Optimood

Transforming Health Data into Practical Advice and Insights

Ivar van Wooning

January 2018
Under the supervision of dr. Mannes Poel
Critical observer: Randy Klaassen
University of Twente

Abstract

In light of the significantly increasing rates of depression, an application was designed that automatically derives mood insights from personal health data. Patterns were found by analysing correlation coefficients and extracting feature importances from a random forest regressor. The effectiveness of the application was examined in a small-scale usability test, where participants were asked to register numerous mood-related variables. In general, participants were able to effortlessly interpret and identify themselves with the generated insights. The extent to which the insights delivered any new or valuable information was less than expected, however. Future research should examine how more advanced and deeper patterns can be discovered and visualised accordingly.

Table of Contents

1 Introduction

Rates of depression have been increasing over the last decades and are currently reaching epidemic proportions (Mojtabai, Ramin, Olfson, & Han, 2016). The World Health Organisation estimates that depression will be the second leading cause of world disability by 2020 (World Health Organization, 2001). The objective of this study is to design an application that can derive mood insights from health data. If mood can be properly registered and analysed, a more personal and adaptable eHealth intervention can be established.

The overarching goal of this study is to determine to what extent health data can be used to provide users with practical advice and mood insights (see definition 1). In order to accurately determine the effectiveness, an eHealth application will be designed that registers important factors that affect mood. Additionally, methods of analysis will be examined that can be used to analyse the gathered data. Finally, the effectiveness of the application will be tested in an experiment.

Ultimately, the application should positively influence the lives of its users by finding out when they are feeling at their best. Correctly interpreting the limited amount of mood data and drawing correct and interesting conclusions is challenging, however. In the context of this research, only the possibilities of automated mood analysis and visualisation will be explored.

“Any underlying patterns and interactions in the health data of users that they were previously unaware of. When users are informed about these patterns, they can adjust their behaviour for the benefit of a happier and more fulfilling life.”

Definition 1: defining the term ‘mood insight’ in the context of a mood analysis application

2 Literature review

A state-of-the-art literature review will be conducted to identify several essential components of an application that can derive mood insights from health data. When these components have been determined, the actual application can be built. In order to find the aforementioned components, research should first point out what factors can positively or negatively affect a person's mood. By gathering a large amount of data with regard to the mood of a user, data analysis can be performed. This collection of data should lead to valuable insights, including the possible causes of mood disturbances for specific users. To generate these mood insights, methods in the field of machine learning and data science will be explored that could be effective for this purpose.

This literature review will initially focus on the construction of a set of variables that can be used for mood analysis. Subsequently, light will be shed on the advances in the fields of machine learning and data science and, more importantly, how these should be incorporated in the application. At last, a conclusion will follow that describes the final application with the combination of the different components.

2.1 Factors that affect mood

The acquisition of health data is an essential component of the proposed application. The question is, however, what health data should be collected specifically. Previous research has already shown that certain lifestyle factors substantially affect mood (Rohrer, Pierce, & Blackburn, 2005; Walsh, 2011). In this review, three broad categories emerged after analysing different lifestyle factors: activity, nutrition and technology. Certainly, there are many more categories of factors that may also have an effect on mood, but in this research only a number of generally well-known and common factors will be pinpointed.

2.1.1 Activity

There are numerous activities that may either improve or depress mood. One of the most profound correlations has been found in social support, or lack thereof, and depression (Wang, Xingmin et al., 2014). In fact, social interaction with peers has even been found to positively affect telomere length and reduce stress levels (Liu, Jia Jia et al., 2017). A distinction can be made between social and emotional loneliness, as some people may feel lonely in spite of of regular social interaction (and vice versa). For this reason, users should not only be required to register the amount of social contact they had on a particular day, but also how lonely they were feeling subjectively.

Similarly, physical exercise is a notable factor that improves mood as well. A study by Blumenthal et al. (2017) showed that 30 minutes of physical exercise for three times a week had the same effect on depressed patients as the antidepressant medication sertraline (Blumenthal et al., 2007). De Moor et al. (2006) also found that regular exercisers were less depressed, anxious and neurotic. A simple polar question can determine whether users had any exercise on a day-to-day basis. In addition, an objective way to measure how much exercise users have had, is via a pedometer attached to a Fitbit Tracker or Apple's HealthKit.

During the night, the human body rests, processes emotions and memories, and attempts to restore any damage (Greer, 2004). If this process is somehow obstructed in the form of sleep deprivation, emotional processing may be altered (Simon et al., 2015). Both the amount of sleep and sleep quality may therefore be interesting variables to include in the analysis.

2.1.2 Nutrition and drugs

Our bodies are to a large extent affected by the substances we do or do not consume. Especially water and sugar intake are factors that also have an effect on the way we feel. Knüppel, Shipley, Llewellyn, and Brunner (2017) point out that the "intake of sweet food, beverages and added sugars has been linked with depressive symptoms in several populations”. In contrast, when people drink more water, it seems to positively affect energy levels and mental clarity (Pross et al., 2014). The application should therefore register the amount of water a user has drunk at the end of the day, as well as the amount of sugar that was ingested.

Drugs, including alcohol, also appear to significantly alter mood. Freed (1978) suggested that “alcoholics experience increasing dysphoria as a consequence of alcohol consumption, while nonalcoholics anticipate — and generally attain — elevated moods as a result of drinking”. Everyday, the application should register whether a user drank any alcohol. Naturally, the amount of alcohol is equally important and should be taken into account as well.

2.1.3 Technology

Technology has an increasingly important role in our society. At the same time, many people have grown attached to their computers and social media feeds. Research by Pantic (2014) and Kross et al. (2013) shows that there is a “significant positive correlation between depressive symptoms and time spent on social networking sites”. The use of Facebook, Instagram or Twitter may lead to a depressed mood and loneliness because of an “altered (and often wrong) impression of the physical and personality traits of other users”. Another study by Thomée, Härenstam and Hagberg (2012) shows that computer use has similar depressive effects on users. Furthermore, computer use was also correlated with an increase in perceived stress and sleep disturbances. Consequently, the factors for the technology category should consist of the amount of hours spent on a computer and on social media.

In the preceding paragraphs, several important factors were identified that may interact positively or negatively with the mood of users. These factors, divided into three main categories, should be filled in on a daily basis and together comprise the first component of the application. This first component provides data that will be utilised in the next component, where data will be analysed to produce beneficial insights.

2.2 Data driven approach to mood analysis

Analysing personal data and providing related health insights has been studied in the recent past. Spanakis, Weiss, Boh, Lemmens, and Roefs (2017) showed that data analysis was highly effective in the domain of problematic eating behaviour. Individual states (mood, location, activity, etc.) were mapped and used to assess their corresponding impact on unhealthy eating patterns. Similarly, successful data analysis has been achieved for health records of elderly diabetic patients, where global patterns were identified (Lin, Orgun, & Williams, 2002). A mood analysis application will require a slightly different approach. In the following paragraphs, several possible methods of analysis are discussed.

2.2.1 Big Data

At this time, it is not necessary for the proposed application to handle extremely large datasets. It is nonetheless valuable to discuss Big Data, however, because the technology may be adopted by health care organisations in the future. De Mauro, Greco, & Grimaldi (2016) define Big Data as high-volume information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation. Raghupathi and Raghupathi (2014) conclude that "Big Data analytics in healthcare is evolving into a promising field for providing insight from very large datasets and improving outcomes while reducing costs". A plethora of tools is presently available for Big Data analysis, but in this research, only the tools and techniques that are potentially useful for a mood analysis application are discussed.

