**Mandaat Afstudeeronderzoek**

Stress herkenning

op basis van

biometrische data

Naam student: Rhea Hau

Studentnummer: 0850154

Naam peer: Juno Verweerd / Roy Buitenhuis

Naam opdrachtgever: ATOS Amstelveen

Naam bedrijfsbegeleider: Michel Metselaar

Datum: 21 September 2018

Status: Herkansing

# Voorwoord

Mijn naam is Rhea Hau en studeer momenteel Technische Informatica aan de Hogeschool Rotterdam. In de afgelopen periode, tussen 1 september 2017 en 31 januari 2018, heb ik voor de minor *Data Science* (datawetenschap) gekozen. Daardoor kreeg ik een steeds grotere interesse voor big data, verschillende wiskundige opdrachten en het leren toepassen van machine learning, wat nu heel erg populair is bij bedrijven. In de toekomst wil ik ook graag goed met machine learning om kunnen gaan, zodat ik als *Data Scientist* de opdrachten van mijn opdrachtgevers beter kunnen uitvoeren.

Dit verslag bevat het onderzoeksvoorstel, die gemaakt is voor de start van de afstudeerstage start. Ik heb mijn afstudeeropdracht gevonden bij het bedrijf ATOS. ATOS is een IT-bedrijf die zich bezighoudt met allerlei systeemontwikkelingen. Het project waar ik voor werk, is momenteel alleen binnen ATOS zelf. In het hoofdstuk ‘2.0 Opdracht omschrijving’ zal ik meer vertellen over wat mijn opdracht inhoudt en wat mijn onderzoeksvoorstel wordt.

Voor de Hogeschool Rotterdam moet ik als student ook de vijf competenties (beheren, ontwerpen, adviseren, analyseren en realiseren) op eindniveau behalen aan het einde van de stage. Deze vijf activiteiten worden later in de scriptie beschreven.

Met veel dank aan mijn docenten en mijn klasgenoten die mij geholpen hebben met het waar maken van een afstudeeropdracht waarin kunstmatige intelligentie voorkomt. Hierbij kan ik beter beeld krijgen voor mijn toekomstige carriere.

# Samenvatting

Deze afstudeeropdracht houdt in dat ik een onderzoek ga doen over stress detecteren op basis van de biometrische data die vanuit de sensoren worden opgevraagd. De data die opgevraagd worden, betekent niet direct stress. Daarvoor is er veel onderzoek nodig. Bijvoorbeeld onderzoeken naar wat stress überhaupt is. Vervolgens wordt er gekeken naar welke biometrische data hebben de bestaande wetenschappelijke onderzoeken gebruikt om stress te detecteren. Daarna wordt er een selectie gemaakt welke sensoren die specifieke biometrische data kunnen leveren.

Deze opdracht past goed bij mijn voorkeur en mijn opleiding. Enerzijds is er een technisch onderdeel waar de docenten aan het einde van de stage mijn opdracht kunnen beoordelen. Anderzijds is het leren van machine learning van belang, omdat die nu steeds populair wordt in de bedrijven wereld. Voor mijn toekomst wil ik meer weten over het toepassen van machine learning tot deep learning. Tijdens mijn opleiding heb ik het vak kunstmatige intelligentie gehad en dat sluit goed aan met deze richting. Deze afstudeeropdracht bevat heel veel verschillende onderdelen naast het verdiepen in Data Science. Bijvoorbeeld het programmeren, communicatie en de overdracht tussen de biometrische data en de ander werkomgeving en het aansluiten van sensoren aan de microcontroller en de huid om metingen te maken. Het vergt heel veel werk voor een korte termijn van vier maanden. In *figuur 3* is te zien welke taken horen bij elk van de vijf competenties. Hiermee kan ik de eindkwalificatie behalen.

De grootste uitdaging van dit project is het stress detecteren met behulp van verschillende classificatie algoritmes, omdat ik de basis van machine learning ken vanuit mijn minor.

Table of Contents

[Voorwoord 2](#_Toc526474126)

[Samenvatting 3](#_Toc526474127)

[Deel 1 | Inleiding 5](#_Toc526474128)

[**1.1** **Aanleiding opdracht** 5](#_Toc526474129)

[**1.2** **Probleemstelling** 5](#_Toc526474130)

[**1.3** **Doel van de opdracht** 6](#_Toc526474131)

[**1.4** **Stakeholders** 6](#_Toc526474132)

[Deel 2 | Opdrachtomschrijving 7](#_Toc526474133)

[**2.1** **Globale opdrachtomschrijving** 7](#_Toc526474134)

[**2.2** **Scope** 7](#_Toc526474135)

[**2.3** **Hoofdvraag en deelvragen** 8](#_Toc526474136)

[**Hoofdvraag: Hoe ziet een algoritme eruit dat, op basis van voldoende biometrische data, stress bij de gebruiker kan detecteren?** 8](#_Toc526474137)

[**Deelvraag 1: Welke biometrische data is er nodig om stress te herkennen?** 8](#_Toc526474138)

[**Deelvraag 2: Welke sensoren zijn er beschikbaar om stress te kunnen herkennen?** 9](#_Toc526474139)

[**Deelvraag 3: Wat voor classificatie algoritme is geschikt om stress te vinden met zo min mogelijk ruis?** 9](#_Toc526474140)

[**Deelvraag 4: Welke privacy gerelateerde aspecten spelen een rol bij de gebruikte data?** 10](#_Toc526474141)

[**2.4** **Deliverables** 10](#_Toc526474142)

[Deel 3 | Theoretisch kader 11](#_Toc526474143)

