaalissa SCR-tapahtumassa amplitudi alkaa huipun jälkeen laskea nopeasti, mutta pian alamäki loivenee. Tästä muodosta voimakkaasti eroavat tapahtumat viittaavat häiriöihin. Nopea EDA-arvon muutos samanaikaisesti henkilön liikkeen kanssa luokiteltiin tutkimuksessa virheeksi. Lisäksi negatiiviset arvot luokiteltiin virheellisiksi, koska signaalin arvo ei voi koskaan laskea nollan alapuolelle. Myös yli viiden prosentin kvantisointivirhe johti näytteen hylkäämiseen. Näitä virheellisiksi tai virheettömiksi luokiteltuja näytteitä käytettiin artefakteja tunnistavan koneoppimismallin opetusdatana.

Yksi ongelma koneellisessa virheiden tunnistamisessa on se, että EDA-signaalit ovat yksilöllisiä. Virhepiikkejä poistavan mallin olisi siis hyvä ottaa huomioon henkilöstä kerätyn datan yksilölliset piirteet ja poistaa virheet nimenomaan tähän nähden. Koneoppimiseen perustuvien algoritmien heikkoutena onkin se, että niiden opetusdatana on usein käytettävä eri henkilöistä kerättyä dataa kuin kenen dataa niillä loppujen lopuksi on tarkoitus analysoida [7].

#### 2.4. Empatica-mittarin esittely

Tässä tutkielmassa käytettävä data on kerätty laitteella nimeltä *Empatica E4 wristband*. Sitä valmistaa yritys nimeltä Empatica, joka kehittää puettavaa teknologiaa terveydenhoidollisiin tarkoituksiin. [13] Sivuillaan yritys luonnehtii Empatica E4 -ranneketta mukavaksi ja tarkaksi tavaksi monitoroida reaaliaikaisia fysiologisia signaaleja [14].

Empatica E4 -ranneke on langaton reaaliaikaiseen tiedonkeruuseen tarkoitettu laite, jossa on neljä sensoria: ihon lämpöä mittaava lämpömittari, kolmella akselilla kiihtyvyyttä mittaava kiihtyvyyttämittari, veritilavuuden muutoksia havainnoiva fotopletysmografi mittari sekä ihon sähkönjohtavuutta mittaava EDA-mittari. EDA-signaalin on tarkoitus antaa tietoa henkilön stressi- ja vireystasosta.

Ranneke mittaa sähkönjohtavuutta kahden hopeoidun elektrodin avulla. Ranneke asetetaan käteen niin, että elektrodit sijoittuvat ranteen kämmenen puolelle. Laite toimii eksosomaattisesti eli se lähettää pienen sähkövirran iholle ja mittaa näin sähkönjohtavuuden. Sähkövirran määrä on kuitenkin niin pieni, että se on käyttäjälle täysin turvallinen. Laitetta voidaan käyttää joko koti- tai sairaalaympäristössä. Terveystieteellisissä sovelluksissa tarkoitus on, että potilas pitää ranneketta ja terveydenhuollon ammattilainen analysoi laitteen keräämän datan. Toisaalta käyttäjä voi kerätä dataa myös omaan käyttöön, kuten tämän tutkielman datan kanssa on tehty. Laite painaa 40 grammaa. [15]

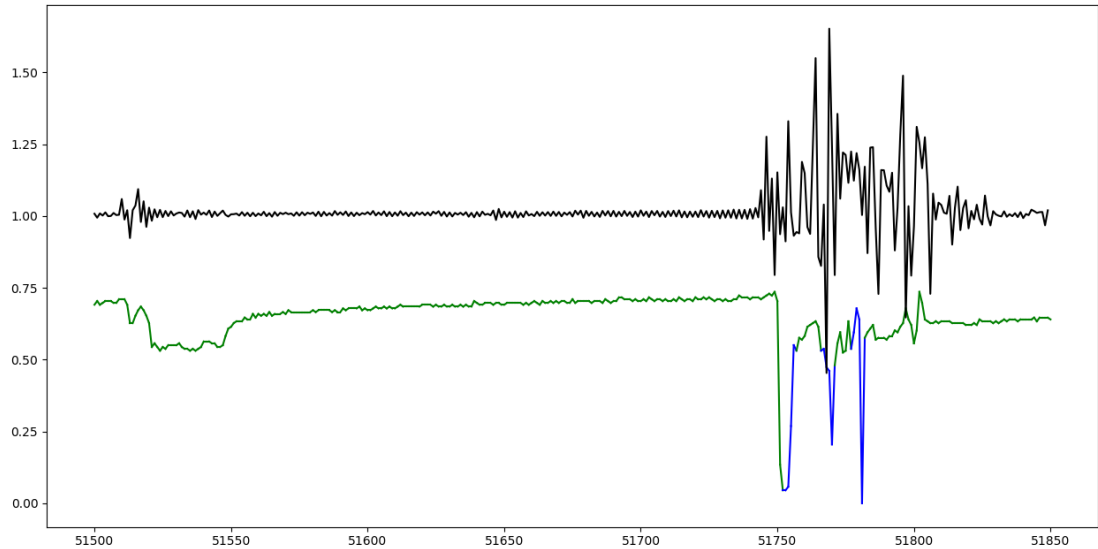
Empatica-ranneketta saa käyttää sekä sisällä että ulkotiloissa. Manuaali ei kuitenkaan suosittele sen käyttämistä kylmemmässä kuin -10 celsiusasteen lämpötilassa. Ranneke on roiskeenkestävä, mutta sitä ei tule upottaa veteen. Laite voi joko lähettää keräämänsä datan reaaliaikaisesti toiselle laitteelle, kuten esimerkiksi kännykkään, tai tallentaa datan omaan muistiinsa. Laitteen mukana tulevalla ohjelmistolla tämä tallennettu data voidaan siirtää muille laitteille. [15]

Manuaalissa neuvotaan asentamaan laite ranteeseen niin tiukasti, ettei se pääse heilumaan. Laitetta ei ole kuitenkaan syytä kiristää niin tiukalle, että se on epämiellyttävän tuntuinen. Vaikka monissa tutkimuksissa on pidetty parempana vaihtoehtona asentaa ranneke ei-dominanttiin käteen liikkeen aiheuttaman häiriön minimoimiseksi, ei Empatica-rannekkeen manuaali anna ohjetta siitä, kumpaan käteen ranneke tulisi mieluummin asentaa. Tämä johtuu siitä, että vaikka dominantissa kädessä pidettävä ranneke onkin herkempi liikkeen aiheuttamille häiriöille, tämän käden EDA-signaali saattaa olla stressitilanteissa vahvempi. Joskus EDA-tapahtumat saattavat olla jopa vain yksipuoleisia. Näiden syiden takia jää käyttäjän valittavaksi, kumpaan käteen hän kiinnittää rannekkeen. [15]

#### **2.4.1. Empatica E4 wristbandin teknisiä tietoja**

Rannekkeen EDA-sensorin näytteistystaajuus on 4 Hz, eikä sitä voi muuttaa. Näytteiden resoluutio on noin 900 pikosiemensä ja mittaushaarukka 0,01 – 100 mikrosiemensä.