Patient profile analytics, where advanced analytics such as segmentation and predictive modelling are applied, is one of these promising techniques. Segmentation of data can be achieved via cluster analysis, meaning similar data points are grouped into similar categories (or clusters). This could lead to the discovery of several variables that frequently accompany each other. In other words, the application might be able to discriminate between days where the user feels great and days where the user does not, based on certain health factors.

Big Data has indisputably come to play a pivotal role in the evolution of healthcare practices. Using the methods described above for analysis of mood data may thus lead to useful mood insights, albeit in a later stage of the project. As of now, datasets are simply not large enough.

2.2.2 Machine learning

Machine learning allows computers to learn tasks without being explicitly programmed. Spanakis et al. (2017) propose a machine learning framework for an eHealth application in the context of eating behaviour, but the framework is just as relevant for mood analysis. In their application, information in the form of several variables was first collected for each user. Classification using a decision tree algorithm was then used to predict an unhealthy eating event in the near future. More importantly, significant rules were extracted that indicated what combinations of variables were predictive of healthy or unhealthy eating.

If the same framework is applied to a mood analysis application, it would be possible to accurately predict when a user's mood is going to be low, so that an intervention can be arranged in advance. It would also allow the extraction of rules, so that patterns in mood can be recognised. Users could, for instance, be given the insight that their low mood is associated with loneliness and lack of exercise. Especially the extraction of rules is therefore a useful addition to the mood analysis component of the application.

2.2.3 Correlation coefficients

Correlations will be an important aspect of mood insights, because they can be used to assess the relationship variables have with the user's mood. Using this statistical method, it is possible to assess a possible linear association between two continuous variables (Mukaka, 2012). Finding relationships between the different mood variables (social interaction, sleep, exercise, etc.) makes it possible to analyse how these variables interact with each other. This could ultimately lead to relevant and important discoveries that are capable of improving the lives of users.

There are two methods that are most commonly applied in correlation analysis. According to Mukaka (2012), Pearson's and Spearman's correlation coefficient are excellent ways to find linear correlations, associations or connections in data. Pearson's correlation coefficient should be used only when variables are normally distributed; in all other cases, Spearman's correlation coefficient is applicable. In addition, Spearman's correlation coefficient is more robust to outliers.

It is important to note that there are a number of complications that may arise when finding correlations. Correlation coefficients do not, for example, provide any information about whether a variable moves in response to another variable. As it is impossible to always correctly discern between dependent and independent variables, relationships should be identified for what they are: associations, and certainly not causal relationships. Non-linear correlations are more difficult to recognise and the methods mentioned above may therefore not be applicable in these cases.

A second essential component was proposed that analyses previously collected mood data. Although datasets are currently small in size and stored locally, Big Data analytics, and segmentation in particular, may provide interesting mood insights in the future. Machine learning techniques can be implemented already and will allow the classification of mood data and the extraction of rules and patterns. Finally, correlation coefficients are a crucial part of the data analysis component as well, as they are capable of demonstrating relationships between mood and the influencing factors that were identified earlier.

2.3 Conclusions and future study

The aim of this state of the art literature review was to establish the essential components of a mood insights application. An effective and congruent application has at least two main components: a data registration component as discussed in section 2.1 and a data analysis component as discussed in section 2.2. The first component consists out of registering health data around (at least) the three categories activity, nutrition and technology, whereas the second component analyses this data using machine learning and correlation coefficients. The data analysis component should also be able to display and visualise relevant patterns or insights that were found. The most relevant data in this review, apart from mood, are respectively:

1. Social activity, loneliness and stress
2. Exercise and number of steps per day
3. Sleep duration and, if possible, quality
4. Sugar, water and alcohol intake
5. Social media usage

There are two notable points of discussion related to the findings. There are probably many more factors that affect mood. Although generally important ones were identified in this research, different users require different sets of factors. Further research should focus on the appropriate and effective declaration of variables for different users: which variables are of interest to which users and why? A mood analysis application has not been researched before, so the reviewed literature may not always have been equally applicable to this particular project. Apart from the quality of the gathered data, the amount of available data is also rather limited. The methods of analysis may therefore not always be accurate under these circumstances. Methods of analysis should be critically reviewed before they are actually inserted into the application, as machine learning may be effective for large datasets, but less effective for smaller datasets. Correlation coefficients may also have restrictions, considering the fact that they are only able to show linear relationships and causality cannot be proven directly.

Knowing the precise requirements is an important aspect of software development. The data registration and analysis components will serve as the basis for future research when the actual application is built and tested. It should be noted that additional mood factors and analysis techniques may be researched and added to the application, although the current components are most essential. With regard to possible future research, many more factors can be found that affect mood.

3 Ideation

The goal of the application, which will be referred to as “Optimood” in this paper, is straightforward: to make users more aware of the specific lifestyle factors (see definition 3) that affect their mood positively or negatively. Variables related to the user’s life are registered (lifelogging) and analysed daily in an attempt to find valuable patterns and insights: some variables may interact with one another without the user actually knowing it. By making the user aware of these patterns, users can adapt their lifestyle accordingly. A user that spends a lot of time on social media websites, as based on the data that he or she registers everyday, may receive an insight such as: “Social media increases the amount of stress you experience. High stress levels also affect your mood negatively to a large degree. By decreasing your social media usage, you may feel happier in general.” With the rise of technologies such as apps, social media and mobile phones, a large variety of data is already being collected. Some devices register the amount of physical activity per day, while others register sleep patterns or even blood pressure. Collecting health data often leads to more awareness and may improve the health and wellbeing of its users.

“Any factor (sleep, stress, diet, etc.) that might interact with other variables such as the mood of the user. Ideally, the user becomes more aware of these variable interactions after using Optimood.”

Definition 3: defining the term ‘lifestyle factor’ or ‘variable’ in the context of a mood analysis application

One important aspect of health and wellbeing is a person's mood. Optimood should be able to determine, exclusively using the collected health variables, what factors contribute to a positive or negative mood. Many apps and devices are already generating an increasingly large amount of data. Optimood should ultimately combine data from these different sources and create one large dataset. In this chapter, an abstract idea and vision will be presented for the application. This idea followed after market research and advances in technology.

3.1 Market research

A survey was conducted among 70 men and women between the ages of 17 and 65 in an attempt to evaluate the demand for an application such as Optimood (see Appendix A: Market research). A large percentage of respondents, nearly 80%, admitted they had suffered from depression or extended periods of low mood. This is an indication that a mood analysis application could be an interesting development for a significant part of the population. Many applications that visualise health data are already available, but these applications do not offer any profound data analysis (see figure 3.1). They do not compare variables from different sources or register patterns in the user's data, even though the survey in Appendix A shows there is a high demand for this feature (see questions F and G).

It should be noted that only 19% of the respondents were familiar with using eHealth technologies. When asked about the possible value of Optimood, however, most people indicated that they would likely find the application valuable. Nine out of ten respondents mentioned they would definitely or maybe be interested in knowing more about the factors that affect their health and happiness. If Optimood were available today, approximately half of all respondents would certainly try the application. Those that had suffered from either depression or periods of low mood in the past were clearly more interested in Optimood than those who had not.