[**3.1 Onderzoek probleemstelling** 11](#_Toc526474144)

[**3.2 Bestaand onderzoek I** 13](#_Toc526474145)

[*“Wearable Sensor Based Stress Management Using Integrated Respiratory and ECG Waveforms” – IEEE Xplore [1]* 13](#_Toc526474146)

[**3.3 Bestaand onderzoek II** 15](#_Toc526474147)

[*“Stress detection in working people” – ScieneDirect [2]* 15](#_Toc526474148)

[Deel 4 | Onderzoeks- en/of implementatiemethode 16](#_Toc526474149)

[Deel 5 | Planning 18](#_Toc526474150)

[**5.1** **Planning Opzet** 18](#_Toc526474151)

[**5.2** **Risicolog** 21](#_Toc526474152)

[Deel 6 | Verantwoording 24](#_Toc526474153)

[Deel 7 | Literatuurlijst 26](#_Toc526474154)

# Deel 1 | Inleiding

## **1.1 Aanleiding opdracht**

ATOS is een detacheringsbedrijf en is continu bezig met oplossingen bedenken van IT-gerelateerde (Information Technology) problemen van zijn klanten. Eén van de problemen die er altijd al een rol speelt, is de stress op de werkvloer. ATOS wil zich daarin verdiepen en meer ontdekken. Het bedrijf heeft momenteel een idee en een bestaande mobiele applicatie, genaamd CHO (Chief Health Officer). Deze CHO-applicatie moet feedback gaan geven aan de gebruiker met stress. De naam “CHO” staat ook voor een visuele assistent in de applicatie zelf. Hij zal met de gebruiker meedenken aan een oplossing om het stressniveau te verlagen. Daarnaast zal “CHO” een voorstel kunnen geven aan de gebruiker, zoals een wandeling maken buiten voor vijf minuten lang. Het idee achter de applicatie is duidelijk, maar deze applicatie krijgt momenteel nog geen biometrische data binnen. Daardoor kan het verder geen informatie verwerken.

Mijn stagebegeleider heeft mij een opdracht gegeven om de biometrische data van stress te vinden. Deze data kan vervolgens gekoppeld worden aan hun applicatie. De biometrische data kan vanuit de sensoren komen of van wearable devices. Er moet een algoritme toegepast worden om stress te vinden vanuit de gemeten data. Op deze manier kan het resultaat dienen als input voor hun mobiele applicatie. De mobiele applicatie is persoonlijk verbonden met alleen gegevens van de gebruiker. Hierbij kunnen er persoonlijke feedback gegeven worden.

## **1.2 Probleemstelling**

ATOS heeft een mobiele applicatie CHO ontwikkeld zonder enige data. De mobiele applicatie kan daardoor nog niet getest worden op eigen medewerkers om hun stressniveau op werk te verlagen. ATOS wil het graag tijdelijk intern houden, omdat dit nog om een “proof of concept” gaat. Stress is niet goed voor de medewerkers en dus niet goed voor het bedrijf. De kwaliteit van de medewerkers gaat achteruit en er kan veel ziekteverzuim ontstaan. Stress maakt mensen dus niet blijer op de werkvloer en zij doen hun werk niet meer met plezier. Er moet een nieuwe aanpak komen om stress te verminderen. Mijn stagebegeleider heeft het als volgt verteld:

*“Voor de employee experience dienstverlening moeten er algoritmen ontwikkeld worden die gedeelde fysiologische gegevens interpreteren.” – Michel Metselaar (stagebegeleider)*

## **1.3 Doel van de opdracht**

Het doel is een wearable device of een smartwatch te gebruiken om biometrische data te verzamelen en daarmee verder stress te laten herkennen door de het algoritme. Er zijn momenteel nog geen smartwatches op de markt die specifiek stress en de accuraatheid van stress kunnen meten. Hierbij is er een onderzoek voor nodig. Het onderzoek moet gericht zijn op verschillende health sensoren die geschikt zijn om stress te kunnen meten. Het resultaat van de stress data kan bij bepaalde percentage goedgekeurd worden voor het volgende onderzoek. ATOS kan hierbij het systeem verder uitwerken en de sensoren zo bouwen dat die in een wearable device passen. Het prototype wordt gemaakt door mij, zodat het bedrijf ook daadwerkelijk iets mee kan werken in de toekomst.

## **1.4 Stakeholders**

Bij dit project zijn er een aantal stakeholders die invloed hebben op het resultaat. De opdrachtgever, ofwel de eigenaar van de CHO-applicatie van ATOS, die de opdracht bedenkt biedt de stagiaire hun bestaande werkomgevingen en informatie aan. Bijvoorbeeld gebruik maken van betaalde software of Cloud platforms en verdere informatie die belangrijk is om aan een goed beeld te krijgen wat de opdrachtgever wil. Mijn stagebegeleider is ook een belangrijk persoon die mij vanaf het begin van de stageperiode tot einde begeleidt. De school heeft naast het stagebedrijf ook belangen aan dit project. Hier wordt op de vijf competenties (beheren, analyseren, ontwerpen, realiseren, adviseren) beoordeeld over de student. De school bepaalt uiteindelijk het eindcijfer voor dit onderzoek en het prototype.

Het prototype die gemaakt wordt door mij, wordt aan het einde van de stage opgeleverd aan het stagebedrijf en de school. Proefpersonen zijn ook van belang voor dit project die willen deelnemen aan het onderzoek.