Kiihtyvyyssmittarin käyttämä mittayksikkö on 1/64 g. Sen näytteistystaajuus on 32 Hz, eikä sekään ole muutettavissa. Kiihtyvyyssmittari kerää kiihtyvyydataa kolmella akselilla: ylös–alas, oikealle–vasemmalle sekä eteen–taakse. Kiihtyvyyssmittarin haarukka on oletusarvoisesti 2 g, mutta sen voi asettaa myös 4/8 g haarukalle. Resoluutio on kahdeksan bittiä valitulla haarukalla, eli mitä pienempää mittaushaarukkaa käytetään, sitä tarkemmin liikettä pystytään mittaamaan. Lisäksi laitteessa on kello, jonka avulla eri sensorien mittaamia dataa voidaan vertailla ajanhetkittäin. Laitteen elinajaksi ennustetaan noin kaksi vuotta oikein käytettynä. Hopeiset elektrodit vaativat kuitenkin vaihtamista joidenkin kuukausien välein. [15]



Kuva 3. Kuvassa näkyy rannekkeen pitäjän liikkumisen vaikutus EDA-signaaliin. Piikit EDA-signaalissa (vihreällä ja sinisellä) sijoittuvat usein kiihtyvyydsdatasta (mustalla) mitattujen piikkien yhteyteen, mikä kertoo siitä, että henkilön liikkuminen on vaikuttanut EDA-signaaliin. EDA-signaali on havainnollisuuden vuoksi kerrottu viidellä, ja sen yksikkönä on mikrosiemens. Kiihtyvyydsdatan yksikkönä on g.

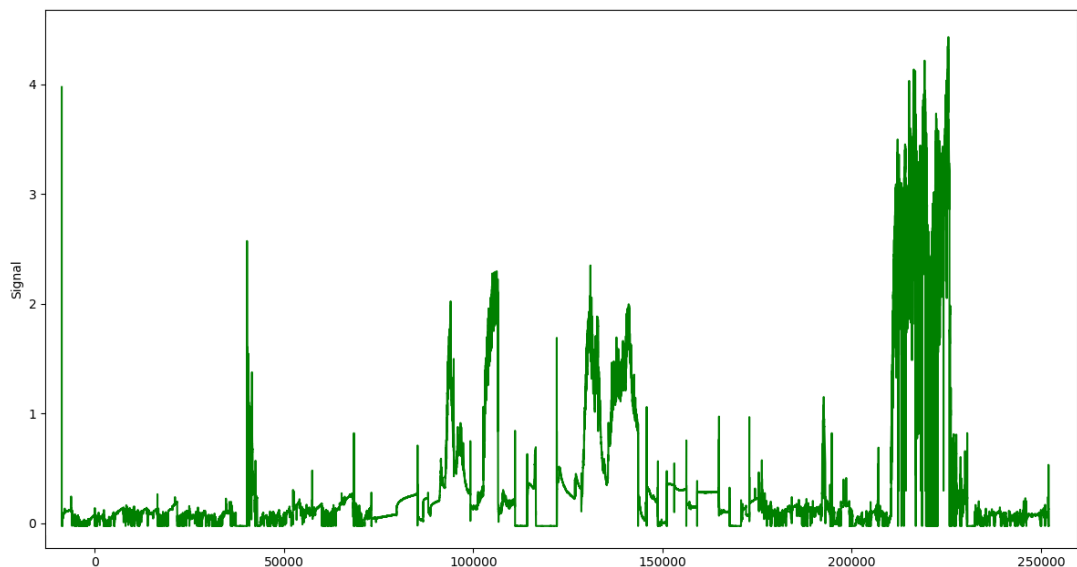
Koska EDA-signaalia monitoroivat rannekkeet on tarkoitettu pidettäväksi kädessä yhtäjaksoisesti pidempiä aikoja, on selvää, että näihin tarkasteluajankausiin tulee sisältymään myös tavanomaista arkista liikuskelua, ehkä jopa urheilusuorituksiakin. Sen takia on tärkeää pystyä tunnistamaan signaalin virheelliset osuudet, jotta ne eivät pilaa koko datan analyysia.

### 3. DATA

Tämän tutkielman pohjana käytetty data on mitattu kappaleessa 2.4 esitellyllä *Empatica E4 wristband* -rannekeella. Data on mitattu kolmelta eri päivältä ranneketta pitävän henkilön eläessä normaalia elämäänsä. Data sisältää siis esimerkiksi nukkumista ja tavanomaista elämään kuuluvaa liikkumista. Data ei kuitenkaan sisällä suoraa tietoa siitä, milloin henkilö on esimerkiksi urheillut tai nukkunut, vaan tämä on pääteltävä pääasiassa kiihtyvyyssanan perusteella. Tämä tuo oman epävarmuustekijänsä datan tulkintaan.

Data sisältää jokaiselta kolmelta mittauspäivältä samanaikaisesti mitatut EDA- ja kiihtyvyyssietodatat. Koska kiihtyvyyssanan näytteistystaajuus on 32 Hz ja EDA-signaalin näytteistystaajuus 4 Hz, on kiihtyvyyssanaa kahdeksankertainen määrä EDA-dataan nähden. Yhteisen kellon ja aikaleimojen avulla datojen rinnakkainen vertailu on kuitenkin mahdollista.

Dataa on ensimmäiseltä näytteistyspäivältä yhteensä 18 tuntia 6 minuuttia, toiselta päivältä 23 tuntia 55 minuuttia ja kolmannelta päivältä 22 tuntia 16 minuuttia. EDA-data on tallennettu yksinkertaisen taulukon muodossa: toisessa sarakkeessa on aikaleima ja toisessa sarakkeessa kyseisenä hetkenä mitattu EDA-signaalin taso. Kiihtyvyys on mitattu kolmelta eri akselilta, joten kiihtyvyyssanataulukossa on aikaleiman lisäksi kolme eri lukua, omansa x-, y- ja z-akseleille.



Kuva 4. Koko 260672 näytettä sisältävä ensimmäisen päivän EDA-signaali kuvattuna vihreällä käyrällä.

## 4. METODOLOGIA

Tämän kandidaatintutkielman tavoitteena oli kehittää ohjelma, joka käy EDA-signaalin läpi ja merkitsee dataan virheelliset kohdat. Tarkoituksena ei ollut, että ohjelma poistaa virheelliset kohdat, vaan se, että se tunnistaa virheelliset kohdat, analysoi virheen laadun ja lisää signaaliin virheestä kertovan leiman. Datan analysointivaiheessa virheellisiksi leimatut osat voidaan käyttötarkoituksen mukaan myöhemmin joko poistaa tai pitää.

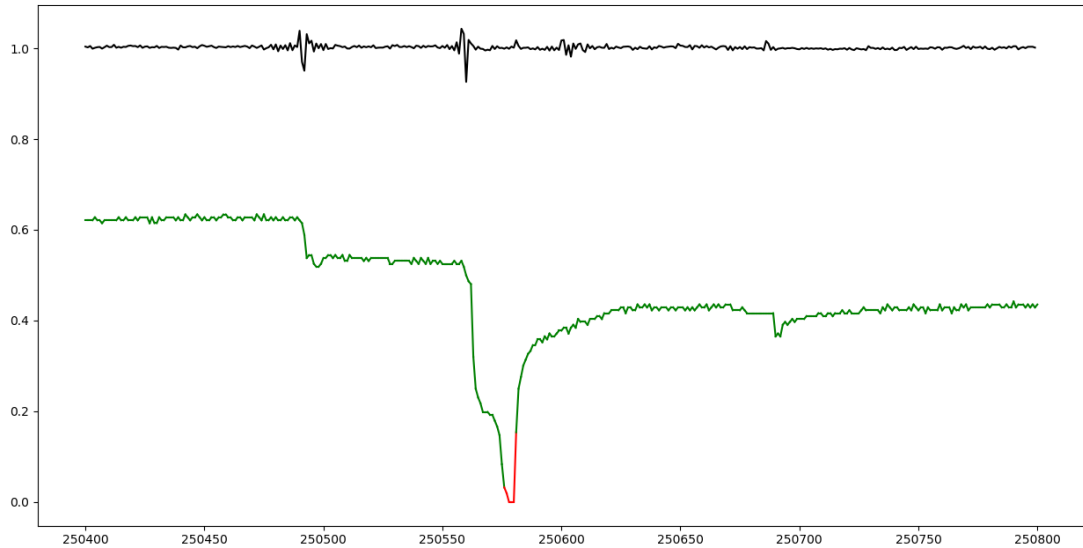
Toteutin ohjelman Python-ohjelmointikielellä. Käytin apuna Pythonin matplotlib- ja scipy-kirjastoja. Matplotlib on kirjasto, jonka avulla Python-ohjelmalla voi piirtää graafisia kuvia ja käyriä [16]. Käytin sitä ohjelmani luomien tulosten tarkasteluun ja havainnollistamiseen. Scipy-kirjastosta käytin sen signaalinkäsittelyosiota, joka sisältää useita valmiita funktioita ja suodattimia signaalien käsittelyyn [17]. En kuitenkaan käyttänyt tämän kirjaston sisältämiä suodattimia vaan ainoastaan funktiota, joka etsii signaalin huiput. Tämä nopeutti hieman ohjelmani toteuttamista.