According to the survey, users are especially interested in knowing the relationship between mood, stress and sleep. Several other factors that users would also like to know the effects of are: leisure time, smoking habits, calorie and salt intake, vitamins and fluctuations in body weight. Optimood should allow users to add their own categorical and continuous variables, as lifestyles differ from person to person and different people will be interested in registering different variables. It should be possible for users to freely add any variable they want.

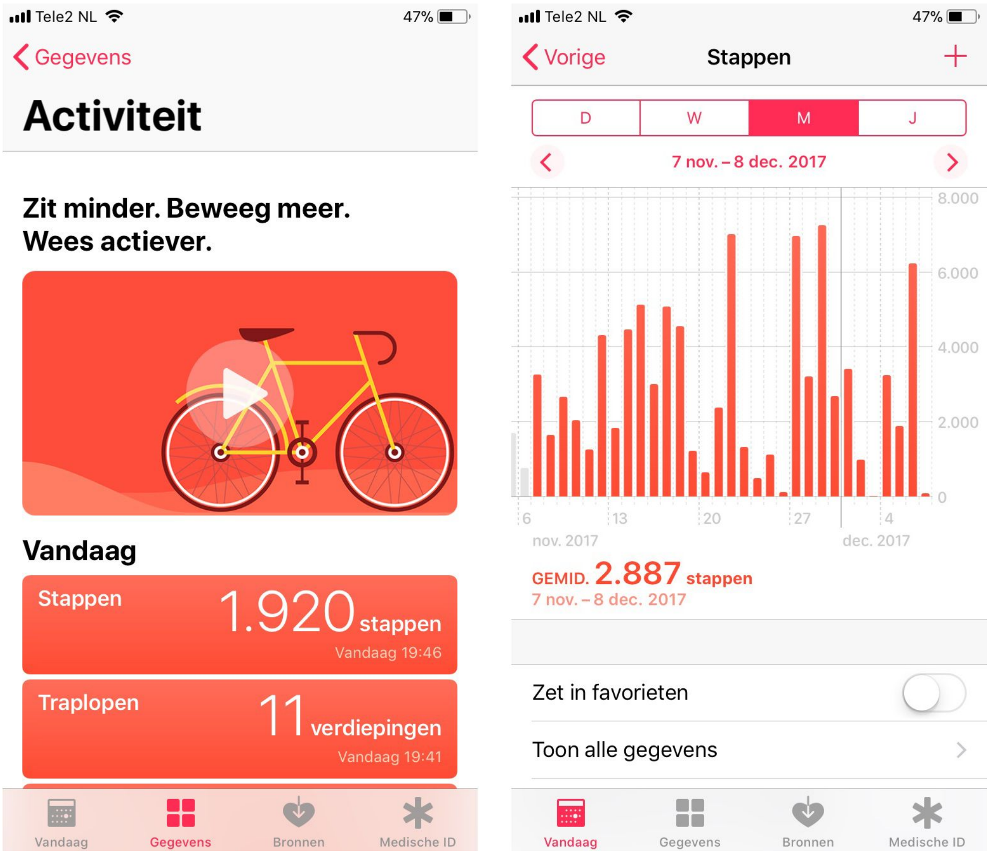

Figure 3.1: Many eHealth applications, in this case Apple's HealthKit, do not incorporate any form of data analysis; they merely visualise information

3.2 Use cases

As part of the ideation process, a use case for the application will be discussed in detail. The primary goal of implementing this technology is to accurately analyse health data from users and inform them about possible patterns related to their mood. Certain lifestyle factors or habits may have an impact on the happiness of the user; Optimood should make users aware of these patterns in an attempt to optimise their mood.

The most important stakeholders for this project are by far the users of the application. Some of these users may simply be interested in knowing more about their health and happiness; others might want to know more about the precise factors that affect their mood. It is also possible that they are interested in storing data about their life in a single and clear overview. These stakeholders will expect a secure system that provides accurate information. Other stakeholders for the system are the developers; their interest lies in building a powerful tool that analyses the mood of users.

By specifically identifying the parts of the system, the scope of the project can be determined. In this project, users will be allowed to enter and save their data in a secure database. Data will be analysed using statistical methods and machine learning and the results from this analysis will be transformed to more readable text. If, for instance, a negative correlation of -0.8 is found between mood and stress, the system could display the insight: "There is a strong negative correlation between your mood and stress. When you experience a lot of stress, your mood is often also lower. A better mood, in contrast, is associated with less stress." The most important data is also plotted and visualised in a graph. The application will not be available on mobile platforms yet; Optimood will initially be developed for Windows, Linux and Mac OS X. The application itself will primarily be a proof of concept: the main functionality will be included, but advanced analysis and insights will not be fully available yet.

The most prominent use case will be for mood insights: people that are interested in knowing more about the factors that affect their mood. Some of these people may have suffered, or are currently still suffering, from depression or other related mood disorders. Others simply want to know more about their health and happiness. These users will fill in their health data on a daily basis for an extended period of time, hoping to gain some knowledge from the insights that are generated by Optimood. In figure 3.2, a storyboard is shown that visualises the process for this use case. Another possible use case is the sole collection of health data. Some users are interested in storing data such as blood pressure, physical activity and sleep quality, so that they can easily track their health. Optimood should store health data in a database and generate corresponding data visualisations. The general flow of the application comes down to the following process:

1. Different variables related to the user's life are filled in everyday. These can include many personalised variables such as the number of cigarettes that a user smoked, the amount of stress a user experienced or whether the user performed any meditation
2. Patterns start to emerge; these are recognised using machine learning and statistical methods
3. The recognised patterns are described and explained to the user as clear and understandable insights
4. The user becomes more aware of certain behavioural patterns and may choose to adjust his or her behaviour accordingly

In the first and second use case, users will start the application and create a new profile. If a valid username is filled in, they will be forwarded to the data entry tab. In this tab, the predetermined variables from section 2.1 (and in the near future also unique personal variables) can be filled in. If, on the other hand, an invalid username is entered, an error message appears on the screen. Inside the data tab, users can select or enter values depending on the type of variable. To register the amount of stress, for example, the user can choose one of three buttons: low, normal or high. When all variables in a certain category have been filled in, the user can press the 'next' button. This process continues until a 'save' button appears. When the user presses the 'save' button, Optimood attempts to save all data in an encrypted database file. A notification should appear that tells the user whether the data was saved successfully or not. The application should then switch to the insights tab, where the user is informed of the status of analysis: a progress bar should show the user how much more data is needed before analysis can start. If enough data has been entered, the progress bar should disappear and insights will be generated and added to the insights tab.

