



**UNIVERSITY
OF OULU**

TIETO- JA SÄHKÖTEKNIIKAN TIEDEKUNTA

Saara Laasonen

Tenttimenestyksen korrelaatio biosignaaleihin

Kandidaatintyö
Tietotekniikan tutkinto-ohjelma
Joulukuu 2024

Laasonen S. (2024) Tenttimenestyksen korrelaatio biosignaaleihin. Oulun yliopisto, tietotekniikan tutkinto-ohjelma. Kandidaatintyö, 33 s.

TIIVISTELMÄ

Tässä tutkinnossa käytetään Empatica E4 ranneketta. Empatica E4 sisältää 4 sensoria; PPG-sensori, EDA-sensori, 3-akselinen kiihtyvyysanturi sekä optinen lämpömittari. Käsitellään Empatica E4:n tuottamaa CSV-muotoista dataa, jonka avulla tutkitaan, löytyykö sensorien avulla saaduista signaaleista minkäänlaista näyttöä tenttimenestyksen kanssa.

Tässä tutkinnossa käytetään viittä eri signaalia; BVP, IBI, EDA, sydämen syke ja ihon lämpötila. Koska data on CSV-muotoista, niitä on helppo käsitellä Python-ohjelmointikielellä. Ensimmäinen vaativa tehtävä on yhdistää datat yhteen tiedostoon. Tässä tutkinnossa olen päättänyt käyttää ylinäytteenottoa, eli kaikki signaalit on venytetty pisimmän signaalin pituisiksi. Seuraavaksi data tulee rajata tenttiajan pituiseksi. Tämän jälkeen voidaan visualisoida signaaleita ja tutkia korreloivatko ne tenttitulosten kanssa.

Stressi on kiinnostava tutkimuksen aihe, koska sitä ei pysty täydellisellä varmuudella päättelemään biosignaalienkaan avulla. Koska ei voi saada varmoja tuloksia, vaatii stressin tutkiminen spekulointia, kun ei pysty itse testihenkilöiltä kysymään olivatko he stressaantuneita.

Avainsanat: Empatica E4, biosignaalit, biomarkkerit, stressi, puettava elektroniikka

Laasonen S. (2024) Correlation of exam performance with biosignals. University of Oulu, Degree Programme in Computer Science and Engineering. Bachelor's Thesis, 33 p.

ABSTRACT

In this thesis Empatica E4 will be used. Empatica E4 includes 4 different sensors: PPG sensor, EDA sensor, 3-axis accelerometer and optical thermometer. In this thesis we will process CSV-data produced by Empatica E4. With this data, we can research if there is any correlation with exam success.

In this thesis we will use five different signals: BVP, IBI, EDA, heart rate (HR) and skin temperature (TEMP). Because the data is in CSV format, it is easy to process with Python language. The first demanding task is to merge all the data into a single file. In this thesis, we have decided to use oversampling, i.e. all signals are stretched to the length of the longest signal. Next, the data should be limited to the length of the exam period. The signals can then be visualized and examined to see if the signals correlate with the exam results.

Stress is an interesting topic of research because it cannot be concluded with complete certainty even with the help of biosignals. Since you cannot get certain results, studying stress requires speculation when you cannot ask the test subjects themselves if they were stressed.

Key words: Empatica E4, biosignals, biomarkers, stress, wearable electronics

SISÄLLYSLUETTELO

TIIVISTELMÄ.....	2
ABSTRACT	3
SISÄLLYSLUETTELO	4
ALKULAUSE	5
LYHENTEIDEN JA MERKKIEN SELITYKSET.....	6
1. JOHDANTO.....	7
2. STRESSI JA BIOSIGNAALIT	8
2.1. Laitteen esittely	8
2.2. Fysiologiset biosignaalit.....	10
2.3. Kirjallisuuskatsaus.....	11
3. STRESSIN TUNNISTUSTA TENTTITILANTEESSA	13
3.1. Tietojen valinta.....	15
3.2. Tutkimuksen tuloksia	16
3.2.1. Sydämen syke.....	16
3.2.2. Ihon lämpötila	18
3.2.3. EDA-signaali	20
3.2.4. BVP-signaali	22
3.2.5. IBI-signaali.....	23
3.2.6. Regressioanalyysi ja R^2 -arvo	25
3.2.7. Lopputulos.....	27
4. MUIDEN TUTKIMUSTEN KÄSITTELY JA VERTAILU	28
4.1. Empatica E4	28
4.2. Muut menetelmät.....	28
4.3. Yhteenveto.....	29
5. YHTEENVETO	30
6. LÄHTEET	31

ALKULAUSE

Haluan kiittää Pekka Siirtolaa työni ohjauksesta ja tuesta. Haluan myös kiittää perhettäni, joka on tukenut minua opinnoissani.

Oulu, joulukuun 11. 2024

Saara Laasonen

LYHENTEIDEN JA MERKKIEN SELITYKSET

EDA	ihon sähkönjohtavuus
BVP	veren tilavuuspulssi
IBI	lyöntien välinen aikaväli
HR	sydämen syke
TEMP	ihon lämpötila
M1	välikoe 1
M2	välikoe 2
F	lopputentti
LDA	Lineaarinen erotteluanalyysi
CT	Central Time Zone (USA)
ER-SCR	Event-Related Skin Conductance Response
PPG	Fotopletysmografia

1. JOHDANTO

Stressin tutkiminen on mielenkiintoista, sillä se on itsessään psykologinen kokemus. Stressi voi myös näyttäytyä fyysisinä oireina, kuten päänsärkyinä, sydämen tykytyksenä tai jännittyneinä lihaksina [1]. Täten, kun halutaan tutkia fyysisiä oireita, on otettava usea eri biosignaali huomioon. Aiemmin mainituista fyysisistä oireista sydämentykytyks on sellainen, jota pystyttäisiin tutkimaan biosignaalien avulla.

Vaikka käytetään useampaa biosignaalia ja tulokset viittaisivat stressiin, ei pysty varmasti sanomaan, onko testihenkilö stressaantunut vai ei. Ainoa tapa saada varmasti tieto stressaantuneisuudesta on kysyä testihenkilöltä itseltään, miltä hänestä tuntui testitilanteessa. Tämäkin on hyvin subjektiivista, sillä eri henkilöillä on eri tasoinen stressinsietokyky. Täten tässä työssä voidaan vain spekuloida dataa tutkiessa, onko testihenkilö ollut stressaantunut vai ei ja onko tällä stressillä ollut vaikutusta testihenkilöiden tenttituloksiin.

Stressin mittauksesta on paljon hyötyä yhteiskunnan terveyden puolesta, esimerkiksi voidaan huomata taudin alkua tai hallita pitkäaikaisen stressin aiheuttamia oireita [2]. Pitkäaikainen stressi voi aiheuttaa ihmisellä monenlaisia vaivoja, jotka heikentävät ihmisen elämänlaatua sekä vähentävät ihmisen elinikää [3].

Tässä työssä käytetään Md Rafiul Amidin, Dilranjan Wickramasuriyan ja Rose T Faghihin luomaa tietokantaa (2022) [4]. Tietokannassa on kymmenen opiskelijan biosignaalien mittaamat tiedot kolmessa eri tenttitilanteessa sekä opiskelijoiden saamat arvosanat tenteistä. Tietokannassa on mitattu kuutta eri biosignaalia ja mittauksessa on käytetty Empatica E4 ranneketta [5], [6].

Tässä työssä käytämme näistä vain viittä; elektrodermaalista aktiivisuutta eli ihon sähkönjohtavuutta, sydämen sykettä, veren tilavuuspulssia, ihon pinnan lämpötilaa sekä sydämen lyöntien välistä aikaväliä. Huomiotta jätetään Empatica E4:n mittaama liikkuminen. Siirtola on artikkelissaan (2019) todennut, että on saatu parempia tuloksia, kun liikkuminen jätetään huomioimatta stressin tunnistuksessa [7]. Tarkoituksena on käyttää Amid et al.[4] luomaa tietokantaa ja verrata eri opiskelijoiden biosignaaleja ja niiden keskiarvoja keskenään.

Ohjelmoinnissa tässä työssä on käytetty ohjelmointikielensä Pythonia. Ohjelmointi on tehty Jupyter Notebookin avulla. Työn tarkoituksena on tutustua erilaisiin biosignaaleihin sekä niiden korrelointiin stressin tunnistuksessa.

2. STRESSI JA BIOSIGNAALIT

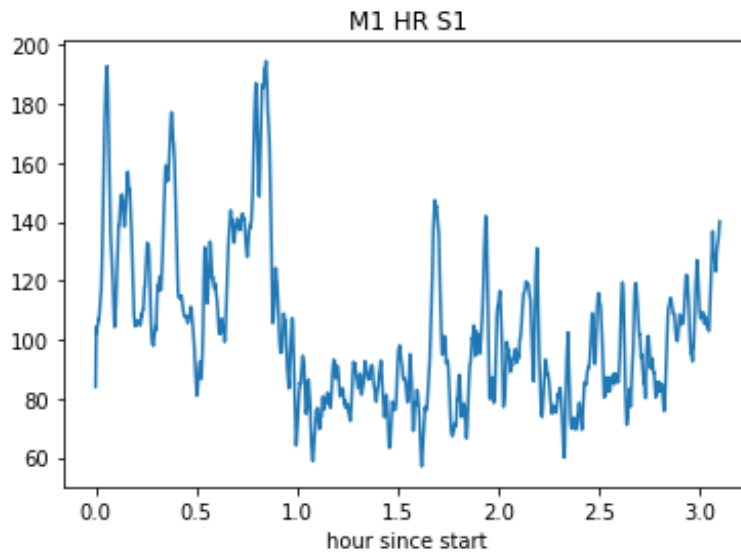
Stressi on fysiologinen reaktio kehossa, joka laukeaa fyysisen tai psykologisen tapahtuman takia [8]. Kun ääreishermoston sympaattinen hermosto (SNS) laukaisee stressireaktion, vapautuu erilaisia hormoneja, kuten kortisolia tai adrenaliinia, joiden takia hengitys nopeutuu, sydämen syke lisääntyy ja lihakset jännittyvät [1]. Biosignaalit, jotka osoittavat stressitilaan, ovat aivosähkökäyrä EEG, sydänsähkökäyrä EKG, veren tilavuuspulssi BVP, ihon sähkönjohtavuus EDA, ihon lämpötila TEMP sekä hengitystiheys RSP [9]. Näistä biosignaaleista parhaimmat tulokset saavutetaan, kun käytetään yhdessä EKG/BVP ja EDA-signaaleja, mutta EKG/BVP ja EDA-signaalit tuottavat myös hyviä tuloksia yksinään [9].

Jatkuva stressi on ihmisen terveydelle haitaksi. Stressi voi aiheuttaa kohonnutta verenpainetta, huonontaa nukkumista sekä heikentää henkistä suorituskykyä [8]. Jatkuva stressi voi myös aiheuttaa myös psykoosia tai jopa itsetuhoisuutta henkilöille, jotka kärsivät muista mielenterveydenongelmista, kuten skitsofreniasta [10]. Stressin tunnistaminen on tärkeää niin työssä kuin opiskelussa, koska liika stressi johtaa toistuviin poissaoloihin ja huonompaan menestykseen niin työssä kuin opinnoissa.

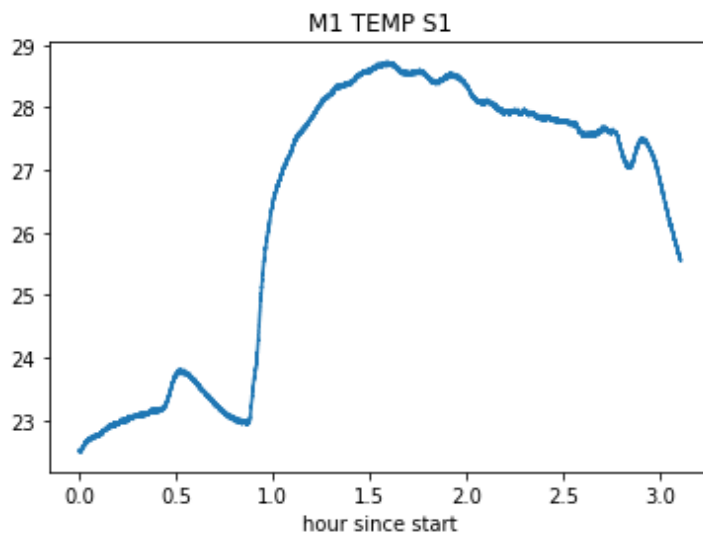
2.1. Laitteen esittely

Tässä tutkimuksessa käytetään Empatica E4:n mittaamia biosignaaleita [6]. Empatica E4 on Empatica yhtiön tuottama ranneke, jonka on tuottava hyvä laatuista ja reaaliaikaista biometristä dataa. Tätä tutkimusta tehtäessä Empatica on julkaissut uudemman rannekkeen, joten Empatica E4:ää ei enää myydä. Empatica:n rannekkeita voidaan käyttää kliinisissä tutkimuksissa, aktigrafiassa eli vuorokausirytmien ja unen tehokkuuden tutkimisessa, potilaiden etävalvonnassa sekä tieteellisissä tutkimuksissa.

