

▶▶  
**UHASSELT**



**Maastricht University**

KNOWLEDGE IN ACTION

## **Faculteit Wetenschappen** **School voor Informatietechnologie**

master in de informatica

### **Masterthesis**

**Slaapkwaliteit en welzijn beter begrijpen: fitbit gegevensonderzoek en interactieve oplossing voor studenten**

**Sean Snellinx**

Scriptie ingediend tot het behalen van de graad van master in de informatica

#### **PROMOTOR :**

dr. Gustavo Alberto ROVELO RUIZ

#### **COPROMOTOR :**

dr. Eva GEURTS

De transnationale Universiteit Limburg is een uniek samenwerkingsverband van twee universiteiten in twee landen: de Universiteit Hasselt en Maastricht University.



**UHASSELT**

KNOWLEDGE IN ACTION

[www.uhasselt.be](http://www.uhasselt.be)

Universiteit Hasselt  
Campus Hasselt:  
Martelarenlaan 42 | 3500 Hasselt  
Campus Diepenbeek:  
Agoralaan Gebouw D | 3590 Diepenbeek

**2022**  
**2023**



**Maastricht University**

# **Faculteit Wetenschappen**

## ***School voor Informatietechnologie***

master in de informatica

### ***Masterthesis***

***Slaapkwaliteit en welzijn beter begrijpen: fitbit gegevensonderzoek en interactieve oplossing voor studenten***

**Sean Snellinx**

Scriptie ingediend tot het behalen van de graad van master in de informatica

**PROMOTOR :**

dr. Gustavo Alberto ROVELO RUIZ

**COPROMOTOR :**

dr. Eva GEURTS



# Dankwoord

Het schrijven van deze masterthesis en bijhorend onderzoek is gerealiseerd voor het behalen van een master binnen de informatica. De hedendaagse tekortkomingen van technologieën voor het evalueren van eigen leefgewoonten was een drijfveer voor het afleveren van dit resultaat. Verder droeg de nieuwsgierigheid van invloeden als alcohol en stress op mijn welzijn ook mee aan het nagaan van correlaties en verbeteringen die toegepast kunnen worden.

Graag wil ik binnen deze sectie mijn dank uitbrengen aan enkele personen die rechtstreeks en onrechtstreeks hun bijdrage hebben geleverd in het tot stand brengen van deze thesis. Allereerst zou ik graag dr. Geurts Eva willen bedanken voor het aanbrengen van ideeën bij de initiële start van deze thesis. Haar kennis in het vakgebied heeft meegeholpen in de afbakening van de studie en heeft een meerwaarde geleverd aan hoe dit onderwerp opgevat kan worden vanuit verschillende perspectieven. Verder zou ik ook graag prof. dr. Luyten Kris willen bedanken. Zijn ervaring en kennis omtrent visualisaties, machine learning en het medische domein hebben samen nuttige feedback kunnen geven die bijdroegen aan het finale resultaat van de applicatie. Als laatste persoon zou ik dan nog graag mijn promotor dr. Roveló Ruiz Gustavo Alberto in het bijzonder willen bedanken. Zijn feedback en input hebben enorm bijgedragen aan het voltooien van deze masterthesis. Hij stond mij bij in het nemen van moeilijke beslissingen en bood de nodige steun gedurende de volledige ontwikkeling van deze thesis.

Verder wil ik ook nog graag mijn dank uitbrengen naar het volledige Expertisecentrum voor Digitale Media en alle testpersonen die vrijwillig deelnamen aan het onderzoek. Hun bijdrage bij het uitlenen van apparatuur, bronnen en toestemming voor het opmeten van data zorgden er samen voor dat ik deze thesis kon uitwerken. Naast hen wil ik ook nog graag mijn familie en vrienden bedanken die mij steunden in het volbrengen van deze thesis. Hun steun en aanmoediging zorgden ervoor dat ik de nodige motivatie vond om deze thesis af te werken.

# Samenvatting

Deze masterthesis presenteert een onderzoek naar oplossingen hoe studenten van 20 tot 25 jaar hun slaapkwaliteit en algemeen welzijn kunnen verbeteren, gebruik makend van gegevens die verzameld werden door Fitbit horloges alsook andere gegevens zoals hun alcohol- en cafeïne-consumptie. Met behulp van een data analyse en machine learning technieken wordt hier een aanpak voorgesteld en toegepast die antwoorden kan bieden op hoe slaappatronen, stress gevoelens en keuzes binnen de levensstijl verbeteringen teweeg kunnen brengen. Door middel van een interactieve applicatie biedt deze studie de mogelijk aan studenten om zelfstandig ongekende complexe correlaties van hun gegevens na te gaan en voorspellingen te maken waardoor hen duidelijk gemaakt wordt welke invloeden hun leefgewoonten hebben. Hiermee willen we hen bijbrengen om betere beslissingen te nemen voor het optimaliseren van hun slaapkwaliteit en hun algemene gezondheid.

Welke exacte methodes en overwegingen gemaakt worden voor het verzamelen en verwerken van de gegevens worden toegelicht. Door het combineren van gegevens verzameld met Fitbit horloges alsook het bijhouden van papieren logboeken voor het opmeten van bijvoorbeeld de alcohol- en cafeïne-consumptie, werd een sterke basis van gegevens opgebouwd waar een doorgronde analyse kon op uitgevoerd worden. Een aanpak via beschrijvende statistiek en het opstellen van voorspellende en verklarende machine learning modellen wordt gebruikt om in dit onderzoek weer te geven op welke manier verborgen verbanden en correlaties onthuld kunnen worden die een oplossing kunnen aangeven hoe de slaapkwaliteit verbeterd kan worden en stress vermindert.

Het resultaat dat voorkomt uit de goed opgestelde analyse is een dynamische applicatie die de bevonden resultaten weergeeft zodanig dat studenten zelfstandig deze kunnen interpreteren. Hiermee probeert deze studie een antwoord te bieden aan de hedendaagse tekortkomingen op het vlak van personalisatie en interpretatie van verbanden tussen de levensstijl en het welzijn van studenten. Deze studie tracht studenten een mogelijkheid te geven om hen bewust te maken hoe zij zelf verbeteringen voor hun slaapkwaliteit en stress kunnen aanbrengen door bepaalde aanpassingen te maken van hun leefgewoonten.

De resultaten van deze studie kunnen een cruciale stap vooruit zijn omtrent onderzoek naar de manier waarop gepersonaliseerde adviezen kunnen overgebracht worden in plaats van louter directe metingen weer te geven. Dit onderzoek biedt een raamwerk voor hoe gegevens verder gebruikt en toegepast kunnen worden, om een verrijkende invloed te kunnen uitoefenen op de complexiteit van factoren die van invloed zijn op de slaapkwaliteit en stress bij studenten. Ondanks de geleverde analyses heeft deze studie enkele beperkingen bij het maken van conclusies rond deze invloed gegeven de beperkte omvang en het gebrek aan een volledig medisch correcte benadering. Echter, de gebruikte kennis, methoden en resultaten kunnen in de toekomst wel een leidraad vormen voor een mogelijke aanpak om met behulp van moderne technieken een positieve impact uit te oefenen in het gebied van zelfevaluatie bij studenten op hun gezondheid. Dit geeft weer dat verder onderzoek op deze studie naast de academische toepassingen ook maatschappelijk praktisch nut kan hebben om mensen de kennis te geven over hoe zij hun slaapkwaliteit, stress en levenskwaliteit kunnen beïnvloeden.

# Inhoudsopgave

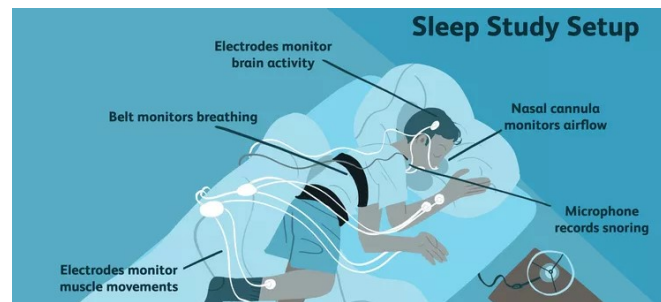
<b>1</b>	<b>Inleiding</b>	<b>5</b>
<b>2</b>	<b>Literatuur</b>	<b>8</b>
2.1	Slaapkwaliteit . . . . .	8
2.1.1	Eigenschappen . . . . .	9
2.1.2	Externe factoren . . . . .	11
2.1.3	Objectieve metingen . . . . .	12
2.2	Machine learning . . . . .	13
2.2.1	Regressiemodellen . . . . .	13
2.2.2	Decision tree modellen . . . . .	15
2.2.3	Gradient boosting . . . . .	17
2.3	Bestaand werk . . . . .	18
<b>3</b>	<b>Methoden</b>	<b>21</b>
3.1	Dataverzameling . . . . .	21
3.1.1	Testpersonen . . . . .	22
3.1.2	Meetinstrumenten . . . . .	22
3.1.3	Procedures . . . . .	23
3.1.4	Ervaringen/problemen . . . . .	23
3.2	Analyse . . . . .	25
3.2.1	Dataverwerking en filtering . . . . .	25
3.2.2	Beschrijvende statistiek en verdeling van parameters . . . . .	28
3.2.3	Identificatie van patronen en correlaties . . . . .	30
3.2.4	Vorbereiding machine learning algoritmen . . . . .	35
3.2.5	Voorspellende modellering voor slaappatronen . . . . .	37
3.2.6	Evaluatie en optimalisatie van modellen . . . . .	38
3.2.7	Implicaties op slaapkwaliteit . . . . .	43
<b>4</b>	<b>Verklarende interactieve oplossingen voor studenten</b>	<b>44</b>
4.1	Kennismaking parameters en gewoonten . . . . .	44
4.2	Ontdekken patronen en correlaties . . . . .	45
4.3	Alcohol en slaapkwaliteit . . . . .	47
4.4	Belangen van parameters onderzoeken . . . . .	48
4.5	Voorspellingen uitvoeren en nagaan . . . . .	49
4.6	Interactieve webapplicatie voor studenten . . . . .	51
<b>5</b>	<b>Conclusies</b>	<b>53</b>
5.1	Limitaties . . . . .	53
5.2	Toekomstig werk . . . . .	54
5.3	Zelfreflectie . . . . .	55
<b>6</b>	<b>Bijlagen</b>	<b>62</b>

# Hoofdstuk 1

## Inleiding

Een van de belangrijkste facetten die een rol speelt in het leven en op een mens zijn of haar gezondheid is slaap[19, 20]. Hierdoor zou het analyseren van slaap een bijdrage kunnen leveren over hoe de mens zijn lichaam beter kan leren begrijpen[9]. Men zou kunnen streven naar een manier waarom een persoon zijn of haar levenskwaliteit zou kunnen verbeteren op basis van zijn slaapgedrag en andere gewoonten. Als er naar slaap gekeken wordt vanuit een wetenschappelijk perspectief probeert men de opmeting van slaap vaak te vertalen vanuit een set van tastbare en objectieve parameters. Enkele voorbeelden hiervan zijn de hartslag, de hersensignalen en hoelang een persoon slaapt[29]. Op basis van een combinatie van verschillende parameters kan men tijdens de slaaperiode ook verschillende slaaphasen onderscheiden. Deze fasen geven op hun beurt al een duidelijker beeld over hoe een persoon heeft geslapen. Om een optimale slaapkwaliteit te bekomen zullen verwachtingen gesteld worden over welke waarden optimaal zijn voor welke van deze parameters. Zo probeert men een combinatie te vinden waarbij kwaliteit het hoogst is. De optimale waarden kunnen sterk verschillen afhankelijk van de persoon waarnaar gekeken wordt. Eigenschappen als leeftijd of sportiviteit van een persoon kunnen hier doorslaggevende parameters zijn die bijvoorbeeld de optimale duratie van de slaap kunnen beïnvloeden[6].

Binnen de medische wereld kan de slaapkwaliteit van een persoon zeer grondig geanalyseerd worden door het opmeten van een reeks aan parameters met behulp van polysomnografie. Dit soort slaap onderzoek is zeer accuraat en kan slaapstoornissen of andere problemen blootleggen[29]. Maar dit soort onderzoeken brengt ook een reeks aan ongewilde effecten met zich mee. Doordat personen in een labo moeten slapen terwijl zij verbonden zijn met een groot aantal sensoren en kabels die metingen uitvoeren zoals zichtbaar op figuur 1.1, vereist dit niet alleen veel middelen maar kan dit het normale slaapritme ook beïnvloeden[53]. Vandaag de dag zijn er nog andere technologische ontwikkelingen in veel verschillende vormen die parameters van de slaap alsook van andere gebieden als fysieke activiteit van een mens kunnen verzamelen. Voorbeelden zijn sensoren die enkel rond de borstkas moeten of polssensoren[50]. De ontwikkeling waar er in dit onderzoek op gefocust zal worden, is de smartwatch. Dit is een elektronisch apparaat dat data van een persoon verzameld door middel van een slim horloge dat extra sensoren gebruikt om zo medische parameters op te meten. Deze data wordt dan vaak met behulp van een connectie met een smartphone of ander elektronisch apparaat verwerkt en opgeslagen. Deze data kan dan vaak achteraf geraadpleegd worden met behulp van een bijhorende applicatie of platform. Onder deze data valt ook de data die tijdens de slaap verzameld wordt van een persoon. De horloges meten bijvoorbeeld op hoelang een persoon slaapt, wat zijn hartslag is of zelfs de saturatie in zijn bloed. Desondanks het verzamelen van deze data zijn er nog steeds een aantal gebreken waardoor het niet zomaar vanzelfsprekend is voor een persoon om na te gaan hoe hij zijn slaap kan verbeteren. Dit komt doordat er nog steeds een aantal externe factoren zijn die een invloed uitoefenen op de slaapkwaliteit van een persoon. Naast de interpretatie van de slaap gegevens is het dus ook nodig dat een persoon kennis heeft over deze externe factoren met



**Figuur 1.1:** De figuur geeft weer welke sensoren allemaal verbonden zijn bij een PSG onderzoek.

betrekking tot zijn slaap. Daarom dat het ook nodig is om rekening te houden met factoren waarbij reeds geweten is dat deze een invloed uitoefenen op de slaapkwaliteit. Voorbeelden van zulke factoren zijn stress, fysieke activiteit, alcohol- of cafeïne-inname.

De externe factoren die hun invloeden uiten op het slaapgedrag zijn niet altijd makkelijk te voorspellen. Dit is te danken aan de complexe relaties die zich kunnen voordoen door de combinatie van al deze factoren. Daarom dat enkel het opnemen ervan vaak niet voldoende is om na te gaan welke factor nu welke invloed heeft op de slaapkwaliteit. Met behulp van machine learning kan er wel geprobeerd worden om een uitleg voor deze relaties te vinden. Door de aard van deze soort algoritmes kan men zonder voorkennis over de verbanden tussen verschillende factoren toch een goed beeld krijgen over mogelijke (complexe) relaties die zich voordoen. Echter is er dan wel vereist dat deze externe factoren, zoals bijvoorbeeld alcohol-inname, op een andere manier opgenomen moeten worden doordat smartwatches hiertoe nog niet in staat zijn.

Het verzamelen en verwerken van data is echter niet voldoende. Het vereist extra inspanningen om deze data begrijpbaar te maken voor mensen. Lange tabellen van rechtstreeks gemeten waarden of van verwerkte resultaten zijn niet praktisch en vertellen weinig over wat de data werkelijk betekent of welke patronen er zich voordoen. De data en verkregen resultaten moeten omgezet worden in een formaat dat snelle interpretatie mogelijk maakt en zodat de gevonden relaties duidelijk weergegeven worden. Grafieken en andere visuele hulpmiddelen zouden hiervoor gebruikt kunnen worden om af te stappen van onoverzichtelijke lijsten van of tabellen. Door mensen zelf de mogelijkheid te geven hun data te interpreteren wordt hen de keuze gegeven met welke diepgang ze willen ingaan op de data. Om te stimuleren dat ze de data zelf verder verkennen zou een interactieve oplossing aangeboden kunnen worden.

Verder bieden hedendaagse oplossingen niet altijd een duidelijk beeld hierover. Als we kijken naar applicaties zoals de gezondheidsapplicatie van Apple of de applicatie die horloge fabrikant Fitbit aanbiedt, worden bepaalde bevindingen niet doorgetrokken. Desondanks deze applicaties en bijhorende toebehoren een accuraat werk leveren in het verzamelen van medische informatie en deze weer te geven, leggen deze applicaties zelf niet de link tussen parameters die mogelijk een invloed op elkaar kunnen hebben. Een interessant gegeven voor mensen zou juist kunnen zijn om een uitleg te krijgen waarom bepaalde resultaten zich voordoen en welke invloeden hun acties kunnen hebben op deze resultaten. Denk hier bijvoorbeeld aan het consumeren van alcohol. Het is reeds bewezen dat de inname van alcohol de slaap van een persoon sterk kan beïnvloeden[13]. Desondanks deze hedendaagse kennis, slagen de bestaande oplossingen er niet in om aan mensen duidelijk te maken dat een slechte nachtrust waarschijnlijk aan de voorafgaande alcohol-consumptie te wijten zou kunnen zijn.

De meeste gezondheidsapplicaties vandaag werken goed als het gaat over de dataverzameling zelf. Echter, maar weinig van deze oplossingen proberen de gebruikers te engageren om hun gezondheid te verbeteren[52]. Binnen dit veld zijn er nog opportuniteiten om de effectiviteit van zulke technologieën te verbeteren. Oplossingen zoals Fitbit slagen er ook niet in om personen



deze gepersonaliseerde en stimulerende adviezen bij te brengen. Op dit vlak zouden nog een heel aantal verbeteringen kunnen komen zoals het suggereren van acties die een persoon kan doen om zijn slaapkwaliteit te doen toenemen. Als we kijken naar die huidige oplossingen die aangeboden worden, zijn deze vaak oppervlakkig en voegen zij weinig toe aan de nieuwsgierigheid van mensen om hun persoonlijke levensstijl te optimaliseren. Deze oplossingen zouden meer de focus kunnen leggen op de parameters waarop mensen rechtstreeks invloed hebben, zoals we vandaag de dag wel reeds zien voor aspecten zoals het aantal stappen dat een persoon zet. Hier stimuleren oplossingen zoals de Apple watches wel mensen doorheen de dag om bepaalde mijlpalen te bereiken. Omdat gebruikers zichzelf zonder medische expertise laten opmeten door deze horloges, moet er ook voor gezorgd worden dat de terugkoppeling enerzijds correct is maar anderzijds ook verbeteringen moet kunnen voorstellen. Diepere analyses kunnen bijvoorbeeld factoren als slaap, beweging en stress opmeten om correlaties hiertussen te ontdekken om zo uiteindelijk misschien een opportuiniteit te vinden om de slaap verder te verbeteren.

Zoals al deze factoren al doen vermoeden is het analyseren van de slaapkwaliteit van een individuele persoon een zeer moeilijke opgave. Niet alleen de parameters die van de persoon zelf allemaal continu verzameld moeten worden, maar ook de algemene eigenschappen en gewoonten van de persoon zelf hebben een sterke invloed. Verder hebben externe factoren of acties die een persoon op zijn lichaam heeft uitgevoerd ook nog een grote invloed. Het is dus de totale huidige staat van een persoon zijn lichaam die overwogen moet worden. Zelfs met al die informatie zijn er ook nog altijd de invloeden van de omgeving die de slaap van een persoon kunnen beïnvloeden. Dit vertoont zich bijvoorbeeld wanneer een persoon uit zijn slaapcyclus gehaald wordt door een luide knal of wanneer de warme weersomstandigheden in de kamer zijn slaap zou onderbreken. Desondanks het volledig verklaren van de slaapgewoonten en kwaliteit van een persoon zeer grote inspanningen en meetinstrumenten zou vereisen, probeert men tegenwoordig toch om met behulp van een slim horloge al een benadering hierop te kunnen uitvoeren.

Het doel van deze thesis is om aan de hand van de bevindingen in de verzamelde data een beeld te scheppen over hoe het slaapgedrag van een persoon zich voordoet en hoe dit overgebracht kan worden naar de studenten zelf. Hierdoor zou dan ingeschat kunnen worden wat een persoon kan doen om zijn slaap te verbeteren en zo dus ook zijn levenskwaliteit. Deze studie zal zich focussen op een kleine doelgroep bestaande uit gezonde studenten met een leeftijd van 20 tot en met 25 jaar. Tijdens dit onderzoek zal er van deze groep data verzameld worden die vervolgens geanalyseerd zal worden met behulp van machine learning algoritmen. De resultaten die hieruit zullen volgen, zullen dan verwerkt en gepresenteerd worden op een manier zodanig dat een persoon waardevolle inzichten kan verkrijgen met betrekking tot zijn eigen slaapkwaliteit en zo ook zijn gezondheid. Dit onderzoek zal proberen een antwoord te formuleren op volgende vragen:

1. Welke algoritmen voor machine learning kunnen bruikbare resultaten opleveren om meer te weten te komen over slaappatronen bij studenten tussen 20 en 25 jaar oud?
2. Wat definieert slaapkwaliteit en stress?
3. Wat zijn de correlaties tussen de verschillende parameters, zoals slaapkwaliteit, alcohol-inname, cafeïne-inname, stress, fysieke activiteit en andere slaapgegevens?
4. Hoe beïnvloedt alcoholgebruik de slaapkwaliteit?
5. Kan het slaapgedrag van een student worden voorspeld op basis van deze parameters?
6. Hoe kunnen correlaties en voorspellingen met betrekking tot deze parameters de slaapkwaliteit van de student verbeteren?
7. Hoe kunnen suggesties en aanbevelingen voor het verbeteren van de slaapkwaliteit en het algemene welzijn aan de studenten worden gepresenteerd?

## Hoofdstuk 2

# Literatuur

Een doorgedreven literatuurstudie ligt aan de basis van deze studie, alsook het verzamelen van de nodige informatie omtrent de specifieke aspecten die deel uitmaken van het onderzoek. Elementen als wat slaapkwaliteit is, wat gekende problemen of beïnvloedende parameters zijn en hoe verbeteringen aangebracht kunnen worden kunnen hier deel van uitmaken. Deze informatie kan gaan van bestaande kennis over het menselijk lichaam tot bestaande technologische ontwikkelingen die vooraf gingen aan deze studie of het nut aanhalen van de studie die wij hier zullen uitvoeren. Deze sectie geeft een toelichting van deze bestaande kennis en technieken toelichten alsook aanhalen waar deze nuttig ingezet kunnen worden. Hierin wordt er een onderscheid gemaakt op drie algemene gebieden, namelijk op de slaapkwaliteit, machine learning algoritmen en de representatie van resultaten aan het doelpubliek waartoe deze studie gericht is. Binnen dit eerste wordt er gezocht naar welke aspecten een belangrijke invloed hebben op de slaapkwaliteit en welke parameters hierin aan bod komen. Ook zal hiervoor gekeken worden welke mogelijke technieken er bestaan om deze gegevens af te nemen van mensen om zo accuraat mogelijke resultaten te verzamelen die een reëel beeld kunnen creëren van de slaapkwaliteit van personen.

Het tweede gedeelte van de literatuur zal zich toelichten op het technische aspect van deze studie. Hierin zal er gekeken worden naar bestaande machine learning algoritmen en modellen die binnen dit domein toepasselijk kunnen zijn. Deze zullen instaan voor de verwerking van de verzamelde data. Verder is het van belang dat deze modellen op een juiste manier gekozen en opgesteld worden zodanig dat deze duidelijkheid kunnen brengen over de specifieke aspecten waar deze studie een antwoord op wil bieden. De keuzes die hier gemaakt zullen worden zullen op hun beurt ook een sterke invloed hebben over de mogelijkheden bij het visualiseren en presenteren van deze resultaten.

Het derde gebied waar reeds bestaand werk onderzocht zal worden, zijn voor gelijkaardige onderzoeken als deze. Hierin wordt er literatuur bestudeerd waarin gelijkaardige manieren worden onderzocht hoe slaapkwaliteit gelinkt is aan parameters en hoe deze verbeterd zou kunnen worden. Het belang hiervan is dat nogmaals het nut van deze studie wordt bevestigd en dat gekende tekortkomingen in acht genomen worden.

### 2.1 Slaapkwaliteit

Hoe kan slaapkwaliteit gedefinieerd worden? En welke parameters hebben allemaal een invloed hierop? Om hier een antwoord op te bieden proberen we een afgebakende set aan parameters te zoeken die nuttig zijn alsook haalbaar zijn om zelf te verzamelen binnen deze studie. Door middel van literatuur zien we dat slaap kwaliteit vaak rechtstreeks afhangt aan een aantal vaste eigenschappen die worden opgemeten. Maar wetenschappelijk onderzoek suggereert dat het daar niet de facto stopt[29]. Omdat mensen geen perfecte invloed kunnen uitoefenen op

hoe lang zij zich bijvoorbeeld in een bepaalde slaapfase willen bevinden, moeten we nagaan welke andere externe factoren zij wel kunnen beïnvloeden om zo onrechtstreeks meer controle te krijgen over hun slaapkwaliteit[29]. Maar dit onderzoek gaat ook verder dan enkel welke parameters we kunnen gebruiken. De manier hoe we waarden aan deze variabelen kunnen koppelen is echter ook van belang. Hoe worden deze parameters allemaal opgemeten? Dit geldt zowel voor de eigenschappen van slaapkwaliteit als de externe factoren. Hierin worden alternatieven naast elkaar vergeleken zodat we in deze studie een correcte keuze kunnen maken over een optimale techniek om deze data te verzamelen. Men moet hierbij verder ook rekening houden dat de slaapkwaliteit van een persoon sterk afhankelijk is van de persoon waarop deze gemeten wordt en hoe nauwkeurig we metingen kunnen doen.

### 2.1.1 Eigenschappen

Er zijn er heel wat eigenschappen die waardevol kunnen zijn als er gekeken wordt naar de kwaliteit van slaap van een persoon. Doordat de kwaliteit van slaap niet rechtstreeks gedefinieerd is als een vaste waarde die gemeten kan worden[29], wordt er meestal aan de hand van deze eigenschappen een inschatting gemaakt over de effectieve kwaliteit. Een reeds bestaande welbekende methode om de slaapkwaliteit te meten is de SQS (Sleep Quality Scale)[70]. Dit is een vragenlijst bestaande uit 28 vragen omtrent het slaapgedrag die men met de bijhorende antwoordenschaal kan beantwoorden. Echter worden de antwoorden op zulke vragenlijst als subjectief gezien en wordt de accuraatheid ervan soms in twijfel getrokken[38]. Een meer objectieve wijze voor het opmeten van de slaapkwaliteit zou het gebruik van sensoren kunnen zijn. Polysomnografie (PSG) zou zo bijvoorbeeld een heel aantal fysiologische parameters kunnen opmeten van een persoon zoals zijn hartslag, bloeddruk, hersensignalen en zuurstof saturatie van het bloed. Met behulp van deze parameters zou men ook een inzicht kunnen krijgen in de algemene slaaptoestand van een persoon. Maar omdat slaapkwaliteit verder gaat op het vlak van de verschillende aspecten van slaap, de quantiteit, de timings waarin alles gebeurt, de fasen structuur, kan er gesteld worden dat traditionele PSG hier te eenvoudig is voor een gevoelige en uitgebreide meting van de slaapkwaliteit[29]. Het uitvoeren van PSG analyses is een zeer complex en duur gegeven dat enkel kan plaatsvinden binnen de daarvoor voorziene laboratoria. Zulke labo's kunnen op hun beurt ervoor zorgen dat de persoon zich niet op zijn gemak voelt of dat hij anders zal gaan slapen door de ongekende omgeving[53]. Dit effect is een gegeven dat deze studie wil vermijden doordat op deze manier bijvoorbeeld stress mogelijks een sterke variatie zou kunnen ondervinden in vergelijking met een bekende slaap locatie.

In de onderstaande secties worden slaap eigenschappen besproken die een meerwaarde kunnen betekenen voor het opmeten van de slaapkwaliteit. Deze eigenschappen kunnen zowel rechtstreeks metingen uit sensoren zijn alsook onrechtstreeks afleidingen van deze metingen. De bijhorende beschrijving zal ook telkens toelichten wat deze exacte eigenschappen zijn, en indien van toepassing ook aanhalen op welke manier ze afgeleid kunnen worden.

**Hartslag** De hartslag van een persoon is een belangrijke factor bij het bespreken van de slaapkwaliteit. Deze weerspiegelt de fysiologische toestand van het lichaam tijdens de slaap. De hartslag daalt gewoonlijk tijdens de diepe stadia van de slaap. Hij neemt toe tijdens de lichtere stadia van de slaap en tijdens de REM-slaap, waardoor hij een nuttige aanduiding is bij de evaluatie van de slaapkwaliteit. Een constant hoge of lage hartslag tijdens het slapen kan wijzen op een onderliggende gezondheidstoestand die de slaapkwaliteit beïnvloedt.[57]

**RMSSD** RMSSD (root mean square of successive differences) is een eigenschap die aangeeft hoe de hartslag verandert van slag tot slag en vertelt ons hoe goed de natuurlijke systemen van het lichaam werken tijdens de slaap. RMSSD is een geaccepteerde manier om de slaapkwaliteit te meten, omdat het ons kan helpen begrijpen of we een rustgevende en herstellende slaap ondervinden. Via RMSSD kunnen slaappatronen beter begrepen worden en eventuele problemen die onze slaapkwaliteit beïnvloeden achterhaald worden.[56]

**NREMHR** Non-REM hartslag verwijst naar de hartslag tijdens de non-rapid eye movement (NREM) fasen van de slaap. Deze eigenschap kan extra informatie voorzien over de slaapkwaliteit doordat het mee een beeld schept over de toestand van het lichaam tijdens de slaap. Net zoals eerder aangehaald zijn er patronen in de variatie van de hartslag in combinatie met de verschillende slaapfasen die een representatie kunnen geven over de slaapkwaliteit en de gezondheid van een persoon.[15]

**Slaap Duratie** De duur van de slaap is slechts één aspect van de slaapkwaliteit, maar het is een belangrijke factor voor geheel de slaapgezondheid en zelfs algemene gezondheid. Voldoende slaap is essentieel voor het behoud van een goede gezondheid, en slaaptekort kan leiden tot een reeks negatieve gevolgen, zoals bijvoorbeeld een verminderde cognitieve functie, een verminderde immuniteit en een verhoogd risico op chronische ziekten. Natuurlijk worden deze effecten niet alleen bepaald door de slaapduur, maar de duur speelt zeker een rol in de slaapkwaliteit en de gezondheid van een persoon.