Virheellisten signaalin osien leimaamisen toteutin ylläpitämällä EDA-signaalin kanssa saman pituista listaa, joka sisältää kyseistä indeksiä vastaavan näytepisteen leiman. Matplotlib-kirjaston avulla signaalin eri osat voitiin myöhemmin piirtää eri väreillä riippuen siitä, millaisella leimalla ne on merkitty. Luokittelin leimat neljään eri tyyppiin, joita listassa merkittiin numeroilla 0–3. Leimat ja niiden selitykset on listattu alla olevassa taulukossa.

Taulukko 1. Virheentunnistusohjelmani käyttämät leimat ja niiden selitykset.

Leima	Selitys
0	Normaalialia signaalia
1	Signaalin taso on pudonnut alle 0,01 mikrosiemensin
2	Signaalissa on liikkumisesta johtuva virhepiikki
3	Signaalissa on tuntemattomasta syystä johtuva mahdollinen virhepiikki

Signaalin palasten analysointi ja leimaus perustuu ohjelmassani kahteen eri kriteeriin. Ensimmäinen kriteeri on se, että signaalin arvon tulee olla yli 0,01 mikrosiemensia. Tämä johtuu siitä, että Empatica E4:n manuaali ilmoittaa tämän arvon minimiarvoksi, joka voidaan mitata. Kyseistä kynnyksarvoa pienempiä mittaustuloksia ei siis voida ottaa mukaan analyysiin. Toinen kriteeri liittyy signaalista löytyneiden huippujen nousukulman jyrkkyyteen eli derivaattaan. Koska signaalin yleinen taso vaihtelee pidemmällä aikavälillä, signaalia ei voida luokitella virheelliseksi vain sen hetkellisen tason perusteella, vaan virhe havaitaan nimenomaan signaalin tason muutosnopeuden perusteella.



Kuva 5. Signaalin arvo on äkisti pudonnut alle kynnsarvon, ja tämä sekä sitä seuraavat näytepisteet on merkitty leimalla 1. Osuus näkyy kuvassa punaisella. EDA-signaalin arvo on kerrottu havainnollisuuden vuoksi viidellä, ja sen yksikkönä on mikrosiemens. Mustalla piirretyn kiihtyvyydsdatan yksikkönä on g.

Ensimmäisen kriteerin perusteella virheelliset osat poistettiin ohjelmassani niin, että koko signaali käytiin läpi ja kaikki alle 0,01 mikrosiemensin minimiarvon jäävät näytepisteet merkittiin leimalla 1.

Toisen kriteerin perusteella virheelliset signaalin pätkät sijoittuivat signaalissa esiintyvien paikallisten huippujen yhteyteen, joten ennen derivaatan laskentaa etsin kaikki signaalin huiput scipy-kirjaston *find\_peaks*-funktion avulla. Kyseinen funktio palauttaa listan, joka sisältää kaikkien signaalista löytyneiden huippujen indekset. Huipuiksi luetaan kaikki kohdat, jotka eroavat ympäröivästä signaalista, eli kyse on paikallisista maksimipisteistä. Syötin signaalin tähän funktioon sekä alkuperäisessä muodossaan että ylösalaisin peilattuna, jolloin sekä ylös- että alaspäin suuntautuvat huiput tulivat mukaan käsittelyyn.

Koska EDA-datan luonteeseen kuuluu, että se nousee ja laskee noin sekunnin mittaisilla intervalleilla, paikallisia huippuja löytyi runsaasti. Ensimmäiseltä mittauspäivältä kerätystä 260672 näytepisteestä löytyi yhteensä 49101 huippua. Toiselta mittauspäivältä kerätystä 344570 näytteen mittaisesta aineistosta löytyi 64980 huippua ja kolmannelta mittauspäivältä kerätystä 320846 näytteen mittaisesta aineistosta 72362 huippua. Keskimäärin aineistossa oli siis huippuja noin viiden näytteen välein. Tämä sopii hyvin kappaleessa 2.1 kuvattuun EDA-signaalin määritelmään.

Kutakin huippua edeltäneen nousun jyrkkyys laskettiin laskemalla huipun ja sitä edeltäneen näytepisteiden amplitudin välinen erotus. Pyrin selvittämään normaalin EDA-signaalin nousukulmia etsimällä datasta silmämääräisesti 15 muutamien sekuntien pituisia näytettä normaalin näköistä EDA-signaalia ja laskemalla sen sisältämien huippujen derivaattoja. Vertailun vuoksi etsin aineistosta myös 15 selvää



henkilön liikkumisesta johtuvaa virhepiikkiä. Näiden virheiden löytämistä varten piirsin kiihtyvyyssdataa ja EDA-signaalin samaan kuvaan Pythonin matplotlib-kirjastolla. Tätä varten kiihtyvyyssdataa näytteistystaajuutta piti madaltaa Pythonin `downsample`-funktioilla, koska se oli alun perin mitattu 32 Hz näytteistystaajuudella ja näin ollen dataa oli kahdeksankertainen määrä EDA-dataan verrattuna.

Löytämiäni normaaleja EDA-huippuja edeltävät nousukulmat olivat välillä 0,00128–0,13 mikrosiemensiä yhden näytteistysvälin aikana eli 0,25 sekunnissa. Liikkumisen vaikutuksesta tarveltyneiden virhepiikkien derivaatat asettuivat välille 0,116–0,84 mikrosiemensiä 0,25 sekunnissa. Päädyin asettamaan kynnsarvoksi 0,05 mikrosiemensin muutoksen kahden näytteen välissä. Jos muutos on tätä suurempi, ohjelman tulee tarkistaa kiihtyvyyssdatasta, onko henkilö liikkunut. Kynnsarvon yläpuolelle jää tällä kynnsarvolla myös normaalista signaalista mitattuja huippuja. Tämä ei kuitenkaan ole ongelma muuta kuin korkeintaan ohjelman suoritusajan näkökulmasta, koska signaalia ei suoraan luokitella virheelliseksi pelkän derivaatan perusteella vaan vasta kiihtyvyyssdataa analysoinnin jälkeen.