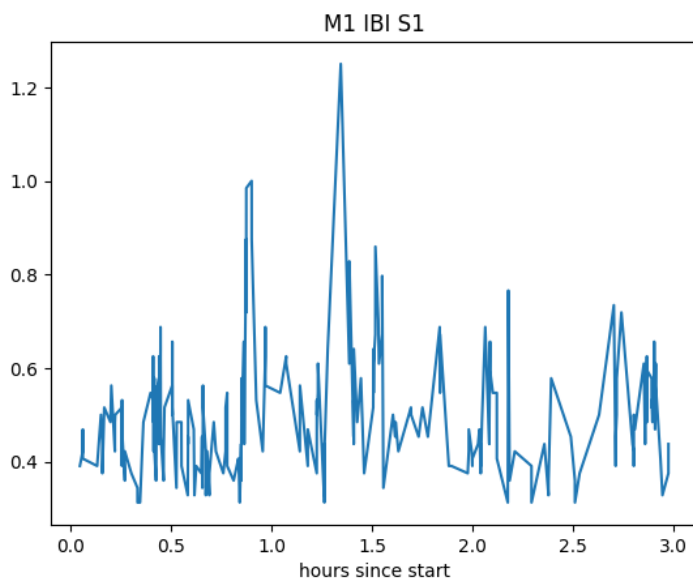
Empatica E4 sisältää neljä eri sensoria; fotopletysmografiasensori eli PPG, 3-akselinen kiihtyvyyssmittari, lämpötilasensori sekä elektrodermaalisuus aktiivisuuden eli ihon sähkönjohtavuuden sensoria. Nämä neljä sensoria tuottavat viittä eri dataa; veren tilavuuspulssia eli BVP, sydämen sykkeen lyöntien välistä aikaväliä eli IBI, elektrodermaalista aktiivisuutta eli EDA, kiihtyvyydataa eli ACC sekä ihon lämpötilaa. Ranneke mittaa myös kokoajan sydämen sykettä. Näillä mitatuilla tiedoilla on erilaisia taajuuksia käytössä. Veren tilavuuspulssin taajuus on 64 Hz, kiihtyvyydataan taajuus on 32 Hz, elektrodermaalisuuden ja ihon lämpötilan taajuus on 4 Hz sekä sydämen sykkeen taajuus on 1 Hz. IBI:llä ei ole mittaustaajuutta, sillä se laskee sydämen sykkeen lyöntien aikaväliä. IBI:n data sisältää täten kaksi saraketta dataa, joista ensimmäinen sarake sisältää ajan sekunteina, jolloin sensori on tunnistanut IBI:n tapahtuneeksi ja toinen sarake sisältää sykkeiden välisen aikavälin sekunneissa.



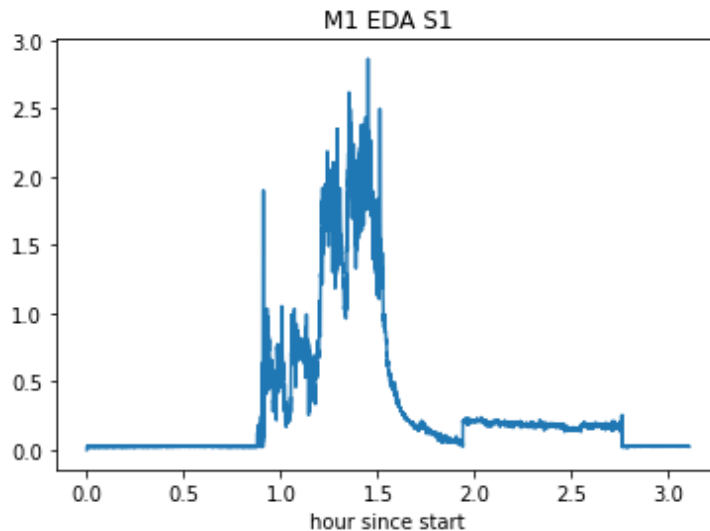
Kuva 1. Välikoe 1, opiskelijan 1 raaka HR-signaali



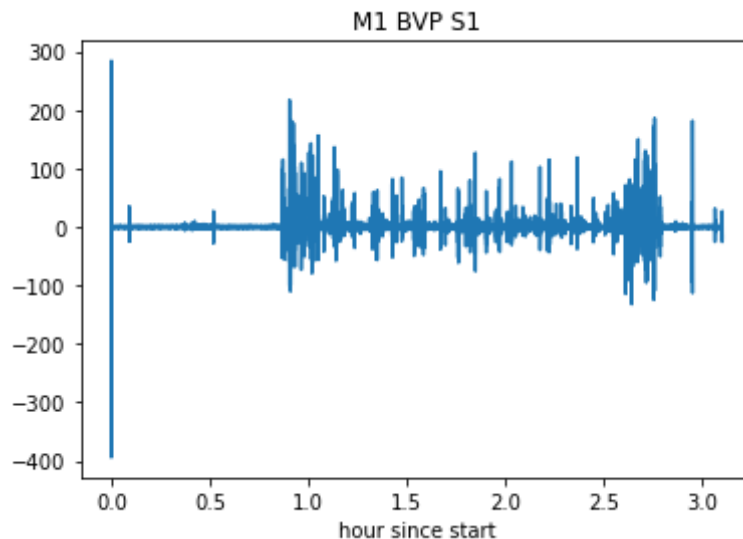
Kuva 2. Välikoe 1, opiskelija 1 raaka ihon lämpötilan signaali



Kuva 3. Välikoe 1, opiskelija 1 raaka IBI-signaali



Kuva 4. Välikoe 1, opiskelija 1 raaka EDA-signaali



Kuva 5. Välikoe 1, opiskelija 1 raaka BVP-signaali

Kuvista 1-5 nähdään opiskelijan 1 välikokeen 1 raat signaalit tässä tutkimuksessa käytettäville signaaleille. Kuvissa 1-5 on käytetty koko mittausajan dataa, jolloin dataa on joka biosignaalille kolmen tunnin verran välikokeelle 1.

2.2. Fysiologiset biosignaalit

Lämpötila yksinään ei anna tarpeeksi hyvää näytettä stressaantuneisuudesta [1]. Mutta antaa hyviä näytteitä, kun lämpötilaa vertaillaan sydämen sykkeen ja BVP-signaalin kanssa [7].

Nopeutunut sydämen syke on tunnettu stressin fyysinen oire, mutta nopea sydämen syke voi myös kuvata yleisesti tunteiden ilmentymää [11]. Tutkimuksissa on myös huomattu, että sydämen sykkeen keskiarvo nousee, kun tutkittava henkilö suorittaa aritmeettisiä tehtäviä [11].

Lähteessä [12] on vertailtu toisiinsa Empatica E4:n ranneketta ja Holter-monitoria keskenään. Holter-monitori on laite, jolla voidaan rekisteröidä EKG-signaalia pitempiaaikaisesti kuten kokonaisen vuorokauden aikana [13]. Laite sijoitetaan potilaan

vyötärölle, johon liitetään elektroditarrat rintakehän kohdalle, joiden avulla saadaan rekisteröityä sydämen sähkökäyrää. Tutkimuksen havaintojen perusteella Empatica E4 tunnistaa vähemmän IBI dataa, kun verrataan Holter-monitoriin. E4:n luotettavuus on myös heikompi kuin Holter-monitorin.

Sympaattinen hermosto säätelee ihon sähkönjohtavuutta, jolloin sähkönjohtavuus eli EDA on herkkä erilaisille kiihottumiselle, kuten stressille [1]. Kuitenkaan EDA-signaalia ei tarvitse käyttää stressin tunnistuksessa, sillä voidaan saada parempia tuloksia, jos EDA-signaali jätetään huomioimatta [7].

Yksinään BVP antaa parhaimmat näytteet stressaantuneisuudesta verrattuna muihin biosignaaleihin [1]. Jos pelkästään huomioidaan BVP-signaali, niin henkilön stressaantuneisuus voidaan tunnistaa noin 80 % todennäköisyydellä, kun käytetään lineaarista erotteluanalyysiä (LDA) [7].

2.3. Kirjallisuuskatsaus

Schmidt et al. [1] käsittelivät artikkelissaan ranteeseen ja rintaan kiinnitettäviä laitteita stressin tunnistuksessa. Koska tässä kandidaatin tutkielmassa käsitellään Empatica E4 ranneketta, keskitymme ranteessa olevan laitteen tuloksiin. Kun tutkittiin kolmella luokalla (lähtötilanne vs. stressitila vs. huvittuneisuus) stressin tunnistusta, niin ranteessa oleva laite tunnisti stressitilanteen noin 75 % todennäköisyydellä. Kun tutkittiin stressin tunnistusta binäärisesti (stressiä vs. ei-stressiä), niin stressitila tunnistettiin noin 87 % todennäköisyydellä.

Siirtola [7] käsitteli Empatica E4 ranneketta artikkelissaan ja tutki E4:n tuottamia signaaleja lineaarista erotteluanalyysiä (LDA) käyttämällä. Parhaimmat tulokset stressin tunnistukselle saatiin, kun käytettiin kehon lämpötilaa TEMP, sydämen sykettä HR ja veren tilavuuspulssia BVP. Tällöin stressitilanne tunnistettiin noin 87,4 % todennäköisyydellä.

Chalabianloo et al. [14] käyttivät Empatica E4 ranneketta ilman EDA-signaalia ja EDA-signaalin kanssa. Tämän tutkimuksen mukaan käyttämällä ExtraTree algoritmiä ja Empatica E4 ranneketta, jossa otetaan EDA-signaali myös huomioon, saadaan korkein stressin tunnistustodennäköisyys, 90.62 %.

De Vries et al. [15] käyttivät tutkimuksessaan Empatica E4:n lisäksi SentiSock-sensoria, joka mittaa EDA-signaalia jaloista. Tutkimuksessa on täten käytetty vain EDA-signaalia. Empatica E4:n stressin tunnistuksen tasapainotettu tarkkuus oli 82,4 % \pm 0,018. SentiSock antoi hiukan paremman tasapainotettu tarkkuuden kuin Empatica E4, 83,4 % \pm 0,019.

Giorgi et al. [16] vertasivat tutkimuksessaan kuluttajakäytössä olevaa Empatica E4:sta laboratorio käytössä olevaan Shimmer-sensoriin, joista vertailtiin EDA-signaalia. Artikkelissa käytettiin kahta stressaavaa tilannetta sensorien vertailuun. Toinen oli Operation-niminen lautapeli ja toinen oli "N-back"-tehtävä, jossa koehenkilöille näytettiin kirjainsarja, joka piti muistaa kahden 'askeleen' päähän; askel on satunnaistettu 500-3000 ms väliltä. Molemmissa tapauksissa Empatica E4:n stressin arvioinnin tulokset olivat pienemmät kuin Shimmer-sensorin[16].

Costantini et al. [17] tutkivat Empatica E4:n sykkeen vaihtelua (HRV), IBI:ä sekä EDA-signaalia. Tässä tutkimuksessa EDA:ta mitattiin kolmesta kohtaa; olkapäistä, sormista ja ranteesta. Jokaisessa näissä mittauspaikoista oli huomattavaa reagoimattomuutta tutkimuksen aikana, joten EDA-signaalin tutkiminen keskeytettiin. Empatica E4:n natiivin IBI:n suorituskyky oli vain 52,3 % (SD = 14,2 %), mutta IBI:n käsittelyn jälkeen IBI:n suorituskyky nousi noin 99 % jokaisessa testiolosuhteessa.

Sanchez-Comas et al. [18] tutkivat galvaanista ihovastetta, joka tunnetaan myös nimellä elektrodermaalinen aktiivisuus eli EDA. Tutkimuksessa verrataan Empatica E4:sta ja Shimmer-sensoria. Empatica-rannekkeet ovat tutkittavan henkilön molemmissa ranteissa. Mielenkiintoisesti vasemman ranteen Empatica-rannekkeen korrelaatio oli tämän tutkimuksen vahvin, jonka keskiarvo r :lle (Pearsonin korrelaatiokertoimelle) oli 0,59. Kuitenkin oikean ranteen Empatica-rannekkeen positiivinen keskiarvo r :lle oli vain 0,38.