# Deel 2 | Opdrachtomschrijving

## **2.1 Globale opdrachtomschrijving**

Er zijn nog weinig bedrijven die aan de stress van de medewerkers werken. ATOS neemt initiatief en komt met een idee om het stressniveau van de medewerkers mogelijk te verlagen. ATOS heeft een CHO-applicatie gemaakt als “proof of concept” om te laten beoordelen of het een goed idee is om uit te breiden en op de markt te brengen. Dat doen ze eerst om eigen medewerkers te laten testen, of de applicatie op die manier stressniveau daadwerkelijk verlaagd kan worden. Hierbij is er hulp nodig om een input te geven voordat zij verder mee kunnen. Ik heb als opdracht om de data te verzamelen via sensoren. Daarmee is het schrijven van een programma van belang, want de het algoritme moet zelf de stress zien te vinden in de toekomst. Na mijn prototype wordt er gekeken naar een verbetering. Er is veel meer te doen dan alleen stress herkennen vanuit de data die ik verzamel. Er moet meer onderzoek gedaan worden. Dat wil ATOS heel graag en ook dat ik dan als eerste hun een input kan geven. Hiermee kunnen ze zelf beoordelen hoe ze mee verder kunnen.

## **2.2 Scope**

Voor dit groot project is er een bespreking over de afbakening noodzakelijk. Mijn deel is slechts een klein stuk van het geheel voor vier volle maanden. Mijn opdracht is het zorgen voor data dat betrekking heeft tot stress, en dit te leveren als input voor hun mobiele applicatie.

Mijn verantwoordelijkheden:

1. Voor het onderzoek wordt er onderzocht naar betaalbare sensoren die aansluitbaar zijn met eenzelfde microcontroller, mocht de wearable device niet de goede oplossing zijn voor dit onderzoek.
2. Voor het onderzoek wordt er naar de minimale biometrische data die er nodig is om stress te kunnen herkennen. Hoe meer sensoren hoe accurater de metingen worden, maar dat kan ATOS in de toekomst zelf instemmen om uit te breiden.
3. Voor het onderzoek wordt er een prototype gebouwd in de vorm van een werkende stressherkenning en de bijbehorende data.
4. Voor dit onderzoek is het herkennen van stress 40% accuraatheid ook acceptabel.

Taken voor anderen:

1. Voor dit project wordt er geen applicatie gemaakt voor de gebruiker, want een exemplaar is er al uitgebracht door ATOS zelf.
2. Voor het onderzoek wordt niet alle interne medewerkers getest. Als prototype pak ik een klein aantal mensen die aan mijn onderzoek willen helpen om de accuraatheid te verbeteren en de conclusie daarvan.
3. Voor het onderzoek wordt er geen rekening gehouden met gebruikers die bepaalde ziekte hebben of welke leeftijdsgroep zij zich bevinden.

## **2.3 Hoofdvraag en deelvragen**

### **Hoofdvraag: Hoe ziet een algoritme eruit dat, op basis van voldoende biometrische data, stress bij de gebruiker kan detecteren?**

Het doel van dit onderzoek is om de stress te kunnen detecteren op basis van de biometrische data die gemeten wordt vanuit de wearable device of van verschillende losse sensoren. Maar heeft een reguliere wearable device genoeg sensoren om stress te kunnen detecteren? Zo niet, welke sensoren kunnen we wel gebruik van maken om biometrische data op te meten? Zo ja, hoe wordt de data dan direct overgebracht naar een andere werkomgeving? Als de data opgevraagd en ontvangen zijn, op welke manier wordt de data dan zo gescheiden dat er een stress patroon te herkennen is door de het algoritme? Allerlei vragen over eenzelfde doel om stress te kunnen detecteren. Hieronder volgen de deelvragen met de bijbehorende onderbouwingen.

### **Deelvraag 1: Welke biometrische data is er nodig om stress te herkennen?**

Iemand met stress heeft bijzondere gevoelens bij. Het gevolg van beginnende stress kan snelle hartkloppingen veroorzaken of een zweterige huid aan de handpalm. Hierbij wordt er onderzocht met behulp van wetenschappelijke artikelen welke sensoren worden gebruikt om stress te kunnen detecteren. Hierbij wordt ook gekeken naar de minimale biometrische data die er nodig is, omdat het niet rendabel qua kosten en tijd om een onderzoek te doen welke uitgebreide sensoren ook stress mogelijk te detecteren. De focus voor dit project is puur stress detecteren en het te versturen naar het bestaande applicatie. Het stagebedrijf kan deze kennis en het idee uitbreiden voor hun verdere onderzoek.

Daarnaast is het onderscheiden tussen wel of niet sportende gebruikers van belang om te vergelijken hoe stress bij hen werken. Uiteindelijk moet er een selectie komen van welke sensoren er gebruikt kunnen worden voor dit onderzoek.

### **Deelvraag 2: Welke sensoren zijn er beschikbaar om stress te kunnen herkennen?**

Hiermee moet er onderzocht worden welke sensoren die beschikbaar zijn om voor het onderzoek te kunnen gebruiken. Er wordt gekeken naar de beschikbaarheid van de data exporteren en er wordt gelet op de prijsklassen. Opdrachtgever wil liefst een smartwatch hebben, maar vanwege te krappe tijd om in te kopen, kan dit niet waarmaken. Er wordt wel een onderzoek gedaan of een smartwatch daadwerkelijk stress kunnen meten in theorie.

Het is niet aangeraden om smartwatches in combinatie te gaan met losse sensoren in verband met de data exporteren. Dat kan namelijk heel vervelend zijn, omdat dit gestructureerd werken is met verschillende data. Het zal geen baat hebben bij deze. Hierin wordt er een selectie gemaakt van losse sensoren voor dit onderzoek om toe te passen.