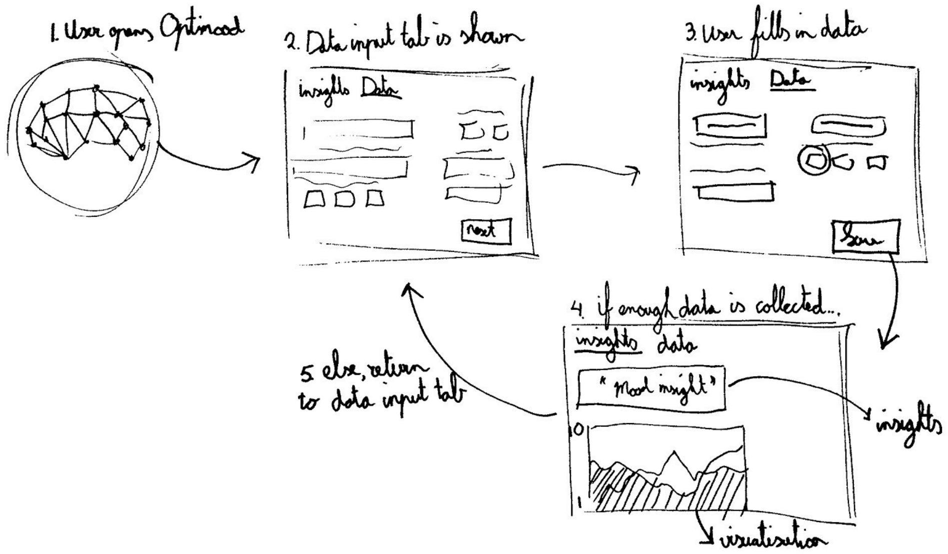

Figure 3.2: Storyboard for Optimood

To further illustrate this process, a user's application flow might look something like the example below. The simulated dataset in this example has slightly exaggerated variables for demonstration purposes; genuine data will most likely be more subtle.

1. Every single day, the user registers, amongst many other variables, his or her mood and the amount of stress that he or she experiences. The data has the following appearance;

Mood	Stress
10	low
7	normal
3	high
9	low
2	high

2. In this case, stress seems to be related to the mood of the user: when the user experiences little stress, his or her mood is excellent (9 or 10). When the user experiences a lot of stress, his or her mood is worse (3). This pattern is also recognised by Optimood
3. The recognised pattern is described to the user as an insight: “Stress is negatively associated with the way you feel. When you are under a lot of stress, your mood seems to suffer. You may benefit from reducing your stress level, for example by performing mindfulness meditation or deep breathing exercises”
4. The user becomes more aware of the particular effect stress has on his or her mood and he or she may choose to take action

3.3 Advances in technology

In this phase of the design process, the goal is to conceptualise and establish the project idea, partially by building upon previous ideas and developments. There are several advances in technology that have paved the way for the ideation of Optimood. Apart from the self-evident increases in computer data storage and computation power, machine learning and data science have, in the last few years, become more popular than ever. The rise in popularity is partially caused by improvements in algorithms, but also by the ever-increasing amount and availability of large datasets. The idea of a mood analysis application was conceived as a possible data science solution to the problem of increasing rates of depression.

Another important development that led to the ideation of Optimood was the *quantified self*, which Lupton (2012) describes as a movement where users collect data about themselves to improve daily functioning. This “self-knowledge through self-tracking with technology” is becoming easier, as many technologies are being invented that unobtrusively register all sorts of health data. Some of these technologies, such as Fitbit, act as a heart rate monitor and pedometer. There are also various digital technologies available for smartphones that register variables such as sleep quality, activity and nutrition. By combining the storage of (health) data with advances in data science, a new type of service is feasible where data is not only stored or visualised in an application, but also intricately analysed. In that way, valuable conclusions can be drawn on an individual level.

Users would, preferably automatically, collect their data from different sources (smartwatches, pedometers, Runkeeper, etc.) in one application. Unique variables can also be added manually and registered daily. After several weeks, when enough data has been acquired, users will be shown the generated data insights in plain and understandable text. The final version of the service should be available both as a desktop and a smartphone application, possibly with data synchronisation. Furthermore, users should be notified when they are supposed to fill in their data or when new insights are found.

3.4 Vision

Many companies track all sorts of variables related to their business and use that data to create models that can optimise their business processes. The vision for Optimood is that a vast amount of information derived from personal data can be just as valuable for human lives. The last few years, artificial intelligence and Data Science have also become ubiquitous. Optimood will be an attempt to use these advances in technology to increase happiness for its users. Even though the application is certainly not the solution to mental health problems, it may still be able to find interesting correlations and enhance mood.

In the future, Optimood could also be a tool for psychologists and psychotherapists to track the mental wellbeing of patients. Doctors could even incorporate certain levels of hormones in the blood, such as testosterone, insulin, cortisol or thyroid hormone, to determine whether this affects any other variables. Optimood has been proposed to the Depressie Vereniging, where the idea of the application was well-received.

Although there are many tools available that allow users to track their mood and health, such as Apple's HealthKit and Google Fit, advanced insights are not yet readily available. Ideally, Optimood would be the missing link that finds the interactions between all the variables that are stored. Data from different users could even be combined in a large database, so that general patterns can be recognised in the population. This Big Data approach would possibly allow data scientists to find certain predicting factors for diseases and other complications.

In conclusion, this chapter described the ideation of a mood analysis application called Optimood. The idea for the application followed from market research and a number of technological advances in the fields of wearables, mobile phones and data science.

4 Specification

Before Optimood can be realised successfully, the exact requirements of the project have to be specified. In this chapter, the ethics of health data storage and analysis will be discussed first: by identifying any ethical issues beforehand, the requirements and specifications can be based on this evaluation. The application's framework will therefore partially follow from this ethical discussion. The design choices are also based on the initial ideation phase and are an attempt to concretise the project idea.

4.1 Ethics of health data storage and analysis

Although there are already several applications and software packages that store health data, the incorporation of machine learning and statistics for optimising mood is a rather new endeavour. Data analysis techniques are often used in companies to optimise sales or other business processes, but may be just as effective at optimising health and happiness in people. This subchapter will discuss the possible ethical objections towards such an approach. It will also elaborate on the feasibility of incorporating machine learning into the application, as well as the effects Optimood or similar services may have on society in the future.

4.1.1 Data storage

There are two significant ethical hurdles regarding Optimood that have to be addressed first. The application strives to maximise values such as health, happiness and knowledge, but does so by gathering a variety of personal data from its users. It stands to reason that this collection of data conflicts with the user's privacy and security. As the application will be used by consumers and possibly even patients, data protection must be impeccable and is of the utmost importance. Only then will users trust the application enough to provide it with their sensitive health data. Nevertheless, some skeptical users might never want to use this technology.

The Optimood database file will, at least for now, be saved locally on the computer's hard disk, so eavesdropping via an unsecure network connection is out of the question. It would theoretically still be possible for a hacker to gain remote access to the database file, however, so thorough encryption is necessary. By encrypting all data according to the Advanced Encryption Standard (AES), the risk of a data leak can be minimised. If the database would somehow still be compromised, it would likely result in an upsetting situation. Personal data such as sleep quality, medication, smoking habits and overall health are unquestionably private to the user. If employers, companies or acquaintances would acquire these details, it may have serious consequences. Leaked data could, for example, be used by advertising agencies to launch targeted campaigns based on health data. Users that do not get enough exercise might receive advertisements for gym memberships. In addition, employers might not want to hire an unhealthy job candidate solely on the basis of their leaked health data.

From a utilitarian point of view, improving people's lives using health data has, nonetheless, greater utility than a possible data breach: even though a data breach is undeniably significant for the person harmed, the probability that such an event happens is very low. The associated risk, which is defined as the calculation of severity of harm multiplied by the probability of occurrence, is therefore almost negligible. Furthermore, all data will be anonymised so that it is exceptionally difficult to relate to a user; Optimood does not store any names, it only stores a set of encrypted variables.