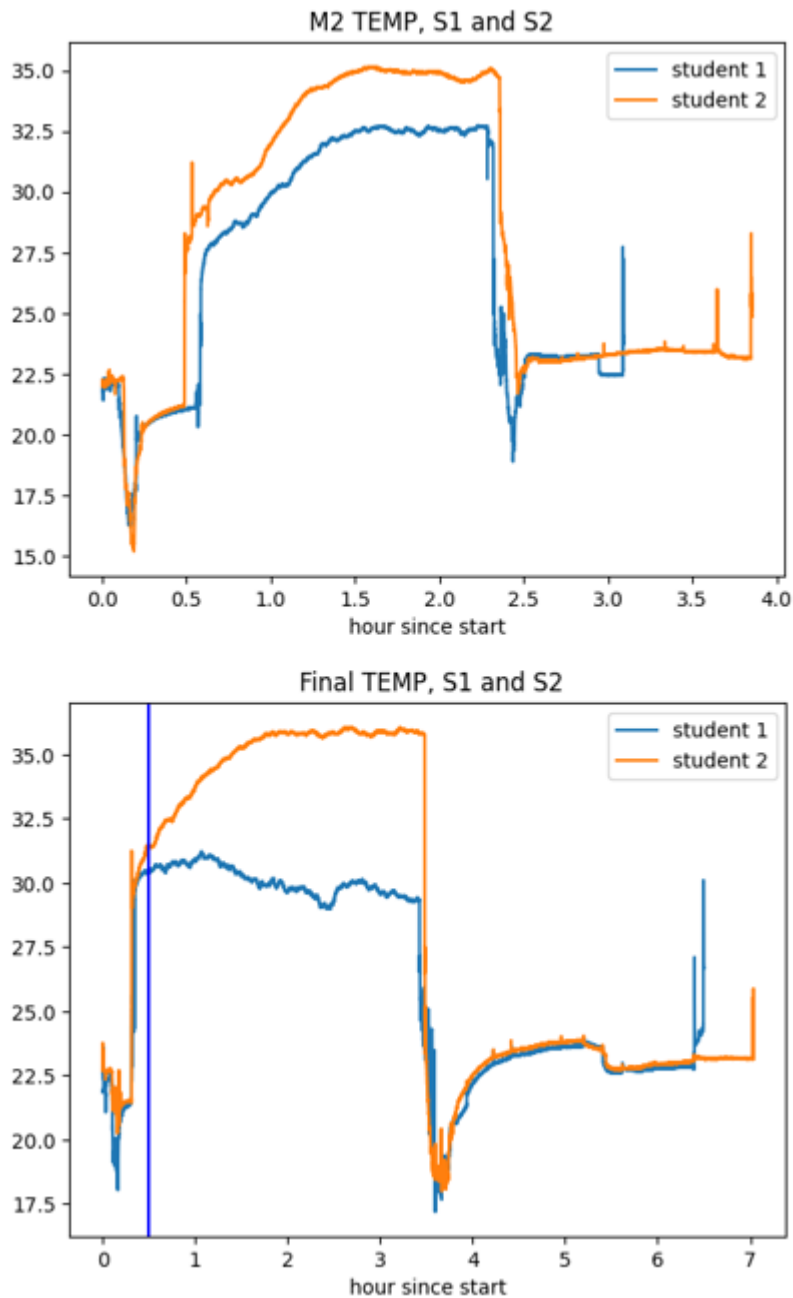
Gil-Martin et al. [19] käyttivät tutkimuksessaan Empatica E4:n lisäksi RespiBAN sensoria ja käyttivät molempien laitteiden signaaleja luokitteluntarkkuuden laskemisessa. Tässä tutkimuksessa oli käytössä kolme luokittelua: 2 eri luokkaa, 3 eri luokkaa ja 5 eri luokkaa. Signaalien prosessoinnissa käytettiin Fourier-muunnosta, kuutiojuurta ja vakio Q -muunnosta (CQT). Paras luokittelutarkkuus saatiin, kun käytettiin 2 luokkaista luokittelua stressi vai ei-stressi ja signaali prosessoinnissa käytettiin Fourier-muunnosta ja kuutiojuurta. Tällöin todennäköisyys oli $96,73 \% \pm 0,11 \%$. Kun lisätään enemmän luokkia, luokittelutarkkuuden todennäköisyys pienenee selvästi. Esimerkiksi jos on 5 eri luokkaa käytössä ja signaalin prosessointi tehdään samalla tapaa kuin parhaalle tarkkuudelle, todennäköisyys laskee $81,21 \% \pm 0,24 \%$.

3. STRESSIN TUNNISTUSTA TENTTITILANTEESSA

Tässä kappaleessa käsittelen oman tutkimuksen tuloksia ja kuinka tutkimus on tehty. Omassa tutkimuksessa olen käyttänyt Python-ohjelmointikieltä ja Jupyter Notebookia koodin ajamista varten.

Tässä tutkimuksessa käsittelen lähteen [4] datasettiä, johon sisältyy elektrodermaalinen aktiivisuus eli EDA, sydämen syke eli HR, veren tilavuuspulssi eli BVP, ihon lämpötila eli TEMP, lyöntien välinen aikaväli eli IBI ja kiihtyvyydata eli ACC. Datasetissä on kymmenen opiskelijaa, joista kahdeksan on miehiä ja kaksi on naisia. Dataa on tallennettu kolmessa eri tentissä, välikokeessa 1, välikokeessa 2 ja lopputentissä. Datasetistä löytää myös opiskelijoiden saamat pistemäärät jokaiselle tentille. Välikokeiden pituudet ovat puolitoista tuntia ja lopputentin pituus on kolme tuntia, mutta datan pituus on noin kaksinkertainen jokaiselle signaalille. Kaikki data on jaettu csv-tiedostoina. BVP:n, HR:n, EDA:n, ACC:n ja TEMP:n tiedostot sisältävät ensimmäisellä rivillä unix-aikaleima koordinoituna yhteisajassa eli UTC:ssä sekä toisella rivillä on näytetaajuus hertseinä. IBI:n tiedostoissa on ensimmäisellä rivillä myös unix-aikaleima, mutta sillä ei ole näytetaajuutta. Koska kyseessä on lyöntien välinen aikaväli, on IBI:n tiedostossa kaksi saraketta. Ensimmäisessä sarakkeessa on aika sekunneissa, jolloin IBI on tunnistettu ja toisessa sarakkeessa on tunnistetun IBI:n kesto aika sekunneissa. TEMP on mitattu celsiusasteissa (°C). EDA on mitattu mikrosiemenseinä (µS).

Kuten Mahadura et al. [20] ovat omassa tutkimuksessaan todenneet, on lopputenttien koodin antama aloitusaika eri kuin tietokannassa [4] oleva tenttien alkuaika. Tietokannassa sanotaan, että kaikki tentit ovat alkaneet aamulla klo 09.00 CT-aikavyöhykkeellä, mutta kun käännetään lopputentille annetut unix-aikaleimat CT-aikaan mittaamisen alkuaika olisi noin 10.28. Tietokannan kuvauksessa sanotaan, että päivänmäärät on siirretty, mutta aikoja ei ole siirretty. Täten lopputentin mittaamisen alkuaika pitäisi myös olla ennen klo 09.00, mutta jostain syystä unix-aikaleima antaa ajan, joka on huomattavasti klo 09.00 jälkeen. Välikokeiden unix-aikaleimat antavat käynnön jälkeen välikokeelle 1 noin 07.55 ja välikokeelle 2 noin 08.16. Koska välikokeiden aikaleimat ovat ennen klo 09.00, käytetään tietokannan antamaa tenttien alkuaikaa välikokeiden alkuaikoina. Lopputenttien unix-aikaleimoista saatavan ajan perusteella lopputentti ei ole voinut alkaa klo 09.00.



Kuva 6. Välikoe 2 ja lopputentti ihon lämpötila, lopputentin kuvaajaan lisätty 0,5 h ajankohta sinisenä viivana.

Kuvasta 6 nähdään, että lämpötilassa tapahtuu samanlainen muutos välikokeessa 2 ja lopputentissä. Välikoe 2 alkaa noin 44 minuuttia mittauksen alun jälkeen, jonka näkee jyrkkänä nousuna ihon lämpötilan visualisoinnissa. Samanlainen jyrkkä nousu näkyy myös lopputentin visualisoinnissa. Kuvan 6 perusteella lopputentissä opiskelijoiden ihon lämpötila on lähtenyt nousuun, ennen 0,5 tunnin kohtaa. Tällöin voidaan todeta, että tentti on alkanut pian mittauksen aloituksen jälkeen. Kuvan 6 ja muiden visualisointien avulla, olen valinnut lopputentin alkuaajaksi 10.45.

3.1. Tietojen valinta

Kuten aiemmin on jo mainittu, tässä tutkimuksessa olen käsitellyt ihon lämpötilaa (TEMP), ihon sähkönjohtavuutta (EDA), veren tilavuuspulssia (BVP), sydämen sykkeen lyöntien välistä aikaväliä (IBI) ja sydämen sykettä (HR). Olen jättänyt pois liikkeen (ACC) tutkimisen, sillä liikkuminen voi johtua jostain muusta kuin kokeen aikaisesta stressistä.

Koska tietokantaa tehtäessä on käytetty Empatica E4 ranneketta, on eri tiedoilla eri näytteenottotaajuuksia eli kuinka monta kertaa sekunnissa mittauksia suoritetaan. HR:n näytteenottotaajuus on 1 Hz. TEMP:llä ja EDA:lla on molemmilla näytteenottotaajuus 4 Hz:ä. IBI:llä ei ole näytteenottotaajuutta, sillä se riippuu HR:n sykkeiden aikavälistä. BVP:n näytteenottotaajuus on 64 Hz.

Tässä tutkimuksessa olen päättänyt skaalata kaiken datan vastaamaan BVP:n taajuutta. Eli HR, TEMP, EDA ja IBI:n data on venytetty BVP:n datan pituiseksi. HR:llä on otettu sama arvo 64 kertaa. TEMP:llä ja EDA:lla sama arvo otetaan 16 kertaa, jotta datasta saadaan yhtä pitkä kuin BVP:n datasta. IBI:lle datan pidennys oli monimutkaisempaa, sillä IBI:n data ei ole jatkuva kuten muilla tutkimuksen signaaleilla. Täten IBI:n datan pidennys tulee tehdä eri tavalla.

```
M1_IBI_long=[]
bvp_name = M1_BVP
ibi_name = M1_IB
ibi_stime = M1_IB_stime
for i in range(10):
    arr_long=[]
    it=0
    for k in range(len(bvp_name[i])): # k on bvp pituus
        bvp_time=k/64
        if it == 0:
            if bvp_time < ibi_stime[i][it]:
                arr_long.append(0)
            elif bvp_time >= ibi_stime[i][it]:
                arr_long.append(ibi_name[i][it])
                it=it+1
        else:
            if bvp_time < ibi_stime[i][it]:
                arr_long.append(arr_long[-1])
            elif bvp_time >= ibi_stime[i][it]:
                if it < len(ibi_name[i])-1:
                    arr_long.append(ibi_name[i][it])
                    it=it+1
                else:
                    arr_long.append(0)

    print(len(arr_long), 'IBI should be same as BVP len, student',i,exam)
    M1_IBI_long.append(arr_long)
```

Kuva 7. IBI:n datan pidennys koodi välikokeelle 1

Kuvasta 7 nähdään, kuinka paljon monimutkaisempi IBI:n pidennys on. Koodista nähdään, että IBI:n aikaa, *ibi_stime*, verrataan BVP:n aikaan, *bvp_time*. Koska IBI:n data kertoo vain millä ajankohdilla on tunnistettu IBI tapahtuneeksi, pitää verrata BVP:n jatkuvan signaalin aikaa IBI:n ensimmäisen sarakkeen arvoihin. Kun *bvp_time* ohittaa *ibi_stime*:n voidaan listaan lisätä kyseisen IBI:n kesto listaan *arr_long*. Listassa *arr_long* toistetaan samaa arvoa, niin kauan kunnes *bvp_time* ohittaa jälleen *ibi_stime*:n.

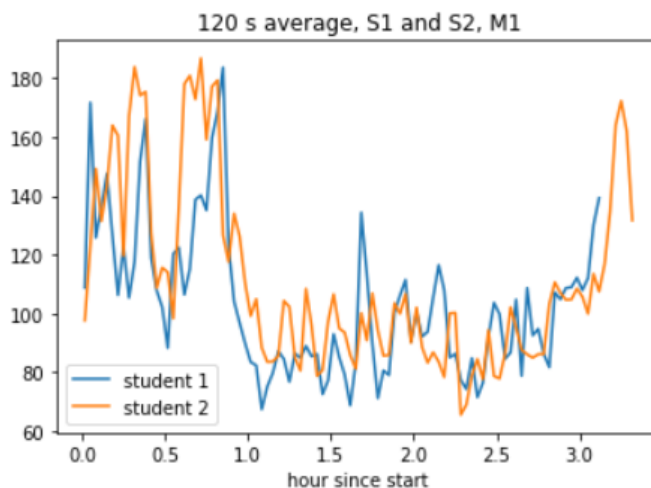
Kun kaikki data on skaalattu BVP:n mittaiseksi, seuraavaksi tulee rajata data kokeen mittaiseksi. Välikokeiden kesto on 1,5 tuntia ja lopputentti on 3 tuntia. Tentin alkuaika saadaan luomalla unix-aikaleima tenteille. Esimerkiksi ensimmäisen välikokeen alkuaika on 13.10.2018 klo 09.00, jolloin ensimmäisen välikokeen alkuajan unix-aikaleima on 1539439200.0. Koska jokaisen opiskelijan mittauksen alkuaika vaihtelee, niin jokaiselle opiskelijalle on laskettava oma tentin alkuajan ja mittauksen alkuajan ero. Tämä tapahtuu yksinkertaisesti, koska meillä on luotuna unix-aikaleimat tenttien aluille ja mittauksen alkuaika saadaan suoraan csv-tiedostosta. Kun nämä kaksi unix-aikaleimaa vähennetään toisistaan, saadaan selville, kuinka monta sekuntia datan alusta tulee poistaa. Tenttien loppuaika saadaan helposti lisäämällä edellä laskettuun eroon tentin kesto sekunteina, välikokeille 5400 sekuntia ja lopputentille 10800 sekuntia.