Kiihtyvyyssdataa analysointia varten laskin erilliseen taulukkoon kokonaiskiihtyvyysslistan, joka sisältää kiihtyvyyssdatasta mitattujen x-, y- ja z-kiihtyvyyksien neliöiden summan neliöjuuren. Tällöin henkilön kokonaisliike tulee huomioiduksi riippumatta siitä, mihin suuntaan hän on liikkunut. Liitin tämän kokonaiskiihtyvyysslistan derivaattatietoon niin, että kokonaiskiihtyvyysslistasta etsitään EDA-datasta löytyneen huipun ajankohtaa vastaava kohta ja tätä ajanhetkeä edeltävän sekunnin ajalta etsitään kaikki huiput kokonaiskiihtyvyydessä. Koska kiihtyvyyssdata on mitattu 32 Hz taajuudella ja EDA-data 4 Hz taajuudella, tuli EDA-datan indeksi siis kertoa kahdeksalla, jotta vastaava kohta kiihtyvyyssdatasta löytyi.

Koska tarkoitukseni oli saada ohjelma määrittelemään, mitkä EDA-signaalin jyrkät huiput johtuivat henkilön rivakasta liikkumisesta, oli rivakalle liikkumiselle määritettävä jokin raja-arvo. Tätä varten käytin samoja silmämääräisesti EDA-datasta eroteltuja normaaleja ja virheellisiä huippuja kuin äskeisessä EDA-signaalin derivaatan raja-arvon määrittelyssä. Etsin kokonaiskiihtyvyysslistan tiedot kyseisiä EDA-huippuja edeltävän sekunnin ajalta ja poimin näistä huiput. Huippujen korkeus kertoo, kuinka suuri kokonaiskiihtyvyys rannekkeen pitäjällä on enimmillään ollut kyseistä piikkiä edeltävän sekunnin aikana.

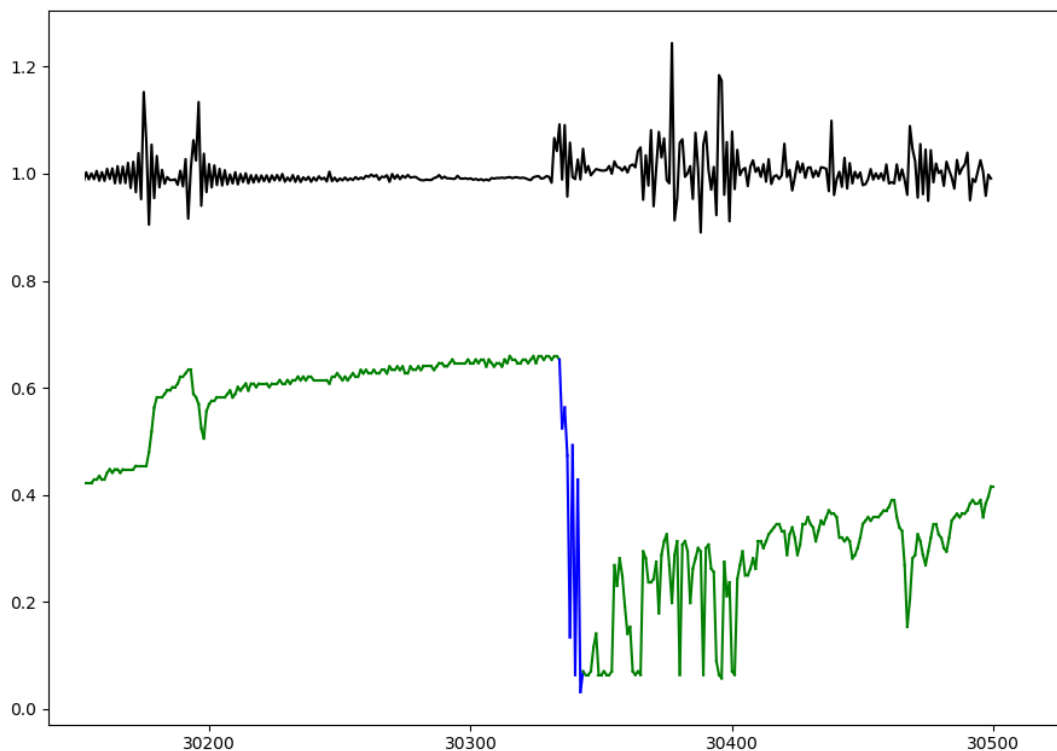
Normaaleja EDA-huippuja edeltävästä kokonaiskiihtyvyyssdatasta löytyneet huiput olivat korkeintaan  $64,75 * 1/64$  g, kun taas virhepiikkejä edelsi aina vähintään  $75,30 * 1/64$  g:n korkuinen piikki. Päädyin asettamaan normaalin liikkeen raja-arvoksi 70, eli jos henkilön kokonaiskiihtyvyys oli käynyt yli  $70 * 1/64$  g:ssä, luokiteltiin sen jälkeen tuleva jyrkkä EDA-piikki liikkumisesta johtuvaksi virheeksi.

Datasta löytyneet huiput voitiin siis jakaa nousukulman ja edeltävän kokonaiskiihtyvyyssdataa perusteella kolmeen eri ryhmään: nousukulma on alle kynnsarvon, nousukulma on yli kynnsarvon mutta kokonaiskiihtyvyys alle kynnsarvon, tai nousukulma ja kokonaiskiihtyvyys ovat molemmat yli kynnsarvon. Tällä periaatteella ohjelma kävi läpi koko signaalin ja merkitsi huiput niille kuuluvilla leimoilla. Jos nousukulma oli alle raja-arvon, kyseisen signaalinpätkän luokitus jätettiin silleen eli se oli normaalia signaalia vastaava 0. Jos taas nousukulma oli yli

raja-arvon 0,05, arvioitiin seuraavaksi kyseistä ajanhetkeä edeltänyt kiihtyvyydata. Jos kiihtyvyydatasta löytyi raja-arvoa 70 suurempia kokonaiskiihtyvyyssiikkejä, luokiteltiin kyseinen EDA-huippu liikkeestä johtuvaksi virhepiikiksi ja se sai leiman 2. Jos kokonaiskiihtyvyy pysyi raja-arvon alla, luokiteltiin EDA-huippu tuntemattomasta syystä johtuvaksi mahdolliseksi virheeksi ja se sai leiman 3.

Jos EDA-huippu oli luokiteltu virheelliseksi, sitä ympäröivä signaali merkittiin virheleimalla sekunnin verran taaksepäin ja yhden näytepisteen eli 0,25 sekuntia eteenpäin. Tämän tavan valitsin toisaalta sen takia, että värin vaihtuessa vain pieneksi hetkeksi kerrallaan kuvio tulee levottoman näköiseksi ja hankalaksi tulkita, ja toisaalta sen takia, että virhepiikkejä välittömästi ympäröivää dataa ei mielestäni voida pitää luotettavana.

Signaalin analysoinnin ja luokittelun jälkeen ohjelmani laski, kuinka paljon kutakin leimaa kyseisen signaalin näytepisteille olin annettu. Suhteelliset osuudet palautettiin prosenttimuotoisina. Silmämääräistä tarkastelua varten signaali vielä piirrettiin niin, että eri leimoilla merkityt osiot piirrettiin eri värisinä: leima 0 vihreällä, leima 1 punaisella, leima 2 sinisellä ja leima 3 magentalla.