**Verskillende slaap fasen** De slaap wordt gewoonlijk onderverdeeld in vier verschillende stadia, die elk worden gekenmerkt door specifieke patronen van hersenactiviteit, oogbewegingen en spieractiviteit. De verschillende slaapstadia zijn belangrijk voor het reguleren van verschillende fysiologische functies en zijn sterk gerelateerd aan het meten van slaapkwaliteit en welzijn. De duur en de verdeling van deze slaapstadia bepalen ook sterk de slaapkwaliteit. Onregelmatigheden van deze stadia, zoals vaak wakker worden of een gebrek aan diepe of REM-slaap, kunnen leiden tot vermoeidheid en een gebrek aan energie doorheen de dag, wat een negatieve invloed kan hebben op het algemene welzijn.[15, 45]

**REM-slaap** REM of Rapid Eye Movement is het stadium waarin de meeste levendige dromen ontstaan. Deze slaapfase wordt gekenmerkt door snelle oogbewegingen, een verhoogde hersenactiviteit en spierverslaving. De REM-slaap is belangrijk voor het mentaal en emotioneel herstel en omvat ongeveer 25% van de totale slaaptijd uit.

**Diepe slaap** Diepe slaap of de N3 slaapfase is ook bekend als de trage-golfslaap, waarbij de hersengolven nog trager worden, waardoor gewekt worden of vanzelf wakker worden moeilijker zal zijn. Deze fase is belangrijk voor het lichamelijke herstel en zou gemiddeld ongeveer 20% van de totale slaaptijd in beslag nemen.

**Lichte slaap** Lichte slaap, ook wel gerefereerd als N1/N2 slaap is de overgangperiode tussen wakker zijn en diepe slaap, waarin het lichaam zich begint te ontspannen en de hersenactiviteit vertraagt. Deze fase is goed voor ongeveer 50% van de totale slaaptijd.

**Wakkere slaap** Dit is wanneer een persoon wakker is en zich bewust is van de omgeving rondom hem. Dit kunnen bijvoorbeeld momenten zijn waar iemand toevallig door geluid in de omgeving of degelijke voor een korte periode, vaak zelfs maar enkele seconden, ontwaakt wordt uit zijn slaap.

**Zuurstofsaturatie** Spo<sub>2</sub> of zuurstofsaturatie, meet de hoeveelheid zuurstof aanwezig in het bloed. Wanneer een persoon slaat kunnen de noden van het lichaam omtrent de zuurstofbehoefte veranderen. Zo kunnen verstoringen van de ademhaling, zoals die veroorzaakt door slaapapneu, leiden tot een verlaging van de saturatie. Deze daling van de zuurstofsaturatie kan de slaapkwaliteit negatief beïnvloeden en kan net zoals een tekort aan slaap zorgen voor een verminderde cognitieve functie en een reeks aan andere gezondheidsproblemen. De Spo<sub>2</sub>-waarden tijdens de slaap kunnen extra info brengen over de slaapkwaliteit van een persoon maar zouden ook kunnen helpen bij het identificeren van mogelijke slaapstoornissen die de algemene gezondheid en het welzijn van deze persoon kunnen beïnvloeden.[11, 12]

### 2.1.2 Externe factoren

Deze sectie zal, in tegenstelling tot de voorgaande sectie ingaan op de externe factoren die een mogelijke invloed kunnen hebben op de slaapkwiteit en het welzijn. Doordat er echter een enorm aantal factoren meespelen die mogelijks op een bepaalde manier gelinkt kunnen worden met de slaapkwiteit, zullen er hier enkele specifieke factoren uitgekozen worden waarop deze studie zich zal focussen. Volgende sectie zal kort toelichten welke deze factoren zijn en hoe bestaande studies deze factoren linken aan de kwaliteit van slaap.

**Fysieke activiteit** Fysieke activiteit van een persoon wordt wel vaker geassocieerd met het slaapgedrag. Ook is er reeds bewijs geleverd dat er wel degelijk een bidirectionele relatie bestaat tussen beiden[28, 37]. Deze relatie verklaart het nut van het opmeten van iemand fysieke activiteit om meer inzicht te krijgen over zijn slaapgedrag en slaapkwiteit. Het opmeten van fysieke activiteit kan op verschillende manieren geïnterpreteerd worden. Men zou bijvoorbeeld kunnen kijken naar de totale tijd dat een persoon per dag sport of fysiek bezig is. Maar zulke waarnemingen nemen natuurlijk weg dat een persoon die weinig sport maar wel de hele dag licht actief bezig is ook fysieke activiteit onderneemt. Om dit laatste in tegen te gaan zou er besloten kunnen worden om naar het aantal stappen van een persoon te kijken, maar dit sluit dan ook weer activiteiten als gewichtheffen of soortgelijke activiteiten zonder stapbewegingen uit. Om een goede representatie van de fysieke activiteit van een persoon te krijgen zullen een aantal aspecten zoals de duratie maar ook de hartslag en zo intensiteit van de activiteit gemeten moeten worden[59]. Hierbij kan dan zware activiteit van normale en lichte activiteit onderscheiden worden. Ook het aantal stappen kan hierbij helpen en zelfs het aantal minuten dat een persoon zittend per dag doorbrengt zou in achtning genomen kunnen worden.[3]

**Stress** Stress en slaapkwiteit hebben net zoals hierboven bij fysieke activiteit een bidirectionele relatie, waarbij stress een negatieve invloed kan hebben op de slaapkwiteit en een slechte slaapkwiteit het stressniveau kan verhogen[7, 18]. Chronische stress kan langdurige veranderingen in het stress responsstelsel van het lichaam veroorzaken, waardoor het moeilijk wordt om in slaap te vallen en in slaap te blijven[21]. Stress kan ook de kans op nachtmerries en verstoorde slaap vergroten. Slaapstoornissen kunnen leiden tot een versterkte fysiologische reactie op stress, waardoor de effecten van stress op het lichaam worden verergerd[8]. Niet goed slapen kan leiden tot moeilijkheden bij het omgaan met stressvolle situaties. Gebrek aan slaap kan uw vermogen om met stress om te gaan verminderen[18]. Daarom is het belangrijk om een goede slaapkwiteit te behouden om zo goed mogelijk met stress om te gaan. Verschillende onderzoeken hebben een significant verband gevonden tussen stress en een slechte slaapkwiteit. De relatie tussen stress en slaapkwiteit is echter wel zeer complex en de een heeft een negatieve invloed op de ander. Het beheersen van stressniveaus en het verbeteren van de slaapkwiteit kan de algehele gezondheid en het welzijn aanzienlijk verbeteren.[18]

**Cafeïne** Cafeïne is een psychoactieve stof met een stimulerend effect die de slaapkwiteit aanzienlijk kan beïnvloeden. Er werd reeds aangetoond dat cafeïne-consumptie de slaap efficiëntie vermindert, de totale slaaptijd verkort en de periode tot men in slaap valt verlengt. Onderzoek heeft aangetoond dat zelfs een lage tot matige cafeïne-consumptie de slaapkwiteit negatief kan beïnvloeden, vooral wanneer die binnen enkele uren voor het slapengaan wordt geconsumeerd[10, 42, 55].

**Alcohol** Alcohol-consumptie heeft een duidelijk verband met betrekking tot de slaapkwiteit. Zo is reeds aangetoond dat alcohol bijvoorbeeld dat grotere hoeveelheden ervoor kunnen zorgen dat er een sterke afname is van de hoeveelheid REM-slaap een persoon die nacht zal hebben[13]. Verder zou alcohol ook nog extra problemen kunnen meebrengen zoals een verslechtering van de ademhaling, een algemeen kortere slaap en kan het slaapstoornissen verergeren[25].

**Water** Water zal voornamelijk als een rechtstreekse parameter gezien worden voor de hydratatie van een persoon. Nu water is een stof die rechtstreeks in verband staat met hydratatie en die gemakkelijker afzonderlijk opgemeten zou kunnen worden in de vorm van hoeveelheid per dag[48, 69]. Indien de hydratatie van een persoon exact opgemeten zou moeten worden komen hier veel meer aspecten bij kijken. Namelijk andere dranken en zelfs voeding bevat nagenoeg altijd ook een percentage water of vocht en het terug uitscheiden van vocht zou hierbij dan ook geïntegreerd moeten worden[48, 49]. Zulke metingen zullen voor deze studie te omvangrijk zijn en daarom dat hier een vereenvoudiging gemaakt wordt tot enkel de inname van water.

### 2.1.3 Objectieve metingen

We hebben nu een aantal factoren gedefinieerd waarbij we hun waardes en mogelijke verbanden met de slaapkwaliteit willen onderzoeken. Enkele van deze zijn makkelijker op te meten dan andere, denk maar aan bijvoorbeeld het opmeten van een hartslag tegenover het opmeten van effectieve stress die iemand ervaart. Voor deze studie zou het praktisch zijn moesten we zoveel mogelijk factoren op een formele en objectieve manier automatisch kunnen opmeten. Hierdoor verlagen we de inzet van de vrijwilligers die zouden deelnemen, doordat zijzelf zo weinig mogelijk zelf zouden moeten doen.

Voor het opmeten van een groot aantal van deze eigenschappen zoals hartslag en andere bio-factoren zijn er wel een aantal opties. Één zo een optie zou het gebruik van een smartwatch of smart wristband zijn. Op de huidige markt zijn er een groot aantal van deze toestellen aanwezig die vaak parameters als hartslag, zuurstofsaturatie en de fysieke activiteit van een persoon nagaan. Voorbeelden van deze smartwatches zijn bijvoorbeeld de Apple smartwatches, de Garmin smartwatches en de Fitbit smartwatches. Binnen deze studie zullen wij ons op deze laatste toeleggen. Deze keuze komt op basis van een aantal afwegingen. Er was voor deze studie een smartwatch gewild die zowel op IOS als Android zou werken, dit om het gebruik voor de testpersonen te vergemakkelijken zodanig dan ze hun eigen toestel kunnen gebruiken tijdens de studie. Ook moet de smartwatch de optie bieden om op een eenvoudige manier de data te ontleiden van zijn platform. Enkel op deze manier is het mogelijk om eigen transformaties en bewerkingen op de data uit te voeren. Voor het opvolgen van de personen zou het ook handig zijn moest hun data vanop afstand opgevolgd kunnen worden. Tot slot is er natuurlijk ook rekening gehouden met de exacte parameters de smartwatches kunnen opmeten. Na een vergelijking tussen verschillende smartwatches die op het moment van schrijven relevant lijken, is een voorkeur voor de Fitbit Sense series gekomen. Binnen deze serie zijn er twee watches waarbij de meest recente voornamelijk een verbetering is van de eerste. De Fitbit Sense 2 is een gezondheid en fitness smartmatch met een groot aantal functies die relevant zijn binnen dit onderzoek. Naast het continue opmeten van de hartslag en het aantal stappen dat iemand zet zijn er nog een heel aanbod aan factoren. De smartwatch onderscheidt de verschillende activiteiten die er persoon dagelijks doet, hij kan ook de zuurstofsaturatie meten en heeft ook de mogelijkheid op stress-niveaus te meten. Verder meet het horloge ook slaapgegevens zoals de duratie, de verschillende slaapfasen en een algemene slaapscore.

Hierboven werden twee eigenschappen aangehaald die toch wat extra uitleg vereisen. Een eerste is de stress metingen die de smartwatch uitvoert. De Fitbit smartwatches maken hiervoor gebruik van een EDA (Elektrodermale activiteit) sensor[74]. Deze sensor meet veranderingen in de geleiding van het huidoppervlak. Deze veranderingen kunnen de emoties en responses van het lichaam weergeven[74]. In combinatie met andere factoren zoals de hartslag zal het horloge dan de metingen proberen te linken aan de emoties van een persoon om zo spanning en stress te kunnen opmeten. De stressscore die zij voorzien steunt in totaal op drie categorieën[17]. Een eerste is deze EDA responsiviteit die zojuist werd aangehaald. De andere twee zijn de invloed van de slaap patronen van een persoon op zijn stress en het evenwicht van de inspanningen van een persoon. Met dit laatste wordt verwezen naar het gegeven of de huidige fysieke activiteit van de persoon bijstaat aan het verminderen van het stressniveau of dat het lichaam te veel inspanningen levert[5]. Naast deze stressscore is er dan ook nog de slaapscore. Deze score wordt

op basis van de volgende drie aspecten berekent: de duratie, de kwaliteit en de restauratie van de slaap[17]. Dit eerste wordt enkel bepaald door de totale tijd dat een persoon slapend doorbrengt. De kwaliteit van de slaap wordt afgewogen op basis van de gependeerde tijd binnen de diepe en REM-slaap, hoe langer hierbinnen gependeed wordt, hoe hoger de score. De restauratie factor komt voort uit de hartslag en slapeloosheid. Een hoge hartslag of continue bewegingen en draaien van het lichaam zorgen voor lagere scores.

Tot slot zijn er ook nog parameters die niet automatisch opgemeten kunnen worden. Waarden voor alcohol-, cafeïne- en water-consumptie kunnen niet door Fitbit horloges gemeten worden. Echter zouden ze wel handmatig bijgehouden kunnen worden binnen de Fitbit applicatie of een extern logboek. Deze self-tracking vereist een inspanning van de deelnemers, maar mogelijke alternatieven zoals bloedafnames voor het nagaan van het alcoholpercentage in het bloed van mensen zou niet haalbaar zijn binnen de omvang van een studie als deze. Dit handmatig bijhouden zal minder nauwkeurig zijn maar is wel een praktische manier om toch deze parameters te kunnen verzamelen.

## 2.2 Machine learning

Deze sectie binnen de literatuurstudie beschrijft reeds bestaande machine learning methoden en technieken die bruikbaar zouden kunnen zijn voor de verwerking van de data die deze studie zal verzamelen. Het gebruik van correcte en juiste modellen hangt sterk samen met wat men exact wil weten over de beschikbare data alsook met de aard van de data zelf. Voor het onderzoeken van deze verschillende modellen kan deze data dus op een verschillende manier gebruikt worden om gewilde resultaten te bekomen. Deze sectie zal hier verder duidelijkheid binnen brengen welke opties er voor welk type mogelijk zijn, wat de mogelijke voorwaarden hiervoor zijn en welke eigenschappen deze modellen mogelijk extra met zich meebrengen. Voor deze eerste soort van data zal deze studie drie soorten modellen onderzoeken namelijk regressie modellen, decision tree modellen en gradiënt boosting modellen.

### 2.2.1 Regressiemodellen

Regressie modellen zijn een soort van machine learning modellen die gebruikt worden om een bepaalde variabele te voorspellen op basis van een gegeven set aan invoer variabelen. Het doel van regressie modellen is om een relatie te vinden tussen de invoer en uitvoer zodanig dat het mogelijk wordt om bij een nieuwe invoer de uitvoer te kunnen voorspellen[2]. Deze relatie kan gezien worden als een wiskundige uitdrukking van variabelen. Hierin zal er een onafhankelijke variabele(n) zijn die overeen zal komen met de invoer alsook de afhankelijke variabele die we kunnen zien als de uitvoer variabele die we proberen te voorspellen. Binnen regressie modellen vallen ook een aantal categorieën te onderscheiden, die elk afhankelijk van het type van de uitvoer en de mogelijke relatie tussen invoer en uitvoer gebruikt zouden kunnen worden.

**Lineaire regressie** Lineaire regressie is een eenvoudige soort van regressie die zoals de naam reeds aangeeft uitgaat van het bestaan van een lineaire relatie tussen de onafhankelijke en de afhankelijke variabele. Deze soort regressie probeert een rechte te vinden die het beste bij de gegeven waarden van de variabelen past. Deze rechte kan omschreven worden door een vergelijking met een bepaalde helling en een punt dat deze rechte snijdt.[2, 67]

**Logistische regressie** Een soort van regressie zeer sterk bij lineaire regressie aanleunt is logistische regressie. Net zoals bij lineaire zal dit soort modellen de relatie tussen de onafhankelijke en de afhankelijke variabele nagaan. Maar in plaats van een rechte te zoeken zal hier de verhouding logistisch zijn waarbij de variabele waarde tussen 2 mogelijke opties zal gaan, zoals bijvoorbeeld waar of onwaar. Dit soort model zal dan als output een kans kunnen voorspellen gegeven de onafhankelijke inputs een bepaalde binaire output waarschijnlijk zal zijn.[68, 71]

**Polynomiale regressie** Polynomiale regressie kan omschreven worden als een meervoudig lineair regressie model. Dit soort model neemt opnieuw de relatie tussen de afhankelijke en onafhankelijke variabelen, maar werkt hier met bepaalde graden. Deze relatie wordt hier omschreven door de machten van een variabele. Deze macht wordt ook wel de graad van het model genoemd. Doordat er met machten gewerkt wordt zal de relatie als een curve omschreven kunnen worden. Deze curve zijn complexiteit zal volgen naargelang de grootte van de macht van het model. Doordat deze functie zich bij grotere machten sterk zou kunnen vormen naar eender welke verdeling van de invoerdata moet er wel rekening gehouden worden met het probleem van overfitting. Dit zou als gevolg kunnen hebben dat het gebruik ervan voor nieuwe data inaccuraat zou kunnen zijn.[60]

**Ridge-regressie** Een andere soort regressie die gebruikt kan worden wanneer de onafhankelijke variabelen een sterke relatie met elkaar hebben, is ridge-regressie. Deze soort van regressie maakt gebruik van een regressie factor. Deze factor beheerst de sterkte van de penalty term die gebruikt wordt om het probleem van de sterke samenhang van de onafhankelijke variabelen aan te pakken, ook wel multi-collineariteit genoemd. Deze penalty kan over het algemeen ook de stabiliteit en de efficiëntie van het model bevorderen. Maar hier kan een probleem mee optreden. Een te hoge penalty zou namelijk kunnen zorgen voor een te eenvoudig model met een zeer hoge bias terwijl een te lage penalty kan zorgen voor overfitting en een hoge variantie binnen het model. Hierdoor is het belangrijk om een optimale regressie factor te zoeken zodat de bias-variantie trade-off optimaal is. Algemeen kan dit soort model zorgen voor extra stabiliteit en het probleem van overfitting tegengaan. Dit zou dan weer als gevolg hebben dat het model robuuster wordt en voorspellingen op nieuwe data accurater zijn.[61, 62]

**Lasso-regressie** Een vorm van regressie die sterk te vergelijken is met Ridge-regressie is lasso-regressie. Dit is een regressie methode die ook gebruik maakt van een regressie factor en ook hierbij aan variabele selectie doet. Dit komt er doordat de penalty hier de som is van de absolute waarden van de regressie coëfficiënten. Hierdoor ontstaat het effect dat sommige van deze coëfficiënten verminderd worden tot nul, waardoor dus eigenlijk deze variabelen wegvallen. Deze soort regressie kan nuttig zijn wanneer er zeer veel onafhankelijke variabelen zijn om zo het model eenvoudiger en beter begrijpbaar te maken.[66]

**Elastische netto-regressie** Elastische net-regressie is een regressie methode die lasso-regressie en ridge-regressie combineert om enkele van de tekortkomingen van elke methode aan te pakken. Het voegt een penalty toe aan de loss-functie van het model die een lineaire combinatie is van de penalties van lasso en ridge. Hierdoor is het met deze soort regressie mogelijk om te werken met veel onafhankelijke variabelen zijn en waarvan sommige sterk gecorreleerd zijn, wat voor lasso- of ridge-regressie alleen een probleem kan vormen. De kracht van beiden soorten penalties kan worden gecontroleerd door regressie factoren, die kunnen worden afgestemd om de beste balans te vinden tussen de variabelen selectie en regressie. Elastische net regressie wordt vaak gebruikt in situaties waarin er veel onafhankelijke variabelen zijn waarbij sommige sterk gelinkt kunnen zijn, en het kan helpen om overfitting te verminderen en de veralgemening van het model te verbeteren.[65, 73]

Nu er een beeld gevormd is van welke soorten regressie er bestaan en welk doel regressie heeft, kunnen we dit makkelijk met voorbeelden terugkoppelen naar de data die binnen dit onderzoek gebruikt zal worden. Regressie modellen zijn een reeds gekende manier voor de analyse van health-data[23, 36, 41]. Met deze modellen kunnen we relaties zoals deze tussen hartslag en cafeïne gebruik of REM-slaap en alcoholgebruik proberen te analyseren om hun onderlinge gedrag beter te begrijpen. Zeker als we zouden kijken naar de verscheidenheid van parameters die we in deze studie willen omvatten kan regressie hier helpen om overfitting te vermijden. Verder zouden we ook kunnen proberen om een model op te stellen dat de slaapkwaliteit van een persoon voorspeld. Op deze manier zouden we mensen dan kunnen bijleren over welke gevolgen er vast hangen aan alcohol-consumptie of intensief sporten met betrekking tot hoe goed ze zullen slapen vannacht. Resultaten uit regressiemodellen zullen dus verwachtingen zijn die de modellen zullen teruggeven op nieuwe invoer variabelen die we er aan meegeven.



Deze verwachten zullen echter nooit 100% accuraat zijn omdat onze modellen nooit zo accuraat kunnen worden. Het is dus nodig om de prestaties van regressiemodellen te evalueren om hun nauwkeurigheid te garanderen. Een mogelijke manier om dit te doen is door de geschiktheid van het model te controleren op de gegeven data. Een andere manier is door de verschillen tussen de voorspelde en werkelijke waarden na te gaan om ervoor te zorgen dat ze normaal verdeeld zijn en een constante variantie hebben. Bovendien kunnen kruisvalidatie technieken worden gebruikt om de prestaties van een model op nieuwe gegevens te controleren. Het is ook belangrijk om rekening te houden met overfitting. Dit kan optreden bij complexere modellen wanneer de focus van het model overgaat van de effectieve relaties tussen waarden naar de ruis die op deze relaties zou zitten. Regressie modellen zoals ridge-, lasso- of elastische netregressie kunnen helpen bij het aanpakken van deze overfitting.[4, 14]

Om nu nog even de regressie modellen algemeen samen te vatten, kunnen we stellen dat deze vaak gemakkelijk te begrijpen en interpreteren resultaten leveren. De modellen zijn goed geschikt om simpele relaties tussen verschillende variabelen te ontleden. Regressiemodellen kunnen ook goed om met zeer grote datasets met veel verschillende variabelen als input. Echter gaan regressiemodellen er wel reeds vanuit dat er een relatie bestaat tussen de onafhankelijke en afhankelijke variabelen (invoer en uitvoer). Ook zijn dit soort modellen redelijk gevoelig aan uitschieters binnen de data en kan het dus nodig zijn om op voorhand deze uitschieters te verwijderen van de data. En tot slot in een andere nadeel van dit soort algoritmen ook dat het mogelijk niet goed zal presteren op datasets met zeer ingewikkelde relaties tussen de variabelen.

### 2.2.2 Decision tree modellen

Een decision tree model of beslissingsboom model is een machine learning model dat opgebouwd is met behulp van een boomstructuur. Bij deze soort bomen wordt in elk knooppunt dan een test uitgevoerd op de attributen en stelt elke tak een de uitkomst van zo een test voor. Tot slot zijn er dan ook nog de bladen van de boom die de resultaten van de uitvoering van het model voorstellen. Deze resultaten kunnen zowel uitgedrukt zijn in klasse labels alsook in numerieke waarden die dan geïnterpreteerd kunnen worden. Het verloop van het model loopt sequentieel van de top van de boom richting een blad, waarbij dan telkens de knooppunten gevalideerd worden om de richting van het verloop aan te beslissen. Een decision tree wordt dus eigenlijk opgebouwd door middel van een reeks regels die voortkomen uit de trainingsdata en deze regels worden dan verder gebruikt om een voorspelling te doen bij het binnenkomen van nieuwe data. Deze reeks regels ontstaan door op een recursieve manier de trainingsdata te splitsen op basis van kenmerken die de data het best in homogene groepen delen tot een bepaalde eind voorwaarde is bereikt. Deze eind voorwaarde kan gedefinieerd zijn als een maximale boom diepte of een minimum aantal bladeren in de boom.[27, 63]

Binnen de medische wereld zijn er reeds een heel aantal gekende voorbeelden waarbij beslissingsbomen helpen bij het maken van beslissingen, de classificatie van medische data en het stellen van diagnoses[1, 47]. Niet alleen wordt het nemen van deze beslissingen steeds belangrijker maar ook de redenering is tegenwoordig een belangrijk gegeven. Op deze manier kunnen artsen samen met hun patiënten of aan anderen duidelijk maken waarom bepaalde keuzes voor andere komen. Binnen deze studie zal het medische nut van deze bomen niet zozeer het stellen van diagnoses zijn. Hier zouden deze bomen voornamelijk keuzes kunnen verantwoorden naar de mensen toe om aan te geven welke acties van hun welk slaapgedrag uitlokken of welke gevolgen dit mogelijks zal hebben. Zou zou er bijvoorbeeld verklaart kunnen worden dat wanneer iemand door de dag veel alcohol consumeert in combinatie met weinig fysieke activiteit dat zijn slaapkwaliteit hieronder zou kunnen leiden waardoor zijn REM-slaap veel minder zal zijn. Doordat een decision tree concrete regels zal opstellen, zal de boom makkelijk visueel voorgesteld kunnen worden. Hierdoor zouden we voorgaande voorbeeld dus bijvoorbeeld via 2 knopen kunnen afleiden waarbij een eerste knoop een test zou geven over het alcoholgebruik, de tweede knoop als kind van de eerste een test over de fysieke activiteit om zo in een blad te komen waar een slechte slaapscore als label getoond wordt.

Alvorens een beslissingsboom op te bouwen, moet er een keuze gemaakt worden over welk soort boom men wil gebruiken. De keuze van dit specifieke algoritme hangt af van een aantal aspecten zijnde: het type gegevens, het probleem dat men wil oplossen alsook hoe de interpretatie van het resultaat eruit zal moeten zien. Enkele bekende voorbeelden van soorten algoritmen hier zijn ID3, C4.5, CART en Random Forest Trees. Doordat de keuze van het beste algoritme niet exact geweten kan zijn op voorhand is het vaak nuttig om een vergelijking te maken van de verschillende algoritmen en hun prestatie te evalueren steunend op welk probleem er opgelost moet worden[32]. Naast de keuze van het algoritme zijn er nog enkele aspecten waarmee rekening moet worden gehouden. De data zal verder soms nog geschaald of genormaliseerd moeten worden. Ook categorische variabelen zullen voor sommige algoritmen op een bepaalde manier numeriek gecodeerd of herleidt moeten worden. Dan is er verder bij het opstellen van de boom ook de splitsingcriteria die sterk van belang is. Deze criteria zal bepalen op welke manier de splitsing van de knopen zal gebeuren. Hier zullen nu enkele bekende criteria toegelicht worden:

**Entropie** De entropie is een criteria dat de chaos binnen de data aanduidt. binnen beslissingsbomen willen we deze entropie juist minimaliseren zodanig dat er minder chaos is na de knoop dan ervoor. Een lagere entropie erna duidt echter ook op een meer homogene subgroep van gegevens, wat in feite ook het algemene doel is van de beslissingsbomen om deze groepen te creëren.

**Gini-index** De Gini-index is een meting die aangeeft hoe onzuiver een dataset is. De waarde hiervan ligt tussen 0 en 1, waarbij 0 overeenkomt met volledige gelijkheid van de data en 1 juist een maximale onzuiverheid aangeeft. Net zoals bij entropie is het hier de bedoeling om deze index zo laag mogelijk te houden zodanig dat de zuiverheid van de subgroepen maximaal wordt.

**Informatiewinst** Een algemene afleiding van het verschil in entropie of de Gini-index is de informatiewinst. Deze criteria evalueert de hoeveelheid informatie die een bepaalde eigenschap meedraagt aan het uiteindelijk labelen van de klasse.

**Informatiewinst Ratio** De informatiewinst ratio is een uitbreiding op de informatiewinst die probeert zijn bias te verminderen. Dit doet hij door ook de intrinsieke informatie van een bepaalde eigenschap te aanschouwen. Hij houdt zo ook rekening met het aantal takken dat zal ontstaan voordat hij de effectieve splitsing doorvoert

**Chi-kwadraat test** De Chi-kwadraat test is een criteria dat vaker gebruikt wordt voor categorische eigenschap selecties. Deze zal een meting uitvoeren over hoe afhankelijk 2 categorische variabelen van elkaar zijn en houdt er rekening mee of de waargenomen verdeling hiervan sterk zou afwijken van de verwachte verdeling. Wanneer de Chi-kwadraat een lage waarde heeft bij een splitsing geeft dit aan dat er een sterke correlatie is tussen de eigenschap en gekozen klasse.