3.2. Tutkimuksen tuloksia

Tässä kappaleessa käsitellän jokaisen signaalin antamia tuloksia ja tutkin, miten nämä tulokset korreloivat tenttitulosten kanssa. Kun jokainen signaali on käsitelty, tutkin vielä tutkimuksen tuloksia regressioanalyysin avulla ja selvitän näyttääkö regressioanalyysi parempaa korrelointia tenttitulosten kanssa. Loppuun kerään vielä tutkimustulokset yhteen.

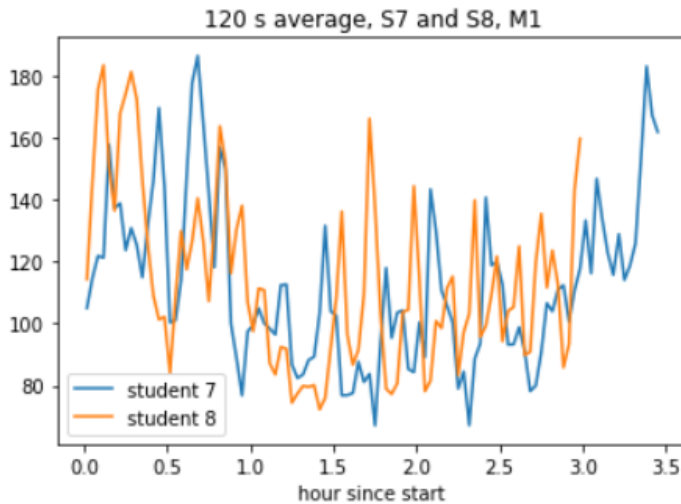
3.2.1. Sydämen syke

Sydämen sykkeen visualisoinnissa huomataan, että sydämen syke on yleensä ennen tenttiä korkeampi kuin kokeen aikana ja laskee kokeen ajaksi, jonka jälkeen syke lähtee taas nousuun.



Kuva 8. Välikoe 1 tasoitettu signaali HR opiskelija 1 ja 2

Kuvasta 8 nähdään selvästi, että ensimmäisen tunnin aikana opiskelijoiden 1 ja 2 sydämen syke on huomattavasti korkeampi kuin kokeen aikana. Opiskelijalla 1 on huomattavia piikkejä kokeen aikana, kun taas opiskelijan 2 sydämen syke pysyy tasaisempana kokeen aikana. Kuvasta 8 huomaa myös, että kokeen loppumisen jälkeen, kun tietojen mittaamisen alusta on kulunut 2,5 tuntia, molempien opiskelijoiden sydämen syke lähtee nousuun.



Kuva 9. Välikoe 1 tasoitettu signaali HR opiskelija 7 ja 8

Kuvassa 9 on mielenkiintoista se, että opiskelijalla 7 on huonoin arvosana välikokeessa 1 kaikista opiskelijoista ja opiskelijalla 8 on kaikista opiskelijoista parhain arvosana. Taulukosta 1 nähdään, että opiskelijan 8 tenttitulos välikokeelle 1 oli 92 pistettä 100 pisteestä ja opiskelijan 7 tenttitulos oli 64 pistettä 100 pisteestä. Molemmilla opiskelijoilla sydämen syke putoaa ensimmäisen tunnin jälkeen, mutta etenkin opiskelijalla 8 on huomattavia piikkejä sydämen sykkeessä. Kuvan 9 perusteella voisi olettaa, että molemmilla opiskelijoilla olisi samat arvosanat tentistä, mutta kuvassa on parhaiten suoriutunut ja huonoiten suoriutunut opiskelija.

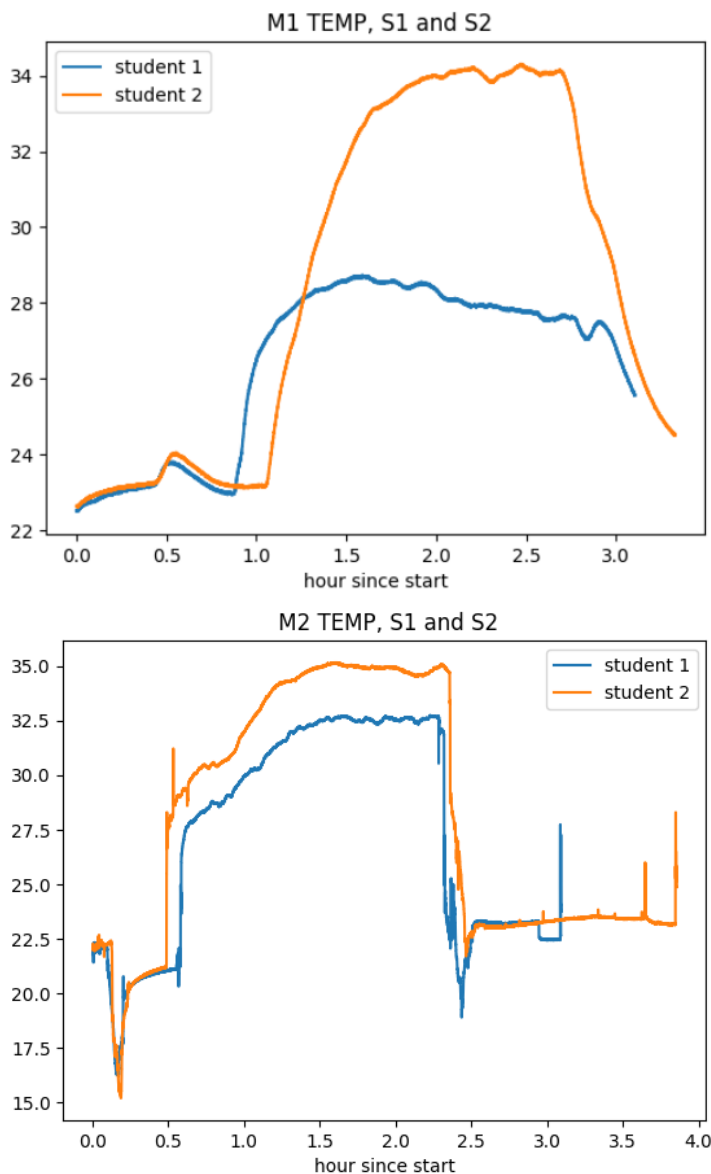
Taulukko 1. Opiskelijoiden tenttitulokset – välikokeiden pistemäärä on 100, lopputentin pistemäärä on 200

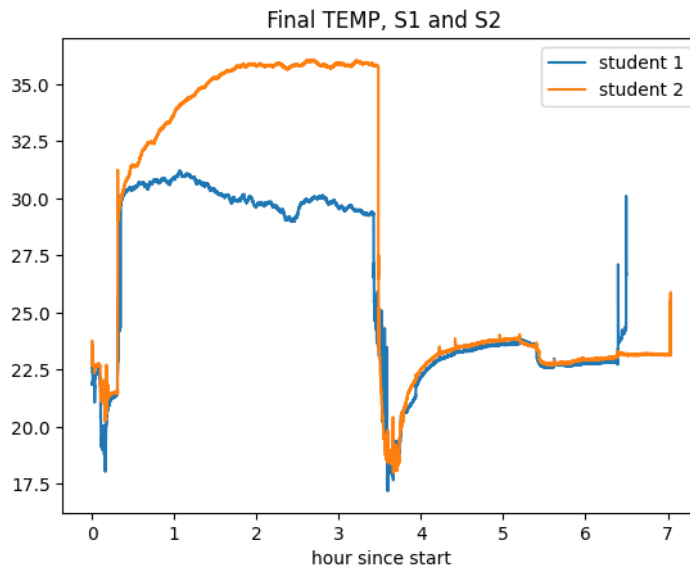
	Välikoe 1	Välikoe 2	Lopputentti
Opiskelija 1	78	82	182
Opiskelija 2	82	85	180
Opiskelija 3	77	90	188
Opiskelija 4	75	77	149
Opiskelija 5	67	77	157
Opiskelija 6	71	64	175
Opiskelija 7	64	33	110
Opiskelija 8	92	88	184
Opiskelija 9	80	39	126
Opiskelija 10	89	64	116

Tämä ero voidaan huomata myös, kun lasketaan sydämen sykkeen keskiarvoissa ja keskihajonnoista ennen tenttiä, tentin aikana sekä tentin jälkeen. Esimerkiksi opiskelijalla 1 sykkeen keskiarvo ennen tenttiä on noin 124 BPM, kokeen aikana sykkeen keskiarvo putoaa noin 89 BPM ja kokeen jälkeen keskiarvo nousee noin 102 BPM. Opiskelijan 1 sydämen sykkeen keskihajonta laskee myös kokeen aikana. Ennen tenttiä sydämen sykkeen keskihajonta on noin 28 BPM ja kokeen aikana sykkeen keskihajonta on noin 16 BPM. Mielenkiintoisesti opiskelijalla 8, jolla on paras tenttitulos välikokeessa 1, sydämen sykkeen keskihajonta välikokeen 1 aikana on noin 23 BPM, joka on kaikista opiskelijoista korkein. Opiskelijan 7, jonka tenttitulos oli välikokeessa 1 heikoin, sykkeen keskihajonta oli opiskelijoista toiseksi korkein, noin 21 BPM.

Näistä tuloksista nähdään, että oppilaiden sydämen sykkeet putoavat, kun he aloittavat kokeen. Suurimmalla osalla opiskelijoista sydämen syke pysyy myös matalampana kuin ennen tenttiä, mutta muutamalla opiskelijalla syke on epätasaisempaa kokeen aikana. Sydämen tykytys on yleinen stressin oire [1], jolloin voidaan olettaa, että suurin osa opiskelijoista ei tunne oloaan stressaantuneeksi tentin aikana, mutta voivat olla stressaantuneita ennen tenttiä ja tentin jälkeen. Mielenkiintoisesti opiskelijoilla 7 ja 8 on jokaisessa tentissä isoja piikkejä HR-signaaleissa. Opiskelijalla 7 on huonoimmat tenttiarvosanat ja opiskelijalla 8 on parhaimpia tenttiarvosanoja. Tämä voi olla näyttöä hyvästä ja huonosta stressistä. Opiskelijalla 7 stressi olisi huonoa ja opiskelijalla 8 stressi olisi hyvää.

3.2.2. Ihon lämpötilä





Kuva 10. Ihon lämpötila, opiskelijat 1 ja 2, 1. ja 2. välikoe sekä lopputentti

Kuvasta 10 nähdään, että ihon lämpötilan visualisoidessa tenttien alkaessa on jyrkkä nousu. Tämän nousun jälkeen ihon lämpötila pysyy likimain samantasoisena tentin aikana. Välikokeen 2 ja lopputentin visualisoinnista nähdään myös selkeästi, kuinka ihon lämpötila putoaa jyrkästi tenttien loppuessa.

Ihon lämpötilan muutokset voidaan myös nähdä, kun lasketaan ihon lämpötilan keskiarvo ennen tenttiä, tentin aikana ja tentin jälkeen.

Taulukko 2. Opiskelijan 1 ihon lämpötilan keskiarvo ennen tenttiä, tentin aikana ja tentin jälkeen

Opiskelija 1	Välikoe 1	Välikoe 2	Lopputentti
Ennen tenttiä	+ 23.6°C	+ 21.6°C	+ 21.0°C
Tentin aikana	+ 28.2°C	+ 31.4°C	+ 29.9°C
Tentin jälkeen	+ 27.2°C	+ 24.3°C	+ 23.1°C

Taulukko 3. Opiskelijan 2 ihon lämpötilan keskiarvo ennen tenttiä, tentin aikana ja tentin jälkeen

Opiskelija 2	Välikoe 1	Välikoe 2	Lopputentti
Ennen tenttiä	+ 23.3°C	+ 23.0°C	+ 21.2°C
Tentin aikana	+ 32.2°C	+ 33.7°C	+ 34.5°C
Tentin jälkeen	+ 30.0°C	+ 24.9°C	+ 23.7°C

Opiskelijalla 2 on korkeampi ihon lämpötila kuin opiskelijalla 1, joka voidaan nähdä myös kuvasta 4. Taulukoista 2 ja 3 nähdään hyvin, että opiskelijoiden 1 ja 2 ihon lämpötilan keskiarvo nousi tenttien ajaksi. Opiskelijan 2 ihon lämpötila nousee joka tentissä korkeammaksi kuin opiskelijan 1. Opiskelijan 2 ihon lämpötilan keskiarvo tenttien aikana on myös kaikista opiskelijoista korkein. Taulukon 1 arvosanojen perusteella opiskelija 2 on 1. ja 2. välikokeessa kolmanneksi paras ja lopputentissä neljänneksi paras. Opiskelijoilla, joilla on paremmat arvostukset kuin opiskelijalla 2, on kuitenkin matalampi ihon lämpötilan keskiarvo kuin opiskelijoilla 2.