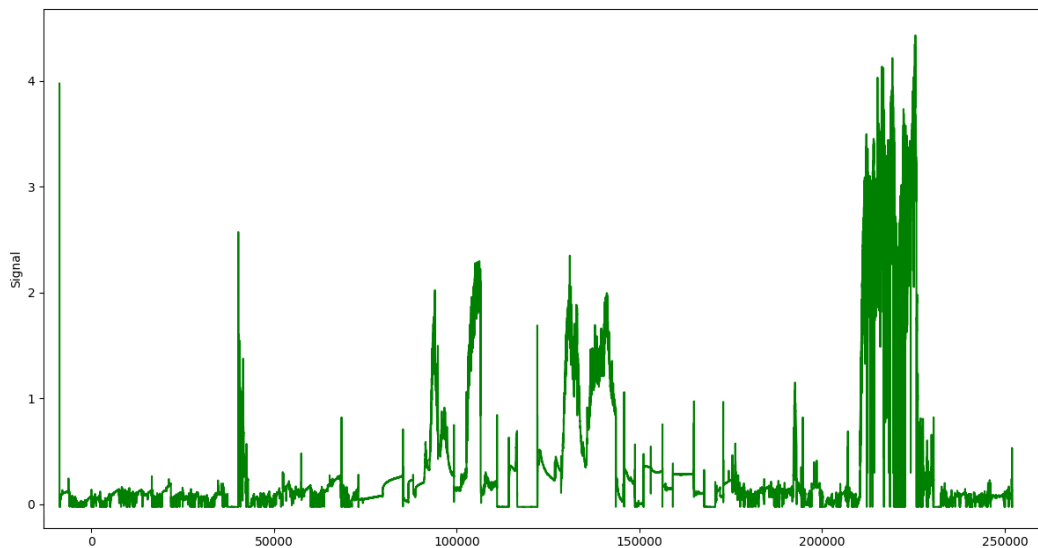
Kuva 6. Normaalista EDA-signaalista (vihreällä) sekä liikkumiseen liittyviä teräviä virhepiikkejä (sinisellä). Samanaikainen kiihtyvyydata on piirretty mustalla. Kuvassa näkyy myös muita EDA-signaalin nousuja ja laskuja, joiden derivaatta ei kuitenkaan ole ylittänyt kynnyksarvoa, joten ne on merkitty normaaleiksi ja piirretty vihreällä viivalla. EDA-signaali on kerrottu havainnollisuuden vuoksi viidellä, ja sen yksikkönä on mikrosiemens. Kiihtyvyydataan yksikkönä on g. X-akselilla on EDA-näytteen järjestysnumero.

## 5. TULOKSET

Seuraavissa taulukoissa esitän ohjelmani luokittelman signaalin metadatan, josta selvää, kuinka suuri osuus signaalista on leimattu milläkin leimalla. Lisäksi taulukoiden alla näkyy koko kyseisen päivän EDA-signaali vihreällä piirrettynä.

Taulukko 2. Ensimmäisen päivän datan tulokset. Datan koko oli 260672 näytepistettä.

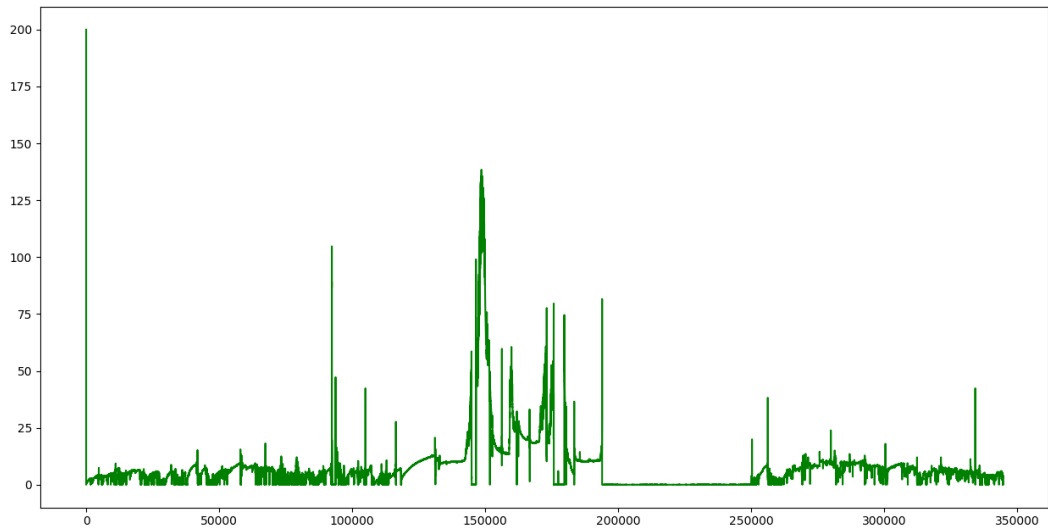
Leima	Selitys	Kappaletta	Prosenttia
0	Normaalia signaalia	229614	88,085410
1	Signaali alle minimiarvon	24439	9,375384
2	Liikkeestä johtuva virhe	6015	2,307498
3	Tuntematon mahdollinen virhe	604	0,231709



Kuva 7. Ensimmäisen päivän EDA-signaali.

Taulukko 3. Toisen päivän datan tulokset. Datan koko oli 344570 näytettä.

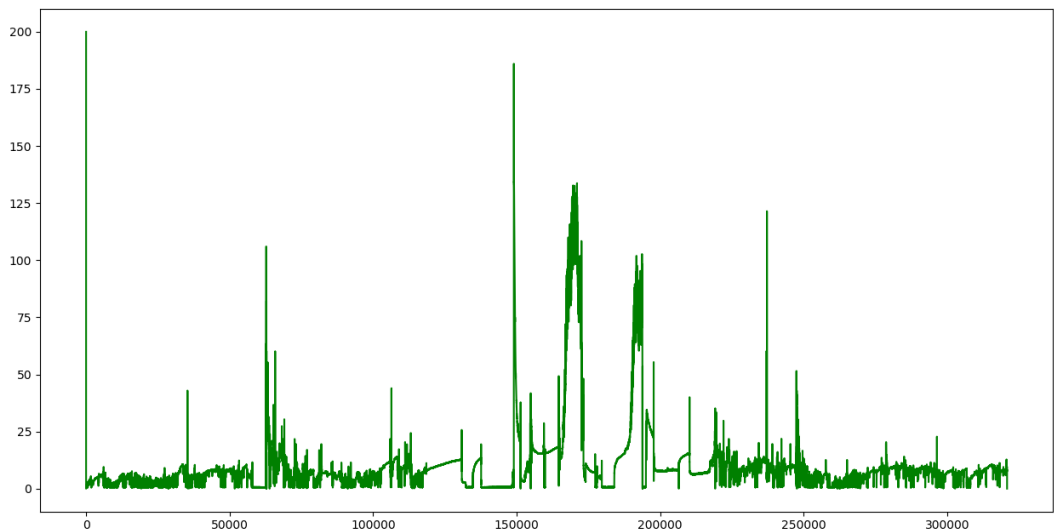
Leima	Selitys	Kappaletta	Prosenttia
0	Normaalia signaalia	269267	78,145805
1	Signaali alle minimiarvon	73537	21,341672
2	Liikkeestä johtuva virhe	1665	0,483211
3	Tuntematon mahdollinen virhe	101	0,029312



Kuva 8. Toisen päivän EDA-signaali.

Taulukko 4. Kolmannen päivän datan tulokset. Datan koko oli 320846 näytettä.