4.1.2 Data analysis

The main focus of Optimood is to process all data and transform this data into convenient and practical insights. Providing users with advice on how to live their lives, however, may lead to resistance. For this reason, a patronising standpoint should be avoided at all costs. The sole purpose of Optimood is to improve the lives of its users. From a deontological perspective, the aforementioned negative consequences, although improbable, are then not significant per se. Regardless, there is still one problem that may pose a threat to users, namely incorrect insights.

Because the software is fallible, it may misinterpret the data and, in rare occasions, do more harm than good. Finding causal links in large datasets can sometimes be problematic, so caution is advised (Illari, & Russo, 2014). Correlations do not, for example, provide any information about whether a variable moves in response to another variable. Sound statistical analysis and precision are by all means a sine qua non for accurate insights. Apart from statistical analysis, techniques in the field of machine learning can, at least under some conditions, also be inaccurate. The models that are often used for predictions can sometimes over- or underfit a dataset. For a mood analysis application, especially overfitting will be a risk, as datasets will not contain a large number of instances. Another problem is that any outliers may have a large effect on the model that is generated. The software should always state how certain it is that the generated insights are correct. It would also be wise to have all users accept a terms of service in order to rule out any absolute guarantees.

Of course, there is a chance that some people will completely disregard the advice and insights that are generated. On the other hand, there will be people that follow the advice and learn in what ways their lifestyle affects mood. In these cases, data insights may very well have an important and positive effect on the wellbeing of the user. In the future, psychotherapists may even incorporate data registration and analysis to gain more insight into the lives of patients. A psychiatrist might want to register the variable ‘medication’ and check whether it influences other variables such as ‘mood’ for depressed patients or ‘hallucinations’ for schizophrenic patients. If a significant correlation emerges for these variables, the psychiatrist may be able to determine the effectiveness of a medicine for a particular patient. Naturally, a patient can also tell the psychiatrist about his or her own experience, but more subtle or advanced patterns might only be recognisable by Optimood (or a similar psychotherapeutic, data-based application).

Personal data insights may not only be useful in a clinical context; there is also a chance that these methods are incorporated in widely used applications such as Apple’s HealthKit and Google Fit. Currently, HealthKit and Fit can store many variables, including sleep quality, diet and the number of steps a user sets during the day. These variables can be visualised in a simple graph, but advanced analysis is not yet available: relationships between variables are not in any way examined. If Apple and Google were to generate personal insights, it would have both positive and negative consequences. On the one hand, users will become more aware of their health and the factors that influence it. On the other hand, automated data analysis is a difficult process that can, at least in some cases, result in incorrect insights. Some users may be disappointed by these insights; others might not want to be involved in it altogether. Possible causes for this are lack of interest or fear, as some users might be frightened by the idea of having their personal data analysed by an AI from Google or Apple.

4.1.3 Conclusions

The proposed application has, in potential, the capability to effectively influence the wellbeing and lives of users in a positive way. Although there are both general and clinical uses for Optimood, there are undeniably a number of ethically constraining factors.

Firstly, there is a trade-off between data storage and privacy. By collecting interesting personal health variables, equally interesting insights can be found. At the same time, this may have repercussions in the event of a data breach. By encrypting and anonymising the database, this risk can be almost entirely mitigated, however.

Secondly, the generated insights may in some cases rely on weak correlations or weak predictive models. Users must be informed that Optimood may come to incorrect conclusions. A possible threshold might help in differentiating between weak and strong correlations and subsequent insights. Taking these points into consideration, as well as the possible future uses for Optimood, the application is still an ethical and conceivably valuable option in determining how an optimal mood can be achieved.

4.2 User requirements

The functional requirements are divided into three main categories: usability, security and accuracy. First and foremost, high usability is essential for Optimood. The application should be easy to understand; it should operate without any delays and require a minimal amount of disk space. Optimood should also be efficient, intuitive and satisfying to use. A usability test can determine to what extent these and other requirements are met. Data entry, an essential part of the application, should be a smooth and effortless experience instead of an otherwise tedious job. The final version of Optimood should unquestionably be able to include data from sensors (Apple Watch, Fitbit Tracker, etc.). For now, however, the concept that will be developed should rely on a self-report method where users fill in their own data. This will significantly reduce the time that is needed for development.

Any graphs and insights that are generated should be easy to interpret and unambiguous: concise, clear and not too technical in nature, while still delivering the user with new knowledge of his or her health. In table 4.2, several survey questions are shown that will be used to measure the usability of the application. To formalise the requirements, a minimally required score is shown in the second column.

Usability test question	Minimal score
How much effort did it take to fill in your data?	≥4
How easy was it to interpret and understand your insights?	≥4
How attractive was the design of the user interface?	≥4
To what extent was the user interface intuitive?	≥4
What is your overall score for Optimood?	≥4
Can you identify with the insights that were generated by Optimood?	≥4
Do the insights deliver any new or valuable information?	≥4

Table 4.2: Usability test questions and their minimally required score

Apart from usability, security is also a fundamental requirement for the project. Optimood is required to encrypt all health data, so that safe data storage can be guaranteed. The encryption method should be based on the Advanced Encryption Standard (AES), which, in theory, is often considered to be uncrackable.

Accuracy is the final requirement for Optimood: the application must be reliable by showing correct insights and informing the user about the statistical significance of generated information.

5 Realisation

In the realisation phase of this project, the application will be designed and programmed according to the earlier specifications and requirements. The application consists out of a back end and a front end. The front end of the application, in turn, consists of an initial user registration screen, a data input and a data insights tab.

5.1 Back end

The back end of Optimood will be written in Python, as Python is an extremely flexible and highly acclaimed programming language for machine learning and data science. By importing Flask, a micro web framework, the user interface can be designed in HTML, CSS and JavaScript, while retaining the powerful functionality of Python. A WebView is then generated that connects to the Flask server, displaying the front end of the application. As Flask will run in the main thread, the WebView and other processes will need to run in a separate thread. Although future versions of Optimood should also be made available to Android and iOS, initial development will focus on a desktop application for Mac OS X, Windows and Linux. A mobile version of Optimood can be built upon the desktop application at a later stage.

Three custom Python scripts are written to determine the weather, temperature, season and day of the week whenever data is entered. These variables will automatically be added when any data is entered in Optimood and may show whether a user's mood is affected by the season or weather. As security is an important aspect of Optimood, any data that is saved will be stored locally in an encrypted CSV file using the PyCrypto library. Encryption and decryption will take place according to the Advanced Encryption Standard (AES). Furthermore, daily desktop notifications will remind the user to fill in their health data.

Although automatic data collection is imperative in the final version of Optimood, the concept that will be realised in this paper requires users, for the most part, to fill in their data manually. When the minimum of 14 days of data has been reached, the application should decrypt the CSV file and start analysing the data.

For the data analysis component, Optimood will use TensorFlow and scikit-learn for any machine learning tasks, NumPy for scientific computing, and pandas for data structure and analysis. Matplotlib, a 2D plotting library, will be used to visualise results while debugging the Python code. This library is obsolete in the final application, as any visualisations are taken care of in the more flexible front end of Optimood.

5.2 Front end

The front end of the application will, as mentioned before, primarily be written in HTML and CSS. By building a web application, scaling to other environments is much easier. HTML is also more adaptable and does not need any compilation, in contrast with many other programming languages, which allows for quick tinkering of the design. By running a local Flask server, Optimood can interact with the Python back end to execute functions or request and submit variables.