Zoals reeds eerder aangehaald is de interpreteerbaarheid van de beslissingsbomen eenvoudig. Ze geven een duidelijke en intuïtieve weergave van het beslissingsproces dat doorlopen wordt en resultaten kunnen zeer gemakkelijk gebacktracked worden. Verder zijn de bomen ook in staat om zowel numerieke als categorische data aan te kunnen. Hierdoor zijn er niet enorm veel beperkingen op de dataset waardoor grote vereisten van preprocessing van de data achterwege blijven. Door hun niet parametrische aard zijn beslissingsbomen ook goed bestand tegen enige uitschieters die zich voordoen binnen de data. Dit maakt dat ze op dit vlak dus robuuster zijn dan bijvoorbeeld de lineaire variant van de regressie algoritmen. Door hun aard van splitsen en keuzes maken zijn deze bomen ook in staat om niet lineaire relaties tussen variabelen te ontdekken. Hierdoor zijn decision trees toepasbaar op een brede set van problemen. Beslissingsbomen hebben verder ook de kwaliteit om het belang van bepaalde eigenschappen in te schatten. Verder zijn ze ook in staat om met ontbrekende waarden toch verder te kunnen. Dit door deze waarden bijvoorbeeld buiten beschouwing te laten bij het splitsen of door deze waarden bij de meest waarschijnlijke klasse toe te voegen.[24, 64]

Naast deze reeks aan voordelen zijn er ook wel enkele nadelen bij het gebruik van beslissingsbomen. Deze bomen zijn gevoelig aan over-fitting. Wanneer de data niet correct gepruned wordt op voorhand of de takken die niet significant bijdragen aan de boom weggelaten worden, kan het model slechte resultaten halen op ongeziene data. Verder zijn bomen ook zeer onderhevig aan kleine aanpassingen binnen de data. Kleine aanpassingen kunnen soms grote gevolgen hebben waardoor er onstabiele voorspellingen gemaakt kunnen worden. Hierdoor is het ook zo dat licht verschillende datasets soms volledig andere bomen zouden kunnen hebben voor eenzelfde algoritme. Verder zijn beslissingsbomen ook niet sterk in het nagaan van complexe relaties tussen een groot aantal variabelen tegelijk. Dit is te danken aan hun eigenschap waarbij ze voornamelijk hun beslissingen maken op basis van 1 eigenschap van de data per knoop. Verder is er bij deze bomen ook de tendens om een bias te ontwikkelen waarbij ze de voorkeur geven aan variabelen met meer verschillende waarden waardoor het belang van eigenschappen met minder verschillende waarden verloren gaat. [24, 64]

### 2.2.3 Gradient boosting

Gradient Boosting algoritmen maken gebruik van machine learning technieken met als einddoel het classificeren van data of het maken van voorspellingen door middel van regressie. Dit soort algoritmen steunt op het combineren van verschillende soorten zwakkere modellen waaronder bijvoorbeeld de eerder besproken beslissingsbomen om hieruit voordeel te halen. Dit probeert men dan te doen door sequentieel modellen toe te passen met telkens de bedoeling om de fouten van het voorgaande model op te lossen. De opbouw van het algoritme is een sterk iteratief proces waar vaak eerst met een zeer simpel gelimiteerd model gestart wordt. Men gaat dat de errors van dit model na en probeert een model op deze afwijkingen te bouwen. Dan worden de errors weer gemeten en worden deze met bepaalde gewichten gecombineerd en herhaald het proces zich. Dit proces blijft zich dan herhalen tot een aangegeven diepte van het aantal modellen wordt bereikt of de prestaties van het model relatief weinig verbeterd. Door de iteraties op voorgaande fouten te baseren zijn dit soort algoritmen zeer effectief in het verbeteren van de gehele nauwkeurigheid van het model. Dit soort modellen staat ervoor bekend om zeer robuuste en accurate modellen af te leveren. [30, 40]

Wanneer we dit soort modellen proberen terug te koppelen aan hun nut voor medisch doeleinden, is dit heel gelijkaardig aan de toepassingen van de beslissingsbomen zoals beschreven in voorgaande sectie 2.2.2. Ze worden zo bijvoorbeeld gebruikt bij het nagaan van slaap aandoeningen of het voorspellen van kankers[35, 54]. Dit geeft aan dat modellen steunend op deze algoritmen zeker ook een bijdrage kunnen leveren als het zou gaan om het voorspellen van de slaapkwaliteit of andere gezondheid gerelateerde parameters. Echter door de omvang en complexiteit die deze modellen vaak met zich meebrengen kan het verklarende aspect van hun resultaten wel moeilijker verlopen. Er zijn dus extra inspanningen nodig om vertalingen te maken of inzichten te krijgen waarom bepaalde uitkomsten nu net dat resultaat geven.

Er zijn voor dit soort algoritmen ook weer een aantal verschillende soorten specifieke boosting technieken die elk hun eigen karakteristieken hebben. Hier zullen we kort toelichten welke de meest relevante soorten zijn en wat deze onderscheidt van de anderen:

**Adaptive Boosting** Adaptive Boosting of Adaboost traint gewoon iteratief voort op de fouten van de zwakkere modellen. Hierbij geeft het meer aandacht of gewicht aan de foutief geklasseerde voorbeelden waarbij de opvolgende modellen geforceerd worden om deze fouten op te lossen. Het eindproduct is een afgewogen combinatie van de verschillende individuele modellen. Dit soort modellen is sterk gevoelig aan ruis op de data en aan uitschieters. Hierdoor kan de prestatie van het model dus sterk lijden. Dit soort algoritmen is eerder eenvoudiger dan de anderen. [43]

**Gradient Boosting Machines** Gradient Boosting Machines focust in plaats van enkel op de fouten van het voorgaande model op de gecombineerde fouten van alle voorgaande modellen. Deze aanpassing ten opzichte van Adaboost modellen zorgt ervoor dat het geheel telkens verbeterde voorspellingen kan genereren. Deze soort modellen kan door zijn volledige sequentiële aard wel veel rekenkracht en tijd vereisen.[30, 40]

**Extreme Gradient Boosting** Extreme Gradient Boosting of XGBoost verbetert de GBM door toepassing van regularisatie. Dit probeert hiermee overfitting te vermijden alsook een efficiënter algoritme voor zijn constructie te gebruiken voor waarden die ontbreken te verwerken. Hierdoor bieden dit soort algoritmen zeer goede prestaties, werken ze efficiënter en hebben ze flexibeler dan GBM's. [35, 54]

**Catboost** Het CatBoost algoritme is ontworpen om op categorische data direct te kunnen gebruiken via de geordende boosting techniek. Hierbij worden er minder vooroordelen genomen en leidt dit tot nauwkeurigere voorspellingen van de categorische parameters. Deze techniek voorziet ook methoden om ontbrekende waarden te behandelen en pakt zelf ook het probleem van overfitting aan. Het is dus een zeer geschikte kandidaat voor het gebruik van categorische data zonder nood aan preprocessing. [22]

Één van de voordelen van gradient boosting is de hoge performance die de algoritmen hebben. Ze buiten verder de voordelen van meerdere verschillende modellen uit door deze samen te voegen en te gebruiken. Hiernaast kunnen ze zeer accurate voorspellingen maken en kunnen ze overweg met zeer complexe verbanden tussen de parameters. Ze kunnen verder ook een inzicht geven in de belangen die ze uiteindelijk hechten aan elk van deze verschillende parameters bij het voorspellen van waarden. Echter speelt hun complexiteit wel niet overal in hun voordeel. Dit soort algoritmen vereisen zeer veel rekenkracht en kunnen veel tijd vergen om deze op te stellen en te trainen. Verder is het voor dit soort modellen ook vaak nodig om een aantal parameters bij te stellen om een optimale prestatie te bekomen. Deze zogenaamde hyperparameters leggen bijvoorbeeld het maximale aantal beslissingsbomen of de evaluatiecriteria vast waarmee de kleinere modellen geëvalueerd worden. Het nagaan van deze hyperparameters is verder ook een intensieve stap doordat er geen optimale eentonige manier is om dit te doen en vaak enkel door middel van het uittesten van verschillende combinaties wordt gedaan. De modellen kunnen echter door hun complexiteit ook snel overfit zijn om hun trainingsdata.

Regularisatie zoals XGboosting dit doet is dus zeker aangeraden om ook op nieuwe ongeziene data hoge prestaties te kunnen halen. Verder is deze complexiteit natuurlijk ook een nadeel wanneer het gaat over het interpreteren van resultaten en het gedrag op de verandering van bepaalde parameters. Hierbij wordt bijvoorbeeld wel vaak gebruik gemaakt van extra hulpmiddelen zoals SHAP waarden. Deze waarden berekenen de bijdrage die elke waarde levert aan de nauwkeurigheid van het model om zo te proberen een uitleg te bieden aan de resultaten die het model genereert [39]. Zo kunnen ze inzicht geven aan welke waarden individueel gunstig zijn alsook welke parameters als geheel het beste bijdragen aan het model.

## 2.3 Bestaand werk

Binnen deze sectie bespreekt reeds bestaande werken binnen het onderzoeksdomein van deze thesis. Hiermee willen we bestaande resultaten aanhalen en mogelijke tekortkomingen van deze studies aan het licht brengen. Deze studies kunnen dienen als basis, leidraad of inspiratiebron in hoe dit onderzoek benaderd kan worden en welke keuzes overwogen moeten worden.

Slaapkwaliteit is een belangrijk begrip dat door mensen gebruikt wordt om aan te geven hoe goed of slecht ze aanvoelen dat ze slapen. Doordat reeds geweten is dat slaap een grote invloed heeft op het welzijn en functioneren van mensen is het belangrijk voor mensen om een zogenaamde goede slaapkwaliteit te hebben. Veel mensen echter geven vaak samen met dit begrip aan dat zij een slechte slaapkwaliteit ervaren of dat zij moeite hebben met het verbeteren hiervan[16]. Met behulp van vragenlijsten kunnen zij dit op een subjectieve manier

doen door een degelijke analyse van hun eigen slaapkwaliteit te onderzoeken. Naast deze subjectieve benaderingen zouden wij echter binnen deze studie een meer objectieve manier willen om verbanden tussen slaapkwaliteit en andere parameters te kunnen vergelijken om naast het opvragen van vragenlijsten werkelijk verbeteringen aan te kunnen bieden aan deze mensen. Bestaande onderzoeken bewijzen reeds ook dat het effectief opmeten van slaapkwaliteit aan de hand van verschillende medische waarden ook een oplossing zou kunnen bieden om meer objectieve inzichten en relaties te kunnen ontdekken [50]. Hierbij is het ook al niet meer nodig om voor de personen een nacht door te brengen in een laboratorium. Draagbare oplossingen zoals sensoren rond de pols of rond de borstkas kunnen ook al zeer nauwkeurige en goede resultaten opleveren waarmee aan de slag gegaan kan worden. We zagen dit ook al bij oplossingen zoals de slimme horloges van Fitbit en Apple die naast de slaap ook de fysieke activiteit erbij opmeten doorheen de dag. Voor hedendaagse oplossingen lijkt het daarom ook toegankelijker om gebruik te maken van deze horloges in plaats van een band rondom de borst. Deze horloges zijn makkelijker de beheren door middel van hun scherm en maken tegenwoordig al een veel groter deel uit van het marktaanbod van draagbare sensoren[33]. Uitgevoerde studies hebben verder ook al onderzoek gedaan naar parameters die de slaapkwaliteit van studenten zouden beïnvloeden. Zo is er een studie waarbij studenten gemonitord werden op basis van een aantal invloeden als cafeïne, stress en fysieke activiteit [58]. De resultaten toonden hier dat stress en onregelmatige slaappatronen een negatieve invloed hadden de slaapkwaliteit. Dit terwijl fysieke activiteit bijvoorbeeld wel voordelig kon zijn. Verder werd ook nog bevonden dat alcohol geen positief effect zou hebben op de slaap en dat er wel degelijk een verband was tussen een slechte slaapkwaliteit en de alcohol-consumptie van een student.

Nu een ander aspect waar deze studie een antwoord op zou willen bieden, is hoe we nu juist de slaap van studenten kunnen verbeteren. Hoe kunnen we verzamelde data van studenten gebruiken op een positieve manier om de resultaten duidelijk te maken aan hen? Een techniek die binnen de bestaande literatuur vaak terugkomt, is om technologie als machine learning te gebruiken in het vinden van verbanden of het maken van voorspellingen[34, 46]. Er zijn reeds studies die gebruik makend van draagbare toestellen met sensoren getracht hebben modellen te maken waarvan de voorspellende resultaten zeker accuraat zouden kunnen zijn om studenten te helpen bij het analyseren van de invloeden van verschillende parameters en voornamelijk de combinaties hiervan.

Deze studies hebben echter vaak een grote beperking waar wij wel een antwoord op willen bieden. Er is reeds werk gedaan naar welke tekortkomingen er zijn naar het verbeteren van de slaapkwaliteit[31]. In een studie van 2016, uitgevoerd door Zilu Liang en Bernd Ploderer, werd aan 12 personen gevraagd om hun slaapgegevens na te gaan met behulp van de oplossing die Fitbit aanbiedt. Hierin wilden zij onderzoeken of het voldoende was om door te zien hoe zij sliepen volgens de horloges ook daadwerkelijk in staat zouden zijn om hun slaapkwaliteit te verbeteren. Echter was dit niet het geval en werden er een aantal obstakels ondervonden die we in deze studie wel willen aanpakken. De deelnemers leken vaak niet te beseffen van zichzelf wat voor hun normale slaap is. Fitbit geeft wel opties over hoe lang zij sliepen, wanneer zij in slaap vielen, ... maar dit biedt geen oplossing op de vraag of deze waarden gemiddeld waren of dat deze hoog of laag waren. Zij haalden zo ook aan dat de grafieken van Fitbit misschien wel veel tonen, maar dat deze eigenlijk weinig zeggen. Er is dus sprake van een gebrek aan referentiepunten voor deze gebruiker over hoe hun slaap normaal verdeeld is. Verder is er ook het probleem rond de identificatie van de werkelijke redenen waarom bepaalde slaapproblemen zich voordoen. Personen krijgen van Fitbit enkel de data te zien maar worden niet begeleidt in het verbinden van hun slaap met de mogelijke factoren die bijdragen aan deze data. Zo vroegen deelnemers zich ook af of het niet interessant zou zijn om juist te kijken of bijvoorbeeld hun cafeïne-inname mogelijks iets te maken zou kunnen hebben met een afwijkende slaapscore. Naast het gebrek aan deze connecties volgde dan ook het probleem dat personen uiteindelijk dus geen idee hadden hoe ze moesten reageren om bijvoorbeeld hun slaapkwaliteit te verbeteren. Hierdoor gaven zij aan dat Fitbit wel graag dingen toont zoals een nieuwe beste score of een samenvatting van een week, maar dat deze aspecten niets bijdragen om acties te ondernemen om zichzelf te verbeteren. Bij het zetten van een doel voor het behalen van een aantal stappen

deed Fitbit dit wel goed. De deelnemers vonden dit motiverend doordat ze makkelijk wisten dat stappen zetten hen dichterbij het volbrengen van dit doel. Maar voor bijvoorbeeld het doel van het behalen van een hoge slaapscore was dit net het omgekeerde. Proberen langer te slapen op commando is niet evident en doordat ze weinig notie hadden rond hoe ze nog hun slaapkwaliteit konden verbeteren, vonden ze weinig motivatie om daadwerkelijk aan dit doel te kunnen werken. Om de werkpunten van deze studie nu samen te vatten en door te trekken naar dit onderzoek, moeten we een oplossing proberen aan te bieden die referenties kan tonen voor verschillende parameters, die connecties kan maken tussen de slaapkwaliteit en factoren waar personen een invloed op hebben en tot slot een oplossing die op persoonlijke maat acties zou kunnen suggereren aan een persoon om te proberen zijn slaapkwaliteit te verbeteren. [31]

## Hoofdstuk 3

# Methoden

Binnen deze sectie zal aangehaald worden hoe de effectieve studie binnen dit onderzoek wordt uitgevoerd. Dit onderdeel kan als een leidraad gebruikt worden bij volgende onderzoeken die dezelfde interesse delen of kan als een framework gezien worden indien een gelijkaardig onderzoek uitgevoerd zal worden. Deze sectie zal in detail beschrijven welke keuzes gemaakt worden, waarom deze keuzes gemaakt worden en welke de verschillende stappen zijn die ondernomen worden om tot de uiteindelijke resultaten van deze studie te komen. Deze sectie kan verder nog onderverdeeld worden in twee algemene subsecties. Hierbij zal het eerste deel zich meer toeleggen op welke manier de data binnen dit onderzoek verzameld werd en wat de waarnemingen hierbij juist zijn. Dit onderdeel zal zo ook proberen een toelichting te geven van de ervaringen die de personen ondervonden bij het opmeten van hun data. Zo wordt er geprobeerd om op een verstaanbare manier de sterktes en zwakten van de dataset aan te halen. In het tweede deel van deze sectie zal dan de effectieve verwerking van deze data aan bod komen. Hier zullen we het exacte verloop beschrijven van hoe de verzamelde data aangepast wordt om bruikbaar te zijn voor de analyse. Verder wordt dat ook de effectieve analyse uitgevoerd waarbij machine learning technieken zullen helpen om de verzamelde data beter te begrijpen. De exacte details van de gebruikte algoritmen en modellen en de evaluatie van deze modellen zullen hier centraal staan. Verder proberen we met deze analyse zo veel mogelijk relaties en variabelen tegenover elkaar te stellen om achteraf zo veel mogelijk van deze resultaten te kunnen interpreteren. Het uiteindelijke doel van deze analyse is om een reeks aan voorspellende en verklarende modellen te hebben zodanig dat we in een verder stadium van deze studie kunnen nagaan welk nut de modellen kunnen bijbrengen aan de mensen en op welke manier dit nut het beste uitgebuit kan worden.

### 3.1 Dataverzameling

Bij het afnemen van de data is er de overweging gemaakt om dit gedurende een periode van 80 dagen te doen voor elke testpersoon. Voor het kiezen van deze termijn is er gekeken naar welke gekende invloeden mogelijks zouden kunnen meespelen binnen het leven van de testpersonen die ook een invloed zullen hebben op deze studie. Doordat we de afname van de gegevens op 13 november 2022 startten, is er volgens ons eerst een iets normalere periode waarbij er niet onmiddellijk veel feestdagen, vakanties of andere invloeden aan bod komen. Wanneer de afname van gegevens dan richting de kerstperiode en het eindejaar gaan, zou men kunnen stellen dat er voor de testpersonen een iets meer feestelijke periode is. Wanneer dit teruggekoppeld wordt naar deze studie zou er een verwachting gemaakt kunnen worden dat de personen meer alcohol zouden consumeren en dat zij door de vakantieperiode mogelijk ook minder stress zouden ervaren. Wanneer deze feestperiode dan eindigt, zullen de testpersonen in de laatste periode komen waarbij examens centraal zullen staan. Doordat de maand januari voornamelijk zal draaien rond het studeren en het afleggen van examens verwachtten wij dat de personen mogelijk

iets vaker met stress te maken zullen hebben. Deze stress factor is net zoals alcohol ook weer een factor die van grote waarde is binnen dit onderzoek. Het einde van de periode viel op 31 januari 2023, waarna het verzamelen van de data voltooid zal zijn en de verwerking ervan zal beginnen plaatsvinden.

### 3.1.1 Testpersonen

Voor deze studie werden in totaal vijf testpersonen uitgekozen. Dit aantal is gekozen op basis van het doel van deze studie. De data van 5 personen is niet genoeg om een studie rond de accuraatheid uit te voeren, maar doordat deze studie voornamelijk als voorbeeld wil dienen voor eventuele grotere of verdere onderzoeken volstaat het om een kleinere set aan gegevens te verkrijgen die alsnog een degelijk beeld kunnen geven over de verdelingen en relaties tussen de parameters zelf en hun invloed op het welzijn. De personen zijn allemaal mannelijke studenten tussen de 18 en 25 jaar oud. Elk van hun getuigt geen gekende gezondheidsklachten of aandoeningen te hebben die voor afwijkingen kunnen zorgen. Ook volgt geen van hen een dieet waarbij er restricties staan op de inname van cafeïne of alcohol. Door deze laatste specificaties wordt er binnen deze studie toch geprobeerd om de verzamelde data zo uniform mogelijk te houden en om voorafgaande gekende verschillen tussen de personen te minimaliseren. De personen werden verder op een vrijwillige basis, specifiek uitgekozen voor deze studie.

### 3.1.2 Meetinstrumenten

Voor het verzamelen van de data zullen er twee methoden gebruikt worden. De eerste methode is het gebruik van de Fitbit Sense 2. De testpersonen zullen dit slimme horloge dragen en op deze manier zal de nodige data automatisch geregistreerd worden. Door middel van de connectie tussen het toestel en de smartphone van de persoon zal de data online opgeslagen worden waarna deze vanop afstand geraadpleegd kan worden. Het toestel zal continu gedragen worden, met uitzondering op momenten dat dit niet toegelaten is zoals tijdens een schriftelijk examen of wanneer het toestel opgeladen moet worden. Het toestel zal zowel overdag als 's nachts gedragen worden. Overdag ligt de focus voornamelijk op het opmeten van de fysieke activiteit van de personen. 's Nachts is het de bedoeling dat de slaap van de persoon opgemeten wordt. Welke parameters exact verzameld worden is terug te vinden in sectie 2.1.3. Door de eenvoud van het horloge moeten de testpersonen zich verder geen zorgen maken de opname van hun data en kunnen zij het op een normale manier gebruiken zonder extra vereisten.

Als tweede methode zullen de testpersonen een logboek gebruiken om de data verzameld door de horloges aan te vullen. Dit logboek is terug te vinden in bijlage 6.3. Van de testpersonen wordt verwacht dat zij dit logboek dagelijks aanvullen met de juiste gegevens van die dag. Ze zullen voor elk van volgende categorieën bijhouden hoeveel consumpties ze respectievelijk in de voormiddag, namiddag, avond en nacht drinken: pint bier, glas wijn, sterke drank, glas water, tas koffie en glas cola. Deze consumpties zijn gekozen op basis van de externe factoren cafeïne, alcohol en hydratatie. Het nut van deze drie in functie van dit onderzoek blijkt reeds uit sectie 2.1.2. Verder heeft dit logboek ook nog één extra parameter namelijk gevoelstoestand. Deze parameter wordt opgevraagd op dezelfde manier als de Fitbit horloges dit doen wanneer zij tekenen van stress of andere sterke gevoelens opvangen. Dit gegeven kan later nuttig zijn om bepaalde momenten of dagen te vergelijken op basis van hoe het horloge dit aangaf tegenover wat de persoon zelf aangaf in het logboek. Dit logboek is een papieren bundel die fysiek aan de testpersonen gegeven wordt. Voor deze fysieke oplossing wordt gekozen in plaats van een elektronisch alternatief, om het invullen ervan te vereenvoudigen. Een papieren versie zou bovendien ook de personen eraan kunnen doen herinneren dat zij hem moeten invullen doordat het fysiek ergens bij hen aanwezig is. Een nadeel ervan is wel dat we erop moeten vertrouwen dat de testpersonen hem dagelijks invullen doordat externe controle hierop niet van toepassing is in tegenstelling tot het dragen van de horloges.



### 3.1.3 Procedures

De testpersonen voor dit onderzoek zijn geselecteerd uit het persoonlijke netwerk van de onderzoekers, gegeven de toegankelijkheid en makkelijke medewerking. Verder vallen deze personen allemaal binnen het gezochte profiel van de testpersonen. Voordat zij officieel als testpersoon gekozen worden, wordt hen eerst mondeling gevraagd of zij interesse hadden voor de deelname aan het onderzoek. Zij worden hier ook kort toegelicht over de verwachtingen, het achterliggend doel van de studie en de verwachte tijd die zij zullen spenderen indien ze hieraan deelnemen. De selectie van de testpersonen verloopt dus op een persoonlijke en gerichte manier, maar blijft voortdurend vrijwillig gedurende de volledige periode van 80 dagen.

Omdat de data die verzameld wordt van hen opgeslagen wordt, ondertekenden zij wel een toestemmingsformulier met betrekking tot het verzamelen en opslaan van hun persoonlijke gegevens. Dit formulier is terug te vinden in bijlage 6.4. Naast dit formulier wordt hen ook een instructie blad bezorgd. Dit dient voornamelijk als een extra toelichting voor wat er van hen verwacht wordt, hoe zij deze verwachtingen kunnen nakomen en welk nut hun bijdrage zal leveren naar het verdere onderzoek. Eens deze formaliteiten overhandigt en ondertekent worden, krijgen zij ook de bijhorende meetinstrumenten. Om de start van de metingen vlot te laten verlopen worden zij geholpen bij het instellen van de horloges en krijgen zij ook een reeds aangemaakt account voor het gebruiken van de Fitbit app. De reden dat zij reeds een gemaakt account krijgen is zodanig dat wij als onderzoekers hun data remote kunnen opvragen en raadplegen. Hiermee wordt het mogelijk om op regelmatige basis te controleren of de horloges wel gedragen worden en of er zich problemen voordoen. Verder verhindert dit de testpersonen niet doordat zij op hun smartphone gewoon hun eigen voortgang ook nog steeds kunnen opvolgen.

Door de persoonlijke band met de groep testpersonen is het mogelijk om op regelmatige basis informeel een gesprek te voeren over hun bevindingen en opmerkingen over de dataverzameling. Dit zorgt ervoor dat de data voor de verwerking besproken kan worden en dat er geprobeerd wordt reeds een verklaring te vinden voor bepaalde uitschieters of onregelmatigheden binnen de data.

### 3.1.4 Ervaringen/problemen

Door de mondelinge gesprekken en opvolging van de testpersonen was het mogelijk om snel ondersteuning te bieden bij vragen, bemerkingen en eventuele problemen. Deze subsectie zal toelichten welke bemerkingen er waren en wat de gevolgen hiervan zijn of welke conclusies getrokken kunnen worden. Een eerste probleem dat zich voordeed was voor een testpersoon waarbij het horloge kort na het starten van de periode slechte metingen leek te geven op de hartslag en slaap gegevens van de persoon. Sporadisch leken de waarden wel te kloppen na verificatie met een ander toestel maar toch bleef dit fenomeen zich op sommige momenten herhalen. Na twee dagen zijn we dan tot de conclusie gekomen dat de grootte van het polsbandje van het horloge het probleem bleek te zijn. Doordat deze te groot bleek voor de persoon, hing het horloge soms te los rond zijn pols waardoor metingen van die periode verwaarloosbaar waren. Dit probleem werd opgelost door de meegeleverde kleinere polsband te gebruiken.

Verder werd er bij een van de testpersonen ook een probleem opgemerkt met het synchroniseren van de horloge met de fitbit app. Dit probleem trad op doordat de bluetooth van deze persoon uitgeschakeld werd. Dit zorgde ervoor dat het horloge niet kon synchroniseren. Dit is normaal geen probleem doordat het horloge zelf ook de data enige tijd lokaal kan bijhouden. Maar doordat de testpersoon enkele dagen zijn bluetooth vergeten was terug aan te zetten, was er data verloren gegaan. Toen werd ook opgemerkt dat de data die het horloge toen synchroniseerde eigenlijk hoorde bij andere dagen dan dat de applicatie aangaf. Het horloge sloeg enkel de eerste 2 dagen aan data zelf op maar voegde deze toe aan de laatste twee dagen van de periode waarin het synchroniseren niet lukte. Hierdoor werd er vereist dat deze data handmatig toegewezen moest worden aan de correcte data. Een gevolg van dit probleem was dat enkele dagen geen geldige data opleverden en deze dagen dus uit de verzamelde data weggelaten werden.

Tijdens de periode van het opmeten werd er ook op enkele momenten een informeel interview gehouden met de testpersonen. Hierin werd onder meer besproken wat de bevindingen waren met betrekking tot de accuraatheid van de data. De manier waarop zij deze data zagen was via de applicatie die Fitbit aanbiedt waarin de personen op elk gegeven moment de waarnemingen van het horloge konden raadplegen. Na enig overleg konden er wel enkele vaststellingen onderscheiden worden die bij de verschillende testpersonen wel terugkwamen. Zo kon er gesteld worden dat het algemene vertrouwen in de kwaliteit van de opgemeten data wel aanwezig was. Op dagen dat ze zich wat minder uitgerust voelden of merkten dat ze slechter geslapen hebben viel meestal wel op te merken dat ze inderdaad ook een lagere slaapscore hadden dan gemiddeld. De duratie van de slaaperiode vanaf het in slaap vallen tot het wakker worden werd ook aanzien als iets wat zeer accuraat leek volgens hen. Slechts bij heel beperkte gevallen werd soms eens aangehaald dat de slaap misschien niet helemaal leek te kloppen maar doordat hiervan geen concreet bewijs is en deze nachten zich zeer zelden voor bleken te doen wordt hier verder geen aandacht aan besteed. Ook wanneer er naar de fysieke activiteit opname van het horloge gevraagd werd, leken hier geen problemen mee ondervonden te zijn. De testpersonen zagen deze waarden als een correcte interpretatie van hoe fysiek zij bezig waren doorheen de dag en de waardes leken goed te kloppen.

Om een beeld te krijgen over hoe de testpersonen het continue loggen van hun dag ervaren stelden we hier rond ook enkele vragen. Zo werd hen gevraagd hoe zij zich voelden bij het feit dat zij soort van voortdurend getracked werden en of zij hier mentaal mee bezig waren. De reacties hierop waren zeer positief. Er werd aangehaald dat zij dit niet zagen als een controle of een verplichting maar eerder als een verrijking voor hun eigen levensstijl beter op te volgen. Er werd verteld dat ze door het dragen van de horloges zich meer bewust waren van hoe goed of slecht ze sliepen en hoeveel ze daadwerkelijk op een dag bewegen. Via de applicatie maakten ze vaker gebruik om zichzelf op te volgen. Vaak gebruikten ze de applicatie ook als een manier om een uitleg te zoeken indien ze zich wat minder energiek of wat slechter voelden om zo achter een uitleg voor dit fenomeen te zoeken.

Wanneer het ging over de gewenning van het horloge had niemand hier een probleem mee. Na enkele dagen was iedereen het wel gewoon om er één te dragen. Doordat de capaciteit van de batterij van de horloges ook zeer hoog was waren ze ook niet vaak genoodzaakt hem op te laden. Dit zagen zij zelf ook als een groot pluspunt. Voor het invullen van het papieren logboek waren de meningen wel eerder verdeeld. Sommigen zeiden dat ze soms wel eens vergaten om een dag aan te vullen waardoor ze dit de volgende dag moesten doen terwijl ze niet altijd 100 % zeker waren van de hoeveelheden of waarden. Ook was er één testpersoon die wel vaker wisselde van plaats waar hij verbleef. Deze moest er dan telkens aan denken om het logboek met zich mee te brengen, wat hem soms ook wel eens ontging. Er werd door sommigen ook de vraag gesteld waarom er niet voor een online oplossing gekozen werd voor het loggen van hun gegevens. Zij stelden dan dat dit het probleem van het meenemen van het papier zou oplossen en dat een applicatie ook meldingen zou kunnen geven indien het logboek aangevuld moeten worden. Na enig overleg over waarom deze keuze voor een papieren versie gekozen werd, waren zij wel van mening dat een papieren versie in het algemeen wel waarschijnlijk de betere optie geweest zou zijn.

Tijdens het dragen van de horloges merkten de testpersonen ook op dat er hen gevraagd werd over hoe zij zich voelde op een bepaald moment. Dit moment ging meestal over een periode juist een tiental minuten eerder. Zij konden deze vraag dan beantwoorden met een van de voorgeprogrammeerde gevoelens die het horloge voorstelde. Deze bevestigingen werden mogelijk gemaakt door de EDA-sensor die op de horloges aanwezig is. Deze detecteren wanneer een persoon een sterke wisseling heeft in zijn zweet niveaus op zijn huidoppervlak. Om een verklaring te vinden voor deze veranderingen vroeg het horloge dan naar hun gevoel op dat moment. Om het moment zelf niet te verstoren gebeurt dit pas enige tijd later. Wanneer de testpersonen gevraagd werden wat zij hier van vonden waren enkele antwoorden wel opmerkelijk. Een testpersoon gaf zo bijvoorbeeld aan dat hij al enige tijd bezig was aan een bepaalde activiteit waar hij het moeilijk mee had. Dan nadat hij de oplossing voor het moeilijke probleem had gevonden, vroeg het

horloge hen hoe hij zich voelde. Hij was zichzelf wel bewust ervan dat zijn lichaam rechtstreeks reageerde op het oplossen van het probleem en dat het horloge dit dus had opgemerkt. Een andere testpersoon merkte ook op dat het horloge soms na intense fysieke inspanningen deze vraag stelde. Alhoewel de functie dus wel handig kan zijn om bepaalde pieken van gevoelens op te meten is het horloge toch niet in staat om zelf de afbakening van de gevoelens voor zich te nemen via de EDA-sensor. De data die de gebruikers ingeven zou in de toekomst wel kunnen bijdragen aan de classificatie van deze emoties.