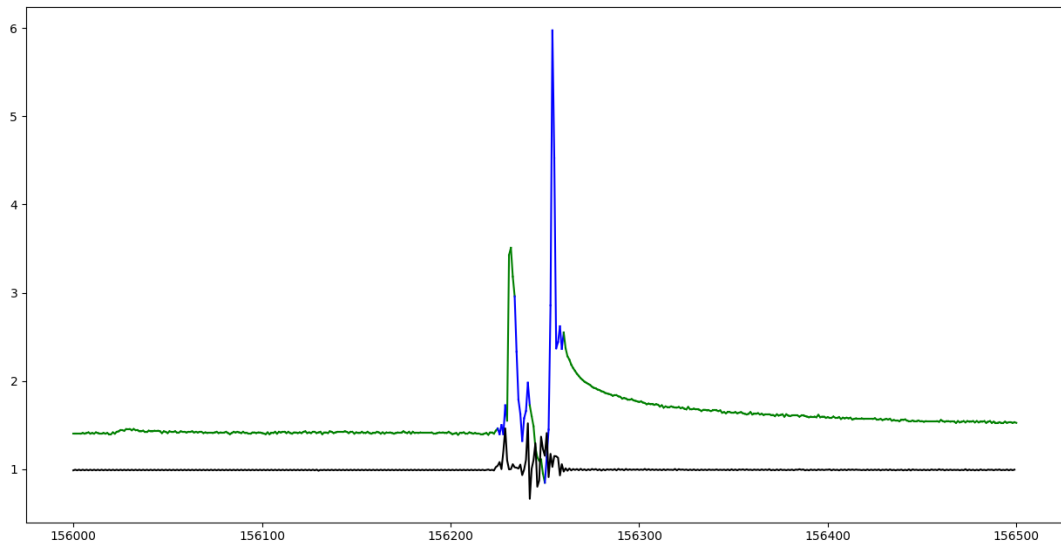
Leima	Selitys	Kappaletta	Prosenttia
0	Normaalia signaalia	316085	98,516111
1	Signaali alle minimiarvon	1730	0,539199
2	Liikkeestä johtuva virhe	2841	0,885472
3	Tuntematon mahdollinen virhe	190	0,059218



Kuva 9. Kolmannen päivän EDA-signaali.

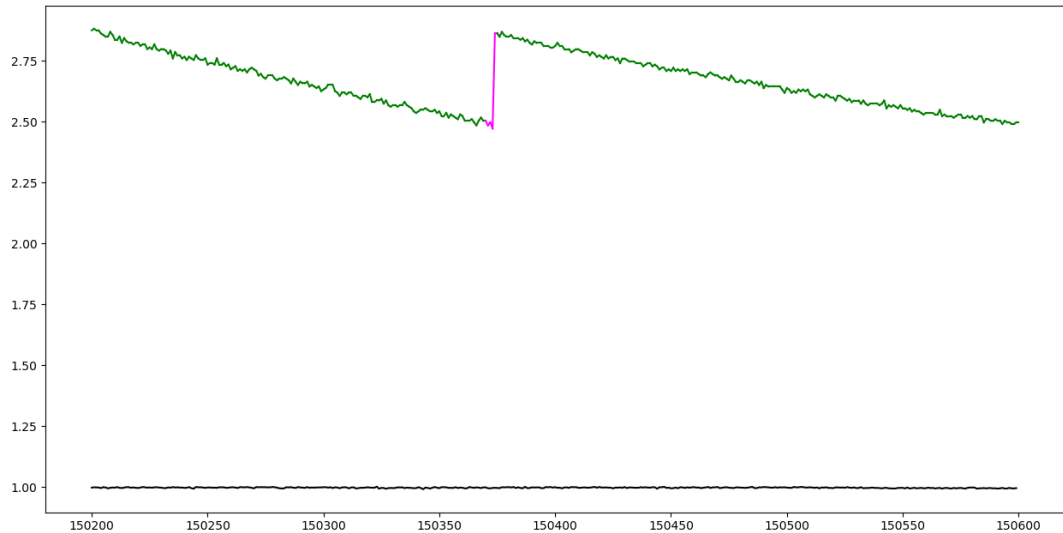
Analysoin kaikkien kolmen päivän datan kappaleessa 4 esitetyn metodin mukaisesti ja tarkastelin sitten luokituksia, jotka ohjelma oli antanut signaalin näytepisteille.

Kaikilta päiviltä kerätystä datasta suurin osa oli luokiteltu normaaliksi signaaliksi, kuten odotettavissa oli.



Kuva 10. Luokiteltu EDA-signaali (vihreällä ja sinisellä) ja kiihtyvyydata (mustalla) piirrettynä samaan kuvaan. EDA-signaali on kerrottu havainnollisuuden vuoksi viidellä, ja sen yksikkönä on mikrosiemens. Kiihtyvyydataan yksikkönä on g. Kuvassa näkyy, että henkilö on liikkunut, mikä on aiheuttanut epäilyttävän jyrkän piikin EDA-signaaliin. Piikki on merkitty leimalla 2 ja piirretty sinisellä. X-akselilla on EDA-näytteen järjestysnumero.

Ensimmäiseltä ja toiselta päivältä kerätystä datasta oli melko paljon kohtia, joissa signaali oli pudonnut kokonaan nolnaan: ensimmäisenä päivänä 9 % ja toisena päivänä 21 %. Kolmantena testipäivänä tällaista nolnaan pudonnutta signaalia oli vain 0,5 %. Liikkeen takia poistettua signaalia oli ensimmäisenä päivänä 2 %, toisena päivänä 0,48 % ja kolmantena päivänä 0,88 %.

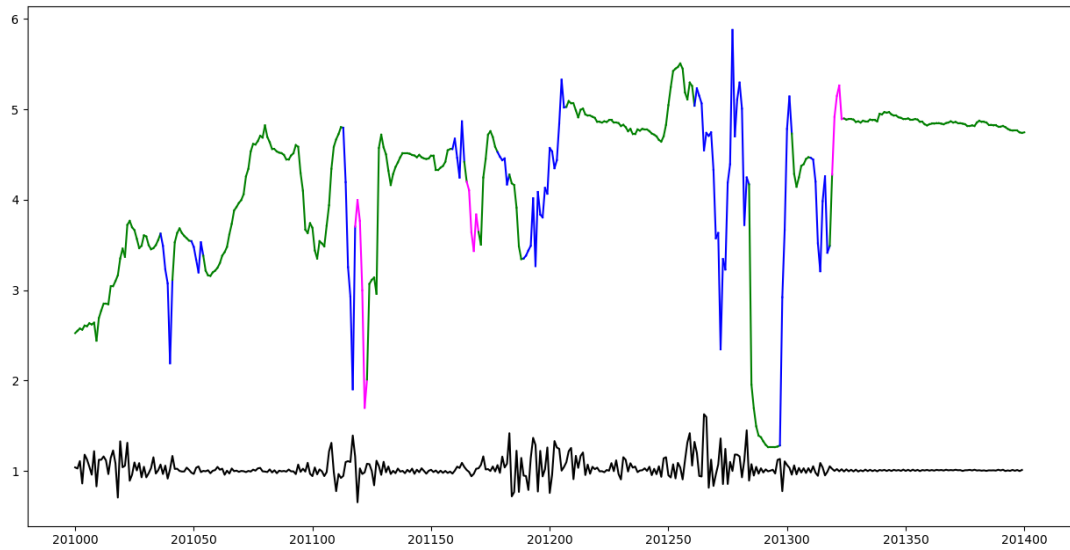


Kuva 11. Jyrkkä EDA-signaalin nousu, joka kuitenkin ei liity henkilön liikkumiseen. Epätavallinen nousu on merkitty magentalla. Kiihtyvyysdata on piirretty mustalla, eikä siinä näy muutoksia EDA-signaalin äkkinäisen muutoksen kohdalla. EDA-signaali on havainnollisuuden vuoksi kerrottu viidellä, ja sen yksikkönä on mikrosiemens. Kiihtyvyysdatan yksikkönä on g. X-akselilla on EDA-näytteen järjestysnumero.