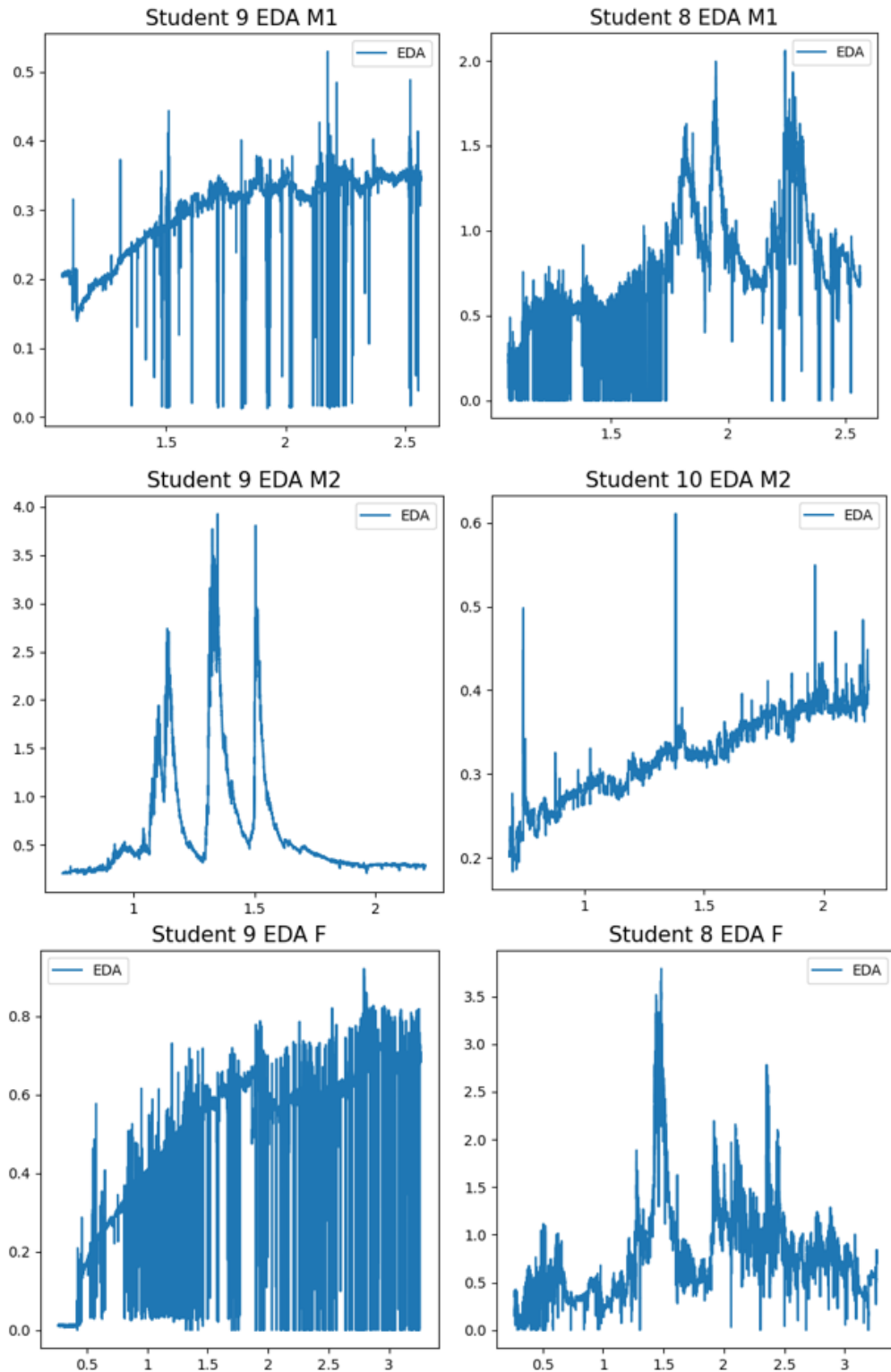
Täten korkea ihon lämpötila ei korreloi hyvän tenttituloksen kanssa. Ihon matalalla lämpötilalla korreloi hieman huonon tenttituloksen kanssa. Opiskelijalla 7, jolla on huonoimmat tenttitulokset, on matalin ihon lämpötilan keskiarvo välikokeessa 2 ja

lopputentissä. Välikokeessa 1 matalimmat ihon lämpötilan keskiarvot ovat opiskelijalla 4, opiskelijalla 7 ja opiskelijalla 10. Taulukosta 1 nähdään, että opiskelijalla 4 ja 7 on huonoimpia tenttituloksia, mutta opiskelijalla 10 on toiseksi paras tenttitulos. Ihon lämpötilan keskiarvot vaihtelevat kokeiden välillä. Välikokeen 2 ihon lämpötilan keskiarvot ovat korkeimmat.

3.2.3. EDA-signaali

EDA voidaan ajatella kuvastavan henkilön tunnetilan voimakkuutta, eli kun ihminen kokee jotain uhkaavaa, riemuisaa tai muita voimakkaita tunteita. EDA-signaalin luettaessa otetaan huomioon merkittävät muutokset. Näitä merkittäviä muutoksia kutsutaan *tapahtumiin liittyviksi ihon johtavuusvasteiksi* eli ER-SCR. [21]

Kun katsotaan tämän tutkimuksessa piirrettyjä signaaleja, suurimmalla osalla opiskelijoista merkittävä ER-SCR on kokeiden alussa. Tämän voidaan ajatella kuvaavaan sitä, että opiskelijat ovat stressaantuneita tentin alussa, mutta rauhoittuvat tentin edetessä. Kuitenkin on opiskelijoita, joilla on useampia ER-SCR-piikkejä tai EDA-signaalissa ei ole selvää piikkiä.



Kuva 11. EDA-signaaleja välikokeessa 1 (M1), välikokeessa 2 (M2) ja loppuentissä (F)

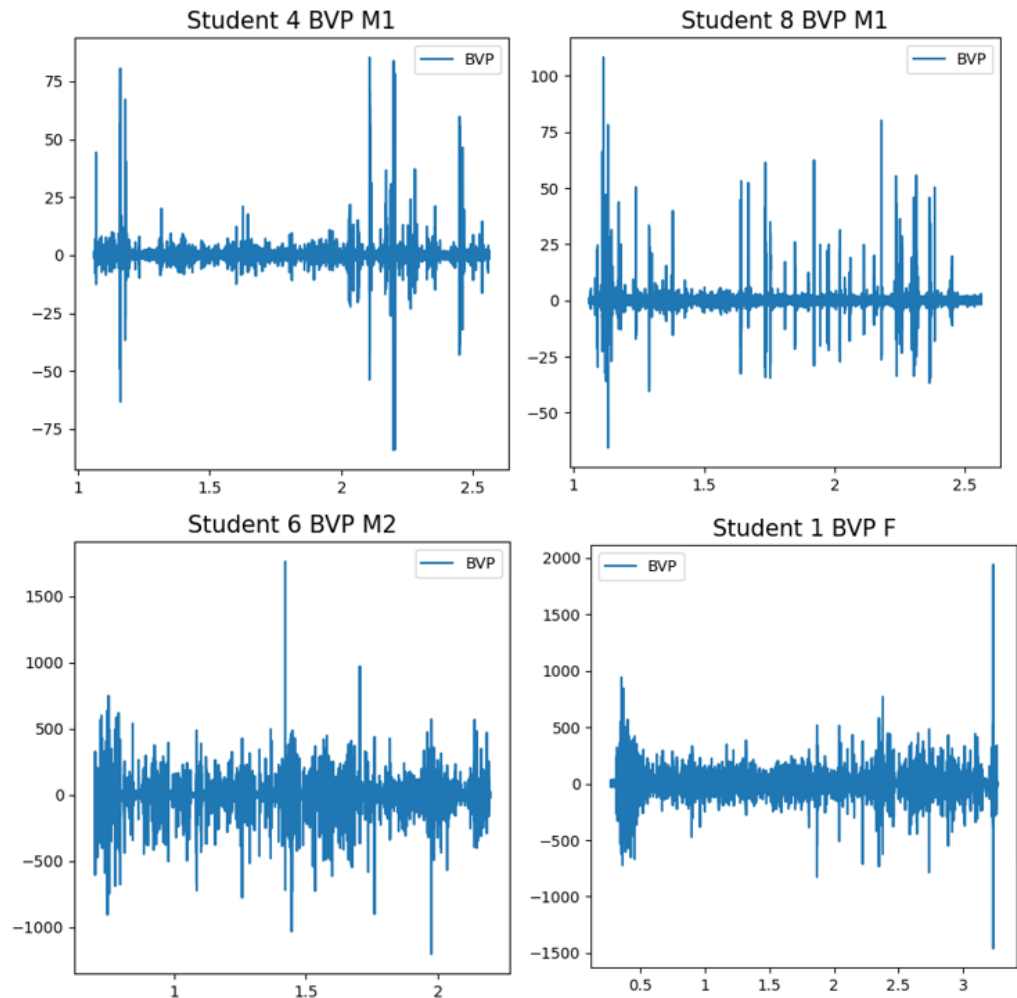
Kuvan 11 signaalit voivat kuvata sitä, että opiskelijat 9, 8 ja 10 tuntevat itsensä stressaantuneiksi tenttien aikana. Etenkin opiskelijalla 9 välikokeessa 2 ja opiskelijalla 8 loppuentissä on ER-SCR-piikki vasta tunnin päästä tenttien alusta. Opiskelijan 9 arvosanat huononevat tenttien edetessä, välikokeessa 1 opiskelijan 9 arvosana on

neljänneksi paras ja loppuentissä opiskelijan 9 arvosana kolmanneksi huonoin arvosana.

Opiskelijan 9 välikokeen 2 signaali on mielenkiintoinen muutos kyseisen opiskelijan signaaleissa. Siinä on paljon vähemmän piikkejä ja muotokin on selvästi erilainen, kun verrataan opiskelijan 9 muihin signaaleihin. Tämä voi kuvastaa sitä, että opiskelija 9 oli stressaantuneempi välikokeessa 1 ja loppuentissä kuin välikokeessa 2. Taulukon 1 perusteella opiskelijan 9 tenttitulos välikokeessa 2 on vain 39 pistettä 100 pisteestä, joka on toiseksi huonoin tenttitulos välikokeessa 2.

3.2.4. BVP-signaali

BVP-signaali on tiheämpää, kun henkilö tuntee olonsa stressaantuneeksi. Tässä tutkimuksessa huomataan, että koska kyseessä on pitkä ajan pätkä, on BVP-signaaleja hankala lukea. Välikokeen 1 signaalit ovat parhaiten luettavissa verrattuna muihin tentteihin. Voidaan ajatella, että opiskelijat olivat stressaantuneimpia välikokeen 2 ja loppuentin aikana kuin välikokeen 1 aikana. Etenkin välikokeiden 1 ja 2 tulisi loogisesti olla yhtä luettavia, sillä kyseiset tentit ovat yhtä pitkiä, mutta välikokeen 2 signaaleista on hankala nähdä selkeitä piikkejä.