## 3.2 Analyse

Eens de vereiste data verzameld was, kan de analyse ervan van start gaan. Zowel de data verzameld door de Fitbit horloges als de logboeken worden hier aanschouwd als de 'verzamelde data' die we verder zullen gaan gebruiken en analyseren om meer inzichten te krijgen in de onderliggende verbanden die zich mogelijks zouden voordoen erop. Maar vooraleer er naar deze diepere analyse gegaan kan worden, zal de data verwerkt en gefilterd moeten worden om eventuele gebreken eruit te halen en het formaat van de data daadwerkelijk bruikbaar te maken voor deze analyse. Beiden het verwerken en analyseren van deze data zal hier verder besproken worden om het volledige verloop van deze analyse duidelijk te maken. Bij de effectieve analyse van de data zijn er wel enkele zaken op te merken. De data die gebruikt wordt is deze verzameld over 5 verschillende personen die een gelijkaardige leeftijd en levensstijl hebben, namelijk die van een doorsnee student. Voor de verwerking van de data zullen we voor de eenvoud van deze studie soms de veralgemening maken dat dit data is van eenzelfde persoon. Hiermee wordt bedoeld dat ieder zijn data binnen een machine learning model aanschouwd zal worden als deze van één persoon om bepaalde patronen te herkennen of correlaties te stellen. In werkelijkheid is dit daarentegen wel niet het geval. Elke persoon zal verschillend reageren op bepaalde facetten als alcohol en fysieke activiteit. Wij zullen hier als neveneffect van de analyse dus niet volledig perfecte resultaten bekomen. De focus van deze studie ligt voornamelijk op hoe deze resultaten zo diep mogelijk gebruikt kunnen worden om een uitleg te geven voor bepaalde fenomenen en een meerwaarde te bieden binnen het welzijn van een persoon en meer specifiek zijn slaapkwaliteit.

### 3.2.1 Dataverwerking en filtering

De verwerking en filtering die de data ondergaan is kan volgens een sequentieel verloop toegelicht worden. Eerst wordt de Fitbit data van de verschillende testpersonen verzameld. Dit gebeurt door op het platform van Fitbit online in te loggen voor elk van de testpersonen hun persoonlijk account en een aanvraag naar hun archief te doen. Eens deze aanvraag dan goedgekeurd werd, is het mogelijk om een .zip-bestand te downloaden van alle data die Fitbit aan het account gelinkt heeft. Dit gaat van de persoonlijke gegevens als mailadressen, naam en bijvoorbeeld datum van het aanmaken van het account tot de effectieve biometrische data die de sensoren binnen de horloges hebben verzameld. Eens alle zip-bestanden werden bemachtigd en ze werden uitgepakt, werd de grotere mappenstructuur bekomen. Deze bestond in total uit meer dan 4000 bestanden die in een 80-tal mappen onderverdeeld waren. Deze grote structuur verklaart de nood voor een vereenvoudiging alvorens bewerkingen en diepgaande analyses op de data uitgevoerd konden worden. Doordat ook niet alle bestanden een meerwaarde hadden voor wat er in deze studie onderzocht zou worden, werd er eerst een selectie gemaakt van de data die we binnen deze studie exact zouden gaan gebruiken.

Allereerst waren er de bestanden in verband met de fysieke activiteit van een persoon. Hier werden drie soorten van bestanden uitgefilterd voor verder gebruik. Elk van deze drie soorten werden opgeslagen in een .json-formaat. Een eerste soort waren de exercise-bestanden. Deze bestanden bevatten de data gerelateerd aan de trainingen of oefeningen die een persoon uitvoerde. Deze werden automatisch gedetecteerd door het horloge en kregen dan ook een label toegewezen naargelang de aard van de activiteit zoals bijvoorbeeld een fietstocht of een loop ses-

sie. Verder werd bij deze trainingen bijgehouden hoeveel minuten deze in totaal duurden. Om nog meer details over de training te geven, was het ook mogelijk om te zien hoeveel minuten de persoon al zittend, licht actief, redelijk actief of zeer actief was gedurende de activiteit.

Ook de begin en eindtijd werden opgeslagen net zoals nog vele andere gegevens die voor dit onderzoek minder relevant waren zoals het aantal verbruikte calorieën of de afstand van de activiteit. Een tweede soort bestanden waren deze die het aantal stappen bijhielden. Hiervoor werd ongeveer maandelijks een nieuw bestand aangemaakt. Soms verschilde dit aanmaken wel eens één of twee dagen. Binnen deze bestanden werden de stappen van een persoon bijgehouden samen met de tijdstippen van de stappen. Deze tijdstippen waren uitgedrukt in datetimes en werden afgerond op minuten, maar ze verliepen niet in regelmatige intervallen. Soms werden de stappen per minuut geregistreerd maar soms konden ook enkele minuten samengenomen worden. De reden van deze irreguliere intervallen blijft onverklaard. Als derde groep waren er voor de fysieke activiteit ook nog de active minutes files. Deze werden net zoals de stappen per maand in verschillende bestanden gezet maar er werd ook per type van intensiteit nog een verder onderscheid gemaakt. Deze types waren zittende, licht actieve, redelijk actieve en heel actieve minuten. Binnen deze soort bestanden werd er een waarde opgeslagen samen met zijn datum. Deze waarde stelde dan het totaal aantal minuten voor dat een persoon die dag in een bepaalde intensiteit heeft doorgebracht, waarbij de intensiteit dus door de naam van het bestand werd aangegeven.

Naast de categorie van fysieke activiteit, is deze van de slaap gegevens zeer belangrijk voor dit onderzoek. Voor de slaap categorie zijn er dan ook een groter aantal bestanden waaruit data genomen werden. Een eerste bestandstype is deze voor de dagelijkse hartslag variabiliteit samenvatting. Hier werd dagelijks een .csv-bestand voor aangemaakt dat slecht 1 rij met data hield. Deze hield volgende parameters die wij zullen gebruiken: de datum van de meting, de RMSSD en de NREMHR. De RMSSD, of ook wel Root Mean Square of Successive Differences, meet kortetermijn variabiliteit in de dagelijkse hartslag van de gebruiker in milliseconden en het is een waarde die binnen medische analysis wel vaker gebruikt wordt als aanvulling van enkel de hartslag van een persoon. De NREMHR is de gemiddelde hartslag van die nacht zonder deze van de REM-slaap mee te beschouwen. Twee andere typen bestanden die gelijkaardig zijn, zijn deze voor de dagelijkse ademhalingsfrequentie en de dagelijkse SP02-waarde. Ook deze hebben dagelijks een nieuw .csv-bestand met 1 rij die de datum houdt alsook de bijhorende waarde. Voor de ademhaling wordt de gemiddelde ademhalingsfrequentie geschat op basis van de diepe slaap indien mogelijk, en van lichte slaap wanneer gegevens over diepe slaap niet beschikbaar zijn. Voor de SPO2-waarde of het zuurstofgehalte in het bloed wordt een gemiddelde waarde berekend voor het zuurstofgehalte tijdens de slaap van die dag. Een ander soort bestanden voor de slaapgegevens is deze met de gegevens van de slaapscore.

De Fitbit horloges berekenen dagelijks ook een slaapscore zoals eerder aangehaald bij de literatuur bij sectie 2.1.3. Deze gegevens voor de scores en subscores werden bijgehouden binnen 1 .csv-bestand dat een rij had voor elke dag, samen met de datum en het id van de overeenkomstige log van de slaap. Tot slot is er voor de slaap gegevens ook nog het type dat het grondige verloop en overzicht van de slaap bijhoudt. Deze .json-bestanden worden ook weer ongeveer maandelijks opgeslagen. Elk element binnen de json stelt hier een slaap sessie voor van de persoon. Deze sessie omvat gegevens als het id van de slaap, de exacte begin-en eindtijd van de slaap, de datum, de duratie info over de verschillende stages van de slaap. Zo wordt er opgeslagen hoe een persoon zijn overgang tussen de verschillende slaapfasen heeft doorgevoerd. Dit gebeurt door elke kleine sectie in een bepaalde fase bij te houden samen met het totale aantal seconden van deze sectie, de begintijd van de sectie en de fase waarin de persoon zich bevond. Deze fasen zijn wakker, lichte, diepe of REM-slaap. Op deze manier zou men dus kunnen berekenen hoelang iemand in totaal heeft doorgebracht in zijn diepe slaap, of ook hoe het verloop tijdens de gehele nacht is geweest.

Wat betreft het archief van Fitbit is er nog een laatste bestand dat gebruikt zal worden, namelijk het bestand omtrent de stresscores. Deze gegevens staan allemaal gebundeld in één .csv-bestand waarbij elke rij een dag voorstelt. Naast de datum valt voor elke dag dan de stressscore

berekend door Fitbit terug te vinden samen met de subscores die gebruikt worden om deze score te berekenen. Belangrijk op te merken is dat de stressscore en het ervaren van stress omgekeerd evenredig zijn. Een hogere stressscore duidt dus op minder stress, en wordt hierdoor dus als beter aanschouwd.

Naast het filteren van welke gegevens gebruikt zullen worden, zullen sommige valse resultaten ook aangepast moeten worden. Fitbit heeft voor sommigen waarden ook velden om aan te geven of de verzamelde data correct is. Dit kan bijvoorbeeld door een bool aangegeven zijn of bij een slaap sessie kan dit bijvoorbeeld door aan te geven met een bool of het over de 'main' slaap van die dag gaat, en niet zozeer over een kort rustmoment doorheen de dag. Ook opvallend binnen de data die Fitbit verzameld is dat zij voor sommigen van de verschillende soorten bestanden niet alleen andere bestandsformaten gebruiken maar soms ook andere formaten voor de gegevens zelf. Zo gebruiken zij soms voor data ook andere formaten waarbij soms het jaartal vooraan staat in plaats van achteraan, of dat er soms in plaats van een koppelteken een backslash wordt gebruikt. Zolang deze formaten niet gelijk zijn, zal het voor het coderen natuurlijk onoverzichtelijk zijn om van verschillende bestandstypes gegevens te matchen op datum. Ook voor sommigen waarden is dit het geval. Sommige bestanden werken met minuten als achterliggende waarde voor een getal terwijl anderen werken met seconden. Om hier ook problemen te vermijden, kunnen deze best allemaal op een gelijke eenheid gezet worden. Verder kunnen er soms natuurlijk ook ontbrekende gegevens zijn waarbij een datum met gegevens voor een bepaalde parameters ontbreekt door bijvoorbeeld het niet dragen van het horloge, of door een ander probleem. Deze waarden moeten dus ook op een correcte manier aangevuld worden zodanig dat deze ontbrekende gegevens later geen probleem kunnen vormen bij een verdere analyse. Tot slot is er ook nog op te merken dat er een onderscheid gemaakt moet worden tussen de slaap gegevens van een dag en de andere fysieke gegevens. Wanneer met over de slaap gegevens van een dag gaat bij Fitbit is het zo dat telkens de slaap van de voorgaande dag aanschouwd zal worden als de slaap van die dag. Dit betekent voor dit onderzoek dat wanneer wij de gevolgen van de fysieke gegevens willen bekijken van een persoon op zijn slaap, de slaap van de volgende dag aanschouwd moet worden en dus niet die met dezelfde datum. Tot slot kunnen sommigen waarden zoals de slaapscore of de stressscore ook nul zijn. Dit geeft dan aan dat Fitbit niet in staat is geweest om voor die dag of slaap een score te berekenen. Deze waarden zijn verder waardeloos voor dit onderzoek en zullen weggelaten worden omdat zij geen correcte meting vertegenwoordigen maar enkel een gebrek aan voldoende data voor die dag.

Nu is de filtering en verwerking van de Fitbit gegevens compleet. Om nu deze gegevens te ontleden en aan te vullen zal er gebruik gemaakt worden van een python programma. Dit programma zal vanuit de .zip-bestanden van Fitbit een nieuw bestand aanmaken als resultaat waarin de eerder aangehaalde gegevens verzameld zullen worden. Omdat we met meerdere testpersonen hebben gewerkt, zal een unieke rij gevormd worden door een combinatie van een datum en de naam van de testpersoon. Elke rij stelt een dag voor met al zijn verschillende gegevens als kolommen. Naast de verzamelde parameters door de Fitbit horloges, is er ook nog het logboek dat elke testpersoon moest bijhouden op dagelijkse basis. Doordat deze op papier werden ingevuld, zullen deze ook nog overgezet moeten worden zodat ook deze data voor de verdere analyse gebruikt kan worden. Dit overzetten gebeurt op een handmatige manier waarbij het .csv-bestand uit het voorgaande programma aangevuld wordt met al de letterlijke data uit de logboeken. Een extra verwerking die hier werd uitgevoerd heeft betrekking op de specifieke groep van alcoholische dranken. Binnen het logboek werd er een onderscheid gemaakt tussen bijvoorbeeld bierglazen, wijnglazen en sterkere dranken. Omdat er binnen dit onderzoek geen aandacht is voor de verschillende effecten van de verschillende soorten alcoholische dranken zullen deze gecombineerd gaan worden. Een perfecte optelling van de consumptie van deze dranken zou inhouden dat er rekening gehouden moet worden met hun alcoholpercentage, hun hoeveelheid en zelfs nog verdergaande met de exacte tijd en verwerkingstijd van de lichamen van de testpersonen om deze alcohol af te breken. Echter het merendeel van deze details ontbreekt en deze exacte toespitsing zou zelfs medische bloedtesten nodig hebben om volledig correcte resultaten hiermee te bekomen. Binnen deze studie zullen enkele simplificaties toegepast worden

om het voor deze studie toegankelijker te maken deze data te gebruiken. Omdat de essentie van de alcohol opmeting vooral aanleunt bij de slaapresultaten, zullen enkel de consumpties die in de avond en nacht gedronken werden aanschouwd worden. Doordat er bij de logboeken ook hoeveelheden aangegeven werden bij de verschillende soorten dranken, kunnen we volgens de volumes en de algemene percentages van de genoemde consumpties een berekening maken. Voor de consumptie van alcohol wordt vaak verwezen naar het units-systeem om een effectieve waarde te koppelen aan de hoeveelheid alcohol die een persoon binnenkrijgt. Hierin kan met enige afronding genomen worden dat de 3 verschillende categorieën een gelijkaardige hoeveelheid alcohol-inname vertegenwoordigen. De afronding die hier wordt gemaakt zou in theorie of voor medische doeleinden om een meer nauwkeurige manier berekend moeten worden, maar voor de verdere doeleinden van deze studie kan er met behulp van deze afronding alsnog waardevolle data bekomen worden om bepaalde patronen in verband met de slaapkwaliteit te herkennen en weer te geven.

### 3.2.2 Beschrijvende statistiek en verdeling van parameters

Bij de verzamelde data zijn meerdere parameters waarover graag meer informatie verkregen zou worden. Het gaat hier dan voornamelijk over de aanwezigheid van correlaties tussen parameters die niet rechtstreeks met elkaar gelinkt zijn, maar die onderling wel elkaar kunnen beïnvloeden. Om deze mogelijke complexe correlaties tussen de verschillende parameters te onderzoeken kunnen een aantal verschillende technieken gebruikt worden. De keuze van welke techniek beter is of een beter resultaat zou kunnen geven hangt van een groot aantal factoren af. Deze factoren kunnen terugslaan op eigenschappen van de verzamelde data zelf, maar ook op gegeven eigenschappen van de technieken. Deze subsectie zal aanhalen welke technieken op de data gebruikt worden en wat de waarnemingen hierbij waren.

Bij de start van deze analyse worden er eerst meer eenvoudige technieken gebruikt om de verdeling van bepaalde variabelen beter te begrijpen. Om een idee te krijgen over de effectieve waarden en hun verdeling werden een aantal simpele waarden berekend. Het gemiddelde en de mediaan geven hier een indicatie over wat normale waarden van deze variabelen zijn binnen de data. We zouden kunnen stellen dat deze waarden gezien kunnen worden als de baseline voor de variabelen van de testpersonen. Verder kan de standaardafwijking een verklaring geven over hoe verspreid de data is of hoever deze af kan wijken van de gemiddelde waarde. Tot slot werden dan ook nog de minimale en maximale waarden ontleedt alsook de kwartiel verdeling van de variabelen. Dit laatste bestaat uit 3 waarden Q1, Q2 en Q3 en kan ook weer extra inzicht geven over de verdeling van de gegevens en helpt ook bij het nagaan van eventuele uitschieters binnen de data. Doordat Q2 overeenkomt met de mediaan wordt deze niet nogmaals expliciet vermeld.

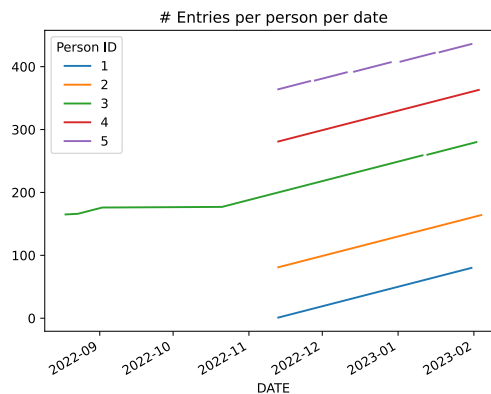
Binnen tabel 3.1 worden deze waarden op verschillende van de parameters weergegeven. Opvallend hieraan is dat voor sommige parameters zoals de RMSSD de minimum waarde nul is. Dit is te verklaren door de restricties die opgelegd werden bij de verwerking en filtering van de dataset zelf zie 3.2.1. Deze waarden zijn in tegenstelling tot de slaap- en stressscore niet weggefilterd omdat andere waarden van die dag wel een meerwaarde kunnen hebben voor de verdere analyse. Dit maakt dat binnen deze tabel voor deze waarden sommigen waarden zoals het gemiddelde een afwijking kunnen hebben. Maar door het beperkt voordoen van deze nul-waarden, namelijk minder dan 1%, blijven deze afwijkingen zeer beperkt en beschouwen we enkel hun minimum waarde als nutteloos. Deze algemene kennis over deze statistische waarden voor elk van de parameters kan bijstaan om bij de verdere analyse bepaalde fenomenen beter te begrijpen, beter met uitschieters om te gaan en om op een zinvolle manier de resultaten van andere bewerkingen te interpreteren.

Nu er een beeld gevormd werd over de verdeling van de waarden van de variabelen, kan er gestart worden met een visuele voorstelling om de data te analyseren. Bij een eerste iteratie werden de verschillende ID's van de verzamelde data weergegeven in functie van de datum en de testpersoon waarbij deze hoorden. Via deze voorstelling kan men duidelijk zien wanneer

de periode was waarbij de data van elke testpersoon werd afgenomen. Deze grafiek is te zien op figuur 3.1. Opvallend aan deze figuur is dat de groene lijn van de testpersoon met ID 3 duidelijk langer is dan de rest. Dit is te wijten aan het feit dat deze testpersoon deel uitmaakte van de onderzoekers zelf en reeds eerder van start kon gaan met het dragen van het horloge. Verder hebben personen met ID 3 en 5 ook beiden een of meerdere missende waarden binnen hun data. Dit is te zien door de onderbrekingen die zich voordoen op hun lijnen binnen de grafiek. Het probleem bij deze missende waarden werd reeds aangehaald binnen sectie 3.1.4 en werd verklaard door het mislopen van het syncen van de horloges.

Variabele	Gemiddelde	Mediaan	Std	Minumum	Maximum	Q1	Q3
Stress score	75.45	75	5.79	61	90	72	75
Sleep score	78.55	80	6.78	51	90	75	80
Steps	7246.16	6322	4337.51	85	30130	4238	6272
Sleep duration	488.77	495	76.93	190	713	443.5	495
Deepsleep	91.03	90	22.11	32	153	75.5	90
Wakesleep	64.62	64	15.95	15	127	54.5	64
Lightsleep	257.87	258	54.96	73	462	223.5	258
REMsleep	75.25	74	29.7	0	160	54	74
Very active mins.	28.3	7	42.42	0	311	0	6
Sedentary mins.	684.19	692	112.56	137	1268	620.75	697.5
Moderately active mins.	20.24	7	34.92	0	192	0	7
Light active mins.	215.55	203	80.63	6	491	155	200
Morning exercise mins.	2.54	0	10.45	0	160	0	0
Noon exercise mins	5.44	0	26.45	0	1212	0	0
Afternoon exercise mins.	11.33	0	30.16	0	199	0	0
Night exercise mins.	1.82	0	8.11	0	119	0	0
RMSSD	84.82	92.18	31.82	0.00	148.32	56.06	92.18
NREMHR	51.21	51.34	9.22	0.00	88.91	44.09	51.34
Daily respiratory rate	14.9	14.8	1.79	0.00	21.60	14.00	14.80
SPO2	95.55	96.4	8.85	0.00	99.10	95.65	96.40

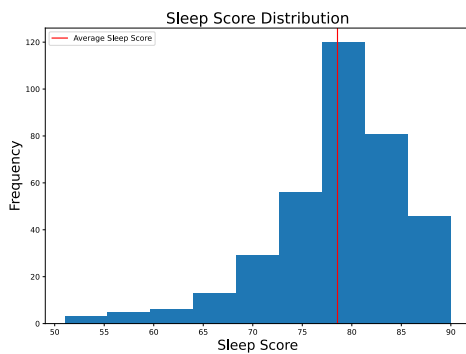
**Tabel 3.1:** De tabel geeft algemene statistische waarden over de verzamelde data variabelen weer.



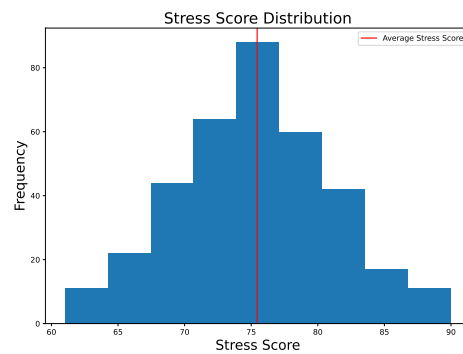
**Figuur 3.1:** De figuur geeft de ID's van de verschillende dagen per testpersoon weer.

Door tabel 3.1 is er nu een beeld gevormd over hoe de verschillende parameters zich verhouden. Binnen deze studie zijn voornamelijk twee parameters waar de algemene focus op zal liggen, namelijk de slaap- en stresscores. Om de verdeling van deze parameters beter te begrijpen zullen we gebruik maken van 2 histogrammen. Op figuur 3.2 kan de verdeling van de slaapscores van de testpersonen waargenomen worden. Hierbij wordt ook met de rode verticale lijn het gemiddelde van deze waarden aangegeven. Voor de verdeling van de scores is het zichtbaar dat

deze naar links scheef verdeeld zijn. Naar onze bevindingen valt dit wel degelijk te verklaren. Het is namelijk makkelijker om slecht te slapen dan om bijzonder goed te slapen. Doordat onze testpersonen ook zelf hen scores konden opvolgen, hebben we dit in samenwerking met hun kunnen onderzoeken. Door naar feesten te gaan, zeer vroege wekkers of door alcoholgebruik werd er vaak opgemerkt dat dit een redelijk lage slaapscore kon als gevolg kon hebben. Dit in tegenstelling tot wanneer de personen gewoon vanzelf op een natuurlijke manier wakker werd en 'uitgeslapen' was, zoals men dit vertelt. Als er dan geweten is dat de gemiddelde slaapscore 78.55 bedraagt, de bovengrens op 100 ligt maar de werkelijk behaalde bovengrens slecht 90 is, is er dus veel minder marge om uitschieters te krijgen die ver boven het gemiddelde liggen. Daarentegen is het wel mogelijk en ook duidelijk aan de data dat uitschieters naar beneden zich wel kunnen en ook vaker zullen voordoen. Wanneer er dan naar het histogram op figuur 3.3 gekeken wordt, zien we hier een typische normale verdeling terug komen. Voor de stressscore kan dus vastgesteld worden dat extreem sterke afwijkingen van het gemiddelde veel minder voorkomen, en dat de meeste scores zich dus evenredig verdelen op een korte afstand van dit gemiddelde.



**Figuur 3.2:** Het histogram geeft de verdeling van de slaapscores van de testpersonen weer, samen met een aanduiding voor de gemiddelde waarde.



**Figuur 3.3:** Het histogram geeft de verdeling van de stresscores van de testpersonen weer, samen met een aanduiding voor de gemiddelde waarde.

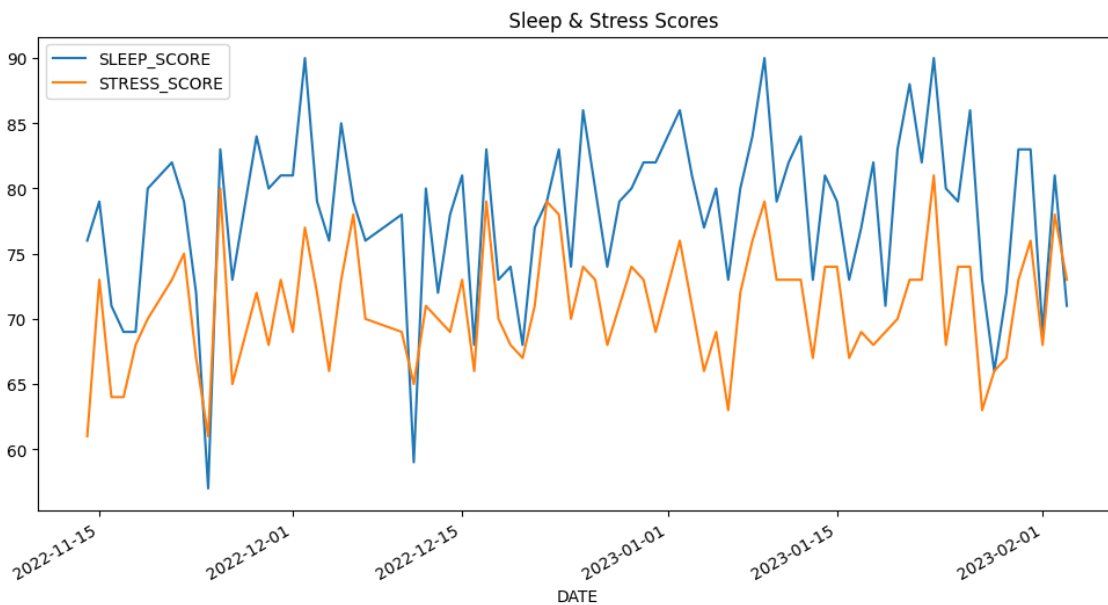
### 3.2.3 Identificatie van patronen en correlaties

Na grafiek 3.1 werd voor elk van de andere variabelen een gelijkaardige grafiek gemaakt met in de plaats van de ID's de waarde van deze variabele in functie van de datum. Om dit overzichtelijk te houden en geen 5 verschillende lijnen op één grafiek samen te zetten, werd er een aparte grafiek gemaakt voor elk van de testpersonen. Deze grafieken stelden zo het verloop van de variabele voor in functie van de tijd. Mogelijke verbanden die hier zichtbaar konden zijn, waren periodes waarbij de slaapscore overduidelijk lager zou liggen of stresscores die gedurende een periode lager zouden liggen. Echter was binnen deze visualisaties in functie van de datum niet dadelijk iets opmerkelijk waarneembaar. De waarden voor de verschillende variabelen leken eerder willekeurig verdeeld. Door de periode van de data afname hadden we hier oorspronkelijk wel andere conclusies verwacht. Het eerste deel van de periode viel in een normalere periode zonder examens of feestdagen, het tweede deel viel rond de eindejaarsperiode en hierop volgde dan de derde periode namelijk de examenperiode voor alle testpersonen. Echter waren op vlak van stress en slaap geen verschillen te detecteren tussen bijvoorbeeld periode twee en drie, terwijl hier eerder verwacht werd dat de stress zou toenemen en de slaapkwaliteit zou verminderen binnen de examenperiode.

Door de manier waarop Fitbit zijn stressscore berekent zouden we kunnen stellen dat er een duidelijk verband zichtbaar moet zijn wanneer de slaapscore naast deze stressscore gelegd zou worden. Om deze hypothese te controleren wordt er een grafiek gebruikt die 2 lijnen voor deze



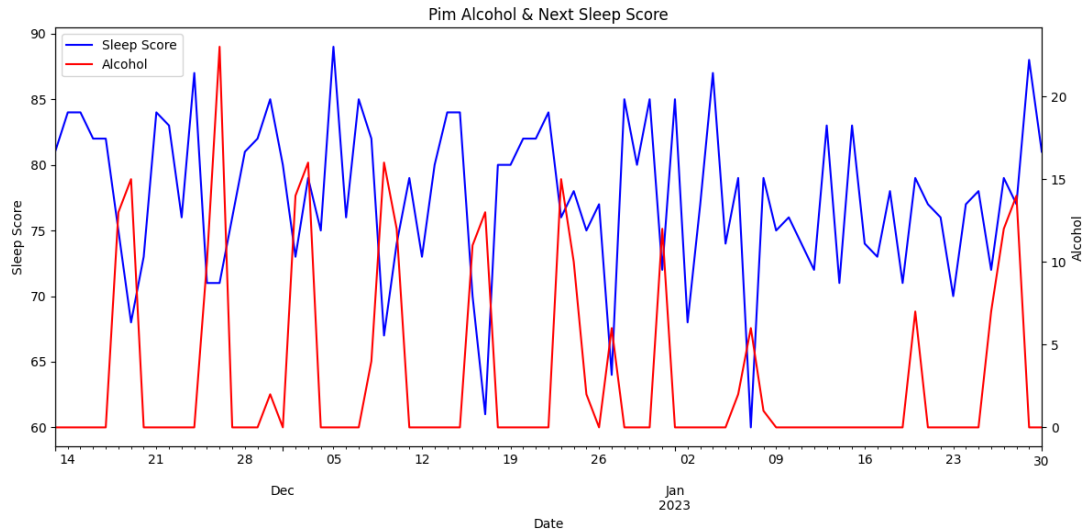
scores weergeeft in functie van de tijd. Het resultaat van deze weergave is te zien op figuur 3.4. Op te merken hier is dat deze figuur slechts de data van één van de testpersonen weergeeft. Dit wordt gedaan ter vereenvoudiging omdat de resultaten bij de andere testpersonen voor deze vergelijking zeer gelijkaardig zijn. Wanneer er naar de figuur gekeken wordt, kan men stellen dat er bij de scores een recht evenredig verband zichtbaar is. Verder is te zien dat de stressscore gemiddeld een stuk lager ligt dan de slaapscore. Deze waarneming werd reeds ook bevestigd binnen de statistische analyse van de variabelen waar de slaapscore ook gemiddeld afgerond acht eenheden hoger ligt. Tot slot moet er wel gesteld worden dat desondanks het verband dat zichtbaar is, nog altijd wel duidelijk is dat andere factoren er nog steeds voor zorgen dat sommigen afwijkingen binnen dit verband niet helemaal evenredig lijken met elkaar. Hiermee wordt dan bedoeld dat de reden waarom de scores soms zeer gelijkaardig zijn of soms juist niet helemaal elkaar trend volgen, te wijten zijn waarschijnlijk aan de andere parameters die Fitbit in acht neemt om de stressscore te bepalen.