### **Deelvraag 3: Wat voor classificatie algoritme is geschikt om stress te vinden met zo min mogelijk ruis?**

Hierin wordt onderzocht hoe stress gedetecteerd wordt vanuit de gemeten biometrische data. Bij sommige wetenschappelijke artikelen hebben de onderzoekers verschillende machine learning algoritmes toegepast, omdat het werken met data met machine learning bepaalde patronen kunnen vinden om stress te kunnen bepalen. Er zijn ook onderzoeken geweest die alles in het grafiek brengen en daarmee de stress detecteren. Deze deelvraag neem ik mee naar mijn onderzoek waar verschillende manieren bestaan om stress te kunnen detecteren.

### ***Deelvraag 4: Welke privacy gerelateerde aspecten spelen een rol bij de gebruikte data?***

*De data die wordt gemeten vanuit een persoon kan heel gevoelig zijn voor een ander die mogelijk misbruik kunnen maken. Dus hierbij is het gedachte over de privacy niet weg te denken. Er wordt hier een onderzoek gedaan over wat de rol binnen dit project kan zijn. Hoe kijken gebruikers en de opdrachtgever tegenaan en welke aspecten moeten zeker besproken worden voor nodige oplossingen?*

*Mogelijk dit weg te halen ivm tijdsgebrek aan dit onderdeel?*

## **2.4 Deliverables**

De opleverset voor ATOS is mijn prototype in de vorm van werkende sensoren of een wearable device die de biometrische data kan opmeten en daarmee de stress kunnen detecteren. Daarnaast zijn mijn programmeercode en allerlei verschillende rapporten ook belangrijk voor de uitbreiding van CHO-applicatie.

De opleverset voor school zijn mijn onderzoeksvoorstel, mijn scriptie met behorende programmeercode en testrapporten. In de scriptie laat ik ook zien hoe ik in staat ben om de vijf competenties zelfstandig kunnen uitvoeren.

# Deel 3 | Theoretisch kader

## **3.1 Onderzoek probleemstelling**

Voor dit onderzoek is er naar de probleemstelling gekeken. Het probleem is dat er nog niet bekend is welke data en welke sensoren nodig zijn voor het onderzoek. Er zijn veel wetenschappelijke onderzoeken geweest met verschillende sensoren om stress te detecteren. Voor het huidige onderzoek zijn slechts drie sensoren gebruikt voor de stressdetectie. De informatie van verschillende wetenschappelijke artikelen kunnen wel een goede bijdrage hebben aan dit onderzoek. Ook worden de minimumeisen aan data gebruikt om stress te detecteren. Daarmee kan dit huidige onderzoek een bijdrage hebben aan een nieuwe conclusie.

*Daarbij gebruiken de onderzoekers ook diverse classificatie algoritmes om een vergelijking te maken welke geschikt is en welke het snelst de stress kan detecteren.*

In *figuur 1* is er een diagram getekend hoe het proces verloopt van het gehele project. Dit is een globaal overzicht en in de scriptie zal de concrete versie staan. Er wordt dus eerst een planning gemaakt voor het project. Vervolgens wordt er heel veel onderzoek gedaan aan het begin van de stageperiode. Wat is stress? Wat houdt stress in? Welke biometrische data speelt een grote rol tijdens de stress? Welke sensoren zijn er op de markt om aan die data te komen? Hoe worden de sensoren aangesloten? Hoe worden de sensoren zo gekalibreerd dat die accuraat kunnen meten, etc. Sensoren moeten tijdig ingekocht worden en data moet overgebracht worden naar een ander werkomgeving. Daarna wordt er onderzocht welke classificatie algoritme geschikt is om stress te detecteren. Hierbij wordt eerst getest met *training set* totdat de stress detecteren steeds accurater wordt. Vervolgens kan de het algoritme met die kennis met een nieuwe dataset als *test set* gebruiken.



Fig. 1 Diagram van gehele proces voor huidige onderzoek stress detectie op basis van sensoren en data

## **3.2 Bestaand onderzoek I**

### *“Smart Wearable Band for Stress Detection” – IEEE Xplore [1]*

Er zijn verschillende parameters die gebruikt kunnen worden om stress te detecteren. Bijvoorbeeld spierspanningen, pupil diameter, hartslagvariabiliteit, elektro-encefalografie om brainwave met betrekking tot stress te meten, cortisol en huidgeleiding. Draagbare apparaten voor de gezondheidszorg die gebaseerd zijn op IoT (Internet of Things) en fitness banden zijn beschikbaar om hartslagvariabiliteit te gebruiken om stressniveau te meten. De HRV (hartslagvariabiliteit) kan de variatie laten zien in meeste gevallen. Bijvoorbeeld de hartslag kan hoger zijn van mensen die staan dan wanneer zij zitten. Daardoor is het gebruik maken van alleen hartslagvariabiliteit als indicator om mentale stress te detecteren tot misclassificatie leiden.

HRV en bloeddruk kunnen de activiteiten van sympathisch en parasympatisch zenuwstelsel monitoren. Huidgeleidingwordt beschouwd als een biomarker voor stress. In de metingen van de huidgeleidingis te zien dat de eccriene zweet activiteit alleen beheerd wordt door sympathisch zenuwstelsel. De variëteit van de huidgeleiding hangt af van de zweetafscheiding. Wanneer een persoon onder stress is, wordt het sympathisch zenuwstelsel in actie geplaatst. Zweetafscheiding van zweetklieren vermindert de weerstand van de huid en wordt de huidgeleiding verhoogd. Dus huidgeleiding kan men beschouwen als een indicator voor sympathische activering als gevolg van stress. Verder kan huidgeleiding samen met hartslag en huidtemperatuur als indicator voor activiteiten van het sympathisch zenuwstelsel worden beschouwd. Dat is te zien in *figuur 1*.