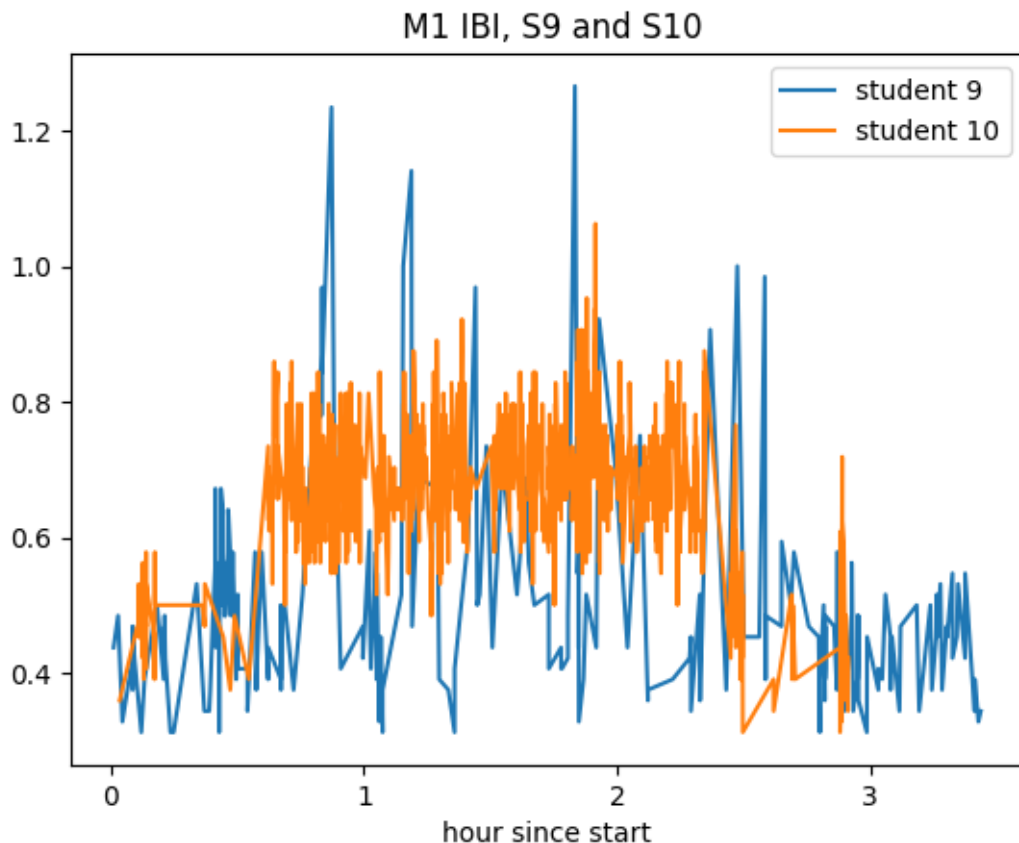
Kuva 12. BVP-signaaleja eri tenteissä

Kuvasta 12 nähdään signaaleja esimerkkejä BVP-signaaleista eri tenteissä. Signaalit on valittu sillä perusteella, että ne ovat helpoiten luettavia signaaleja. Kuvasta 12 huomaa, että välikokeen 1 signaaleissa on helpommin nähtäviä piikkejä, kun taas välikokeessa 2 ja loppukokeessa selviä piikkejä on huomattavasti vähemmän.

Opiskelijoiden 4 ja 8 signaaleissa on vähän piikkejä verrattuna muiden opiskelijoiden välikokeen 1 signaaleihin, mutta opiskelijoiden tenttitulokset ovat hyvin erilaiset. Opiskelija 8 tenttitulos on paras välikokeessa 1 ja opiskelijan 4 tenttitulos on neljänneksi huonoin tenttitulos. Opiskelijoiden 6 ja 1 signaalit ovat kapeimmat signaalit y-akselin perusteella välikokeessa 2 ja loppulentissä. Kapea BVP-signaali ei myöskään korreloi selvästi tenttituloksen kanssa, sillä opiskelijalla 6 välikokeessa 2 tenttitulos oli kolmanneksi huonoin ja opiskelijan 1 loppulentin tenttitulos on kolmanneksi parhain.

3.2.5. IBI-signaali

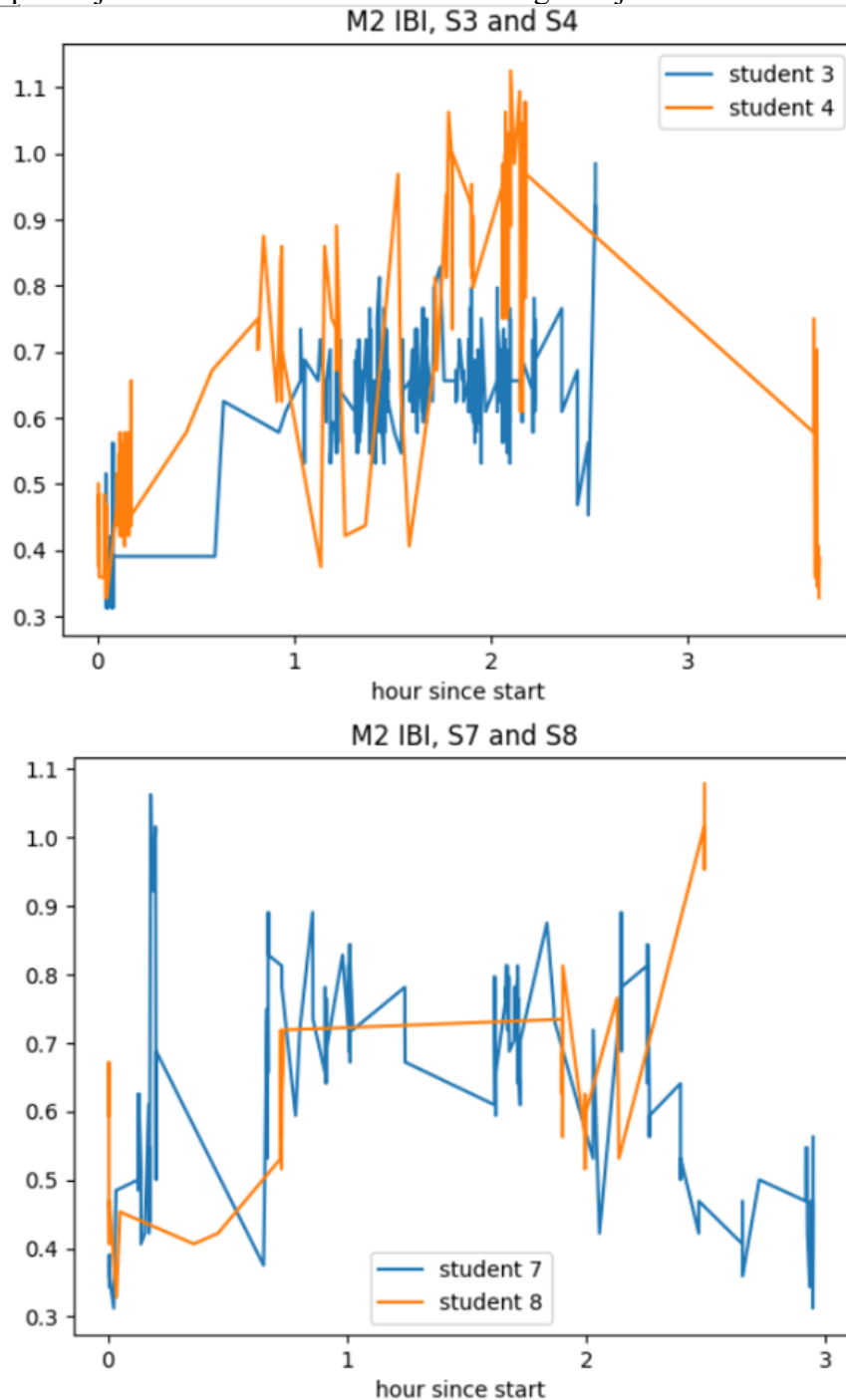
Koska IBI:n pidennyksessä on käytetty samaa arvoa niin kauan kunnes seuraava IBI-piikki tapahtuu, piirretty signaali on laatikkomainen. Kun ihminen on stressaantunut, stressi voi näkyä IBI:n signaalin pienempinä arvoina [22]. Jaafarin ja Chung Xianin artikkelin [22] kuvasta 4 nähdään, että jokaisen neljän koehenkilön IBI-arvot pienenevät ja arvojen vaihteluväli kutistuu. Tässä tutkimuksessa piirretyt IBI-signaalien arvot kasvavat tenttien aikana.



Kuva 13. IBI-signaali, välikoe 1, opiskelija 9 ja 10

Koska IBI-signaalien arvot kasvavat tenttien ajaksi, voidaan todeta, että opiskelijat tuntevat itsensä stressaantuneilta ennen tenttiä, mutta rauhoittuvat tenttien ajaksi.

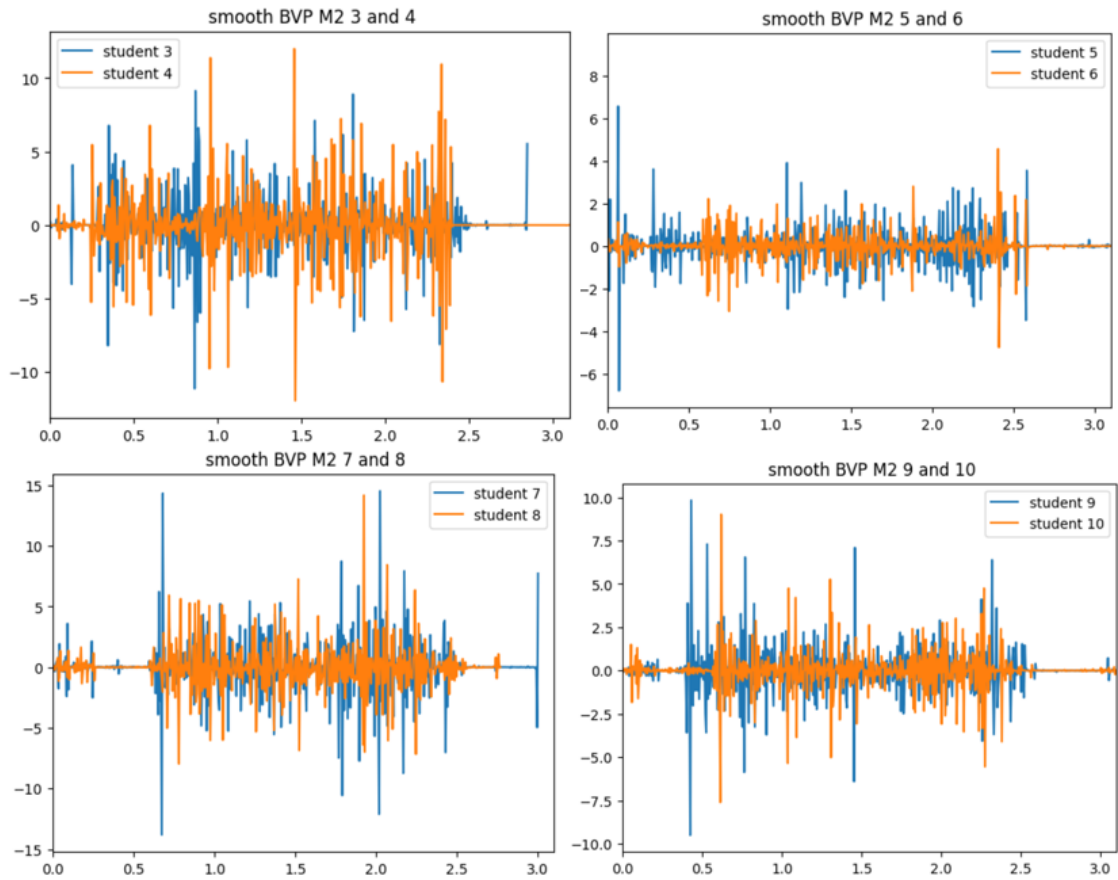
Tämä korreloi hyvin sydämen sykkeen kanssa saaduista tuloksista. Useimpien opiskelijoiden IBI-signaalit ovat samanlaisia kuin kuvassa 13, mutta muutamilla opiskelijoilla on erikoisemman näköisiä signaaleja.



Kuva 14. Erikoisia IBI-signaaleja välikokeessa 2

Kuvan 14 mittaukset ovat selkeästi epäonnistuneet sillä signaaleja ei ole käsitelty millään tapaa. Kyseinen epäonnistuminen johtuu luultavimmin laitteesta itsestään. Joko sensori ei ole ollut oikealla kohdalla tai sensorissa on ilmestynyt jonkinlainen ongelma välikokeen 1 jälkeen. Empatica E4:n IBI saadaan E4:n mittaamasta BVP-signaalista [23], joten IBI:ä prosessoitaessa on voinut käydä jonkinlainen virhe, joka on sitten johtanut kuvan 14 signaaleihin. Olander et al. [24] ovat tutkimuksessaan

todenneet, koska IBI mittaa BVP-signaalista aikaväliä, niin jos BVP-signaali on heikko jää IBI:en määrä vähäiseksi.



Kuva 15. tasoitetut BVP-signaalit välikokeesta 2 (koko data) opiskelijoille 3-10

Opiskelijoilla 5, 6, 9 ja 10 kaikilla on normaali IBI-signaali tentin aikana. Opiskelijoilla 3, 4, 7 ja 8 on kuvan 14 mukaiset IBI-signaalit. Kun verrataan kuvan 15 opiskelijoiden 3, 4, 7 ja 8 BVP-signaaleja opiskelijoiden 5, 6, 9 ja 10 BVP-signaaleihin, jokaisessa kuvassa on runsaasti selkeitä piikkejä. Täten Ollander et al. [24] väite, ei selitä miksi opiskelijoiden 3, 4, 7 ja 8 IBI-signaalit näyttävät niin erilaisilta verrattuna muihin opiskelijoihin.