**Figuur 3.4:** De figuur geeft de slaap- en stressscore weer van één testpersoon in functie van de tijd.

Een ander interessant gegeven binnen dit onderzoek is de invloed van alcohol op de slaapkwaliteit van de testpersonen. Er kan hierbij gekeken worden naar de alcohol-inname van een persoon in verhouding tot de behaalde slaapscore van de volgende dag. We zouden hier dus graag willen zien of er een zichtbaar verband is tussen het innemen van bepaalde hoeveelheden alcohol en slaapkwaliteit na deze consumptie. Net zoals voor de vergelijking van de stressscore en de slaapscore wordt het verband tussen beiden eerst op een zeer eenvoudige manier bekeken. Voor de twee parameters werd opnieuw een lijn grafiek opgesteld in functie van de tijd. Deze grafiek is terug te vinden op figuur 3.5. Doordat de parameters beiden verschillende eenheden uitdrukken, is bij deze grafiek wel rekening te houden dat de verticale hoogteverschillen niet rechtstreeks te vergelijken zijn op basis van hun sterkte in stijging of daling. Het is aangeraden om voornamelijk te kijken of de dagen met meer alcoholische consumpties of juist deze zonder een bepaald terugkerend patroon vertonen met betrekking tot de lijn van de slaapscore. Op deze figuur 3.5 willen we graag twee bevindingen opmerken. Enerzijds als er enkel gekeken wordt naar de dagen waarin er geen alcohol geconsumeerd werd, merken we dat hier geen onmiddellijke trend in terug te vinden is. De slaapscores variëren hier van een kleine 70 tot bijna 90 in wat zonder extra gegevens een willekeurig patroon lijkt. Wanneer er anderzijds naar de dagen met meer dan vijf alcoholische consumpties gekeken wordt, in hier wel een duidelijk patroon zichtbaar. Namelijk, de sterke stijgingen van de alcohol brengen als reactie zeer vaak

een sterke daling van de slaapscore met zich mee. Vanuit dit standpunt en deze bevindingen wordt de vraag of alcohol de slaapkwiteit doet afnemen nog niet zozeer bevestigd, maar het suggereert wel dat een mogelijk verband tussen de twee niet uitgesloten wordt.



**Figuur 3.5:** De figuur geeft de slaapscore en alcohol-inname weer van één testpersoon in functie van de tijd.

Voor verdere relaties binnen de verschillende variabelen te ontdekken, kan op een gelijkaardige wijze verdergegaan worden. Binnen dit onderzoek zijn er zo nog een aantal combinaties van parameters tegenover elkaar gezet om te zoeken naar bepaalde verbanden maar deze waren niet altijd duidelijk zichtbaar. Patronen herkennen op deze manier is niet eenvoudig doordat er een sterke relatie moet zijn tussen de twee variabelen alvorens hij op deze manier duidelijk wordt weergegeven. Maar omdat hier niet onmiddellijk verbanden worden gevonden, sluit niet uit dat er achterliggend geen relatie zou kunnen gelden onder de andere combinaties van variabelen. Om een stap verder te gaan in het ontdekken van deze evenredige verbanden, wordt er gekozen om correlatie plots toe te passen tussen enkele van de parameters.

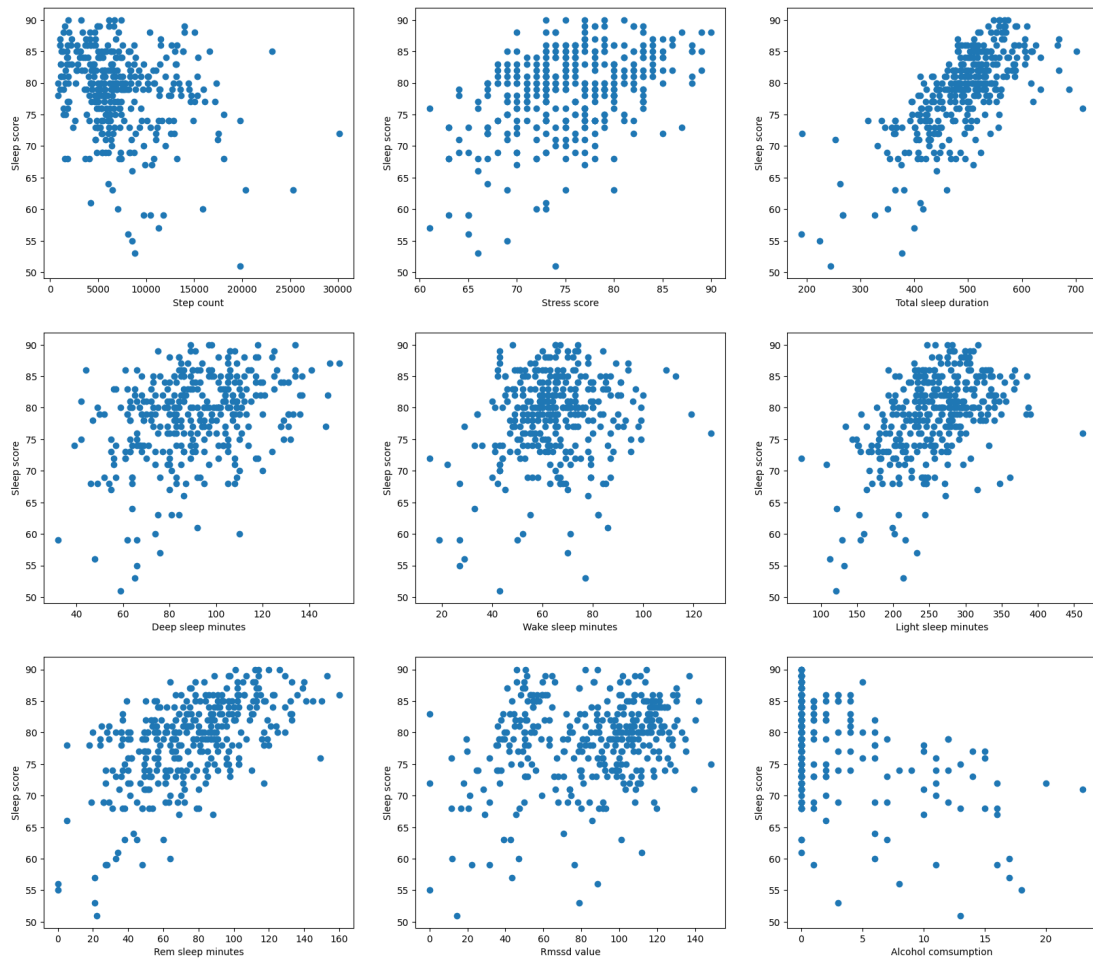
Één van de belangrijkste parameters binnen de dataset is de slaapscore. Door enkele van deze correlatie plots op te stellen in functie van andere parameters wordt er geprobeerd vanuit een andere invalshoek een mogelijk verband met de parameters te vinden. De uitwerking van deze plots is zichtbaar op figuur 3.6. Binnen zulke soort visualisaties moet er voornamelijk met twee aspecten rekening gehouden worden. Een eerste heeft te maken met de verdeling van de verschillende punten. Indien de verschillende punten zich meer positioneren richting een denkbeeldige lijn, des te sterker de correlatie tussen de variabelen aanwezig is. Anderzijds wijst een meer ongelijk verdeelde figuur erop dat er een zwakkere relatie is tussen de variabelen. Het tweede aspect waar er hier op gelet zal worden, is de richting van de mogelijke correlatie. Deze richting kunnen we als positief of als negatief beschouwen. Een positieve correlatie doet zich voor indien bij een toeneming van één variabele, de andere ook zal toenemen. We zouden dan de eerder aangehaalde denkbeeldige lijn omschrijven als een rechte die stijgt op de grafiek. Zo een voorspellende lijn wordt in wiskunde ook een regressielijn genoemd. Als we dan een negatieve correlatie zouden aanschouwen, dan zal deze regressielijn een dalende trend volgend. Dit geeft dat aan dat de stijging van één van de variabelen een daling bij de andere als gevolg zou hebben en andersom.

Wanneer we nu naar figuur 3.6 kijken, kunnen we voor enkele van de plots wel een relatie opmerken. Wanneer we de slaapscore in functie van de slaap duratie bekijken, kan er bevonden worden dat er hier zeker sprake is van een positieve correlatie tussen beiden. Er is namelijk een stijgende relatie en de punten bevinden zich allemaal dichtbij elkaar en vormen zo een

denkbeeldige lijn. Wanneer we naar de relatie van de slaap- en stresscores kijken, zien we ook duidelijk een positieve trend tussen beiden. Echter is de verdeling hiervan wel meer verspreid, wat zou kunnen duiden op een iets minder sterke correlatie tussen beiden. Verder zijn voor beiden de plots met de REM-slaap minuten en de lichte slaap minuten in functie van de slaapscore ook weer lichte positieve relaties zichtbaar. Indien we naar de plot kijken die ook het aantal stappen als tweede variabele heeft, zijn deze waarden ook weer wijd verspreid over de plot. Maar desondanks deze spreiding zou er beargumenteerd kunnen worden dat deze waarden al dan niet een negatievere trend lijken te hebben met de slaapscore. Maar door de spreiding van de punten kunnen hier zeker geen harde conclusies getrokken worden. Hetzelfde zou men kunnen stellen over de relatie met de alcohol-consumptie. Zoals we eerder al zagen bij figuur 3.5, stelden we vast dat alcoholgebruik een negatief effect had op de slaapscore. Nu bij de correlatie plot is er sprake van een wijde spreiding van de punten, maar net zoals bij deze van het aantal stappen zou men kunnen stellen dat er wel degelijk een eerder negatieve tendens aanwezig is.

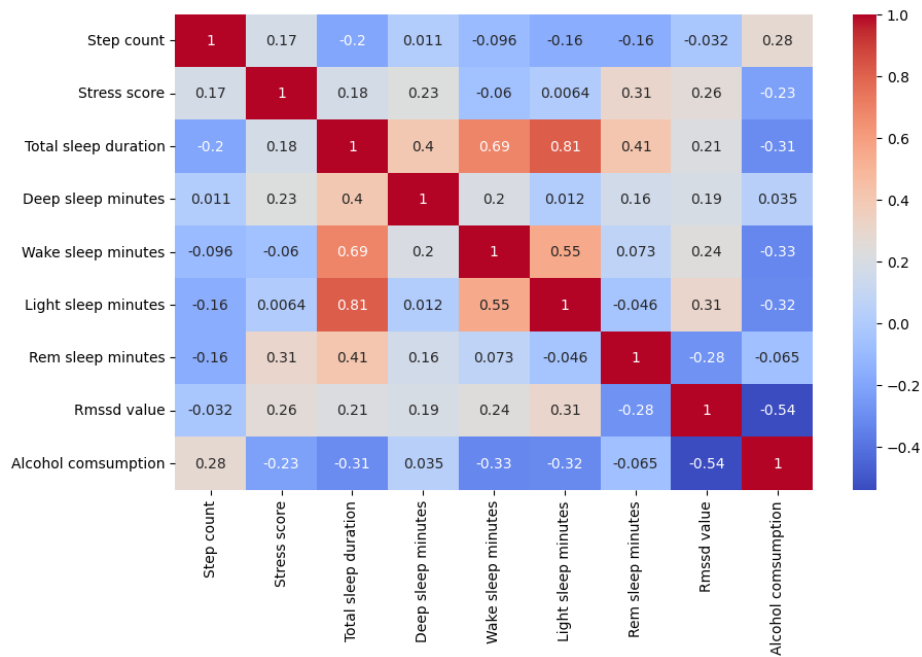
Tot slot voor de resterende plots zoals deze van de wakkere slaap met de slaapscore is er een sterke spreiding van de waarden waar niet onmiddellijk een duidelijke trend opgemerkt kan worden. Men zou alsnog een mening kunnen hebben of er toch geen positieve of negatieve correlatie zich voordoet, maar met deze onzekerheid is de meerwaarde van die bevinding toch van weinig waarde. Ten slotte is het voor deze plots belangrijk op te merken dat het opmerken van een correlatie niet betekent dat hier een oorzaak gevolg relatie optreedt tussen de variabelen. Zelfs bij een zeer sterke en duidelijke correlatie geeft dit geen bewijs voor een causale relatie tussen de twee. Er kunnen nog altijd andere achterliggende parameters zijn die een invloed uitoefenen op deze zichtbare correlaties. Ook is het nog nodig om aan te halen dat relaties niet altijd zicht voordoen als een rechte. Dit geldt enkel voor lineaire relaties, maar andere meer complexe correlaties kunnen zich ook voordoen die met deze techniek niet onmiddellijk ontdekt kunnen worden. Hiervoor zullen verdere technieken of bewerkingen nodig zijn om de achterliggende patronen of relaties te ontdekken. Verder binnen dit onderzoek zullen correlatie plots nog terug komen om correlaties tussen andere variabelen aan te halen maar door de omvang van het aantal variabelen binnen de dataset werd hier enkel voor de voorgaande combinaties een gedetailleerde toelichting gegeven. Andere combinaties kunnen echter wel op een gelijkaardige manier benaderd en toegelicht worden.

Voorgaande correlatie plots kunnen een sterk vermoeden geven voor mogelijk correlaties die zich zouden voordoen, maar voor een correcte interpretatie wordt er voornamelijk gesteund op het inzicht en de kennis van de waarnemer zelf. De waarnemingen zijn vaak geen zwart-wit situaties waarbij met 100% zekerheid gezegd kan worden wat het verband tussen twee variabelen is. Een andere manier om deze correlaties tussen variabelen te ontdekken is met behulp van een heatmap. Dit soort visualisaties kan metingen doen waarbij de sterkte en de richting van de mogelijke lineaire relatie tussen twee variabelen gekwantificeerd wordt. Een correlatie coëfficiënt wordt berekend voor elk paar variabelen van de dataset. Deze waarden kunnen variëren van -1 tot 1, waarbij -1 een perfecte negatieve correlatie aangeeft en +1 aanschouwd kan worden als een perfecte positieve correlatie. Een waarde van 0 duidt op het gebrek aan een lineaire correlatie tussen beide variabelen. Hiermee wordt het dus duidelijker om als persoon zonder ervaring om exacte kennis over de correlatie plots wel een opvatting te bekomen over de mogelijke lineaire correlatie die zich voordoet tussen twee variabelen. Verder geven heatmaps ook een kleuren schaal mee voor deze berekende waarden. Door het gebruik van toepasselijke kleuren kan het voor de waarnemers sneller duidelijk worden welke variabelen een sterkere positieve of sterkere negatieve correlatie met elkaar hebben. Op figuur 3.7 is een heatmap voor een beperkte set van onze variabelen zichtbaar. Het eerste wat opvalt is natuurlijk de diagonaal van rode vakken. Deze stellen de correlaties van de variabelen met zichzelf voor, wat natuurlijk een perfecte positieve correlatie is doordat de waarden exacte mappings zijn van zichzelf. Verder zien we ook een sterk positieve correlatie tussen de totale duratie van de slaap en het aantal lichte minuten slaap. Dit is te verklaren doordat het gedeelte lichte minuten van een persoon zijn slaap het grootste deel van zijn totale slaapcyclus uitmaakt. Dit is terug te vinden in tabel 3.1, waarbij de gemiddelde totale slaap 487 minuten bedraagt en de lichte slaap hier gemiddeld 257 minuten



**Figuur 3.6:** De figuur geeft een aantal correlatie plots weer van de slaapscore van alle testpersonen in functie van enkele andere parameters die betrekking zouden kunnen hebben op deze score.

ofwel meer dan de helft van deze totale slaap uitmaakt. Hierdoor is het logisch dat beiden sterk afhankelijk zijn van elkaar. Eenzelfde soort correlatie is op te merken bij de totale slaap en de wakkere slaap. Dit is eerder opvallend doordat het aantal wakkere minuten gemiddeld genomen slechts 64 minuten is ofwel afgerond een 13% is van de totale slaap. Nu we deze twee sterke correlaties opgemerkt hebben die beiden een link hebben met de totale duratie van de slaap, zou er gekeken kunnen worden naar de onderliggende correlatie van deze 2 variabelen zelf. Hier is te zien dat de wakkere minuten en de minuten lichte slaap ook een redelijk positieve correlatie vertonen. Hier geldt echter nog steeds de opmerking dat een sterke correlatie geen causale relatie vastlegt tussen de twee variabelen. Hiervoor zouden extra medische proeven uitgevoerd moeten worden om zekerheid te scheppen. Op de heatmap van figuur 3.7 is ook een nogal sterke negatieve correlatie op te merken, namelijk deze tussen de alcohol-consumptie en de RMSSD waarde van die nacht. Na verder onderzoek achter deze combinatie wordt door bronnen ook bevestigd dat deze negatieve correlatie geen toeval is, maar reeds gekend is binnen de medische wereld[26]. Verder kunnen er nog enkele zwakkere correlaties opgemerkt worden met absolute waarden rond 0.4 en 0.3. Deze kunnen zelf ook wijzen op lineaire correlaties die aanwezig zouden zijn tussen de verschillende variabelen, maar hier gaan we momenteel niet verder op in omdat de zekerheid hiervan niet onmiddellijk te garanderen is.



**Figuur 3.7:** De figuur geeft een aantal correlatie plots weer van de slaapscore van alle testpersonen in functie van enkele andere parameters die betrekking zouden kunnen hebben op deze score.

### 3.2.4 Voorbereiding machine learning algoritmen

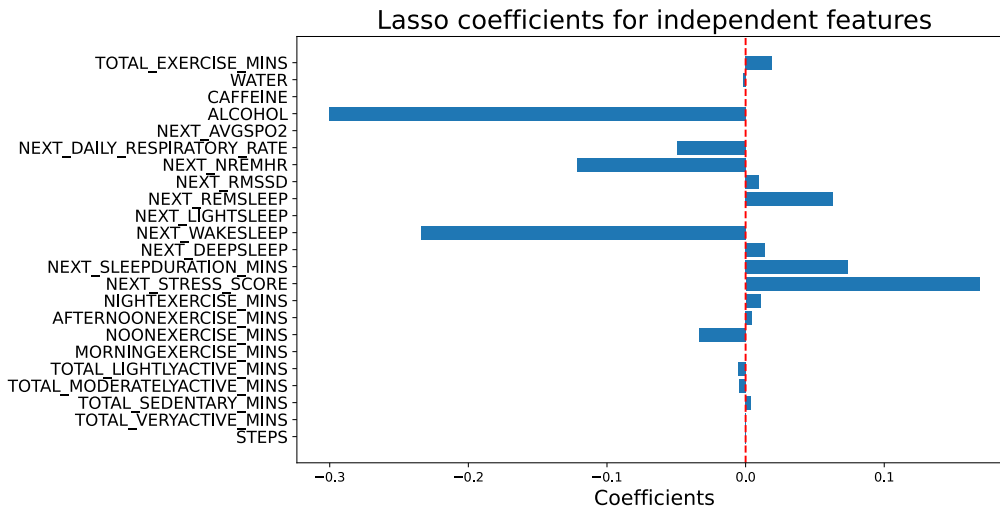
Naast het toepassen van beschrijvende statistiek van de data of het nagaan van lineaire correlaties tussen verschillende variabelen zijn er nog andere meer geavanceerde technieken mogelijk voor een verdere analyse naar patronen en eigenschappen van onze dataset. Binnen deze studie zullen we enkele machine learning algoritmen modellen opstellen die een diepgaande analyse van de data mogelijk zullen maken. Deze modellen worden gekozen op basis van hun eigenschappen op de soort data die beschikbaar is voor deze studie alsook op de resultaten die zij kunnen afleveren. Deze exacte modellen, hun resultaten en welke conclusies getrokken kunnen worden uit deze modellen zullen verder in de volgende secties besproken worden. Omdat voor deze studie de slaap- en stressscore als belangrijkste parameters aanschouwd worden, zullen de modellen voornamelijk gefocust zijn op het voorspellen van deze scores. Op deze manier willen we proberen te achterhalen welke invloeden de andere parameters op deze twee hebben om zo uiteindelijk te kunnen werken naar een verbetering van de slaap en een verlaging van de stress. Dit doel zal dan uiteindelijk ook bijdragen naar wat deze studie nu net wil beantwoorden, namelijk hoe we de levenskwaliteit van ons doelpubliek zouden kunnen verbeteren. Verder binnen deze en volgende secties zullen we refereren naar deze voorspellende scores als de afhankelijke parameters van de modellen, waarbij de overige dan de onafhankelijke parameters voorstellen. Omdat een model slechts 1 parameter tegelijk als afhankelijk zal nemen, zal de andere van de twee soms wel als onafhankelijke gebruikt worden.

Modellen vereisen voor hun gebruik een opsplitsing van de dataset en andere voorbereidingen. Dit is om een accurate representatie van de prestaties van de modellen te kunnen berekenen en te valideren. We kunnen dan onze modellen trainen, testen en valideren ten opzichte van elkaar maar hiervoor zullen wij dus een splitsing moeten toepassen op de dataset [51]. Hiermee willen we het neveneffect van duplicaten bij het trainen en testen vermijden. Wij zullen onze data allereerst in een trainingsset en een testset splitsen. Deze eerste zullen we uiteindelijk dan nog eens verder onderverdelen in de echte trainingsset en de validatieset. De trainingsset zal uiteindelijk nog altijd de grootste subset zijn van de drie en zal gebruikt worden om de machine learning modellen effectief te trainen. Deze zal de dataset krijgen waarbij hij zowel

de onafhankelijke parameters als de overeenkomstige afhankelijke parameters, ofwel de ground truth meekrijgt. De validatie subset zal dan een kleiner deel uit de dataset zijn dat gebruikt kan worden om de hyperparameters van het model bij te werken. Dit bijwerken gebeurt echter niet tijdens de training van het model maar deze subset wordt effectief als validatie gebruikt om de beste configuratie van de hyperparameters te vinden op nieuwe ongeziene data. Dit zorgt er ook voor dat overfitting verholpen wordt indien het model zeer goed op de training data werkt maar slecht werkt op nieuwe data.

Tot slot is er de test set die nogmaals een complete afscheiding is van de voorgaande twee die zal gebruikt worden om de effectieve finale prestaties van het model te evalueren, waarbij de beste hyperparameters al vast liggen en er dus eigenlijk op een echt scenario het model wordt uitgevoerd.

Nu de dataset in verschillende subsets is onderverdeeld zal deze nog geschaald worden. Hiervoor gebruiken we Min-Max scaling of normalisatie. Hierbij zorgen we ervoor dat alle verschillende waarden op eenzelfde schaal worden herleid van 0 tot 1. Dit heeft enkele voordelen zoals het sneller kunnen trainen van de modellen, maar voor sommigen soorten modellen haalt het ook weg dat deze meer belang hechten aan de parameters met grotere waarden.



**Figuur 3.8:** De figuur geeft het resultaat van het lasso regressie model waarbij de belangcoëfficiënten van de onafhankelijke parameters voor de slaapscore worden weergegeven.

Nu onze data onderverdeeld is in subsets en geschaald is, is deze klaar om als input gebruikt te worden binnen modellen zelf. Maar wanneer we op onze volledige trainingsset nu lineaire regressie zouden aanbrenge, dan zou dit de verliesfunctie minimaliseren door coëfficiënten te kiezen voor elke parameter. Echter geeft dit het probleem dat grotere coëfficiënten zullen zorgen voor overfitting. Hierbij zal het model zich te sterk om de trainingset willen laten matchen waardoor de prestatie zeer slecht zal zijn op ongeziene nieuwe data. Om dit probleem op te lossen kunnen we gebruik maken van regularisatie. Met deze methode willen we een straf koppelen aan de algemene kostenfunctie van het lineaire regressie model om zo de grootte van de coëfficiënten te doen verminderen. Zoals eerder besproken in sectie 2.2.1 kunnen we nu lasso regressie gaan gebruiken om een selectie van meest belangrijke parameters te selecteren uit de dataset. We kiezen hier voor lasso regressie boven bijvoorbeeld de gelijkaardige ridge regressie omdat deze de coëfficiënten van de minder relevante parameters zal reduceren tot bijna nul. Verder zal dit ook het probleem van multicollineariteit oplossen doordat we deze parameters met een zeer kleine coëfficiënt zullen laten vallen bij het opstellen van onze modellen. Op deze manier kunnen we ons groot aantal verschillende parameters doen afnemen om berekeningen

en de interpretatie van de modellen zeer sterk te vereenvoudigen op een manier waarbij we de prestaties wel proberen te bewaren. Wanneer we enkel onze modellen nodig hebben voor hun voorspellende waarde zou deze multicollineariteit geen probleem zijn, maar om dit model correct te interpreteren en toe te lichten achteraf zou nagenoeg onmogelijk zijn. Voor deze studie willen wij het beste van deze twee werelden combineren zodanig dat onze modellen goede voorspellingen maken maar toch interpreteerbaar blijven. Wanneer we nu naar figuur 3.8 kijken zien we hier de coëfficiënten van al de verschillende parameters. Deze waarden geven de belangen aan van de parameters binnen het lasso regressie model. Wij zullen nu enkel de volgende parameters selecteren voor onze verdere modellen voor het voorspellen van de slaapscore op basis van hun coëfficiënten: de alcohol-inname, de wakkere slaap, de non-REM hartslag, de stressscore, de totale duratie van de slaap en de REM-slaap. Indien er specifiek over een van de nu niet gebruikte parameters toch meer geweten zou willen zijn, kunnen deze toch opgenomen worden of gewisseld met bijvoorbeeld een andere parameter die minder interessant is. Hierbij zullen de modellen dan mogelijk moeilijker te begrijpen worden of een iets slechtere prestatie leveren, maar dit zijn afwegingen die door de situatie en hun impact verantwoord kunnen worden. Deze manier van parameter selectie zou ook gebruikt kunnen worden bij andere keuzes van de afhankelijke parameters zoals bijvoorbeeld de stressscore.

### 3.2.5 Voorspellende modellering voor slaappatronen

Nu de dataset verdeeld is in subsets en we in voorgaande sectie een set van parameters hebben uitgekozen voor onze verdere modellen kunnen we starten met het kiezen en opstellen van modellen voor het voorspellen van onze afhankelijke parameter, namelijk de slaapscore. Binnen deze sectie zullen er drie machine learning algoritmen gebruikt worden op dezelfde data. Op deze manier kan later de performance van de verschillende technieken vergeleken worden met elkaar. De drie verschillende soorten modellen die hier opgesteld worden, zijn een meervoudig lineair regressie model, een random forest regressie model en een model dat extreme gradient boosting regressie zal toepassen.

Het meervoudig lineair regressie model zal gebruikt worden om de meer eenvoudige recht evenredige correlaties tussen de onafhankelijke en de slaapscore te ontdekken. Dit model zou hier toepasselijk kunnen zijn omdat parameters als de totale duratie en de REM-slaap waarschijnlijk wel zulke relatie ten opzichte van de totale score zullen beschrijven. Ook bij sectie 3.2.3 werden al reeds lineaire correlaties opgemerkt bij verschillende van de variabelen die al dan niet een positieve of negatieve verhouding hadden. Verder is dit soort modellen ook makkelijker te interpreteren als het gaat over de individuele bijdrage die elk van de parameters heeft aan de slaapscore. Natuurlijk zorgt de aard van dit model er wel voor dat complexere relaties tussen de slaapscore en de andere parameters niet volledig ontdekt of uitgebuit zullen worden, wat wel als voordeel heeft dat de training en voorspellen van het model zeer snel zijn. Verder levert dit soort modellen normaal een goede prestatie op kleine datasets en is hij minder vatbaar voor overfitting door deze kleinere omvang. Tot slot is er ook wel nog op te merken dat het model wel gevoelig is voor uitschieters. Deze kunnen een sterke invloed hebben op de coëfficiënten van het model en kunnen zo leiden tot minder robuuste voorspellingen.

Om nu een model te hebben dat in plaats van het lineaire model wel overweg zou kunnen met niet lineaire relaties kiezen we voor het random forest algoritme. Dit model maakt zichzelf op door meerdere beslissingsbomen te bouwen en deze hun voorspellingen te combineren voor de finale voorspellingen. Elke boom wordt zo getraind op een willekeurig stuk van de trainingsdata en parameters om zo variatie te voorzien en overfitting te reduceren. In tegenstelling tot het meervoudige lineaire model zijn dit soort modellen ook minder gevoelig voor uitschieters binnen de data. Ze zijn verder ook zeer geschikt voor eventuele grote datasets, maar vooral hun goede prestaties op een grotere set van parameters kan voor onze toepassing wel een voordeel geven. Verder maakt hun complexiteit het wel meer ingewikkeld om de resultaten te begrijpen.

Wel kunnen ze individuele belangen van de parameters goed weergeven, maar het gecombineerde effect hiervan begrijpen is zeer complex. Desondanks door de meerdere beslissingsbomen het risico op overfitting gereduceerd wordt, blijft dit soort algoritmen nog steeds hiervoor vatbaar bij een gebrek aan data of bij zeer diepgaande boomstructuren.