Kaikkien päivien datoihin sisältyi myös pieniä määriä selittämättömiä jyrkkiä piikkejä. Käyrää tarkastellessani huomasin, että suuri osa näistä selittämättömiksi virheiksi luokitelluista piikeistä sijoittui suoraan liikkeestä johtuvien virheiden perään. Jos kiihtyvyysdataa laskettiin taaksepäin sekunnin sijaan puolentoista sekunnin ikkunalla, leimalla 3 merkittyjen alueiden määrä putosikin melko huomattavissa määrin. On siis hyvinkin mahdollista, että osalle tämän leiman saaneista näytteistä kuuluisi oikeasti leima 2. Ikkunan venytys entisestään kahden sekunnin mittaiseksi vähensi tämän ryhmän osuutta entisestään, mutta sekään ei poistanut kaikkia 3. tyyppin leimoja.

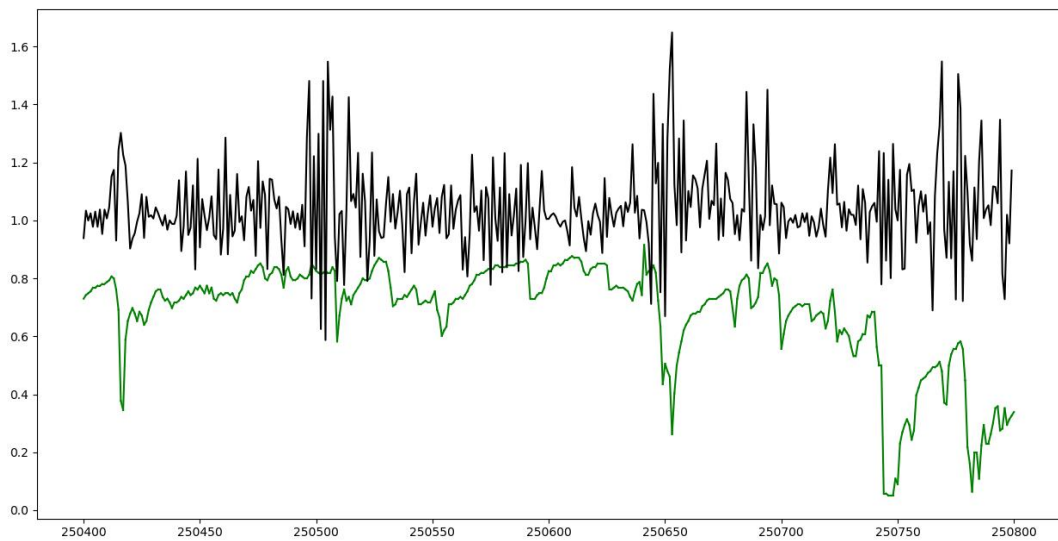
Kiihtyvyyssdatan haarukointi yhden sekunnin verran taaksepäin ei siis välttämättä riitä paljastamaan kyseistä piikkiä edeltänyttä liikettä. On mielenkiintoinen kysymys, voidaanko määritellä jotain raja-arvoa, miten pitkään henkilön liikkuminen voi vaikuttaa EDA-signaaliin.

Joissakin tilanteissa leima 3 näytti selittyvän sillä, että signaali oli ensin jonkin aikaa nollassa ja sen jälkeen nopeasti hypähti ylöspäin. Tämä johtunee siitä, että laitteen kontaktipinta on ollut jonkin aikaa kokonaan irti ihosta, ja kun sen jälleen koskettaessa ihoa mittaustuloksia taas saadaan, signaali näyttää nousevan äkisti ilman, että tähän liittyisi henkilön äkkinäistä liikettä.

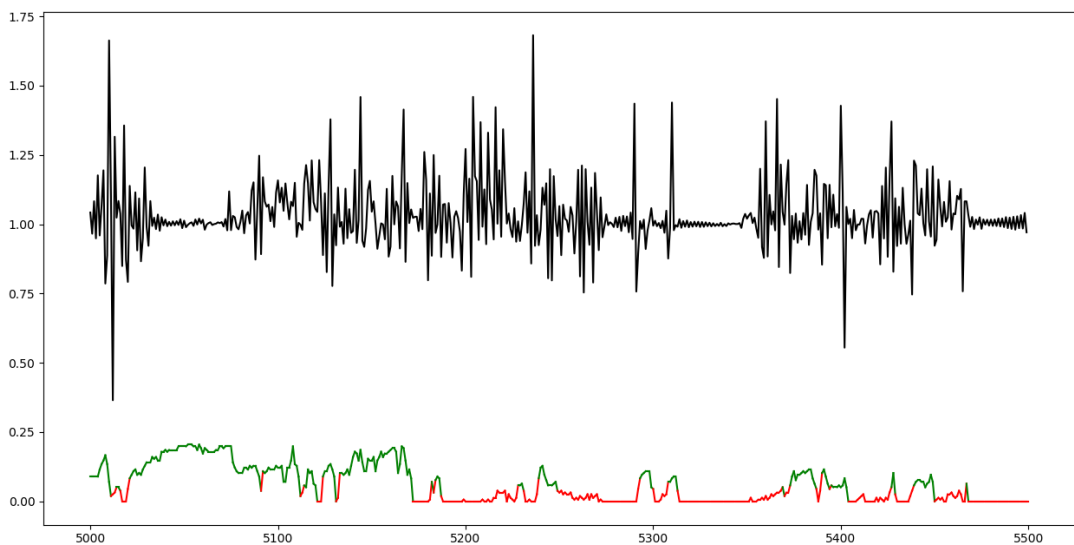


Kuva 12. Heittelehtivä EDA-signaali. Kuvassa näkyy, että magentalla (leima 3) merkityt signaalien osat sijoittuvat aikatasossa suoraan sinisellä merkittyjen osuuksien (leima 2) jälkeen. Vihreällä merkitty signaali on luokiteltu normaaliksi. Kiihtyvyydsdata on piirretty mustalla, ja sen yksikkönä on g. EDA-signaali on havainnollisuuden vuoksi kerrottu viidellä, ja sen yksikkönä on mikrosiemens. X-akselilla on EDA-näytteen järjestysnumero.

Jokaisen kolmen päivän signaalien ensimmäiset pisteet oli merkitty joko leimalla 2 tai 3, ja näissä kohdissa signaalissa olikin jyrkkä nousu nollassa. Tämä johtunee siitä, että laitetta on vasta asetettu testihenkilön käteen. Myös datojen loppupäässä signaalissa oli rajua ylösalaista heittelyä, mikä saattaa johtua siitä, että laitetta on otettu pois päältä tai irrotettu kädestä. Nämä osat olivat kuitenkin kestoaltaan varsin lyhyitä, eivätkä ne siten aiheuta huomattavaa vääristymää eri leimojen suhteisiin.



Kuva 13. EDA-signaali (vihreällä) ja kiihtyvyydata (mustalla) piirrettynä samaan kuvaan. EDA-signaali on kerrottu viidellä havainnollisuuden vuoksi, ja sen yksikkönä on mikrosiemens. Kiihtyvyydataan yksikkönä on g. Kuvassa näkyy, että henkilö on liikkunut, mutta liike ei ole ollut riittävän suurta, jotta se olisi tarvellyt EDA-signaalin. X-akselilla on EDA-näytteen järjestysnumero.



Kuva 14. EDA-signaali (vihreällä ja punaisella) on pudonnut lähelle nollaa. Todennäköisesti mittalaitteen kontaktipinta on ollut huonosti kiinni, ja sen takia mittaustulosta on saatu katkonaiseksi. Kiihtyvyydata on piirretty mustalla, ja sen yksikkönä on g. EDA-signaali on havainnollisuuden vuoksi kerrottu viidellä, ja sen yksikkönä on mikrosiemens. X-akselilla on EDA-näytteen järjestysnumero.