3.2.6. Regressioanalyysi ja R^2 -arvo

Seuraavaksi verrattiin regressioanalyysin avulla opiskelijoiden saamia tentti arvosanoja ja eri signaaleja. Tässä tutkimuksessa on käytetty lineaarista regressiota, jonka kaava on $y = a \cdot X + b$, jossa a on lineaarisen regression kaltevuus ja b on lineaarisen regression leikkauspiste. Koodissa olen käyttänyt komentoa *linregress*, jonka avulla olen saanut arvot helposti a :lle ja b :lle sekä r -arvon. Komento *linregress* käyttää Pearsonin korrelaatiokerrointa r -arvon saamiseksi. Pearsonin korrelaatiokerroin kaava on $\rho_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$, jossa lasketaan X :n ja Y :n kovarianssi, joka sitten jaetaan X :n ja Y :n keskihajonnalla, σ_X ja σ_Y [25]. R^2 -arvo saadaan helposti kertomalla r -arvo itsellään eli $R^2 = r \cdot r$. R^2 -arvoa käytetään usein

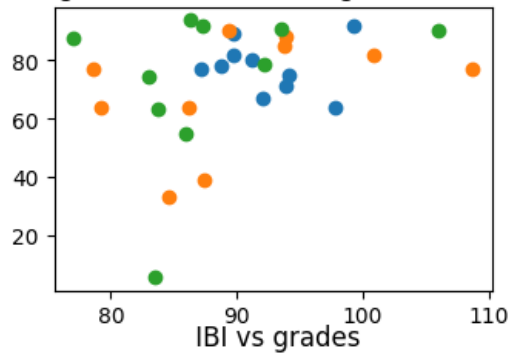
selittämään biologisen datan varianssiosuutta, kun käytetään lineaarisia malleja [26]. R^2 -arvo myös helpottaa biologisen datan vaihtelun syiden ymmärtämistä [26]

Regressioanalyysin perusteella biosignaaleista parhaiten tenttien arvosanoihin korreloi BVP-signaali. Parhain R^2 -arvo on toisen välikokeen tenttitulosten ja BVP-signaalin korrelointi, joka on noin 0,42. Tämä arvo ei kuitenkaan näytä vahvaa korrelointia tenttitulosten ja BVP-signaalin välillä, sillä saatu R^2 -arvo on alle 0,50. Muiden biosignaalien korrelointi tenttituloksiin on vielä heikompaa. Taulukon 4 tulosten perusteella voidaan todeta, että tässä tutkimuksessa käytetyissä biosignaaleilla ei ole minkäänlaista korrelaatiota tenttitulosten kanssa.

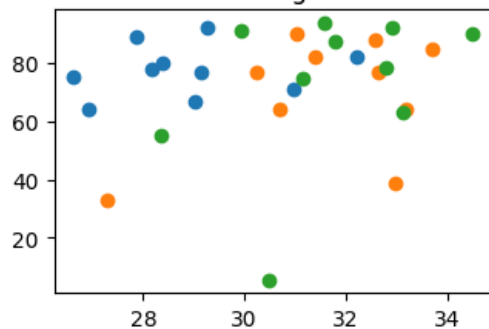
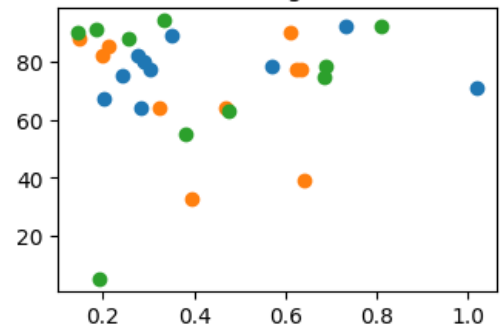
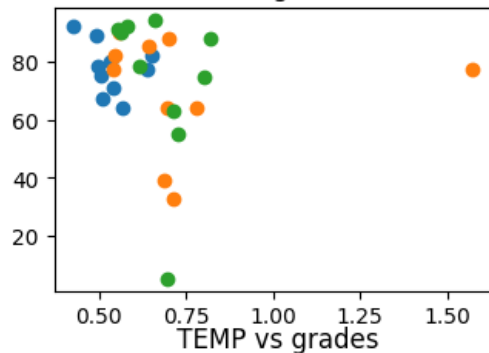
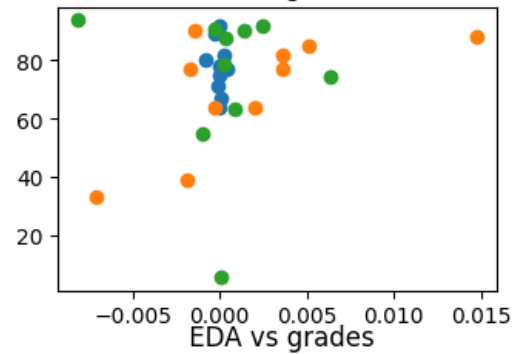
Taulukko 4. Biosignaalien ja tenttitulosten R^2 -arvot tenttien aikana

Biosignaalit	R^2 -arvo (Välikoe 1)	R^2 -arvo (Välikoe 2)	R^2 -arvo (Loppupentti)
HR	0,0077	0,1553	0,0964
BVP	0,0278	0,4226	0,0089
IBI	0,1000	0,0003	0,0874
EDA	0,0283	0,0714	0,0392
TEMP	0,0318	0,1809	0,1539

HR vs grades, blue M1, orange M2, Green Final



BVP vs grades



Kuva 16. Signaalien keskiarvo ja tenttitulosten korrelaatio, x-akseli on signaalien tenttien aikaiset keskiarvot ja y-akseli on tenttitulokset, välikoe 1 merkitty sinisellä, välikoe 2 merkitty oranssilla, lopputentti on merkitty vihreällä

Kuvasta 16 nähdään selkeästi, että mikään signaaleista ei muodosta selkeää suoraa, kun tenttien aikaisia keskiarvoja korreloidaan tenttitulosten kanssa. Etenkin IBI:n keskiarvot ovat pakkautuneet yhteen nurkkaan eivätkä muodosta minkäänlaista suoraa. Edes BVP-signaali 2. välikokeessa ei muodosta selkeää suora viiva tenttitulosten kanssa, vaikka tällöin R^2 -arvon perusteella tilanne oli todennäköisin.

3.2.7. Lopputulos

Omien tutkimuksieni perusteella en pysty vahvasti näyttämään kunnollista korrelointia biosignaalien ja tenttitulosten välillä. Täten yksinkertainen datan tutkimus ei riitä korreloinnin löytämiseen, vaan tarvitaan joko koneoppimista tai muita haastavimpia datan tutkimuksen menetelmiä.

4. MUIDEN TUTKIMUSTEN KÄSITTELY JA VERTAILU

Tässä kappaleessa käsittelen muita aiheeseen liittyviä artikkeleita ja vertailen saamiani tuloksia muihin tutkimuksiin. Tutkimukset on jaettu niihin, jotka ovat käyttäneet Empatica E4 ja niihin, jotka ovat käyttäneet jotain muuta sensoria tai muunlaisia datan keräysmenetelmää, esimerkiksi sylkinäytteitä tai kyselyitä.

4.1. Empatica E4

Indikawati ja Winiarti [27] käyttävät tutkimuksessaan Empatica E4 sensoria. Heillä oli tiedossa neljä luokitusta testihenkilöiden tilalle: lähtötila, stressi, huvi ja meditaatio. Minun käyttämässäni tietokannassa ei ollut annettu minkäänlaista luokittelua, että olivatko opiskelijat tunteneet itseään stressaantuneiksi, joten en pysty sanomaan olivatko opiskelijat stressaantuneita. Indikawati ja Winiarti käyttivät logistista regressiota, Decision Tree ja Random Forest algoritmia luokitteluun. Parhaimmat luokittelutulokset antoi Random Forest algoritmi, jonka luokittelun tarkkuus vaihteli 88 – 99 % välillä.

Graft et al. [28] käyttivät tutkimuksessaan logistista regressiota ja Naive Bayes'in algoritmia. Tutkimuksessa käytettiin kolmea eri koulutus-testaus skenaariota, osallistujien välistä ristiin validointia, kerrostettua 10-kertaista ristiin validointia sekä korvausta. Käyrän alla oleva pinta-alan (AUC) mukaan korvausmenetelmää sekä logistista regressiota käytettäessä saadaan parhain luokittelukyky, noin 86 %, kaikista mahdollisista algoritmi ja koulutus-testaus skenaario yhdistelmistä. En edes ajatellut käyttää koulutus-testaus skenaarioita omassa tutkimuksessani, mutta tietokannan datamäärä oli pieni, 10 opiskelijaa ja 3 tenttiä. Graft et al. [28] ensimmäisessä tutkimuksessa tuotti yhteensä 60 tehtävää viidelle osallistujalle, toisessa tutkimuksessa oli vain yksi osallistuja, joka teki kahden viikon ajalla 10 tehtävää.

Siirtola ja Röning [29] käyttivät tutkimuksessaan Random Forest algoritmia luokitteluun, sillä se oli osoittautunut parhaimmaksi luokittelualgoritmiksi stressin tunnistuksessa Siirtolan [7] aiemman tutkimuksen perusteella. Tässä tutkimuksessa käytettiin myös regressiomallia Bagged Tree. Tutkimuksessa verrattiin käytetyn regressiomallin ja käytetyn luokittelualgoritmin todennäköisyyttä, kun käytettiin testikohteiden BVP:n ja TEMP:n ominaisuuksien yhdistelmää. Tämän perusteella regressiomalli Bagged Tree:llä on parempi stressin tunnistustarkkuus. Itse käytin yksinkertaista regressioanalyysiä, mutta tulokset olivat hyvin huonoja tenttitulosten ja fysiologisten signaalien välillä. Ehkä olisin saanut parempia korrelaatiotuloksia, jos olisin käyttänyt jotain muuta regressioanalyysimenetelmää.

Kumar et al. [30] käyttivät samaa tietokantaa kuin Gil-Martin et al. [19] tutkimuksessaan. Tässä tutkimuksessa käytettiin konfuusiomatriisia, jonka avulla laskettiin suorituskyvyn mittarit F1-arvo, sisäinen tarkkuus (precision), herkkyyys (recall) ja tarkkuus (accuracy). Itse en edes ajatellut, että olisin voinut käyttää konfuusiomatriisia ja laskea suorituskyvyn mittareita. Tässä tutkimuksessa oli käytetty kolmea luokkaa: perustaso, stressi ja huvi.

4.2. Muut menetelmät

Hearn ja Stocker [31] eivät käyttäneet tutkimuksessaan sensoreita vaan opiskelijat täyttivät itsearviointilomakkeen FFMQ (Five Facet Mindfulness Questionnaire) ja

antoivat sylkinäytteen, josta mitattiin kortisoliarvot perustasossa eli ryhmätyössä ja stressitilanteessa eli ennen tenttiä. Kortisoliarvojen ja FFMQ vastausten perusteella opiskelijat olivat selvästi stressaantuneimpia ennen tenttiä kuin ryhmätöissä. Omassa tutkimuksessani totesin myös sydämen sykkeen, IBI:n ja EDA:n perusteella, että opiskelijat luultavimmin olivat enemmän stressaantuneita ennen tenttiä kuin tentin aikana.

Petrescu et al. [32] tutkivat ahdistuksen tunnistusta Shimmer sensorien avulla virtuaaliodellisuusympäristöissä. Tässä tutkimuksessa keskityttiin EDA ja HR tutkimukseen. Fysiologiset signaalit jaettiin eri alueisiin. EDA jaettiin aika- ja taajuusalueelle sekä tapahtumiin liittyviin ominaisuusalueisiin. HR jaettiin vain aika- ja taajuusalueisiin. Nämä alueet jaettiin vielä 32 ominaisuuteen, jotka voi nähdä tutkimuksen taulukosta 1 [32]. Tämän tutkimuksen perusteella mitä useampia ominaisuuksia käytetään, sitä parempi arvion tarkkuus. Kun käytetään seitsemää ominaisuutta, tarkkuus on noin 98 %, kun käytössä on kaksi tasoa ahdistukselle, matala ja korkea. Jos käytössä on kolme ahdistuksen tasoa, matala, lievä ja korkea, seitsemän ominaisuuden tarkkuus on noin 90,5 %.

Mozos et al. [33] tutkimus on selvästi vanhempi julkaisu verrattuna muihin lähteisiin. Tutkimuksessaan he ovat käyttäneet kahta sensoria, bioback:in langaton sensori ranteessa ja sosiometrinen anturi, jota pidetään kaulanauhassa. Langaton sensori mittaa EDA-signaalia, fotopletysmografiasignaalia PPG ja sykkeen vaihtelua HRV. Sosiometrinen anturi mittasi kehon liikkumista, kehon asento ja puhetta. Tutkimuksen perusteella paras luokittelumenetelmä oli AdaBoost, kun $T=300$, joten tätä menetelmää käytettiin signaaleja tutkittaessa. Mielenkiintoisesti Mozos et al. [33] saivat parasta näyttöä, kun käytettiin molempien sensorien signaaleja yhdessä. Heidän tutkimuksensa perusteella, jos käytetään vain fysiologisia signaaleja, saatiin huonoimmat tulokset, tarkkuus oli 79 %. Jos käytetään vain sosiologisia signaaleja, tarkkuus nousi 89 %. Kun käytettiin molempia signaaleja, tarkkuus nousi 94 %. Kuitenkin uusimmissa tutkimuksissa on todettu, että liikkuvuus johtaa huonompiin luokittelutuloksiin kuin fysiologiset signaalit [1], [7]. Tässäkin tutkimuksessa oli käytetty konfuusiomatriisia ja sen avulla laskettu tarkkuus, sisäinen tarkkuus ja herkkyys. Olisi mielenkiintoista nähdä, että olisinko itse pystynyt käyttämään konfuusiomatriisia ja sen mittareita.