Als laatste model hebben we nog het extreme gradient boosting model of XGBoost. Dit model is een diepgaande ontwikkeling van een gradient boosting algoritme en bouwt meerdere beslissingsbomen in een sequentiële orde. Elk van deze bomen heeft als doel de fouten of afwijkingen van zijn voorgangers te verminderen. Verder maakt het ook gebruik van regularisatie technieken om overfitting tegen te gaan en zo de prestaties op ongeziene data beter te veralgemenen. Dit algoritme maakt verder ook zeer goede voorspellingen en doet het dikwijls beter dan de andere alternatieven [72]. Dit soort modellen zijn net zoals het random forest model ook in staat om niet lineaire complexe relaties en interacties tussen de verschillende onafhankelijke parameters en de slaapscore te ontdekken en op te stellen. Een nadeel van dit soort algoritmen is wel dat dit algoritme een afstelling van zijn hyperparameters vereist voor een optimale prestatie. Wat deze parameters inhouden wordt in volgende sectie 3.2.6 verder toegelicht. Dit afstellen vereist echter wel extra inspanningen om het algoritme zo goed mogelijk te krijgen. In tegenstelling tot bijvoorbeeld het lineaire regressie model, is XGBoost vaak zwaarder om te trainen en te gebruiken. Deze vereiste van extra rekenkracht voor dit algoritme zal voor ons onderzoek echter geen groot probleem vormen doordat de omvang van onze dataset eerder beperkt is voor zulke momenten. Indien dit onderzoek echter op een grotere schaal herhaald zou worden met meer testpersonen en over een langere periode zou hier zeker rekening mee gehouden moeten worden tijdens het opstellen en trainen van dit model.

### 3.2.6 Evaluatie en optimalisatie van modellen

Om de verschillende modellen te vergelijken met elkaar zullen we doen op basis van enkele te berekenen maatstaven. Allereerst zullen we de nauwkeurigheid van de modellen bekijken. Deze waarde wordt normaal eerder gebruikt om de prestaties op te meten als het gaat om de classificatie van data en niet zozeer regressie of het voorspellen van continue waarden. Binnen deze studie zullen wij deze nauwkeurigheid echter licht anders gaan gebruiken. Wij zullen deze eerder zien als een waarde die kan aangeven hoe ver de voorspellingen van de werkelijke waarden afzitten. We verduidelijken dit even kort met een voorbeeld. Stel dat een werkelijke slaapscore 75 bedraagt en ons model een nauwkeurigheid heeft van 95%. Dit model zal dus gemiddeld genomen 5% mis zitten met zijn voorspelling. Dus op de waarde van 75 zou dit dus een range geven van  $(75 - (75 \times 0.05)) = 71.25$  tot en met  $(75 + (75 \times 0.05)) = 78.75$ . Naast deze nauwkeurigheid zullen we ook de R-kwadraat score of determinatie coëfficiënt gebruiken. Deze waarde meet hoe goed het regressie model past. Dit geeft een indicatie van hoe goed de onafhankelijke parameters de variabiliteit in de afhankelijke parameter rond zijn gemiddelde waarde kunnen uitleggen. Deze score meet dus eigenlijk de grootte van de variantie in de afhankelijke parameter dat voorspelbaar is vanuit de onafhankelijke parameters. Deze score bevindt zich meestal ergens tussen 0 en 1 waarbij 0 aangeeft dat het model geen uitleg kan bieden op iets van de variantie in de afhankelijke parameters. Een score van 1 zou dan aangeven dat het model juist perfect alle variantie kan verklaren. Desondanks een zeer hoge score misschien wel vaak betekend dat het model zijn voorspellingen heel goed zijn, is dit niet altijd met zekerheid. Dit zou namelijk ook een indicatie kunnen zijn van overfitting. Als extra op deze twee waarden voor de evaluatie zullen we ook de gemiddelde afwijking of error van de modellen berekenen.

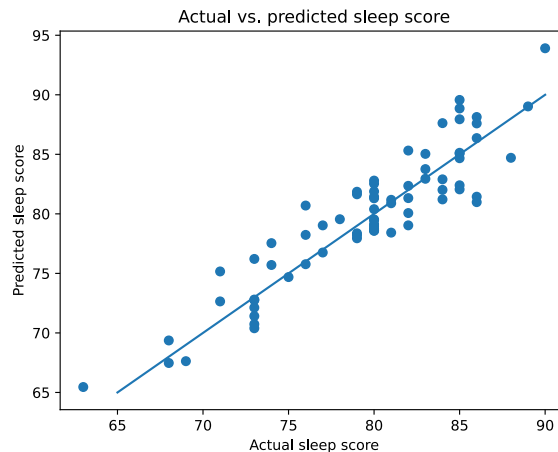


Alvorens we de modellen vergelijken, zullen we ook een basisgrens opstellen die kan aangeven of de modellen zelfs enigszins nut zouden hebben. Deze basisgrens zal altijd de gemiddelde waarde voor de slaapscore voorspellen. Als we dan naar onze modellen kijken, zouden deze altijd het beter moeten doen dan deze basisgrens om enigszins extra waarde te hebben. Binnen tabel 3.2 zijn de waarden van deze basisgrens te zien. Opvallend is de R-kwadraat score die zelfs juist negatief blijkt te zijn. Dit geeft aan dat deze manier van voorspellen zeer slecht is om zelfs maar enige afwijking te voorspellen die afwijkt van het gemiddeld, wat heel logisch is aangezien deze methode juist altijd het gemiddelde zal voorspellen.

	Accuracy	R-squared	Mean Absolute Error	Mean Squared Error
Baseline	94.42%	-0.014	4.3611	29.6389
Multiple Linear Regression	97.76%	0.8316	1.7840	4.9214
Random Forest	97.54%	0.8103	1.9231	5.5458
XGBoost	97.45%	0.7618	2.0109	6.9619

**Tabel 3.2:** De tabel geeft voor de verschillende modellen de metingen weer in verband met hun prestaties. Deze metingen zijn gebeurd op basis van de originele eerste splitsing van de data, zonder kruisvalidatie of afstelling van de hyperparameters.

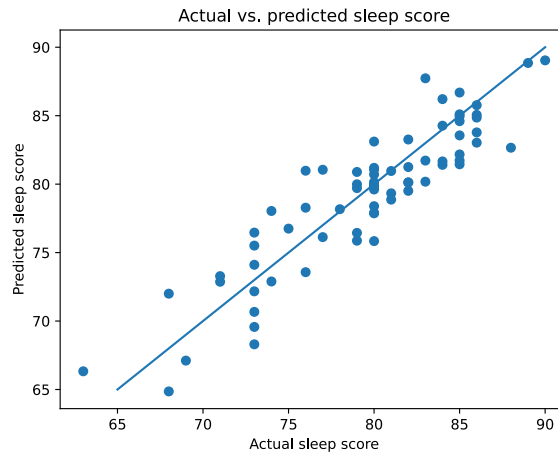
Als we nu dan zouden kijken naar de prestaties van het meervoudige lineaire model zien we dat deze op elk van de evaluatiecriteria uit tabel 3.2 beter scoort dan de basisgrens. Op figuur 3.9 worden de resultaten van de model op de validatieset ook weergegeven. Hier is zichtbaar dat het model redelijk accuraat lijkt te zijn alsook extreme uitschieters lijken hier niet op te treden. Het lineaire model presteert dus zeker wel goed op de data die hij kreeg en kan mogelijk dus zeker wel een meerwaarde leveren bij het voorspellen van de slaapscore. Bij dit model valt echter wel nog op te merken dat hier ook rekening gehouden moet worden met de aanwezigheid van multicollineariteit. Doordat de slaapscore van Fitbit gebaseerd is op bepaalde parameters waaronder bijvoorbeeld de REM-slaap is het niet uit te sluiten dat hun relatie beschreven zou kunnen worden als lineair. Dit beïnvloedt natuurlijk de prestaties van het model op een positieve manier.



**Figuur 3.9:** De figuur geeft de voorspellingen weer van het meervoudig lineaire regressie model in functie van de effectieve waarden van de slaapscore. De waarden die hier getest worden zijn deze uit de validatieset. De blauwe rechte beschrijft de recht evenredige relatie tussen beide assen en kan als indicatie gebruikt worden voor de prestaties van het regressie model.

Wanneer we nu naar het random forest model gaan zullen we naast de evaluatie criteria ook de belangen van de parameters in acht nemen. In tabel 3.2 is zichtbaar dat het model het beter doet dan de basisgrens. Er zit dus zeker wel potentieel in dit model om accurate voorspellingen

te maken voor de slaapscore. Verder kan ook geconstateerd worden dat het model met zeer weinig verschil slechter presteert dan het lineaire regressie model. Gelijkaardig zoals bij dit andere model, tonen we voor het random forest model op figuur 3.10 opnieuw de voorspelde waarden van het model op de validatieset in functie van de effectieve waarden. De voornaamste waarneming die we hier maken is te zien bij de waarden vanaf 76 en lager. In vergelijking met het lineaire model en figuur 3.9 is te zien dat bij het random forest model de waarden verder van de rechte liggen en dus slechtere voorspellingen doen voor deze lagere waarden. Dit zou kunnen komen doordat zeer slechte slaapscores vaker sterker afwijken van de gemiddelde slaapscore in tegenstelling tot zeer goede slaapscores, zoals reeds bemerkt werd in sectie 3.2.2.



**Figuur 3.10:** De figuur geeft de voorspellingen weer van het meervoudig lineaire regressie model in functie van de effectieve waarden van de slaapscore. De waarden die hier getest worden zijn deze uit de validatieset. De blauwe rechte beschrijft de recht evenredige relatie tussen beide assen en kan als indicatie gebruikt worden voor de prestaties van het regressie model.

Als derde model hebben we dan nog het extreme gradient boosting. Wanneer we via tabel 3.2 dan weer de verschillende evaluatiecriteria vergelijken zien we dat dit model het wel beter doet dan de basisgrens maar van de drie echte modellen wel het minst accuraat is in zijn voorspellingen en het minst goed presteert. Omdat zowel dit model als het random forest model belangen of waarden hechten aan hun onafhankelijke parameters, gaan we eens dieper ingaan op de onderlinge verschillend. In tabel 3.3 worden deze belangen weergegeven. De waarden geven een indicatie over hoe sterk een parameter de onzuiverheid van de beslissingsbomen doet afnemen. Zoals eerder besproken binnen 2.2.2 moet elke knoop binnen een boom de data zo splitsen zodanig dat gelijkaardige waarden van de afhankelijke parameters bij elkaar blijven na deze splitsing, en zo de onzuiverheid doen dalen. Dit wordt in de literatuur de Gini-index genoemd voor de classificatie problemen. Wanneer we over regressie problemen spreken wordt deze onzuiverheid de variantie genoemd. Voor deze twee modellen is een parameter met een hogere waarde dus een parameter die zorgt voor een grote daling van de variantie van het model. Voor beide modellen is de duratie van de slaap de parameter met het meeste invloed op de voorspelling van de slaapscore. Maar waar bij het random forest model de REM-slaap de opvolger is, is bij het boosting model de inname van alcohol de tweede meest invloedrijkste parameter. Dit is wel een opmerkelijk verschil omdat bij het random forest model alcohol-inname als minst belangrijk werd aanzien van de zes parameters. Een mogelijke oorzaak hiervan zou de aard van het random forest model kunnen zijn als we spreken over correlaties tussen de parameters. Zoals in figuur 3.7 te zien is heeft de duratie van de slaap een relatief hoge correlatie met meerdere van de andere parameters die gebruikt werden bij dit model. Vanaf dat de parameter van duratie dus als grootste voorspeller wordt gebruikt binnen het model zal het belang van deze andere gecorreleerde parameters automatisch dalen. Dit komt doordat

het merendeel van de onzuiverheid dat deze parameters kunnen verhelpen gemeenschappelijk is en deze onzuiverheid al verwijderd zal zijn door deze eerste parameter. Daarom dat we de belangen die het random forest model geeft aan deze parameters met een kritisch oordeel opvatten.

	Stressscore	Sleep Duration	Wake Sleep	REM Sleep	Non-REM Heart rate	Alcohol intake
Random Forest	0.070	0.528	0.079	0.162	0.093	0.069
XGBoost	0.038	0.268	0.092	0.188	0.187	0.227

**Tabel 3.3:** De tabel geeft voor de belangen van de onafhankelijke parameters voor het random forest en het extreme gradient boosting model.

Door de relatief beperkte omvang van de dataset die wij binnen deze studie gebruiken, kan de splitsing van data in een trainingsset en testset zorgen voor over- of underfitting. Dit zou kunnen geven dat een bepaald model te goed of juist niet goed genoeg zou passen bij de trainingset. Verder testen we bij één enkele splitsing ook niet grondig de generalisatie van een model op nieuwe ongeziene data. Een mogelijke oplossing voor deze problemen is om kruisvalidatie uit te voeren. Kruisvalidatie is een gelijkaardige methode als het splitsen van de data in de verschillende subsets, maar in plaats van eenmalig splitsen doen we nu een herhaalde splitsing met daarop volgend testen op de nauwkeurigheid van het model. Voor elk van de splitsingen zullen we de trainingsset en de testset verschillend nemen. Algemeen kunnen we deze methode zien als het opsplitsen van de data in een aantal subsets, wat we vouwen noemen. Dan zullen we 1 van deze vouwen vasthouden en trainen we het model op al de resterende vouwen. Achteraf testen we het model dan op deze vouw die afgezonderd werd. Deze stap wordt dan herhaald maar telkens waarbij telkens een andere vouw wordt vastgehouden. Het totaal aantal stappen is gelijk zijn aan het aantal vouwen dat origineel gekozen werd bij het opsplitsen. In tabel 3.4 zijn de resultaten te zien op onze modellen wanneer we een onze data door 4 delen. Wanneer we nu deze waarden vergelijken met onze eerdere bevindingen bij slechts één opsplitsing van de data in figuur 3.2, zien we dat de modellen elk iets lager scores na de kruisvalidatie. De oorspronkelijke splitsing kwam dus met andere woorden gunstiger uit voor de prestaties van de modellen en overfitting was dus lichtelijk aanwezig bij de originele modellen. Om hier nog even dieper op te focussen, kan er ook gekeken worden naar tabel 3.5 waarbij de R-kwadraat scores voor elk van de verschillende stappen wordt weergegeven. We kunnen hier bemerken dat er redelijk veel fluctuatie zit tussen de keuze van de subsets. Het toepassen van kruisvalidatie is hier dus zeker wel nuttig om de accuraatheid van de evaluatie zelf na te gaan.

	Accuracy	R-squared	Mean Absolute Error	Mean Squared Error
Multiple Linear Regression	97.45%	0.8290	2.0033	7.4734
Random Forest	97.20%	0.8184	2.1996	8.3721
XGBoost	97.08%	0.7950	2.2987	9.3062

**Tabel 3.4:** De tabel geeft de waarden voor de evaluatie criteria van de verschillende modellen na uitvoering van viervoudige kruisvalidatie op elk van deze modellen.

	1st Fold	2nd Fold	3rd Fold	4th Fold	Average
Multiple Linear Regression	0.6632	0.8636	0.8789	0.9102	0.8290
Random Forest	0.7557	0.8185	0.8368	0.8625	0.8184
Extreme Gradient Boosting	0.6901	0.7919	0.8216	0.8764	0.7950

**Tabel 3.5:** De tabel geeft de waarden voor de R-kwadraat score bij elke stap voor elk van de modellen alsook het uiteindelijke gemiddelde van deze scores.

Bij sommige machine learning modellen kunnen we verder nog hyperparameters aanpassen om hun prestaties te verbeteren. Deze optie is zowel bij ons random forest model als bij het XGboost model mogelijk. Het optimaliseren van deze parameters is echter een zeer moeilijke

operatie. Vaak is het makkelijker en efficiënter om deze door middel van testen te achterhalen. Net zoals de kruisvalidatie is dit aanpassen van de hyperparameters een oplossing om overfitting tegen te gaan. Het aanpassen van deze parameters moet dus ook niet enkel getest worden op eenzelfde splitsing van de data. Binnen dit onderzoek hebben wij zo gebruik gemaakt van 'random search cross validation'. Bij deze manier voor het afstellen van de hyperparameters kiezen we een range van voorgedefinieerde waarden voor de verschillende hyperparameters die we willen optimaliseren. Hierna zal het algoritme dan willekeurig waarden uit al deze ranges kiezen en deze toepassen als hyperparameters op de modellen. Dan wordt hierop kruisvalidatie toegepast en wordt dit model op basis hiervan geëvalueerd. Een extra optie hierbij is dat wij zelf kunnen kiezen hoeveel van deze willekeurige kruisvalidaties we willen doen. Een hoger aantal zal de kans op het vinden van de optimale hyperparameters vergroten maar zal ook meer rekenkracht vergen.

Voor het random forest model hebben wij volgende parameters uitgekozen om te optimaliseren:

**Aantal beslissingsbomen** Deze parameter staat voor het aantal beslissingsbomen dat binnen het random forest model gebruikt zal worden.

**Maximum diepte** Dit geeft aan hoe diep een beslissingsboom zich maximaal mag nesten.

**Splitting criteria** Dit geeft aan op basis van welke waarde het algoritme zijn waarden zal splitsen. Hier gaven we als opties ofwel de absolute error als criterium te gebruiken ofwel de gekwadrateerde error.

**Minimum split aantal** Met deze waarde wordt bedoeld hoeveel datapunten er minimaal nodig zijn om een node nog verder te kunnen splitsen

**Minimale daling onzuiverheid** Deze waarde geeft aan hoe sterk de daling van de onzuiverheid moet zijn alvorens de split gemaakt mag worden.

Bij elk van deze parameters gaven we een range van standaardwaarden mee waaruit de willekeurige zoekfunctie zal mogen kiezen. Na het dan toepassen van de 'random search cross validation' krijgen we als resultaat de exacte beste waarde terug voor de hyperparameters. Wanneer we deze waarden zouden toepassen op ons model, zullen we een meer optimale prestatie van ons model bekomen. Deze hypothese wordt ook bevestigd. Na toepassing van verkregen waarden voor de hyperparameters zien we nu in tabel 3.6 dat de nauwkeurigheid, r-kwadraat score en fouten beter zijn dan in tabel 3.4 waarbij enkel kruisvalidatie werd toegepast op het originele random forest model.

Op een zeer gelijkaardige manier pasten we deze optimalisatie techniek 'random search cross validation' ook toe op ons extreme gradient boosting model. Ook hier werden de hyperparameters voor het aantal beslissingsbomen en maximale diepte gebruikt. Verder waar er hier ook nog enkele andere hyperparameters als de leersnelheid van het algoritme, het constructiealgoritme voor de beslissingsbomen en de minimale verliesreductie nodig voor verdere opdeling, die we hier probeerden te optimaliseren. In tabel 3.6 is ook bij deze resultaten vast te stellen dat ons model verbeterd is ten opzichte van de iteratie na het enkel toepassen van de kruisvalidatie, zoals in tabel 3.4 te zien is.

	Accuracy	R-squared	Mean Absolute Error	Mean Squared Error
Multiple Linear Regression	97.45%	0.8290	2.0033	7.4734
Random Forest	97.51%	0.8325	1.8561	6.3729
XGBoost	97.25%	0.8110	2.0769	7.1940

**Tabel 3.6:** De tabel geeft de finale waarden voor de evaluatie criteria van de verschillende modellen na uitvoering van viervoudige kruisvalidatie op het MLR model en toepassing van 'random search cross validation' op het random forest model en het XGBoost model.

### 3.2.7 Implicaties op slaapkwaliteit

In voorgaande sectie hebben we enkele modellen opgesteld die in staat zijn om de slaapscore van een persoon te voorspellen aan de hand van enkele parameters. Met deze resultaten hebben we niet direct een eenduidig antwoord op hoe we de slaap van een persoon kunnen verbeteren. Echter kunnen deze modellen wel bijstaan in het weergeven van de impact die deze parameters zouden kunnen hebben op de slaapscore. Een direct voorbeeld hiervan kan bijvoorbeeld de impact van alcohol gebruik zijn op de slaap van een persoon. Wanneer we aan een model nu meegeven dat voor een bepaalde avond het alcoholgebruik een bovengemiddelde waarde aanneemt en voor de andere parameters een gemiddelde waarde, kan dit model mij al waarschuwen dat mijn slaapkwaliteit zal lijden hieronder. De voorspellende modellen kunnen zo op een nauwkeurige manier bevonden correlaties onrechtstreeks vertalen naar een persoon. Hetzelfde zou gelden wanneer we op een zeer gelijkaardige manier voor onze modellen de stressscore als afhankelijke parameter zouden kiezen. Een persoon zou dan daadwerkelijk zijn stress al op voorhand kunnen voorspellen door de andere parameters op te meten en de modellen te gebruiken. Men zou dan kunnen werken aan het verminderen van persoonlijke stress door ervoor te zorgen dat de combinatie van de parameters die men wel in controle heeft een betere stressscore voorspellen binnen deze modellen. Het vinden van deze onderliggende verbanden of correlaties tussen de onafhankelijke en afhankelijke parameters is hier van groot belang. Maar we mogen tot slot wel niet vergeten dat deze verbanden of relaties tussen deze parameters hier geen rechtstreekse band definieerde tussen de twee. De verbanden die de opgestelde modellen aangeven vertellen ons enkel dat er een statistisch verband aanwezig is tussen beiden, maar dit duidt niet noodzakelijk op een causaliteit waarbij een effect op een parameter een rechtstreeks gevolg is van de andere. Desondanks kan dit statistische verband wel al bijstaan om in ons voorbeeld de slaapscore beter te begrijpen en er meer controle over te krijgen.

## Hoofdstuk 4

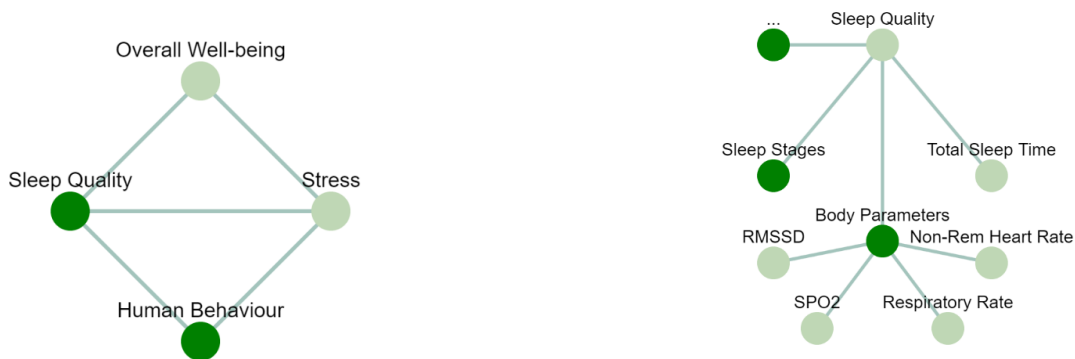
# Verklarende oplossingen voor gegevensanalyse via machine learning voor studenten

Het doel van dit onderzoek is om studenten een oplossing te bieden die hun slaap- en levenskwaliteit positief kan beïnvloeden. Sectie 3.2 beschrijft hoe we parameters uit het leven van deze studenten kunnen gebruiken om correlaties en voorspellingen te maken voor bijvoorbeeld de slaap en stress. Oplossingen worden opgesteld in de vorm van statistische beschrijvingen en machine learning modellen. Deze oplossingen zijn echter niet makkelijk interpreteerbaar voor deze studenten en brengen in deze vorm weinig bij aan dit doelpubliek. Binnen deze sectie zal er een antwoord geboden worden aan deze tekortkomingen. We proberen hier rechtstreekse oplossingen te bieden zodanig dat studenten op een interactieve manier met behulp van hun medische data inzichten kunnen verwerven over de invloeden van hun gedrag op hun lichaam. Hier zullen we proberen om een antwoord te bieden op de vragen van dit onderzoek. Als resultaat van dit onderzoek werd een webapplicatie opgesteld voor studenten. Via deze applicatie kunnen zij op een zelfstandige en interactieve manier kennis opdoen over hun medische data en de verbanden die te ontdekken vallen.

### 4.1 Kennismaking met parameters en eigen gewoonten

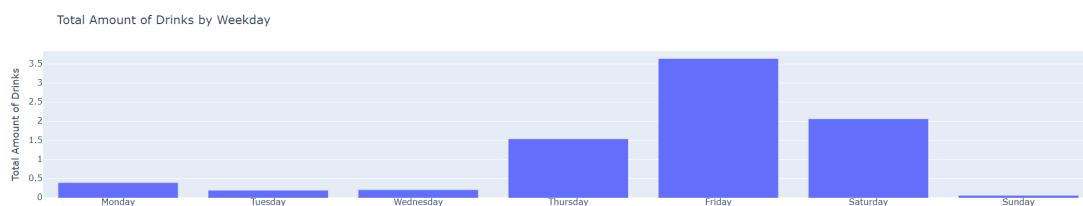
Deze studie maakt gebruik van een brede set aan parameters zoals slaap data, fysieke data maar ook zaken als alcohol-inname. Binnen deze parameters zijn er zeker gekende relaties en hiërarchische onderverdelingen mogelijk om structuur te brengen. Door bestaande literatuur weten we reeds dat tussen parameters zoals slaap en stress een bepaalde relatie bestaat. Om deze structuur en kennis over te brengen naar de gebruikers geeft de applicatie een interactief schema weer. Dit schema licht de verschillende parameters toe zoals bijvoorbeeld hoe onze stressscore werkt of wat zuurstof saturatie betekent. Twee voorstellingen van hoe dit schema zich vertoont zijn te zien op figuur 4.1. De verschillen in kleur van de cirkels geeft hier de gebruiker een indicatie dat deze parameter nog een verdere onderverdeling heeft. De uitleg van elke parameters wordt zichtbaar wanneer de cursor over deze cirkels bewogen wordt aan de zijkant van het scherm. Met deze visualisatie proberen we volgende vraag te beantwoorden: “Wat definieert slaapkwaliteit en stress?”.

Een vraag die een student zich kan stellen bij de definities is waarschijnlijk hoe zijn eigen data eruit ziet. Studenten hebben niet altijd notie over hoeveel slaap zij krijgen op een nacht, hoeveel zij drinken of hoeveel zij nu exact sporten op een dag. Mogelijk vragen zij zich ook af of zij de laatste weken minder actief zijn, of hoe hun slaapcyclus verdeeld is. Om deze vragen



**Figuur 4.1:** De figuren geven weer hoe het schema van de verschillende parameters in de applicatie aanwezig kan zijn in verschillende vormen. De donkere cirkels geven aan dat er op deze parameters geklikt kan worden om een diepere nesting van parameters te krijgen. Zo is de figuur rechts het resultaat van de selectie van 'Sleep Quality' op de linker figuur.

te beantwoorden en de student zijn data zelf te laten ontdekken, biedt onze applicatie enkele algemene grafieken aan. Zo krijgt hij staafdiagrammen om de verdeling van zijn activiteitslevels na te gaan, zijn aantal stappen per dag te tracken of om te kijken hoeveel hij drinkt op welke dagen in de week. Ook vindt hij hier grafieken terug in verband met de verdeling van zijn slaap en het verloop van zijn slaap en stressscores. Een voorbeeld van hoe zo een grafiek eruit kan zien, is te vinden op figuur 4.2. Hier wordt bijvoorbeeld weergegeven met behulp van een staafdiagram hoeveel de student gemiddeld drinkt op welke weekdag over een geselecteerde periode. Zo kan hij zien dat zijn alcoholgebruik hier op vrijdag en zaterdag bijvoorbeeld veel hoger ligt en dat hij op maandag gemiddeld het minste alcohol consumeerde. Andere figuren die zo ook andere data weergeven van de student zijn te vinden in bijlagen 6.1 en 6.2. Omdat studenten een nogal wisselende levensstijl kunnen hebben door bijvoorbeeld examenperiodes of vakanties, wordt hen hier ook de optie geboden om de periode waarvan ze hun data bekijken aan te passen. Zo kunnen ze met één simpele interactie zelf nagaan hoe binnen een bepaalde specifieke periode hun data eruit zag.



**Figuur 4.2:** De figuur geeft een grafiek weer uit de applicatie waar de gebruikers hun alcoholgebruik kunnen nagaan per weekdag voor een geselecteerde periode.

## 4.2 Patroon- en correlatie-analyse tussen de parameters

Hoe kunnen we overbrengen aan studenten wat conceptueel zeer vereenvoudigd de verhouding is tussen slaap, stress, alcohol, cafeïne en bijvoorbeeld een specifieke parameters als REM-slaap om een eerste beeld te scheppen hiervan? Om dit te doen op een interactieve manier zouden we simpelweg een evenredige verhouding kunnen aanschouwen tussen deze parameters voor de eenvoud. Zoals we in de literatuur reeds hebben gezien hebben alcohol, stress en cafeïne een negatieve correlatie ten opzichte van de slaapkwaliteit in tegenstelling tot de hoeveelheid REM-slaap. Op figuur 4.3 is hiervan de uitwerking binnen de applicatie te zien hoe we deze

vraag proberen te beantwoorden. Dit doen we aan de hand van sliders die de student kan bedienen om bijvoorbeeld te gaan van weinig stress naar een sterk stress gevoel. Aan de linkerkant zal dan live de kleur van de cirkel die de slaapkwaliteit voorstelt groener worden indien de sliders naar de betere kant gedragen worden en rood indien deze de slechtere richting opgaan. Veel REM-slaap in combinatie met geen alcohol, stress en cafeïne zal dus zorgen voor een sterk groene kleur terwijl het omgekeerde zal zorgen voor een sterk rode kleur. Met behulp van een disclaimer boven de figuur waarschuwen we de gebruiker ook over de grote veralgemeningen hierbij gemaakt worden en welke conclusies hij hier zeker niet uit mag trekken. Zo een foutieve conclusie waarvoor gewaarschuwd wordt is bijvoorbeeld dat het opheffen van een parameter in de vorm van overcompensatie bij een andere in werkelijkheid niet op deze manier zal werken. Een teveel aan alcohol kan bijvoorbeeld niet opgelost worden door volledig geen stress te ervaren.

### Conceptual Sleep Quality Model

Since Sleep Quality is not a directly measurable factor, we try to give it a value based on the factors we actually can measure or could possibly provide ourselves to get an estimation of our sleep quality.

Take the colored circle below as an indication of your sleep quality. We qualify the color green as good sleep quality, and the color red as bad sleep. Now while controlling the sliders on the right you can get an idea on how our sleep quality is influenced by these factors.