Fig. 2 Sympathische zenuwstelsel aciditeit van het onderzoek “Smart Wearable Band for Stress Detection” (blz. 2)

In dit onderzoek gaat over een smartband waarin huidgeleiding sensor, 3-axis accelerometer, bluetooth en microcontroller zijn aangebracht. De huidgeleiding sensor wordt op de onderkant van de pols geplaatst om de fluctuatie te meten. Het signaal indiceert het korte termijn fluctuatie van de huidgeleiding. De accelerometer helpt het classificeren van activiteiten zoals het zitten en het lopen. Lichaamsbeweging stimuleert ook zweetklieren die variatie in huidgeleiding veroorzaken. Hiermee kan men de spanningsdetectie accuraat meten.

De smartband monitort en meet de huidgeleiding van de gebruiker door middel van twee elektroden. De twee elektroden geven een klein schok als ware in dc-voltage door de huid heen en wacht dan op respons. De smartband houdt de uitvoer in de vorm van voltage bij. Vervolgens wordt er verschillende algoritmes gebruikt om de data te analyseren, terwijl andere parameters in overweging worden gebracht om de stress conditie van de gebruiker te voorspellen.

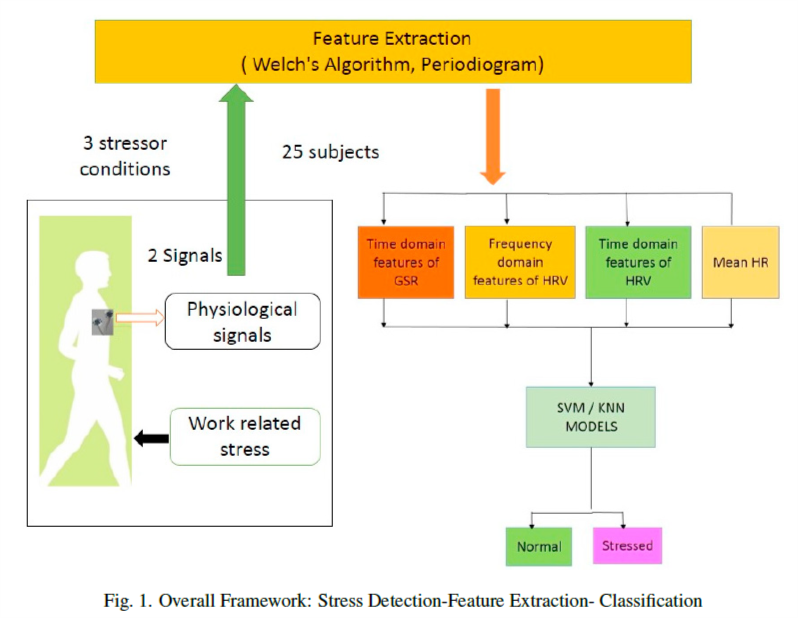
Tot slot is deze smartband geschikt om het stressniveau continu te monitoren tijdens de dagelijkse activiteiten. Bijvoorbeeld alert berichten versturen via de mobiele telefoon, stress hanteren en de data van de gezondheidszorg met de dokters en familieleden te delen. [1]

*Nemen we temperatuur ook mee dan aan het onderzoek?*

## **3.3 Bestaand onderzoek II**

### *“Stress detection in working people” – ScieneDirect [2]*

In dit onderzoek wordt met GSR (galvanische huidreactie), HRV (hartslagvariabiliteit), HR (hartslag) gewerkt om stress te detecteren. Volgens dit onderzoek hebben er verschillende literatuurstudies besproken over stressdetectie, omdat stress een belangrijk factor is voor de leefstijl van elk individu. Ghaderi et al. heeft voor zijn onderzoek stress kunnen detecteren met behulp van ademhaling, hartslag, elektromyografie, galvanische huisreactie voor voeten en handen. De ademhaling heeft aanzienlijk betrekking tot stress. En dat is terug te zien in zijn onderzoek.

fig. 3 Framework voor stress detectie, classificatie voor kenmerken extraheren.

Er worden dus GSR-, HRV- en HR-sensoren geplaatst op de persoon. GSR-sensor kan verschillende niveaus van de huidgeleiding detecteren wanneer een persoon in stress is. Het moment dat een persoon onder stress is, het zenuwstelsel reageert direct met zweten. Voor dit onderzoek zijn er elektroden geplaatst onder de vingers om de weerstand te meten. Terwijl ECG (elektrocardiografie) de dominante identificatie is voor stressdetectie voor zowel lang als korte termijn. ECG is het meten van elektrische activiteit van het hart op basis van de HRV.

De benodigde kenmerken worden opgemeten via de ECG en GSR sensoren die worden weer toegepast als niet-parametrische en parametrische classificatoren. Voor de niet-parametrische classificator wordt de KNN (K-nearest Neighbour) gebruikt. Daarmee kan die het stressniveau vaststellen via de Euclidische afstand tussen de nearest neighbour van de getrainde en geteste feature vectors. Feature vector is een vector die verschillende elementen bevat van eenzelfde soort object.