When the application is first started, the user is prompted for his or her name. This name is then saved and stored inside a configuration file. The application is designed using Rocket, a responsive admin dashboard theme, which in turn is based on the bootstrap framework. The design itself should be simple and intuitive for all users, requiring little, if any, explanation.

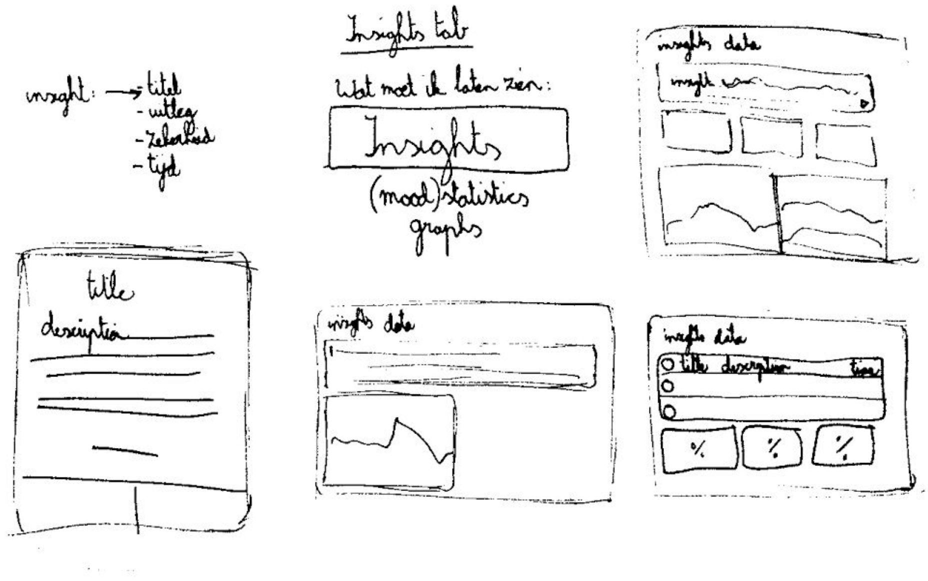

Figure 5.2: Designing the insights tab

Optimood will respectively feature a data and an insights tab. Inside the data tab, users should have the ability to quickly input their health data. At this point, adding custom variables is difficult and, unfortunately, has different repercussions. For this reason, users will only be allowed, at least for this version of Optimood, to enter the variables that were identified in chapter 2. When data is saved, users are forwarded to the insights tab. If any essential data entries are missing, an error message should appear and notify the user.

The data tab registers both objective and subjective variables. Some of the dichotomous categorical variables are, for instance, whether the user drank any alcohol or had any exercise. Ordinal categorical variables include variables such as stress levels and sugar intake. Some variables, including those for water intake, hours of sleep and number of steps, are measured quantitatively. Mood could possibly be measured on a scale of one to ten, where the lowest possible score indicates depression or an otherwise undesirable mood and the highest score indicates an optimal mood. Selecting mood on such a scale may be a somewhat constraining factor for more advanced insights, but it also simplifies data analysis. An alternative would be the Affect Grid as proposed by Russel, Weiss and Mendelsohn (1989), where not only the pleasantness of the user's mood is assessed, but also his or her state of arousal.

Lastly, the insights tab will import and display any insights that were found by the Python back end of Optimood. Data will also be visualised in a graph using D3, a JavaScript library that is often used for producing dynamic, interactive data visualizations in web browsers. These visualisations will start appearing after one week, whereas insights can take up to two weeks. Users will receive regular progress updates and a message is displayed that shows the number of days until insights are available.

5.3 User registration design

The first time Optimood is run, the user can choose to fill in his or her name (see figure 5.3). This username is stored in a configuration file and allows the user to create a unique profile. By creating a profile, the user can be addressed with his or her name to further personalise insights. It also paves the way for backups and possible synchronisation between different devices in the future. When the profile is saved, Optimood automatically switches to the data entry tab.

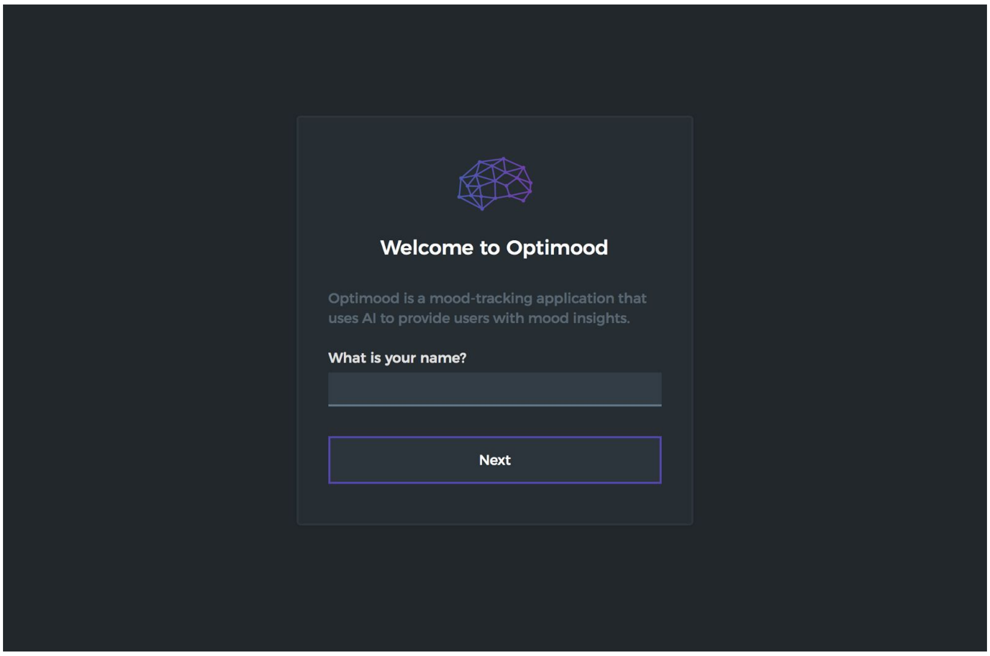

Figure 5.3: Prompting the user for his or her name

5.4 Data tab design

In figure 5.4, the final design for the data input tab is displayed. Note that the program slightly deviates from the categories that were defined in section 2.1, even though all variables from this section were included in the data tab. By grouping the variables into the categories mental health, activities and nutrition and alcohol, three groups of an even four variables could be formed. The user has already filled in whether she had any exercise that day. After pressing the ‘next’ button, all data will be validated. If any data is missing, an error message will appear.

5.5 Insights tab design

The insights tab will perform different types of analysis on the user's health data. The implementation of these methods is described below. The final screen for insights is illustrated in figure 5.5.