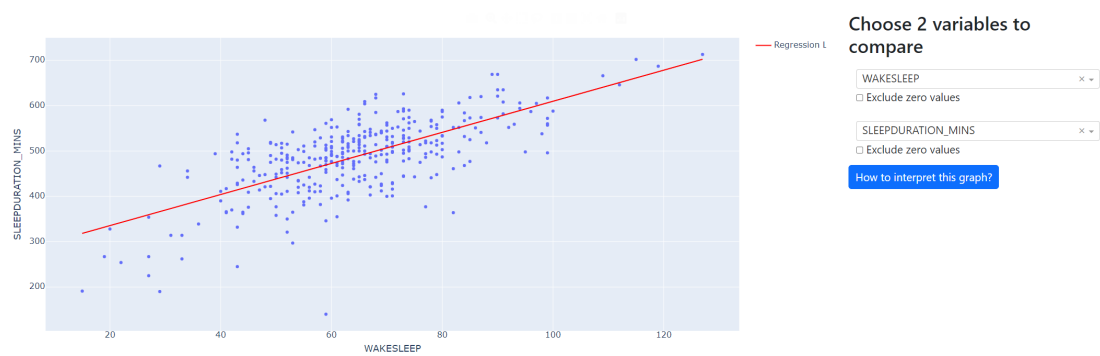
Caution has to be taken with this model! Since it's a conceptual model the weights of the different parameters is likely not accurate when it comes to practice. In this example every parameter has the same impact. Also in real life, compensating hard for one bad parameter with another one doesn't fix the initial problem. Please don't rely on the exact coloring and values shown here!



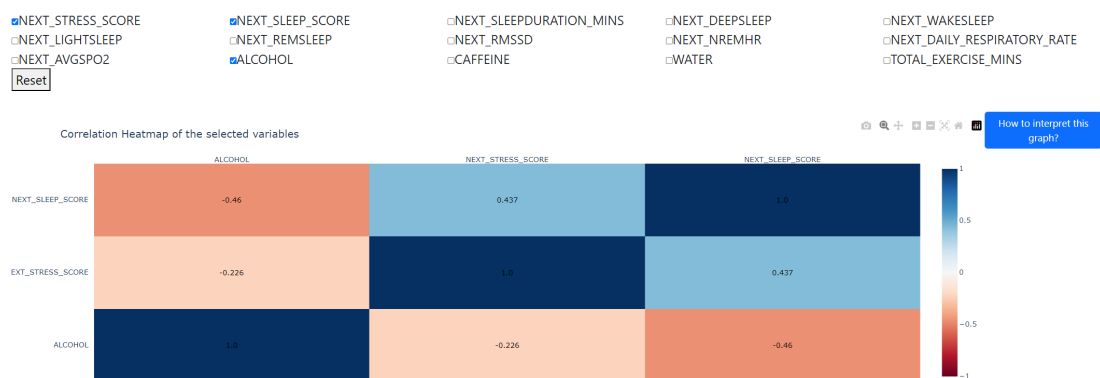
**Figuur 4.3:** De figuur geeft de visualisatie weer binnen de applicatie waarin we zeer veralgemeend een interactief conceptueel model aanbieden om de verhouding tussen bepaalde parameters te ontdekken.

Een groot deel van dit werk ontleedt zich over het nagaan van verbanden tussen de verschillende parameters die aan bod komen. Zoals sectie 3.2.3 reeds heeft aangetoond, kunnen dit soort verbanden soms verduidelijkingen brengen of inzichten geven in hoe twee parameters voorkomen en omgaan met elkaar. Binnen de inleiding in sectie 1 stelden we ons als derde vraag: Wat zijn de correlaties tussen de verschillende parameters, zoals slaapkwaliteit, alcohol-inname, cafeïne-inname, stress, fysieke activiteit en andere slaapgegevens? Om hierop nu een antwoord te bieden, voorzien we een generieke oplossing binnen onze applicatie. Via een gelijkaardige benadering als in onze analyse willen we zowel visueel de gebruiker zelf correlaties laten ontdekken alsook numerieke exacte waarden tonen over hoe zijn gegevens zich verhouden ten opzichte van elkaar. Door de omvang van het aantal verschillende parameters dat we beschikbaar hebben, moet er wel opgelet worden dat gebruikers niet overspoeld worden met grafieken en diagrammen. Daarom bieden we de optie om met behulp van tabs ofwel een scatter plot weer te geven, ofwel een heatmap die verschillende parameters als inhoud kan nemen. Bij beiden deze figuren kan de gebruiker zelf kiezen welke parameters hij wilt vergelijken. Voor het scatter plot kan hij zo twee parameters selecteren waarna dan het plot zich aanpast. Om de interpretatie van dit plot duidelijk te maken geven we ook een regressielijn weer, indien van toepassing voor beiden parameters, om het makkelijker te maken om een mogelijk aanwezige correlatie te ontdekken zoals te zien is op figuur 4.4. Ook kan de gebruiker extra uitleg krijgen via een dialoog dat hij kan aanklikken. Hierin wordt op een passieve en eenvoudige manier toegelicht hoe hij dat plot en de regressielijn moet interpreteren. Een voorbeeld van deze heatmap is terug te vinden op figuur 4.5, waar bovenaan de gebruiker de parameters kan selecteren, de parameters kan resetten en voor extra uitleg kan vragen voor deze visualisatie. Hiermee kan de gebruiker de exacte correlatiewaarden tussen de verschillende parameters nagaan en interpreteren.





**Figuur 4.4:** De figuur geeft de grafiek weer uit de applicatie waar de gebruikers zelf correlaties kunnen nagaan tussen twee parameters met behulp van een scatter plot met een regressielijn.

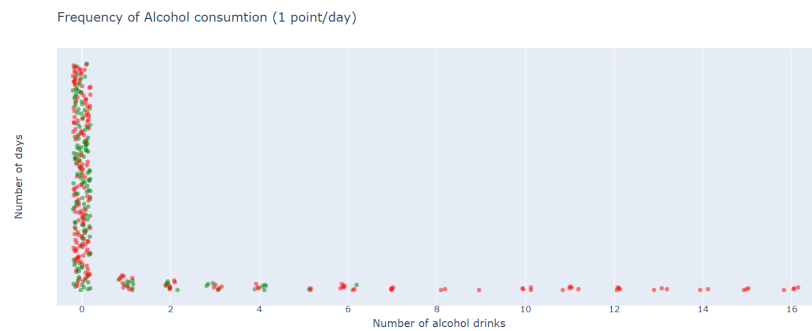
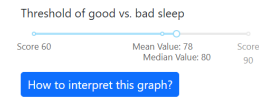


**Figuur 4.5:** De figuur geeft de heatmap weer zoals deze aanwezig is binnen de applicatie. Bovenaan zijn de parameters te vinden alsook de knop voor het informeren naar de werking van deze figuur.

### 4.3 Verhouding tussen alcoholgebruik en slaapkwaliteit

Studenten zijn een doelgroep waar reeds bekend is dat zij vaak alcohol consumeren en waarbij alcoholgebruik negatieve gevolgen kan hebben voor hun gezondheid [44]. In onze inleiding in sectie 1 werd reeds de vraag gesteld of alcoholgebruik de slaapkwaliteit zou beïnvloeden. Daar deze studie geen medisch antwoord kan formuleren op deze stelling, werd er wel geprobeerd om de impact van alcohol duidelijk te maken door het verband tussen de opvolgende slaapscore weer te geven. Op figuur 4.6 wordt getoond hoe dit resultaat eruit ziet. Er wordt door middel van een scatter plot aangegeven hoeveel dagen de student hoeveel alcoholische consumpties drinkt. Verder door middel van de kleur van de punten kan hij dan de relatie met de opvolgende slaapscores nagaan. De kleur geeft aan of hij die dag goed of slecht heeft geslapen op basis van een schaal die hij zelf kan kiezen. Een rode kleur geeft aan dat zijn slaapscore onder deze schaal ligt, en dus als slecht aanzien wordt, terwijl een groene kleur een goede slaap aangeeft die zich boven deze schaal bevindt. Door de slider aan te passen, kunnen studenten de verdeling van de kleuren observeren. Sterker rode stippen aan de rechterkant geven aan dat alcohol de slaapscore negatief kan beïnvloeden, terwijl groenere stippen een positieve correlatie zouden suggereren. Deze visualisatie helpt studenten bijstaan de relatie tussen alcohol-consumptie en hun slaapkwaliteit te begrijpen, waardoor ze weloverwogen beslissingen kunnen nemen en conclusies kunnen trekken over de impact van alcohol op hun slaapkwaliteit. Op te merken is dat deze waarnemingen kritisch gemaakt moeten worden doordat een causaal verband tussen beiden hier niet genomen kan worden, maar enkel de frequentie of het statistisch voorkomen weergegeven wordt.

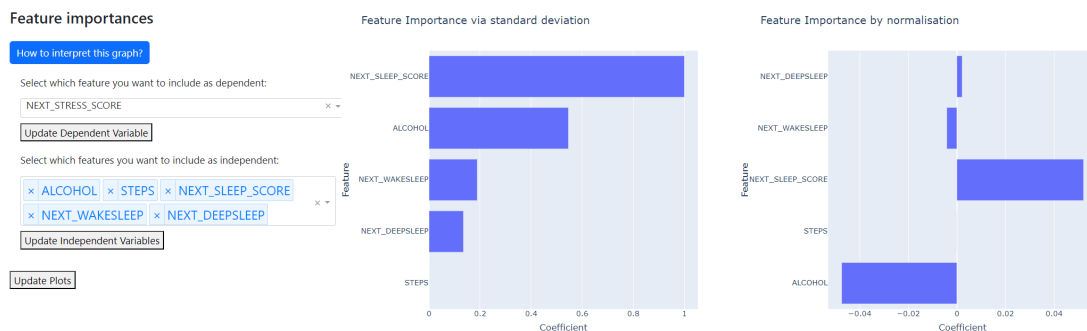
### Change the sleep quality threshold



**Figuur 4.6:** De figuur geeft een scatter plot weer van het aantal dagen dat een bepaalde hoeveelheid alcohol geconsumeerd werd. De kleuren binnen dit plot geven aan of de slaapscore die hierop volgde ofwel goed was in het groen, of dat deze slecht was in het rood. De maatstaf voor wat een goede of slechte score is, kan de gebruiker zelf bepalen door middel van de slider aan de linkerkant van de visualisatie.

## 4.4 Belangen van parameters onderzoeken

Binnen dit onderzoek hebben we al enkele keren de correlatie tussen slaap en stress besproken en vergeleken. Verder is ook geweten dat Fitbit zijn slaap gegevens ook gebruikt als hulpmiddel bij het definiëren van een stressscore. Maar welke parameters nu het meeste belang hebben bij deze scores is iets wat we ook reeds ontdekten binnen de analyse van deze studie. Nu is het de bedoeling dat we deze kennis overbrengen naar onze doelgroep van studenten. Hiervoor kiezen wij ervoor om hen zelf de mogelijkheid te geven om een selectie van bepaalde parameters te kiezen, deze door een lasso regressie model te laten gaan en de resultaten hiervan terug weer te geven aan hen. Het resultaat van deze oplossing is zichtbaar op figuur 4.7. Aan de linkerkant zien we dat de gebruiker via dropdown menu's de afhankelijke en onafhankelijke parameters kan selecteren. In de twee diagrammen die hiernaast getoond worden zal de student dan de belangen van deze parameters terugvinden wanneer het gaat over het voorspellen van de afhankelijke parameter. Enerzijds met behulp van de standaardafwijkingen van de parameters en anderzijds na toepassing van normalisatie op deze coëfficiënten. Opnieuw wordt hier geprobeerd om de student veel vrijheid te geven om de parameters, verbanden en belangen zelfstandig te ontdekken. Bij deze visualisatie kan hij indien nodig ook extra uitleg krijgen om de waarden en achterliggende berekeningen beter te begrijpen.



**Figuur 4.7:** De figuur geeft een visualisatie weer van de applicatie waarde belangen van de gekozen parameters weergeven die door een lasso regressie model berekend worden. De linker grafiek geeft deze weer via hun standaardafwijking terwijl de rechter grafiek deze weergeeft na een toepassing van normalisatie.

## 4.5 Voorspellingen uitvoeren en nagaan

Doordat er binnen sectie 3.2.5 enkele machine learning modellen ontworpen werden, zijn we in staat om op basis van een aantal parameters, een andere afhankelijke parameters te voorspellen. Op deze manier kan er inzicht verworven worden naar welke invloeden bepaalde acties hebben waarbij een parameters verhoogd of verlaagd zou worden. Wanneer we dit terugkoppelen naar hoe een student dit zou kunnen waarnemen, zou hij zich bijvoorbeeld kunnen afvragen niet alleen wat bijvoorbeeld zijn REM-slaap effectief betekend maar ook hoe sterk het zijn slaapscore zou kunnen beïnvloeden.

Binnen de inleiding haalden we in vragen drie en vier ook aan dat we naast enkel de slaap gegevens ook de verbanden van parameters als alcohol maar ook cafeïne of water willen nagaan. Wat zijn de correlaties tussen de verschillende parameters, zoals slaapkwaliteit, alcohol-inname, cafeïne-inname, stress, fysieke activiteit en andere slaap gegevens? En hoe beïnvloedt alcohol-gebruik de slaapkwaliteit? Een goede manier om op deze vragen een antwoord te bieden, is om de mogelijkheid zelf te geven aan de studenten om zulke parameters af te stemmen en de resultaten van de modellen waar te nemen. Binnen onze applicatie wordt hier een oplossing geboden in de vorm van acht histogrammen. Op figuur 4.8 kunnen we zien dat het bovenste en ook grootste histogram de verdeling van de student zijn slaapscore weergeeft. Al de andere histogrammen geven de verdeling van een parameter mee die de gebruiker zal kunnen aanpassen om een voorspelling te maken. Het aanpassen van deze parameters zal hij kunnen doen door gebruik te maken van de sliders die zich telkens onder de bijhorende parameter bevinden. Hier wordt ook weergegeven van de exacte waarden zijn zodat hij nauwkeurig zijn voorspellingen zal kunnen doen. De manier hoe de student zijn voorspellingen om de slaapscore hier zal kunnen opvatten, is door middel van 3 verticaal gekleurde lijnen op het histogram bovenaan. Voor elke kleur kan hij ook nogmaals extra uitleg krijgen van wat deze betekend. Elk van deze lijnen geeft een voorspelling weer van de waarden van de de student deze koos op een verschillend machine learning model. Zo vertegenwoordigt elke kleur een ander model, namelijk blauw staat voor een meervoudig lineair model, geel voor een lasso regressie model en zwart voor een random forest model.

Zoals we reeds in sectie 3.2.5 hebben ondervonden hebben zijn onze modellen niet altijd 100% nauwkeurig. Doordat we hier nu drie verschillende modellen gebruiken, zal de student zien dat deze niet altijd eenzelfde voorspelling maken. Binnen de extra uitleg zal hij ook kunnen lezen dat modellen inderdaad een afwijking kunnen hebben. Desondanks de onenigheid tussen deze modellen kan dit er wel door zorgen dat het mentale beeld dat de student heeft van de modellen niet blindelings gevolgd moet worden. Wanneer de modellen zeer ver uit elkaar zullen gaan en dus bepaalde fouten binnen enkele modellen groter zullen worden, zal de student dit ook kunnen gebruiken om te begrijpen dat zulke voorspelling dan met minder zekerheid gemaakt kan worden. Anderzijds wanneer de drie modellen exact eenzelfde voorspelling maken, vergroot dit de kans dat de voorspelling wel accuraat zal zijn. Binnen deze visualisatie proberen we de gebruiker genoeg vrijheid te bieden. Echter behouden we de balans tussen de interactiviteit van de applicatie en de rekentijd die vereist is om deze visualisatie te genereren. Indien de rekentijd verwaarloosd kan worden zonder de bruikbaarheid van de applicatie te verminderen, zouden de studenten de optie geboden kunnen worden om zelf ook andere variabelen zoals die van hun fysieke activiteit hierbinnen te betrekken.

Normale studenten hebben waarschijnlijk weinig notie bij hoe machine learning modellen werken. Echter zouden we binnen dit werk studenten toch oppervlakkig willen aanbrenge hoe de complexe algoritmen achterliggend werken. Wanneer zij namelijk in staat zijn om duidelijk te begrijpen hoe deze werken, zullen zij kritischer en beter onze visualisaties en voorspellingen gaan begrijpen. Dit laatste kan dan hopelijk bijdragen aan het nut waarrond deze studie draait, namelijk het proberen verbeteren van de slaapkwaliteit maar zelfs het algemene welzijn van deze doelgroep. Om studenten een denkbeeld aan te brengen voor machine learning modellen, kunnen we gebruik maken van SHAP-waarden. Binnen de literatuur in sectie 2.2.3 werden deze soort waarden reeds toegelicht. Wanneer we een zeer complex soort model zouden nemen

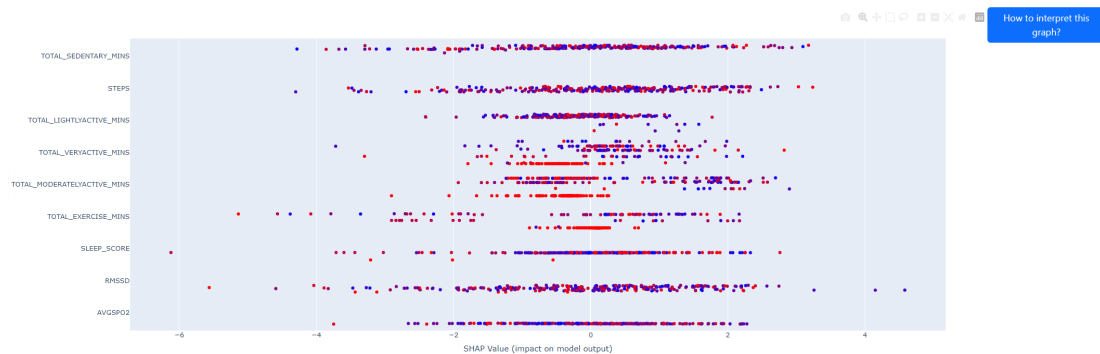


**Figuur 4.8:** De figuur geeft een visualisatie weer van de applicatie waar voorspellingen rond de slaapscore in de vorm van histogrammen geïnterpreteerd kunnen worden. Het bovenste histogram heeft zo de verdeling van de slaapscore weer samen met drie verticale lijnen die voor drie verschillende modellen een voorspelling van deze score geven. De onderstaande histogrammen samen met hun sliders kunnen dan gebruikt worden om deze voorspellingen aan te passen door andere inouts aan de sliders te geven.

als ons XGB model binnen sectie 3.2.5 willen we graag weten hoe dit model tot zijn voorspellingen komt. De SHAP-waarden kunnen hier dan laten zien welke eigenschappen bijdragen aan de afwijking tussen de gemiddelde waarde van de te voorspellen parameter en de effectieve voorspelde waarde van het model. We zouden dus kunnen waarnemen welke waarden binnen onze trainset van data ons model bevorderen, en welke de voorspelling slechter maken. Zoals binnen de literatuur wordt aangehaald binnen sectie 2.2.3, zit het verschil hem zeer sterk in het positief of negatief zijn van de SHAP-waarde. Wanneer we een resultaat van een model willen verklaren zou men kunnen kijken naar deze waarde met de grootste impact. Zo kunnen we dan beter gaan begrijpen op basis van welke parameters het model deze voorspelling maakte. In onze applicatie zullen wij zo aan de studenten toelichten hoe hun fysieke activiteit een impact heeft op hun stress. Dit zullen we doen door middel van de resultaten van een XGB model dat achterliggend berekend wordt. Op figuur 4.9 is te zien hoe deze uitwerking eruit ziet. Voor elke parameter kan de student op een bepaalde verticale hoogte de SHAP-waarden voor deze parameters waarnemen. Verder wordt aan de verschillende datapunten ook een kleur gegeven. Deze toont hoe hoog of hoe laag die bepaalde waarde is in vergelijking met de gemiddelde waarde voor die parameter. Dit zou eventueel ook kunnen helpen om te bepalen of groepen met ongeveer gelijke waarden gezamenlijk een grote of kleine impact zouden hebben. Voor deze visualisatie wordt ook een uitgebreide uitleg beschikbaar gesteld voor de studenten om hem beter te begrijpen.

Hoe kunnen we een student onmiddellijk de invloed van bijvoorbeeld zijn slaap en fysieke activiteit op zijn stress laten zien voor een range van verschillende waarden? Om dit te doen voor één combinatie van twee parameters is natuurlijk eenvoudig, hiervoor laten we een machine learning model gewoonweg een voorspelling maken. Maar er is net gewild dat we dit voor een range van waarden onmiddellijk kunnen krijgen zonder talloze individuele voorspellingen te moeten maken. Een mogelijke oplossing op deze vraag is door gebruik te maken van een partial

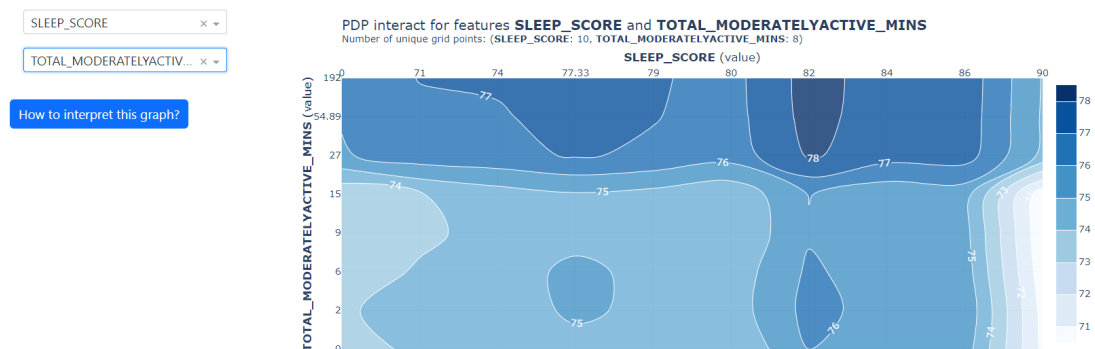
SHAP and variable value



**Figuur 4.9:** De figuur geeft een visualisatie weer van de SHAP-waarden van een XGB model voor het voorspellen van de stressscore op basis van de fysieke activiteit van een persoon.

dependency plot. Dit soort plots toont dan de afhankelijkheid tussen de twee onafhankelijke parameters samen met de afhankelijke parameter. Omdat hier een restrictie ligt op het aantal onafhankelijke parameters marginaliseren we al de andere parameters van het gebruikte model. Bij de visualisatie op figuur 4.10 geven we studenten binnen onze applicatie de mogelijkheid om zelf te selecteren welke parameters ze willen gebruiken bij hun vergelijking. In één beeld kan dan waargenomen worden hoe het model reageert op verschillende waarden van de parameters en waar de 'beste' voorspellingen zouden liggen. Zo kan een student dan nagaan voor welke hoeveelheden van fysieke activiteit en welke scores van slaap hij het minste stress zou ervaren of dus een beste score zou hebben. Of anderzijds zou hij kunnen kijken hoe hij meer stress zou kunnen vermijden door bijvoorbeeld zijn trainingsschema aan te passen naargelang zijn bevindingen hier. Opnieuw krijgt de student hier de mogelijkheid om meer informatie op te vragen voor deze visualisatie zodat hij zeker is dat zijn eigen bevindingen correct geïnterpreteerd zouden zijn.

Changing 2 variables, what is their impact on my stress?



**Figuur 4.10:** De figuur geeft een visualisatie weer van de SHAP-waarden van een XGB model voor het voorspellen van de stressscore op basis van de fysieke activiteit van een persoon.

## 4.6 Interactieve webapplicatie voor studenten

Binnen dit hoofdstuk werden verschillende visualisaties aangeboden om de vragen van deze studie te beantwoorden. Door een bundeling van deze visualisaties binnen één webapplicatie worden deze resultaten beschikbaar gemaakt voor de studenten. Zelfstandig kunnen zij dit platform gebruiken en raadplegen volgens hun eigen noden. De webapplicatie is opgebouwd

met notie voor het brengen van een compleet, toelichtend en vloeiend te interpreteren oplossing voor het doelpubliek van studenten dat aangenomen wordt wel overweg te kunnen met eenvoudige technologie. Deze webapplicatie kan gezien worden als een uitbreiding en verbetering van platformen zoals Fitbit deze aanbiedt en de gebreken die hiermee gemoeid zijn zoals aangehaald door Zilu Liang en Bernd Ploderer [31]. Zo probeert deze oplossing gebruikers meer te informeren naar mogelijke verbanden, kan deze wel gepersonaliseerde oplossingen aanbieden en kunnen acties voor het verbeteren van de slaapkwaliteit en aspecten als stress wel overgebracht worden.

# Hoofdstuk 5

## Conclusies

Deze masterthesis is vertrokken vanuit de gebreken die aanwezig zijn bij bestaande gezondheidsapplicaties, en de vraag hoe studenten hun slaapkwaliteit zouden kunnen verbeteren. Er werd hier voornamelijk gefocust om met behulp van technieken zoals machine learning, resultaten te bekomen die studenten zouden kunnen helpen bij het begrijpen van complexere relaties of verbanden die zich voordoen op hun data. Het maken van een webapplicatie probeert hier deze brug te maken tussen het diepgaand analyseren van de data van studenten en het laten begrijpen van deze diepgaande analyse. Naar mijn persoonlijke mening denk ik dat dit bekomen resultaat zeker als een verrijking aanzien kan worden op de oplossingen die bijvoorbeeld door Fitbit of andere bedrijven vandaag de dag commercieel worden aangeboden. Er zit veel potentieel in het gebruik van machine learning modellen dat studenten of andere mensen meer inzicht kan geven in welke acties, welke reacties zullen hebben als het gaat over alcoholgebruik, fysieke activiteit en slaapkwaliteit. Ook stress management is iets wat binnen dit onderzoek werd aangehaald. Oplossingen hier zouden ook enorm kunnen bijdragen zodanig dat studenten hun lichaam beter leren gebruiken en hoe ze een betere levensstijl kunnen aannemen. Desondanks deze studie als een voorbeeld gezien kan worden hoe mogelijk studenten al deze info bijgebracht kan worden, heeft deze studie ook een aantal limitaties en gebreken. Deze zullen verder binnen dit hoofdstuk aangehaald en besproken worden. Om hierop een antwoord te vinden, zullen er extra inspanningen en verbeterde studies moeten plaatsvinden. Tot slot zal ook besproken worden wat mijn eigen bevindingen zijn over de bekomen resultaten, waar ikzelf verbeteringen kon aanbrengen en wat het mij heeft bijgebracht.

### 5.1 Limitaties

Voor het onderzoek dat deze masterthesis beschrijft zijn er wel een aantal limitaties die dit onderzoek heeft. Wanneer we kijken naar het verzamelen van de data, werd dit enerzijds automatisch gedaan door middel van Fitbit horloges. Doordat de software van deze horloges kan veranderen over de tijd en zij ook mogelijks minder accuraat zijn dan sommigen andere methoden zoals PSG-onderzoeken, is het zo dat deze data niet altijd uniform en volledig accuraat zal geweest zijn. Deze inconsistenties hebben natuurlijk een al dan niet kleinschalige invloed op de analyse en machine learning modellen gebruikt binnen de diepgaande analyse. Anderzijds geeft het gebruik van logboeken die de testpersonen manueel moesten invullen ook als bijgevolg dat hier fouten bij gemaakt werden. Naast deze fouten is er voor de parameters als alcohol een sterke vereenvoudiging gebruikt om de totale hoeveelheid van alcohol-consumptie van een dag samen te vatten. Door het simpelweg optellen van de hoeveelheid alcoholische dranken zal ook dit resultaat niet volledig consistent geweest zijn. Ook doordat we met vijf verschillende testpersonen werken voor het verzamelen van de data zullen resultaten niet altijd uniform zijn. Elke persoon heeft een andere weerstand tegen alcohol of cafeïne. Ook slaapt iedere persoon op een andere manier of met andere hoeveelheden. Door een beperking te zetten op de leef-

tijdsrange van deze personen en andere aspecten zoals dat zij nog studeren proberen we meer gelijkaardige profielen te gebruiken, maar alsnog zijn hier ongetwijfeld verschillen tussen. In het algemeen kan er dus gesteld worden dat de data niet ideaal is om samen voor machine learning modellen gebruikt te worden. Ook door de nogal korte periode van 80 dagen, de beperkte set van testpersonen en het onbruikbaar worden van sommige data door foutieve metingen, slechte waarden, etc. werd er eigenlijk te weinig data verzameld om zeer goede, sterke en robuuste data analyses uit te voeren. Ideaal genomen zouden we data van één persoon gebruikt moeten hebben over een zeer lange periode van bijvoorbeeld enkele jaren. Op deze manier wordt het makkelijker om mogelijke relaties en verbanden duidelijker en met meer zekerheid te ontdekken. Dit zorgt dat voor betere resultaten op het vlak van de analyses en zo kunnen dan duidelijker deze correcte resultaten overgebracht worden via de opgestelde webapplicatie. Verder binnen de data waren parameters als cafeïne ook vrij waardeloos. Door het gebrek aan cafeïne houdende dranken dat de testpersonen consumeerden werd het vrijwel onmogelijk om resultaten te vinden met deze data.

Een andere sterke assumptie die binnen deze studie ook aangenomen werd, was de manier hoe Fitbit zijn stress en slaapscores berekent. Hier zijn wij altijd uitgegaan dat Fitbit dit op een zeer correcte manier doet. Het bedrijf geeft echter niet zijn exacte berekeningen prijs, maar geeft enkel een indicatie waarop zij deze scores min of meer baseren. Maar zonder exacte bronnen of medische verantwoording is het natuurlijk onmogelijk om te besluiten dat dit black box model dat zijn toepassen achterliggend daadwerkelijk correct is. Om deze limitatie aan te pakken, zou er zelf met een berekening gekomen moeten worden om deze scores te berekenen. Dit vereist echter een medisch verantwoorde studie om hier correcte keuzes binnen te maken die een werkelijke representatie zijn van de slaapkwaliteit of de stress waarde.

Tot slot kon er binnen deze studie ook een terugkoppeling gemaakt worden van de ontworpen applicatie naar de testpersonen. Een gemiste kans binnen dit onderzoek is het laten testen en gebruiken van de applicatie door de studenten zelf. Hier kon dan feedback verzameld mee worden die mogelijke verbeteringen kon suggereren.