4.3. Yhteenveto

Useassa muissa tutkimuksissa on käytetty jotain luokittelualgoritmia, jotka ovat antaneet hyvää näyttöä stressin tunnistuksesta. Usein tulokset olivat esitetty konfuusiomatriisin avulla. Olisi ollut mielenkiintoista kokeilla omassakin työssäni luokittelualgoritmien käyttöä sekä koneoppimista, mutta todennäköisesti dataa oli liian vähän tähän tarkoitukseen. Näiden menetelmien käyttäminen olisi mahdollistanut tulosten paremman analysoinnin konfuusiomatriisin avulla.

5. YHTEENVETO

Tässä työssä käsiteltiin fysiologisia signaaleja, stressin tunnistusta kyseisistä signaaleista sekä yritettiin löytää korrelaatiota tenttitulosten ja fysiologisten signaalien välillä. Stressin tunnistusta käsiteltiin muiden tutkimusten perusteella. Fysiologisia signaaleja ja niiden korrelaatiota käsiteltiin sekä käytännöllisesti, että muiden tutkimusten perusteella. Kandidaatintyötä varten tutustuin aiheeseen liittyvään kirjallisuuteen, käytettyyn laitteistoon sekä signaalien käsittelyyn tarvittavaa ohjelmointia Pythonilla.

Työssä käytettiin Amin et al. [4] julkaisemaa fysiologisten signaalien Open Access tietokantaa. Tietokannassa käytettiin nyt jo vanhentunutta Empatica E4 ranneketta, jonka antamiin signaaleihin tässä työssä tutustuttiin. Tässä työssä käsiteltiin tietokannasta viittä eri signaalia; elektrodermaalista aktiivisuutta EDA, sydämen sykettä HR, ihon lämpötilaa TEMP, veren tilavuuspulssia BVP sekä lyöntien välistä aikaväliä IBI. Tietokannan data olivat kaikki CSV-tiedostoja, joten niiden ohjelmointi Pythonin kanssa oli yksinkertaista.

Omien tulosten perusteella tenttituloksilla ei ole minkäänlaista korrelaatiota biosignaalien kanssa. Oman tutkimuksen tulosten perusteella osien fysiologisten signaalien, HR, IBI ja EDA, perusteella opiskelijat ovat stressaantuneita ennen tenttiä eivätkä tentin aikana. TEMP ja BVP antavat taas näyttöä sille, että opiskelijat olivat stressaantuneimpia tenttien aikana. Löytyi myös näyttöä hyvästä ja huonosta stressistä, esimerkiksi kuva 9, jossa on välikokeen 1 parhaimman ja huonoimman tuloksen saaneet opiskelijat, mutta signaalit ovat hyvin samanlaiset. Parempia tuloksia olisi voinut saada, jos olisin käyttänyt haastavampia signaalien käsittelymenetelmiä.

Muista tutkimuksista nähtiin, että on mahdollista saada hyvää näyttöä stressin tunnistuksesta fysiologisista signaaleista. Muissa tutkimuksissa oli käytetty useita eri luokittelumenetelmiä, jotka tuottivat hyviä tuloksia. Tutkimuksissa oli myös käytetty konfuusiomatriisia, joiden mittareista oli saanut hyvää näyttöä stressin tunnistukselle.

6. LÄHTEET

- [1] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, “Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection,” Oct. 2018, doi: 10.1145/3242969.3242985.
- [2] M. K. Moser, B. Resch, and M. Ehrhart, “An Individual-Oriented Algorithm for Stress Detection in Wearable Sensor Measurements,” *IEEE Sens J*, vol. 23, no. 19, pp. 22845–22856, Oct. 2023, doi: 10.1109/JSEN.2023.3304422.
- [3] G. Vos, K. Trinh, Z. Sarnyai, and M. Rahimi Azghadi, “Ensemble machine learning model trained on a new synthesized dataset generalizes well for stress prediction using wearable devices,” *J Biomed Inform*, vol. 148, Dec. 2023, doi: 10.1016/j.jbi.2023.104556.
- [4] Md. Rafiul Amin, D. S. Wickramasuriya, and R. T. Faghih, “A Wearable Exam Stress Dataset for Predicting Grades using Physiological Signals,” in *2022 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)*, IEEE, Mar. 2022, pp. 30–36. doi: 10.13026/wn66-m424.
- [5] “E4 wristband,” Empatica Inc. Accessed: Nov. 30, 2024. [Online]. Available: <https://support.empatica.com/hc/en-us/categories/200023126-E4-wristband>
- [6] “E4 wristband technical specifications,” Empatica Inc. Accessed: Dec. 03, 2024. [Online]. Available: <https://support.empatica.com/hc/en-us/articles/202581999-E4-wristband-technical-specifications>
- [7] P. Siirtola, “Continuous Stress Detection Using the Sensors of Commercial Smartwatch,” 2019, doi: 10.1145/nnnnnnn.nnnnnnn.
- [8] M. Gjoreski, H. Gjoreski, M. Luštrek, and M. Gams, “Continuous stress detection using a wrist device - in laboratory and real life,” in *UbiComp 2016 Adjunct - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Association for Computing Machinery, Inc, Sep. 2016, pp. 1185–1193. doi: 10.1145/2968219.2968306.
- [9] S. Saganowski *et al.*, “Consumer wearables and affective computing for wellbeing support,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Dec. 2020, pp. 482–487. doi: 10.1145/3448891.3450332.
- [10] L. Rachakonda, S. Mohanty, E. Kougianos, and P. Sundaravadivel, “Stress-Lysis: A DNN-Integrated Edge Device for Stress Level Detection in the IoMT,” *IEEE TRANSACTIONS ON CONSUMER ELECTRONICS*, VOL. 65, NO. 4. Accessed: Jan. 17, 2023. [Online]. Available: <https://ieeexplore-ieee-org.pc124152.oulu.fi:9443/stamp/stamp.jsp?tp=&arnumber=8830476>
- [11] S. Campanella, A. Altaleb, A. Belli, P. Pierleoni, and L. Palma, “A Method for Stress Detection Using Empatica E4 Bracelet and Machine-Learning Techniques,” *Sensors*, vol. 23, no. 7, Apr. 2023, doi: 10.3390/s23073565.
- [12] E. E. Van Voorhees *et al.*, “Ambulatory Heart Rate Variability Monitoring: Comparisons between the Empatica E4 Wristband and Holter Electrocardiogram,” *Psychosom Med*, vol. 84, no. 2, pp. 210–214, 2022, doi: 10.1097/PSY.0000000000001010.
- [13] H. Eerola, “Sydämen EKG:n pitkäaikaisrekisteröinti eli Holter-tutkimus - Terveyskirjasto,” Duodecim. Accessed: Oct. 04, 2024. [Online]. Available: <https://www.terveyskirjasto.fi/snk99009>
- [14] N. Chalabianloo, Y. S. Can, M. Umair, C. Sas, and C. Ersoy, “Application level performance evaluation of wearable devices for stress classification with

- explainable AI,” *Pervasive Mob Comput*, vol. 87, Dec. 2022, doi: 10.1016/j.pmcj.2022.101703.
- [15] S. De Vries *et al.*, “Accurate Stress detection from Novel Real-Time Electrodermal Activity Signals and Multi-task Learning Models,” in *Cognitive Computing and Internet of Things*, AHFE International, 2022. doi: 10.54941/ahfe1001845.
- [16] A. Giorgi *et al.*, “Wearable technologies for mental workload, stress, and emotional state assessment during working-like tasks: A comparison with laboratory technologies,” *Sensors*, vol. 21, no. 7, Apr. 2021, doi: 10.3390/s21072332.
- [17] S. Costantini *et al.*, “Wrist-Worn Sensor Validation for Heart Rate Variability and Electrodermal Activity Detection in a Stressful Driving Environment,” *Sensors (Basel)*, vol. 23, no. 20, Oct. 2023, doi: 10.3390/s23208423.
- [18] A. Sanchez-Comas, K. Synnes, D. Molina-Estren, A. Troncoso-Palacio, and Z. Comas-González, “Correlation analysis of different measurement places of galvanic skin response in test groups facing pleasant and unpleasant stimuli,” *Sensors*, vol. 21, no. 12, Jun. 2021, doi: 10.3390/s21124210.
- [19] M. Gil-Martin, R. San-Segundo, A. Mateos, and J. Ferreiros-Lopez, “Human Stress Detection with Wearable Sensors Using Convolutional Neural Networks,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 37, no. 1, pp. 60–70, Jan. 2022, doi: 10.1109/MAES.2021.3115198.
- [20] S. Mahadura, K. Kammangoda Mudalige, A. Tasnim Preoty, S. Munasinghage, and A. Fernando, “Data Mining Project: Stress Dataset for Predicting Cognitive Performance.”
- [21] B. Farnsworth, “What is EDA? And how does it work?,” iMotions. Accessed: Aug. 15, 2024. [Online]. Available: <https://imotions.com/blog/learning/research-fundamentals/eda/>
- [22] R. Jaafar and O. Chung Xian, “Analysis of Heart Rate Variability Using Wearable Device,” in *Lecture Notes in Electrical Engineering*, Springer Science and Business Media Deutschland GmbH, 2021, pp. 453–461. doi: 10.1007/978-981-33-4069-5_37.
- [23] “E4 data - IBI expected signal,” Empatica Inc. Accessed: Nov. 30, 2024. [Online]. Available: <https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal>
- [24] S. Ollander, C. Godin, A. Campagne, and S. Charbonnier, “A comparison of wearable and stationary sensors for stress detection,” in *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings*, Institute of Electrical and Electronics Engineers Inc., Feb. 2017, pp. 4362–4366. doi: 10.1109/SMC.2016.7844917.
- [25] Y. Liu, Y. Mu, K. Chen, Y. Li, and J. Guo, “Daily Activity Feature Selection in Smart Homes Based on Pearson Correlation Coefficient,” *Neural Process Lett*, vol. 51, no. 2, pp. 1771–1787, Apr. 2020, doi: 10.1007/s11063-019-10185-8.
- [26] S. Nakagawa, P. C. D. Johnson, and H. Schielzeth, “The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded,” *J R Soc Interface*, vol. 14, no. 134, Sep. 2017, doi: 10.1098/rsif.2017.0213.
- [27] F. I. Indikawati and S. Winiarti, “Stress Detection from Multimodal Wearable Sensor Data,” in *IOP Conference Series: Materials Science and Engineering*,

- Institute of Physics Publishing, Mar. 2020. doi: 10.1088/1757-899X/771/1/012028.
- [28] J. Graft, W. Romine, B. Watts, N. Schroeder, T. Jawad, and T. Banerjee, “A Preliminary Study of the Efficacy of Using a Wrist-Worn Multiparameter Sensor for the Prediction of Cognitive Flow States in University-Level Students,” *Sensors*, vol. 23, no. 8, Apr. 2023, doi: 10.3390/s23083957.
- [29] P. Siirtola and J. Rönning, “Comparison of regression and classification models for user-independent and personal stress detection,” Aug. 02, 2020, *MDPI AG*. doi: 10.3390/s20164402.
- [30] A. Kumar, K. Sharma, and A. Sharma, “Hierarchical deep neural network for mental stress state detection using IoT based biomarkers,” *Pattern Recognit Lett*, vol. 145, pp. 81–87, 2021, doi: 10.1016/j.patrec.2021.01.030.
- [31] J. H. Hearn and C. J. Stocker, “Mindfulness practice correlates with reduced exam-induced stress and improved exam performance in preclinical medical students with the ‘acting with awareness’, ‘non-judging’ and ‘non-reacting’ facets of mindfulness particularly associated with improved exam performance,” *BMC Psychol*, vol. 10, no. 1, Dec. 2022, doi: 10.1186/s40359-022-00754-3.
- [32] L. Petrescu *et al.*, “Integrating biosignals measurement in virtual reality environments for anxiety detection,” *Sensors (Switzerland)*, vol. 20, no. 24, pp. 1–32, Dec. 2020, doi: 10.3390/s20247088.
- [33] O. M. Mozos *et al.*, “Stress Detection Using Wearable Physiological and Sociometric Sensors,” *Int J Neural Syst*, vol. 0, no. 0, pp. 1–17, 2005.