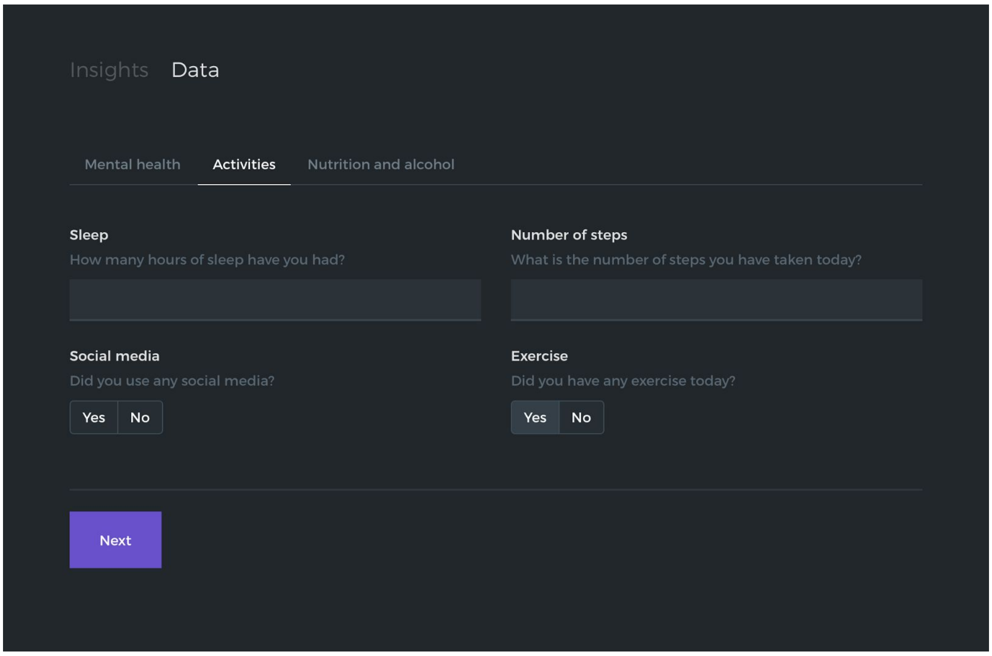

Figure 5.4: Final design for the data input tab

5.5.1 Hard-coded insights

Optimood will have a number of 'hard-coded' insights and advices that do not require any form of advanced analysis (see table 5.5.1). These advices will be generated whenever certain variables fall, on average, below a predefined threshold. Advices are largely based on scientific research and recommendations from the World Health Organization.

When users repeatedly report they set less than 6,000 steps per day, they will be advised to walk more; preferably around 7,000 to 10,000 steps per day (see figure 5.5.1). This advice is based on research by Tuder-Locke et al. (2011). If users report less than 150 minutes of physical activity per week, they are also advised to increase their amount of physical activity, as based on recommendations by the World Health Organization (2010). Research by Yuenyongchaiwat (2016) shows that 10,000 steps per day are associated with a lower blood pressure and a positive mood.

```
steps_average = df['steps'].mean()

if(steps_average < 7000):
     steps_average = int((steps_average + 100 - (steps_average % 100)))
     steps_advice = "The amount of steps you set per day is approximately " + str(steps_average) + ". For an optimal health, you are advised to set at least 7,000 to 10,000 steps per day."

insights_list.append(steps_advice)
```

Figure 5.5.1: Implementing the hard-coded insights. If the user sets less than 7,000 steps per day, an advice is generated that is appended to the list of insights

Loneliness and social isolation, particularly in the older adult, have been shown to influence psychosocial well-being (Alpass & Neville, 2010). When users have been feeling lonely for an extended period of time, they should be given tips and advice on how to adequately cope with loneliness. If users report low amounts of social activity, similar advice should be given. In this way, Optimood attempts to reduce feelings of loneliness by stimulating users to reach out to friends or family.

The average adult needs around seven to nine hours of sleep every night (Centers for Disease Control and Prevention, 2011). As users fill in how many hours of sleep they have had on a particular day, Optimood takes notice of users that are at risk for sleep deprivation. Enough sleep is important for an optimal mood, which was shown by Simon et al., (2015). When users continuously sleep less than 7 hours per night, an advice is automatically generated.

The Institute of Medicine recommends that adults drink between 2.7 to 3.7 litres of water per day (2005). If Optimood notices the user drinks, on average, less than this recommended value, the user is advised to drink more water.

A final hard-coded insight is generated when the user experiences a lot of stress on a regular basis. If stress levels are continually high, Optimood will advise the user to reduce stress by means of meditation or deep breathing exercises. These generated advices are not exceptionally valuable per se, but can be of interest to users nonetheless.

Variable	Threshold	Representation
steps	7,000 and below	The amount of steps you set per day is approximately *steps*. For an optimal health, you are advised to set at least 7,000 to 10,000 steps per day.
loneliness	3 and above	You have been feeling lonely for an extended period of time. To reduce your feelings of loneliness, you may try volunteering, contacting a relative or doing something creative.
sleep	7 and below	You sleep *sleep* hours per night on average. For an optimal mood, you should sleep at least 7 to 9 hours per night.
water_intake	2.7 and below	On average, you drink about *water_intake* litres of water per day. The Institute of Medicine recommends drinking at least 3.7 litres of water per day.
stress	3 and above	Your stress levels have been relatively high. Chronic stress can contribute to long-term health problems. To reduce your stress levels, try deep breathing exercises or mindfulness meditation.

Table 5.5.1: Overview of the hard-coded insights, their predefined thresholds and how they are presented in the GUI

5.5.2 Correlation coefficients

Whenever Optimood is run, correlation analysis is performed on all variables. In this paper, the strength of the correlations wil be described using the guide that Evans (1996) suggested for the value of r:

- Any correlation between 0 and 0.19 is very weak and immediately disregarded by Optimood

- Correlations between 0.2 and 0.39 are also considered weak; these correlations are also disregarded by Optimood. At a later stage, the user should be able to decide whether these correlations are shown or not, but this is not included in the initial proof of concept
- Correlations between 0.4 and 0.59 are considered to be moderate in strength
- Any correlation greater than 0.6 is considered to be strong or very strong

After all correlations are calculated, insights are generated accordingly and appended to the list of insights. To illustrate, if a correlation of >0.6 is found for the variables mood and exercise, the following insight would be generated: "For your health data, exercise seems to be strongly associated with a positive mood." An overview of all possible mood correlations and their respective insights can be found in Appendix C.