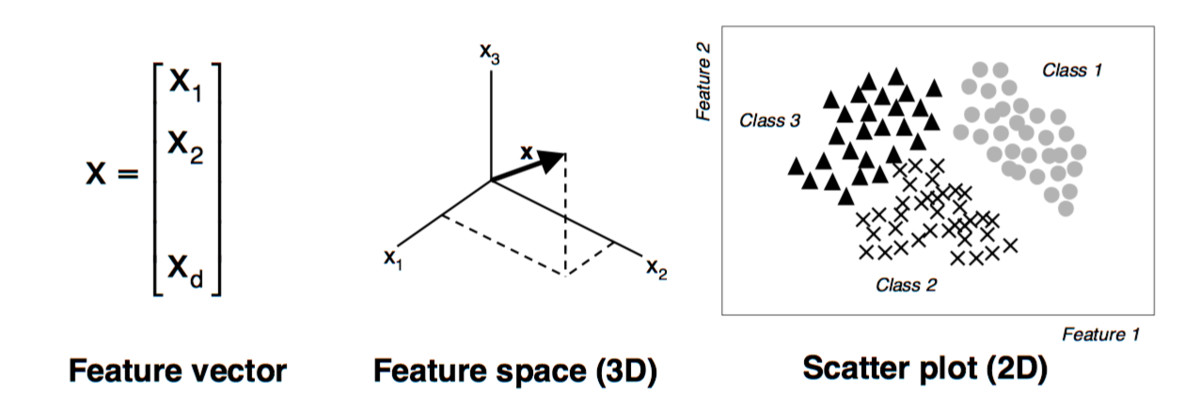


Fig. 4 De basis van feature vectors. Bron: <https://brilliant.org/wiki/feature-vector/>

Voor de parametrische classificator wordt er een SVM (Support Vector Machine) gebruikt waarin twee type kernel methode worden gebruikt. Een kernel methode wordt gebruikt om problemen die niet-lineair zijn met een lineair classificatie. De data die daarvoor wordt gebruikt, wordt dan in een hogere dimensie gebracht en uitgescheiden. In dit wetenschappelijke onderzoek worden de lineair kernel en RBF (Radial basis Function), die een niet-lineair kerel is, gebruikt. RBF wordt gebruikt om groot aantal data te verwerken naar een grafiek die gladder laat lopen. Het kan de maximum en de minimumwaarden voorspellen.

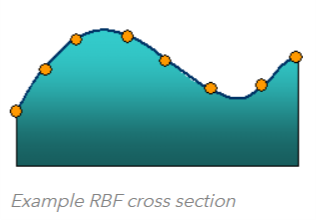
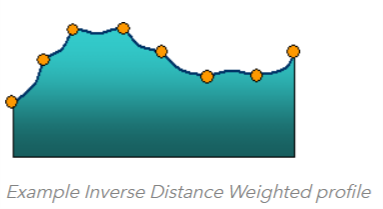


Fig. 5 Voorbeeld hoe RBF het kan helpen met de voorspelling van de waarden.

Voor dit wetenschappelijke onderzoek wordt er een groot deel van een zogenaamd SWELL-KW dataset gegenereerd vanuit de kennis van de medewerkers. Die zijn gebruikt voor het ontwikkelen van een stress detectie algoritme. De gegevens worden verzameld van werkende medewerkers ter plekke. Hierbij wordt de stress in een het algoritmelog bijgehouden door de sensoren die zij dragen, gezichtsuitdrukkingen, lichaamshoudingen via de Kinect 3D sensor, HR, HRV en de ECG-sensoren die huidgeleidingsniveau meten.

Uit het resultaat is gebleken dat de data die via SVM ten opzichte van KNN hogere percentage heeft behaald voor het accuraat meten. SVM in combinatie met een RBF kernel, heeft het onderzoek 92.75% behaald om stress te kunnen detecteren. De data die in combinatie met HR, HRV en GSR zijn dan ook noodzakelijk voor de stress detectie. [2]

# Deel 4 | Onderzoeks- en/of implementatiemethode



Fig. 3 Vijf competenties op volgorde met bijbehorende taken die worden uitgevoerd voor huidige onderzoek. Dit is ook terug te vinden in de scriptie voor meer informatie.

Voor de eerste deel van het onderzoek gebruik ik voor de sensoren en data onderzoeken een literatuuronderzoek methode. Daarmee kan er gekeken worden welke sensoren die ooit waren gebruikt voor onderzoek en welke resultaten uitkwamen. Literatuuronderzoek zoals wetenschappelijke artikelen uit de website van *ScienceDirect* of *IEEE* zijn meest betrouwbaarste bronnen op internet. Hierin worden veel geschreven waarom ze voor die sensoren zijn gegaan en welke methode zij gebruikt hebben om stress te kunnen herkennen.

Vervolgens worden er dus sensoren gekocht aan de hand van de wetenschappelijke artikelen. Daarmee kan er een ontwerp gemaakt worden voordat de sensoren aankomen. Zodra de sensoren er aanwezig zijn, wordt die aangesloten per sensor om te kijken of die werkt en of die goed werkt. Hierbij is kalibreren van belang. Daarvoor zijn er dus ook andere sensoren, bijvoorbeeld smartwatch / fitness band, aanwezig om mee te vergelijken of de data wel ongeveer overeenkomt met de metingen, zoals de hartslag metingen. De data wordt verwerkt in een aparte omgeving waar je ook een grafiek ziet in de vorm van een lijn plot. Hiermee kan er bekeken worden wanneer ik als proefpersoon niet stress op verschillende momenten achter elkaar.