## 5.2 Toekomstig werk

Deze studie biedt door zijn tekortkomingen en aard een grote opportuniteit voor toekomstige onderzoeken. Het verloop van deze studie kan voor volgende onderzoeken gebruikt worden om een gelijkaardige aanpak te hanteren, maar dan op een grotere schaal. Door dit werk op schaal te herhalen kan dit kansen creëren om daadwerkelijk het effect ervan te testen in realiteit. Dit kan dan bijvoorbeeld gedaan worden door de testpersonen de applicatie mee te laten gebruiken terwijl hun gegevens verzameld worden. Er kan dan onderzocht worden of een interactieve oplossing zoals de webapplicatie binnen dit onderzoek ook werkelijk de slaapkwaliteit van mensen zal verbeteren of dat dit geen invloed zou hebben hier op. Ook zouden verdere stappen kunnen zijn om het doelpubliek van studenten uit te breiden naar bijvoorbeeld volwassenen zoals de groep van arbeiders, die in hun werksituatie waarschijnlijk meer met stresserende situaties te maken zullen hebben. Of voor atleten zouden gelijkaardige onderzoeken als deze kunnen aantonen hoe de atleet optimaal zou moeten trainen om op tijd te herstellen of om zijn slaapkwaliteit niet te doen lijden onder zijn trainingen. Verder kunnen ook andere aanpakken zoals het analyseren van het daadwerkelijke specifieke verloop van de nachtrust geanalyseerd kunnen worden, zouden de reeds gebruikte modellen nog verder geoptimaliseerd kunnen worden of zouden alternatieve en mogelijk betere algoritmen getest kunnen worden. En tot slot zou een meer nauwkeurige dataverzameling er ook voor kunnen zorgen om correctere resultaten te bekomen. Deze masterthesis kan dus zeker gebruikt worden als een raamwerk of een fundering om verdere onderzoeken binnen dit domein te ondersteunen of hun nut te verklaren.



## 5.3 Zelfreflectie

Wanneer ik zelf naar de bekomen resultaten kijk, ben ik ervan overtuigd dat de bekomen visualisaties binnen de webapplicatie een zeer interessant en verduidelijkend beeld kunnen geven voor bepaalde fenomenen. Mogelijks komt het ook door mijn interesse, maar wanneer ik de applicatie gebruik, ben ik zelf heel nieuwsgierig naar het aanpassen van de grafieken of interactieve tools om nieuwe ontdekkingen te doen op bepaalde invoeren. Het geeft mij persoonlijk echt het gevoel een soort 'playground' te gebruiken om te leren hoe mijn lichaam werkt en wat onderliggende relaties zouden kunnen zijn. Een nadeel hiervan is misschien wel dat wanneer ik hiernaar kijk, ik wel ook na ga hoe ik bijvoorbeeld mijn slaap kan optimaliseren maar ook hoe ik het andere uiterste zou kunnen opzoeken. Soms probeer ik eerder de limitaties van de visualisaties op te zoeken in plaats van wat nu eerder nuttig zou zijn om te weten.

Naast het eindresultaat had ik ook op andere gebieden van dit onderzoek wel beter kunnen doen. Zo heb ik bijvoorbeeld naar mijn mening weinig gedaan met de informatie rond de gevoelens van de testpersonen die ik heb opgevraagd via de logboeken alsook met de stress die Fitbit voor mij verzamelde. Naast het voorspellen en vergelijken van de stressscore met andere waarden komt deze hier niet veel meer aan bod. Het was beter geweest moest ik ook hier mee gedaan hebben binnen de analyse alsook het eindresultaat. Voor de analyse heb ik verder ook verschillende voorspellende machine learning algoritmen gebruikt. Een andere toepassing hierop was geweest dat ik meer aan patroon herkenning gedaan zou hebben. Ik had namelijk ook data beschikbaar die het volledige verloop van de slaap van een persoon bijhield gedurende elke nacht. Met deze data kon ik ook analyses gedaan hebben naar patronen zoals het verloop van de verschillende slaapfasen of de relatie tussen verschillende slaapcycli waarbij het aantal blokken diepe slaap ten opzichte van de wakkere slaap wordt afgemeten. Het volledig gebruiken en benutten van de slaap data is dus zeker wel een aspect dat naar mijn mening beter kon en wat ik graag ook verder had benut.

Wanneer ik mijn bevindingen van deze studie vergelijk met wat er in bestaande studies reeds bevonden werd, zag ik hier wel veel overeenkomsten. Wanneer het gaat over de verbanden tussen de verschillende parameters is zeer duidelijk zichtbaar dat deze binnen de literatuur al sterk gekend zijn, maar dit werd met dit werk dan nogmaals bevestigd. Binnen de literatuur werd door enkele bronnen aangehaald dat Fitbit horloges soms slechte metingen zouden doen of dat ze niet betrouwbaar zouden zijn. Dit aantal bronnen wat deze stelling maakte was echter beperkt maar naar mijn eigen bevindingen denk ik hier eigenlijk helemaal anders over. Ik denk dat de sensoren al zeer goed zijn in het nagaan van correcte gegevens. Toen ik in de literatuur over Fitbit las en ik vond dat ik de data van Fitbit makkelijk kon opvragen om zelf mee aan de slag te gaan, had ik wel meer de verwachting dat de data beter gestructureerd zou zijn. Terug kijkend naar hoe Fitbit op een niet uniforme en zeer complexe manier zijn data opslaat en splitst, vraag ik me wel sterk af of dit echt hun bedoeling is. De inconsistentie en in mijn ogen slechte bestandsstructuur is niet iets wat ik van een bedrijf van zulke omvang zou verwachten.

Als ik denk over wat ik bij het uitvoeren van deze masterthesis heb geleerd, zou ik dit op meerdere gebieden kunnen beschrijven. Ik heb geleerd om kritisch met bestaande sensoren, technologieën en software om te gaan. Op het vlak van hoe een groter onderzoek wordt uitgevoerd, heeft mijn promotor en begeleider me veel bijgeleerd over hoe verschillende inzichten en aanpakken een mogelijkheid geweest waren voor dit onderzoek. Hierdoor ging ik sommige problemen anders aanpakken vanuit het grotere beeld dat deze studie gaf in plaats van te sterk op kleine geforceerde oplossingen te zoeken. Ook als het gaat over de effectieve technische kennis die ik heb opgedaan, ben ik wel trots op mezelf dat ik zonder enige echte voorkennis van machine learning algoritmen deze technologie heb kunnen bestuderen om hem te gebruiken voor een werkelijk onderzoek. Ook werd doorheen mijn studie wel vaker verwezen naar het uitvoeren van zelfstandige data analyses of hoe het verzamelen van data moest gebeuren, maar om dit zelf daadwerkelijk uit te voeren heeft mij zeker een ervaring bijgeleerd over het exacte proces, welke problemen er allemaal kunnen optreden alsook met wat er allemaal rekening gehouden

moet worden.

Tot slot wil ik graag hier nog aanhalen dat ikzelf me ook zeer bewust ben geworden van de werkpunten die ik op persoonlijk vlak heb wanneer het gaat over het uitvoeren van onderzoeken op zulke grote schaal. Doorheen de proef zijn er vaak veranderingen opgetreden in wat er net binnen deze studie gedaan zou worden en bij het iteratieve proces werd ik me wel vaker bewust dat ik sommige opportuniteiten liet gaan door persoonlijk voor andere oplossingen te kiezen. Moest ik opnieuw kunnen starten aan dit onderzoek zou ik toch graag beter nadenken over deze gemiste opportuniteiten en andere benaderingen toch niet voor betere meer interessante resultaten gezorgd zouden hebben. Desondanks ben ik wel zeer tevreden met de keuzes die ik uiteindelijk gemaakt heb en ben ik enthousiast over de bekomen resultaten van deze thesis.

# Bibliografie

- [1] Ahmad Taher Azar and Shereen M El-Metwally. Decision tree classifiers for automated medical diagnosis. *Neural Computing and Applications*, 23:2387–2403, 2013.
- [2] David K Blough, Carolyn W Madden, and Mark C Hornbrook. Modeling risk using generalized linear models. *Journal of health economics*, 18(2):153–171, 1999.
- [3] Lisa A Cadmus-Bertram, Bess H Marcus, Ruth E Patterson, Barbara A Parker, and Brittany L Morey. Randomized trial of a fitbit-based physical activity intervention for women. *American journal of preventive medicine*, 49(3):414–418, 2015.
- [4] Narciso Cerpa, Matthew Bardeen, Barbara Kitchenham, and June Verner. Evaluating logistic regression models to estimate software project outcomes. *Information and Software Technology*, 52(9):934–944, 2010.
- [5] Taryn Chalmers, Blake Anthony Hickey, Phillip Newton, Chin-Teng Lin, David Sibbritt, Craig S McLachlan, Roderick Clifton-Bligh, John Morley, and Sara Lal. Stress watch: The use of heart rate and heart rate variability to detect stress: A pilot study using smart watch wearables. *Sensors*, 22(1):151, 2021.
- [6] Jean-Philippe Chaput, Caroline Dutil, and Hugues Sampasa-Kanyinga. Sleeping hours: what is the ideal number and how does age impact this? *Nature and science of sleep*, pages 421–430, 2018.
- [7] Luenda E Charles, James E Slaven, Anna Mnatsakanova, Claudia Ma, John M Violanti, Desta Fekedulegn, Michael E Andrew, Bryan J Vila, and Cecil M Burchfiel. Association of perceived stress with sleep duration and sleep quality in police officers. *International journal of emergency mental health*, 13(4):229, 2011.
- [8] Ho-Sung Cho, Young-Wook Kim, Hyoung-Wook Park, Kang-Ho Lee, Baek-Geun Jeong, Yune-Sik Kang, and Ki-Soo Park. The relationship between depressive symptoms among female workers and job stress and sleep quality. *Annals of occupational and environmental medicine*, 25(1):1–9, 2013.
- [9] Chin Moi Chow. Sleep and wellbeing, now and in the future, 2020.
- [10] Ian Clark and Hans Peter Landolt. Coffee, caffeine, and sleep: A systematic review of epidemiological studies and randomized controlled trials. *Sleep medicine reviews*, 31:70–78, 2017.
- [11] BG Cooper, D Veale, CJ Griffiths, and GJ Gibson. Value of nocturnal oxygen saturation as a screening test for sleep apnoea. *Thorax*, 46(8):586–588, 1991.
- [12] Gerald Dubowitz. Effect of temazepam on oxygen saturation and sleep quality at high altitude: randomised placebo controlled crossover trial. *Bmj*, 316(7131):587–589, 1998.
- [13] Irshaad O Ebrahim, Colin M Shapiro, Adrian J Williams, and Peter B Fenwick. Alcohol and sleep i: effects on normal sleep. *Alcoholism: Clinical and Experimental Research*, 37(4):539–549, 2013.

- [14] Frank Emmert-Streib and Matthias Dehmer. Evaluation of regression models: Model assessment, model selection and generalization error. *Machine learning and knowledge extraction*, 1(1):521–551, 2019.
- [15] T Enzel, JW Kantelhardt, CC Lo, RP Bartsch, PC Ivanov, T Penzel, and K Voigt. Dynamics of heart rate and sleep stages in normals and patients with sleep apnea. *Neuropsychopharmacology*, 28(Suppl 1):S48–S53, 2003.
- [16] Marco Fabbri, Alessia Beracci, Monica Martoni, Debora Meneo, Lorenzo Tonetti, and Vincenzo Natale. Measuring subjective sleep quality: a review. *International journal of environmental research and public health*, 18(3):1082, 2021.
- [17] Fitbit, Inc. Fitbit. <https://www.fitbit.com/>, accessed 2023.
- [18] Vincent J Fortunato and John Harsh. Stress and sleep quality: The moderating role of negative affectivity. *Personality and Individual Differences*, 41(5):825–836, 2006.
- [19] Miriam Rodéhn Fox. The importance of sleep. *Nursing Standard (through 2013)*, 13(24):44, 1999.
- [20] Kamalesh K Gulia and Velayudhan Mohan Kumar. Importance of sleep for health and wellbeing amidst covid-19 pandemic. *Sleep and vigilance*, 4(1):49–50, 2020.
- [21] Martica H Hall, Melynda D Casement, Wendy M Troxel, Karen A Matthews, Joyce T Bromberger, Howard M Kravitz, Robert T Krafty, and Daniel J Buysse. Chronic stress is prospectively associated with sleep in midlife women: the swan sleep study. *Sleep*, 38(10):1645–1654, 2015.
- [22] John T Hancock and Taghi M Khoshgoftaar. Catboost for big data: an interdisciplinary review. *Journal of big data*, 7(1):1–45, 2020.
- [23] Rave Harpaz, Krystl Haerian, Herbert S. Chase, and Carol Friedman. Mining electronic health records for adverse drug effects using regression based methods. In *Proceedings of the 1st ACM International Health Informatics Symposium, IHI '10*, page 100–107, New York, NY, USA, 2010. Association for Computing Machinery.
- [24] Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.
- [25] Sean He, Brant P Hasler, and Subhajit Chakravorty. Alcohol and sleep-related problems. *Current opinion in psychology*, 30:117–122, 2019.
- [26] Victor M Karpyak, Magdalena Romanowicz, John E Schmidt, Kriste A Lewis, and John M Bostwick. Characteristics of heart rate variability in alcohol-dependent subjects and nondependent chronic alcohol users. *Alcoholism: Clinical and Experimental Research*, 38(1):9–26, 2014.
- [27] Carl Kingsford and Steven L Salzberg. What are decision trees? *Nature biotechnology*, 26(9):1011–1013, 2008.
- [28] Christopher E Kline. The bidirectional relationship between exercise and sleep: Implications for exercise adherence and sleep improvement. *American Journal of Lifestyle Medicine*, 8(6):375–379, 2014.
- [29] Andrew D Krystal and Jack D Edinger. Measuring sleep quality. *Sleep medicine*, 9:S10–S17, 2008.
- [30] Cheng Li. A gentle introduction to gradient boosting. URL: [http://www.ccs.neu.edu/home/vip/teach/MLcourse/4\\_boosting/slides/gradient\\_boosting.pdf](http://www.ccs.neu.edu/home/vip/teach/MLcourse/4_boosting/slides/gradient_boosting.pdf), page 30, 2016.

- [31] Zilu Liang and Bernd Ploderer. Sleep tracking in the real world: a qualitative study into barriers for improving sleep. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction*, pages 537–541, 2016.
- [32] Tjen-Sien Lim, Wei-Yin Loh, and Yu-Shan Shih. A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine learning*, 40:203–228, 2000.
- [33] Fabio V Lima, Vishnu Kadiyala, Alice Huang, Kartik Agusala, David Cho, Andrew M Freeman, and Regina Druz. At the crossroads! time to start taking smartwatches seriously. *The American Journal of Cardiology*, 179:96–101, 2022.
- [34] Xiaoli Liu, Satu Tamminen, Topi Korhonen, Juha Rönning, and Jukka Riekk. Prediction of sleep efficiency from big physical exercise data. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '19 Adjunct, page 1186–1189, New York, NY, USA, 2019. Association for Computing Machinery.
- [35] Baoshan Ma, Fanyu Meng, Ge Yan, Haowen Yan, Bingjie Chai, and Fengju Song. Diagnostic classification of cancers using extreme gradient boosting algorithm and multi-omics data. *Computers in biology and medicine*, 121:103761, 2020.
- [36] Gunasekaran Manogaran and Daphne Lopez. Health data analytics using scalable logistic regression with stochastic gradient descent. *International Journal of Advanced Intelligence Paradigms*, 10(1-2):118–132, 2018.
- [37] Michael P Mead, Kelly Baron, Morgan Sorby, and Leah A Irish. Daily associations between sleep and physical activity. *International journal of behavioral medicine*, 26:562–568, 2019.
- [38] Fábio Mendonça, Sheikh Shanawaz Mostafa, Fernando Morgado-Dias, Antonio G Ravelo-Garcia, and Thomas Penzel. A review of approaches for sleep quality analysis. *Ieee Access*, 7:24527–24546, 2019.
- [39] Karim El Mokhtari, Ben Peachey Higdon, and Ayşe Başar. Interpreting financial time series with shap values. In *Proceedings of the 29th annual international conference on computer science and software engineering*, pages 166–172, 2019.
- [40] Alexey Natekin and Alois Knoll. Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7:21, 2013.
- [41] Eduardo Nunez, Ewout W Steyerberg, and Julio Nunez. Regression modeling strategies. *Revista Española de Cardiología (English Edition)*, 64(6):501–507, 2011.
- [42] Frances O’Callaghan, Olav Muurlink, and Natasha Reid. Effects of caffeine on sleep quality and daytime functioning. *Risk management and healthcare policy*, pages 263–271, 2018.
- [43] Poonam Pandey and Radhika Prabhakar. An analysis of machine learning techniques (j48 & adaboost)-for classification. In *2016 1st India International Conference on Information Processing (IICIP)*, pages 1–6. IEEE, 2016.
- [44] Crystal L Park. Positive and negative consequences of alcohol consumption in college students. *Addictive behaviors*, 29(2):311–321, 2004.
- [45] Aakash K Patel, Vamsi Reddy, and John F Araujo. Physiology, sleep stages. In *StatPearls [Internet]*. StatPearls Publishing, 2022.
- [46] Dinh-Van Phan, Chien-Lung Chan, and Duc-Khanh Nguyen. Applying deep learning for prediction sleep quality from wearable data. In *Proceedings of the 4th International Conference on Medical and Health Informatics, ICMHI '20*, page 51–55, New York, NY, USA, 2020. Association for Computing Machinery.

- [47] Vili Podgorelec, Peter Kokol, Bruno Stiglic, and Ivan Rozman. Decision trees: an overview and their use in medicine. *Journal of medical systems*, 26:445–463, 2002.
- [48] Barry M Popkin, Kristen E D’Anci, and Irwin H Rosenberg. Water, hydration, and health. *Nutrition reviews*, 68(8):439–458, 2010.
- [49] Sai Priya, AM Mahalakshmi, Sunanda Tuladhar, Bipul Ray, BS Sushmitha, S Shivashree, B Saravanan, Muhammed Bishir, Abid Bhat, and Saravana Babu. Sleep and body fluids. *International Journal of Nutrition, Pharmacology, Neurological Diseases*, 10(2):65–68, 2020.
- [50] Javad Razjouyan, Hyoki Lee, Sairam Parthasarathy, Jane Mohler, Amir Sharafkhaneh, and Bijan Najafi. Improving sleep quality assessment using wearable sensors by including information from postural/sleep position changes and body acceleration: a comparison of chest-worn sensors, wrist actigraphy, and polysomnography. *Journal of Clinical Sleep Medicine*, 13(11):1301–1310, 2017.
- [51] Zuzana Reitermanova et al. Data splitting. In *WDS*, volume 10, pages 31–36. Matfyzpress Prague, 2010.
- [52] Preethi R Sama, Zubin J Eapen, Kevin P Weinfurt, Bimal R Shah, and Kevin A Schulman. An evaluation of mobile health application tools. *JMIR mHealth and uHealth*, 2(2):e3088, 2014.
- [53] Wilhelm Daniel Scherz, Daniel Fritz, Oana Ramona Velicu, Ralf Seepold, and Natividad Martínez Madrid. Heart rate spectrum analysis for sleep quality detection. *EURASIP Journal on Embedded Systems*, 2017:1–5, 2017.
- [54] Yewen Shi, Lina Ma, Xi Chen, Wenle Li, Yani Feng, Yitong Zhang, Zine Cao, Yuqi Yuan, Yushan Xie, Haiqin Liu, et al. Prediction model of obstructive sleep apnea-related hypertension: Machine learning-based development and interpretation study. *Frontiers in Cardiovascular Medicine*, 9:1042996, 2022.
- [55] Jan Snel and Monicque M Lorist. Effects of caffeine on sleep and cognition. *Progress in brain research*, 190:105–117, 2011.
- [56] May St-Onge, Derk-Jan Dijk, and Michael A. Grandner. Heart rate variability as an objective tool to assess sleep quality in older adults: A systematic review. *Sleep Medicine Reviews*, 24:86–97, 2015.
- [57] Emilio Vanoli, Philip B. Adamson, Edward Ba-Lin, Gian Domenico Pinna, Ralph Lazzara, and William C. Orr. Heart rate variability during specific sleep stages: A comparison of healthy subjects with patients after myocardial infarction. *Circulation*, 92(1):143–146, 1995.
- [58] Feifei Wang and Éva Bíró. Determinants of sleep quality in college students: A literature review. *Explore*, 17(2):170–177, 2021.
- [59] Klaas R Westerterp. Assessment of physical activity: a critical appraisal. *European journal of applied physiology*, 105:823–828, 2009.
- [60] Wikipedia contributors. Polynomial regression — Wikipedia, the free encyclopedia, 2022. [Online; accessed 15-May-2023].
- [61] Wikipedia contributors. Ridge regression — Wikipedia, the free encyclopedia, 2022. [Online; accessed 15-May-2023].
- [62] Wikipedia contributors. Bias–variance tradeoff — Wikipedia, the free encyclopedia, 2023. [Online; accessed 15-May-2023].
- [63] Wikipedia contributors. Decision tree — Wikipedia, the free encyclopedia, 2023. [Online; accessed 15-May-2023].

- [64] Wikipedia contributors. Decision tree — Wikipedia, the free encyclopedia, 2023. [Online; accessed 16-May-2023].
- [65] Wikipedia contributors. Elastic net regularization — Wikipedia, the free encyclopedia, 2023. [Online; accessed 15-May-2023].
- [66] Wikipedia contributors. Lasso (statistics) — Wikipedia, the free encyclopedia, 2023. [Online; accessed 15-May-2023].
- [67] Wikipedia contributors. Linear regression — Wikipedia, the free encyclopedia, 2023. [Online; accessed 15-May-2023].
- [68] Wikipedia contributors. Logistic regression — Wikipedia, the free encyclopedia, 2023. [Online; accessed 15-May-2023].
- [69] Matthew Wilkinson. *Influence of Hydration on Sleep, Nocturnal Heart Rate Variability, and Cognition*. PhD thesis, Ohio University, 2017.
- [70] Hyeryeon Yi, Kyungrim Shin, and Chol Shin. Development of the sleep quality scale. *Journal of sleep research*, 15(3):309–316, 2006.
- [71] Aleš Zamuda, Christine Zarges, Gregor Stiglic, and Goran Hrovat. Stability selection using a genetic algorithm and logistic linear regression on healthcare records. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, GECCO '17, page 143–144, New York, NY, USA, 2017. Association for Computing Machinery.
- [72] Rongtao Zhang, Binbin Li, and Bin Jiao. Application of xgboost algorithm in bearing fault diagnosis. In *IOP Conference Series: Materials Science and Engineering*, volume 490, page 072062. IOP Publishing, 2019.
- [73] Zheng Zhang, Zhihui Lai, Yong Xu, Ling Shao, Jian Wu, and Guo-Sen Xie. Discriminative elastic-net regularized linear regression. *IEEE Transactions on Image Processing*, 26(3):1466–1481, 2017.
- [74] Lili Zhu, Pai Chet Ng, Yuanhao Yu, Yang Wang, Petros Spachos, Dimitrios Hatzinakos, and Konstantinos N Plataniotis. Feasibility study of stress detection with machine learning through eda from wearable devices. In *ICC 2022-IEEE International Conference on Communications*, pages 4800–4805. IEEE, 2022.

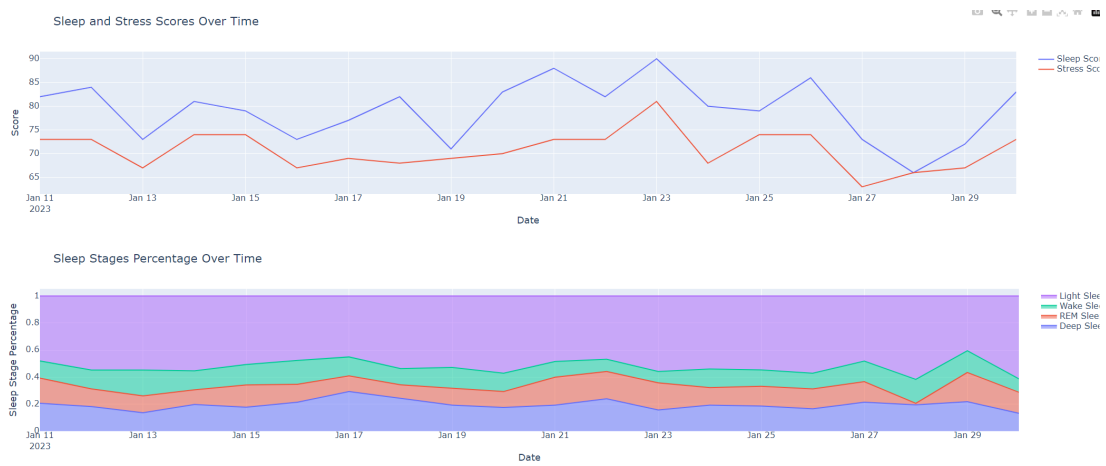
# Hoofdstuk 6

## Bijlagen



**Figuur 6.1:** De figuur geeft een gedeelte van de applicatie weer waarin de gebruikers hun activiteitsniveaus kunnen nagaan alsook hun aantal stappen met behulp van verschillende grafieken.





**Figuur 6.2:** De figuur toont twee grafieken uit de applicatie waarin de gebruikers informatie krijgen over hun slaapscore ten opzichte van hun stressscore alsook de verdeling van hun slaap in verschillende slaapfasen kunnen opvolgen.

Persoonlijk Logboek Alcohol en Cafeïne inname

Naam : \_\_\_\_\_

Date: _____	Voormiddag: 6u - 12u	Namiddag: 12u - 17u	Avond: 17u - 23u	Nacht: 23u - 6u
<i>Pint Bier (250ml)</i>				
<i>Glas wijn (100ml)</i>				
<i>Sterke drank (&gt; 15%, 30ml)</i>				
<i>Glas Water</i>				
<i>Tas koffie</i>				
<i>Glas Cola</i>				
<b>Algemene gemoedstoestand</b>	<input type="checkbox"/> Opgewonden   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Tevreden   <input type="checkbox"/> Kalm   <input type="checkbox"/> Droevig   <input type="checkbox"/> Bezorgd   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Gefrustreerd   <input type="checkbox"/> Gestrest			

Date: _____	Voormiddag: 6u - 12u	Namiddag: 12u - 17u	Avond: 17u - 23u	Nacht: 23u - 6u
<i>Pint Bier (250ml)</i>				
<i>Glas wijn (100ml)</i>				
<i>Sterke drank (&gt; 15%, 30ml)</i>				
<i>Glas Water</i>				
<i>Tas koffie</i>				
<i>Glas Cola</i>				
<b>Algemene gemoedstoestand</b>	<input type="checkbox"/> Opgewonden   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Tevreden   <input type="checkbox"/> Kalm   <input type="checkbox"/> Droevig   <input type="checkbox"/> Bezorgd   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Gefrustreerd   <input type="checkbox"/> Gestrest			

Date: _____	Voormiddag: 6u - 12u	Namiddag: 12u - 17u	Avond: 17u - 23u	Nacht: 23u - 6u
<i>Pint Bier (250ml)</i>				
<i>Glas wijn (100ml)</i>				
<i>Sterke drank (&gt; 15%, 30ml)</i>				
<i>Glas Water</i>				
<i>Tas koffie</i>				
<i>Glas Cola</i>				
<b>Algemene gemoedstoestand</b>	<input type="checkbox"/> Opgewonden   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Tevreden   <input type="checkbox"/> Kalm   <input type="checkbox"/> Droevig   <input type="checkbox"/> Bezorgd   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Gefrustreerd   <input type="checkbox"/> Gestrest			

Date: _____	Voormiddag: 6u - 12u	Namiddag: 12u - 17u	Avond: 17u - 23u	Nacht: 23u - 6u
<i>Pint Bier (250ml)</i>				
<i>Glas wijn (100ml)</i>				
<i>Sterke drank (&gt; 15%, 30ml)</i>				
<i>Glas Water</i>				
<i>Tas koffie</i>				
<i>Glas Cola</i>				
<b>Algemene gemoedstoestand</b>	<input type="checkbox"/> Opgewonden   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Tevreden   <input type="checkbox"/> Kalm   <input type="checkbox"/> Droevig   <input type="checkbox"/> Bezorgd   <input type="checkbox"/> Vrolijk   <input type="checkbox"/> Gefrustreerd   <input type="checkbox"/> Gestrest			

**Figuur 6.3:** De figuur toont het template van het logboek voor het manueel opmeten van parameters voor de testpersonen.

### **Informed consent**

You are asked to voluntarily participate in a research study as part of a Master's Thesis. This study is organized by the Expertise Centre for Digital Media (EDM), the ICT research institute of Hasselt University. By participating in this study, you are helping us gather data that will help conduct this thesis. The main goal of this thesis is to improve sleep quality and well-being based on the insights that we will learn with this gathered data.

In this study, you will be provided with a Fitbit Sense smartwatch and a paper log book. The instructions and expectations of this research study can be found in the separate "Instructions and Expectations" file.

In all, the time required to participate in this experiment is four weeks. During this time, small daily efforts will need to be performed, which will only take approximately 5 minutes.

Participation in this experiment is entirely voluntary. Refusal to participate will involve no penalty, and you may discontinue participation at any time. As such, you are free to stop the study at any point and to decline to answer any of the questions without further explanation. We only ask you to let us know as quickly as possible via email at [sean.snellinx@student.uhasselt.be](mailto:sean.snellinx@student.uhasselt.be). All data gathered from you will also be removed and not used in this research.

To address privacy concerns, your personal information collected during this study will never be disclosed. Your contributions will remain pseudonymous and will be treated confidentially. Because of the applied pseudonymization, your right to request access to your personal data will be technically possible only right after you have finished your participation. Identical timing constraints apply to your right to alter your personal data.

We will collect the following personal information as part of the study:

- Your demographics: name, gender, age, and signature

In addition, we will collect the following non-personal information:

- Your sleeping information: start/end times, duration and change of different sleep stages, wrist temperature changes, times awoken
- Your heart pulsation and oxygen saturation (continuous every few seconds)
- Your physical activity: step count, light/moderate/active minutes
- Your stress score (a score out of 100 calculated by Fitbit based on sleep, physical activity, mood logs and heart rate)
- Your fluid intake: caffeine and alcohol intake.

Your data will be stored in a private cloud-based repository as provided by the infrastructure of Hasselt University until five years after the publication of the results of our study. If you don't want your data to be used in potential follow-up studies, please explicitly indicate so by ticking the checkbox below.

Please sign this form to indicate that (a) you have read the above information and consent to participate in this study based on the stated terms and (b) you agree with the specified data collection for this study.

I do not want my data to be used in potential follow-up studies

Name: \_\_\_\_\_ Signature: \_\_\_\_\_

Date and place: \_\_\_\_\_

**The Hasselt university research team thanks you in advance for your assistance in this research project!**

Principal investigators: Sean Snellinx ([sean.snellinx@student.uhasselt.be](mailto:sean.snellinx@student.uhasselt.be)), Gustavo Rovelo ([gustavo.roveloruiz@uhasselt.be](mailto:gustavo.roveloruiz@uhasselt.be)), Eva Geurts ([eva.geurts@uhasselt.be](mailto:eva.geurts@uhasselt.be))

**Figuur 6.4:** De figuur toont het Informed Consent formulier dat de testpersonen ondertekenden alvorens hun deelname aan de studie.