## 6. YHTEENVETO JA POHDINTA

Kehittelin tässä tutkielmassa ratkaisua ihmisiholta mitatun sähköjohtavuussignaalin eli EDA-signaalin virhepiikkien automaattiseksi suodattamiseksi. Käytössäni oli kolmelta eri päivältä mitattua EDA-dataa ja sen kanssa samaan aikaan mitattua kiihtyvyydataa. Ratkaisuni perustui toisaalta signaalille asetettuun minimiamplitudiin ja toisaalta signaalin nousukulman mittaamiseen. Analyysin apuna käytin EDA-signaalin kanssa samaan aikaan kerättyä kiihtyvyydataa.

Silmämääräisen tarkastelun perusteella oli mahdollista todeta, että EDA-signaalin esiintyvät epätavallisen jyrkät piikit olivat useimmiten yhteydessä rannekkeen pitäjän samanaikaiseen liikkumiseen. Tällaisten virhepiikkien nousukulma oli selvästi normaalista aaltoilevasta EDA-signaalista eroava, joten oli mahdollista asettaa kynnyсарvo, jonka yli menevät signaalin osat mitä todennäköisimmin ovat virheellisiä. Samoin kiihtyvyydataassa esiintyneet piikit olivat virheellisten EDA-piikkien kohdalla korkeampia kuin muualla dataassa.

Kehitin Python-ohjelmointikielellä ohjelman, joka analysoi signaalin sen amplitudin ja nousukulman perusteella ja luokitteli näin kunkin näytepisteen neljään luokkaan, joista yksi tarkoitti normaalia signaalia ja kolme muuta eri tavoin virheellistä signaalia. Ohjelma luokitteli suurimman osan käytössäni olevasta datasta normaaliksi. Suurin virheellisen signaalin luokka oli nollaan pudonnut signaali. Seuraavaksi eniten oli henkilön liikkumisesta johtuvia virheitä ja kaikkein vähiten jyrkkiä nousu- tai laskukulmia, jotka viittaisivat henkilön liikkumiseen, mutta joihin ei liittynyt epätavallisen korkeita piikkejä kiihtyvyydataassa sekunnin mittaisella takautuvalla aikaikkunalla.

Käyttämäni raja-arvot normaalille EDA-signaalin nousukulmalle sekä häiritsevän voimakkaalle liikkeelle olisi voinut määritellä toisinkin. Valitsemani kynnyсарvot olivat suurin piirtein havaitsemieni normaalien ja epänormaalien arvojen puolivälissä. Kynnyсарvon madallus olisi varmistanut kaikkien virheiden havaitsemisen, mutta toisaalta liian matala kynnyсар aiheuttaisi sen, että myös tosiasiallisesti normaalia signaalia leimattaisiin virheelliseksi. On mahdollista, että käyttämälläni kynnyсарvolla osa liikkumisesta johtuviksi virhepiikeiksi luokitelluista virheistä ei sitä tosiasiallisesti ole, mutta koska näiden virheiden osuus luokitellusta datasta on joka tapauksessa pienekkö, analysointikelpoista dataa tuskin menetetään liikaa.

Suurin tämän tutkielman tuloksiin liittyvä epävarmuus liittyi niihin signaalin osiin, jotka näyttivät liikkumisesta johtuvilta virhepiikeiltä, mutta joita edeltävä henkilön liikkuminen oli tapahtunut yli sekunti ennen kyseistä EDA-signaalin rajua muutosta. Toistaiseksi avoimeksi kysymykseksi jää, voidaanko määritellä jokin raja-arvo, miten pitkään henkilön liikkuminen voi vaikuttaa EDA-signaaliin. Tämän tutkimuksen perusteella näyttäisi siltä, että yli sekunnin verran.

Olisi myös mielenkiintoista vertailla tämän ratkaisun tulosten eroavaisuuksia muihin, kappaleessa 2.3.1 esitelyihin virheentunnistustapoihin. Kiinnostavaa olisi tutkia sekä sitä, suodattavatko ne virheitä yhtä hyvällä tarkkuudella, että sitä, kuinka ne toimivat eri tyyppisten virheiden kanssa suhteessa toisiinsa. Eri tekniikoiden yhdisteleminen on myös pohdinnan arvoinen kysymys.

## LÄHTEET

- [1] Boucsein, W (2012) *Electrodermal Activity*. Boston, MA, Springer.
- [2] Suomen Standardoimisliitto (2019) *SI-opas*. URL: [https://www.sfs.fi/files/8774/SI-opas\\_2019.pdf](https://www.sfs.fi/files/8774/SI-opas_2019.pdf). Viitattu 29.11.2019.
- [3] VTT (2019) *SI-järjestelmä*. URL: <https://www.mikes.fi/mittayksik%C3%B6t>. Viitattu 24.10.2019.
- [4] Gashi, S, Lascio, E & Santini, S (2019) Using Unobtrusive Wearable Sensors to Measure the Physiological Synchrony Between Presenters and Audience Members. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, volume 3, issue 1, March 2019. New York, USA, ACM.
- [5] Weller, Richard P.J.B, Hunter, Hamish J.A & Mann, Margaret W (2015) *Clinical dermatology*. Hoboken, New Jersey, John Wiley & Sons Inc.
- [6] Rook, A (2010) *Rook's textbook of dermatology*. Hoboken, New Jersey, Burns, Tony, FRCP, Wiley-Blackwell.
- [7] Hernandez, J, Morris, R & Picard, W (2011) Call Center Stress Recognition with Person-Specific Models. Teoksessa: D'Mello S, Graesser A, Schuller B & Martin JC (toim.) *Affective Computing and Intelligent Interaction. ACHI 2011. Lecture Notes in Computer Science*, vol 6974. Berlin, Heidelberg, Springer: 125–134.
- [8] Martinez-Rodrigo, A, Zangróniz, R, Pastor, J & Sokolova, M (2017) Arousal Level Classification of the Aging Adult from Electro-dermal Activity: From hardware Development to software architecture. *Pervasive and Mobile Computing*, Volume 34, January 2017: 46–59.
- [9] Poh, M. Z, Loddenkemper, T, Swenson, N. C, Goyal, S, Madsen, J. R, & Picard, R. W (2010) Continuous monitoring of electrodermal activity during epileptic seizures using a wearable sensor. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. Piscataway, NJ, IEEE.
- [10] Siirtola, P, Koskimäki, H, Mönttinen, H, & Röning, J (2018) Using sleep time data from wearable sensors for early detection of migraine attacks. *Sensors*, 18(5): 1374.
- [11] Taylor, S, Jacques, N, Chen, W, Fedor, S, Sano, A & Picard, R (2015) Automatic Identification of Artifacts in Electrodermal Activity Data. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Piscataway, NJ, IEEE.

- [12] Kappeler-Setz, C, Gravenhorst, F, Schumm, J, Arnrich, B & Tröster, G (2013) Towards long term monitoring of electrodermal activity in daily life. *Personal and Ubiquitous Computing* 17(2): 261–271.
- [13] Empatica (2019) About Us. URL: <https://e4.empatica.com/about>. Viitattu 14.10.
- [14] Empatica (2019) E4 wristband. URL: <https://e4.empatica.com/e4-wristband>. Viitattu 24.10.
- [15] Empatica (2018) Empatica E4 User Manual. URL: <https://empatica.app.box.com/v/E4-User-Manual>. Viitattu 14.10.2019.
- [16] Matplotlib (2019) URL: <https://matplotlib.org/>. Viitattu 14.10.
- [17] The SciPy Community (2019) Signal Processing (scipy.signal). URL: <https://docs.scipy.org/doc/scipy/reference/signal.html>. Viitattu 14.10.