Vervolgens wordt er gekeken welke methode gebruikt kan worden om stress op te wekken. Hiermee wordt er weer gekeken naar de data die zeer waarschijnlijk stress kan betekenen. Er is dan ook een extra persoon nodig om te noteren wanneer de proefpersoon in stress is. De data die omgezet wordt in een grafiek vorm is dan te zien bij welke parameters als stress gezien kan worden. De kans is groot als de lijngrafiek een piek heeft dat er daadwerkelijk stress heeft plaatsgevonden. Daarna wordt de data verzameld en verstuurd naar de cloud toe. In de cloud wordt er verder met de data verwerkt door middel van classificatie algoritme van machine learning.

Voor het tweede deel van het onderzoek wordt er ook literatuuronderzoek gebruikt. De data die binnenkomt vanuit de sensoren zijn gestructureerd en ruis bestaat er altijd tussen de data. Er zal een oplossing moeten komen voor het in structuur aan te brengen, zodat het makkelijk wordt met de verwerkingen. Vervolgens wordt er gekeken naar de bestaande oplossingen die stress kunnen herkennen. Er zijn verschillende classificatie algoritmen in machine learning, zoals *decision tree, k-nearest neighbors*, *artificial* *neural networks* en *naive bayes*, die daarop toegepast kunnen worden. Daarbij worden de meest geschikte classificatie algoritmen die gebruikt zijn in wetenschappelijke onderzoeken om stress te herkennen toegepast aan dit huidige onderzoek. Daarmee wordt er een vergelijking gemaakt. De data wordt dan in de vorm van een tabel met percentage aangegeven gezet. Hiermee wordt er bekeken naar de nauwkeurigheid van de stress detectie. Ten eerste wordt er een *training set* gebruikt. Dit is gebruikelijk bij classificatie algoritmen, want de het algoritme zal het zelf moeten calculeren in de toekomst. Hiermee kan er een deel van de dataset gebruikt worden om handmatig stress te detecteren. Er bestaat altijd wel ruizen in de dataset. Dat wil zeggen dat de data veel afwijkende waarden bevatten die eigenlijk niet veel doet, maar wel invloed kan hebben op conclusie. Wanneer er sprake is van minimaal 40% accuraatheid, dan kan een deel van de dataset voor als *test set* gebruikt worden. Dat wil zeggen dat ander deel van de dataset bewaard wordt voor de het algoritme om die zelf de stress uit de data te detecteren.

# Deel 5 | Planning

## **5.1 Planning Opzet**

|  |  |  |  |
| --- | --- | --- | --- |
| WBS NUMBER | TASK TITLE | START DATE | DUE DATE |
|
| 100% | 90% | 60% | 40% |
| 1 | Algemeen |  |  |
| 1.1 | Gesprek met docenten | 9/12/2018 | 1/30/2019 |
| 1.1.1 | Gesprek met stagebegeleider | 9/4/2018 | 1/29/2019 |
| 1.2 | Gesprek met andere afdelingen school | 9/10/2018 | 1/21/2019 |
| 1.3 | Mandaat schrijven + afronden | 9/3/2018 | 9/21/2018 |
| 1.4 | Scriptie schrijven + afronden | 9/3/2018 | 1/23/2019 |
| 1.5 | Presentaties aan Hogeschool Rotterdam | 11/7/2018 | 1/16/2019 |
| 1.6 | Presentaties aan ATOS | 10/17/2018 | 10/17/2018 |
| 1.7 | Stagebezoek | 10/17/2018 | 10/31/2018 |
| 1.8 | Gesprek met stagebegeleider en manager | 10/1/2018 | 1/28/2019 |
| 2 | Onderzoeksfase |  |  |
| 2.1 | Requirements opstellen. | 9/5/2018 | 9/7/2018 |
| 2.2 | Literatuuronderzoek naar stress herkenning op basis van biometrische data. | 9/10/2018 | 9/14/2018 |
| 2.3 | Selectie van benodigde sensoren of die geschikt zijn om op de microcontroller aan te sluiten. (Omdat niet alle sensoren geschikt zijn voor eenzelfde microcontroller.) | 9/17/2018 | 9/21/2018 |
| 2.4 | Selectie maken van de microcontroller of een smartwatch. (Omdat ik dan weet hoe mijn ontwerp moet uitzien en welke ontwikkelomgeving ik zal gebruiken. Ik weet dan ook welke programmeertaal ik moet gaan focussen.) | 9/17/2018 | 9/21/2018 |
| 2.5 | Onderzoek welke algoritmen er bestaan voor kunstmatige intelligentie. | 10/8/2018 | 10/12/2018 |
| 3 | Ontwerpfase |  |  |
| 3.1 | Ontwerp eventueel van de aansluiting aan de microcontroller. | 9/24/2018 | 9/28/2018 |
| 3.2 | Data verzamelen en daarvan voor elke sensor een grafiek maken. Want in de grafiek kan ik dan zien wanneer stress hoe het uit zou zien. | 9/24/2018 | 9/28/2018 |
| 3.3 | Oplossingsmethode ontwerpen hoe de patroon van stress te herkennen wordt door de het algoritme. | 9/27/2018 | 10/12/2018 |
| 4 | Bouwfase |  |  |
| 4.1 | Sensoren aansluiten op de microcontroller en op een proefpersoon (ik) | 10/1/2018 | 10/5/2018 |
| 4.2 | Realiseren van prototype o.a. programmeren. | 10/15/2018 | 11/2/2018 |
| 5 | Testfase |  |  |
| 5.1 | Stress herkenning testen bij mijzelf. | 10/15/2018 | 10/26/2018 |
| 5.2 | Bij andere proefpersonen opnieuw data verzamelen en grafiek maken. Vervolgens patroon van stress laten herkennen door de het algoritme. | 10/29/2018 | 11/2/2018 |
| 7 | Afronding |  |  |
| 7.1 | Prototype volledig bouwen | 12/3/2018 | 12/14/2018 |
| 7.2 | Examenzitting | 1/23/2019 | 1/30/2019 |

## **5.2 Risicolog**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Risico Beschrijving** | **Kans** | **Impact** | **Risico\*** | **Maatregel** | **😶** | **Status Omschrijving** | **Datum** |
| R1 | Stagebegeleider kan ziek worden. | 2 | 5 | 10 | Zo snel mogelijk en zoveel mogelijk de vragen stellen. Niet te laat ermee beginnen. | N | Stagebegeleider heeft een agenda waar iedereen in kan zien wanneer hij afspraak heeft en niet aanwezig zal zijn op het hoofdkantoor. | Hele periode |
| R2 | Ik kan ziek worden. | 5 | 8 | 40 | Genoeg kleding, eten meenemen voor extra energie. Anders thuis doorwerken als ik echt niet lekker voel. | ? | Er is contact tussen mijn stagebegeleider en ik via het mobiele nummer of eventueel e-mail als er iets is om door te geven. | Hele periode |
| R3 | Niet genoeg data kunnen verkrijgen vanuit de sensoren. | 7 | 8 | 56 | Ontwerp van het prototype aanpassen zodat het niet makkelijk los van de huid gemeten wordt. | ? | Er zijn genoeg gereedschappen om de sensoren zo goed aan de huid te plaatsen voor de juiste meting. Het ontwerp komt nog en zal worden omschreven in de scriptie. | Hele periode |
| R4 | Niet de goede sensoren of sensoren gaan stuk. | 5 | 8 | 56 | Zo snel mogelijk de geschikte kopen en de oude terugsturen. | ? | Zodra de sensoren binnen zijn, worden de sensoren op tijd getest. | Hele periode |
| R5 | Verkeerde keuze maken waardoor ik te lang aan eenzelfde onderzoek hang. | 7 | 9 | 63 | Bespreken met de stagebegeleider of docent voor feedback. | ? | Momenteel zijn er 2-wekelijkse vergadering met mijn stagebegeleider en mijn manager om over de voortgang van het project te bespreken. | Hele periode |
| R6 | Niet accuraat kunnen meten met de sensoren. | 7 | 8 | 56 | Kalibreren met een bestaande sensor of een smartwatch / fitness band die bepaalde sensor ook heeft. | ? | Momenteel is er een alternatieve smartwatch aanwezig om als indicatie te hebben voor het kalibrern. | Hele periode |
| R7 | Niet accuraat stress kunnen detecteren met behulp van kunstmatige intelligentie. | 7 | 8 | 56 | Hierbij heb ik begeleiding nodig van mijn stagebegeleider. | ? | Er zijn meerdere Data Scientists aanwezig op de werkvloer voor vragen van mij. Zij zijn ook bereid om mij te helpen als ik hulp nodig heb. | Hele periode |
| R8 | Verkeerde algoritme kiezen waardoor tijd te kort ontstaat voor verdere onderzoek. | 5 | 8 | 40 | Na het testen van een gekozen algoritme en het niet meer nuttig lijkt te zijn, tijdig aangeven en bespreken met mijn stagebegeleider. | ? | Er zal eerst een ontwerp gemaakt worden met bijbehorende onderbouwing. Die zal ook eerst worden nagekeken door mijn stagebegeleider. |  |
| R9 | … |  |  |  |  | ? |  |  |
| R10 | … |  |  |  |  | ? |  |  |

# Deel 6 | Verantwoording

Deze onderstaande tabel kan afwijkingen bevatten. Er wordt verder aan gewerkt om een concrete indicatoren aan te geven in welke hoofdstukken van de scriptie terug te vinden zijn.

|  |  |  |  |
| --- | --- | --- | --- |
| **Hoofdstuk** | **Onderwerp** | **Activiteiten** | **Indicatoren** |
| 1 | Inleiding |  |  |
| 2 | Opdrachtomschrijving | Analyseren | A1, A2, A3, A4, A5, A6, A7 |
| 3 | Theoretisch kader | Analyseren | A8, A9 |
| 4 | Methode | Beheren, Realiseren | B1, B2, R2 |
| 5 | Planning | Realiseren |  |
| 6 | Resultaten | Ontwerpen, Realiseren | O1, R1, R2, R4, R6 |
| 7 | Conclusie | Adviseren | D1 |
| 8 | Aanbevelingen | Adviseren | D2, D3 |
| 9 | Literatuurlijst |  |  |
| Bijlage 1 | Technisch ontwerp + functioneel ontwerp | Ontwerpen | O2, O3, O4, O5 |
| Bijlage 2 | Testplan + testrapport | Beheren, Realiseren | B3, B4, R5 |
| Bijlage 3 | Gebruikershandleiding |  |  |
| Bijlage 4 | Risicomanagementplan | Beheren | B5 |



Fig. 4 Vijf competenties op volgorde met bijbehorende taken die worden uitgevoerd voor huidige onderzoek. Dit is ook terug te vinden in de scriptie voor meer informatie.

# Deel 7 | Literatuurlijst

[1] Kemeng Chen, Wolfgang Fink, Janet Roveda, Richard D. Lane, John Allen, Johnny Vanuk. (2015). Wearable sensor based stress management using integrated respiratory and ECG waveforms. (Ongepubliceerd eindwerk). University of Arizona, Tucson.

[2] Sriramprakash.S ­, Prasanna Vadana. D, O. V. Ramana Murthy. (2017). Stress Detection in Working People (Ongepubliceerd eindwerk). Department of Electrical and Electronics Engineering Amrita School of Engineering Coimbatore, India.