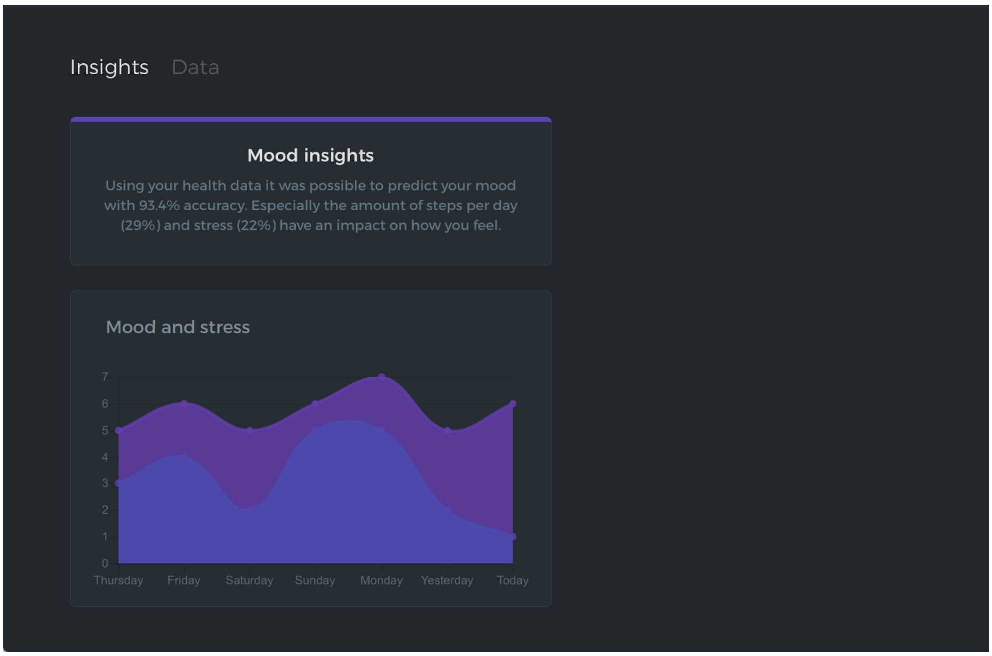

Figure 5.5: Final design for the insights tab

5.5.3 Machine learning

Many machine learning models perform extremely well at classification and regression tasks. Though the theory behind these models is often well-understood, they are generally regarded as black boxes because the weights of the model remain unknown. It would be an achievement if a user's mood could be accurately predicted from the health variables that Optimood stores. It would not, however, explain what variables contribute to a positive or negative mood state. Decision trees (as a predictive model) have a somewhat unique property: they are relatively easy to interpret and allow us to understand why the classifier or regressor makes particular decisions. Using a decision tree, it is possible to assess the feature importance (as stated earlier in section 2.2.2). This property will prove useful in a mood analysis application.

After three weeks of data collection, a random forests regressor is used to predict the user's mood score. To put it simply, random forests are a collection of decision trees that can also be used on smaller datasets. They do not only have the advantages of a decision tree that were discussed previously; random forests are also less prone to overfitting, which is a definite risk when modeling smaller datasets. When the random forests regressor is accurately trained, the importance of the different features will be extracted. Considering the rather small size of the datasets, an accuracy of more than 90% will be considered accurate. The extraction of feature importances will lead to a list of features ranked by the effect they have on a user's mood. Finally, the most important features are converted to insights.

A system architecture diagram for Optimood can be seen in figure 5. The latest development release of Optimood can be downloaded from *https://github.com/ ivarvanwooning/optimood.*

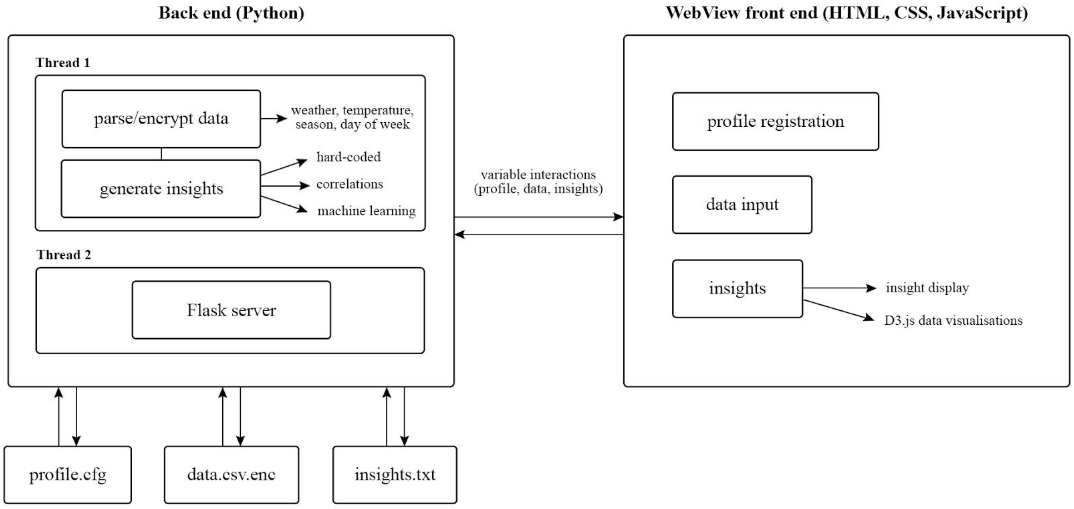

Figure 5: system architecture diagram for Optimood

6 Evaluation

In the evaluation phase of this project, the realised concept from chapter 5 will be evaluated. The requirements for Optimood were divided into three categories: usability, security and accuracy. To measure the usability and effectiveness of Optimood, a small-scale usability test was conducted with ten participants (six men and three women). The participants were asked to register, at least for two weeks, the variables that were earlier defined in section 2.1. Users were also allowed to register their own unique variables if they wanted to receive more advanced insights. After five weeks or so, the users were asked to review Optimood on several usability aspects (see table 6.1). The application itself operated on an extremely fast HTML framework and required just five megabytes of disk space. Entering data was a very efficient process; users gave an average score of 4.4 for data entry and a 4.6 (out of 5) for the attractiveness of the design. The majority of the users rated the application as intuitive and the overall score that was given to Optimood was also sufficient. Two users stated explicitly that they enjoyed filling in their data, because it made them more aware of what they had done during the day. One user reported she was disappointed the usability test had ended.

Apart from these positive results, there were also aspects that require improvement. When users did not enter a lot of data, the quality of the insights was substantially worse. Some people filled in many different variables with great accuracy, while others did not. If possible, users should add more variables and register their data for a longer period of time. By incorporating automated data registration via wearables and apps into Optimood, users might find it easier to register their data. A precursor for this, however, would be a mobile version of Optimood.

Some users reported that it was difficult to interpret and understand their insights, partially because the program was written in English. For others, the insights were not as advanced as they had hoped. Although most users could identify with the insights to some degree, the amount of new information it conveyed was less than they had expected. There are several reasons why the insights were not as clear or advanced as users had hoped. Firstly, due to time constraints, it was impossible to build in more advanced insights. In future research, more time should be invested in the development of advanced insight generation. Secondly, the amount of available data was fairly limited; most users did not fill in any unique personal variables or registered only two weeks of data. Thirdly, the insights could possibly have been clearer if they were more visual: confidence levels should, for example, be represented by stars instead of percentages. Taking this feedback into consideration, a new design for insights is proposed in figure 6.1. This example demonstrates a strong positive correlation between mood and exercise.

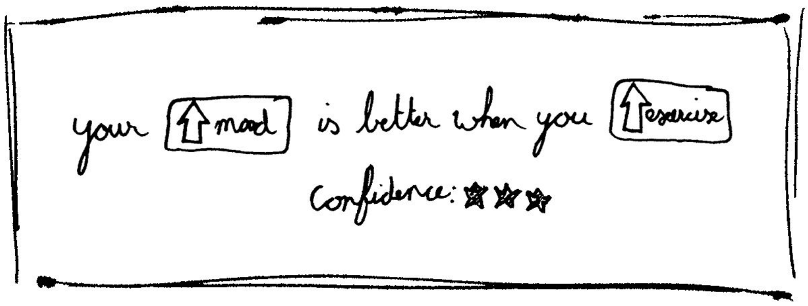

Figure 6.1: A more visual insight that displays the confidence in stars instead of percentages: your mood is better when you exercise more

Another essential requirement for Optimood was security: all health data should be encrypted and stored securely. Optimood fulfils this requirement by encrypting and decrypting all data according to the Advanced Encryption Standard (AES) protocol. All data that the user saves is stored in an encrypted CSV file. When the data threshold of two weeks has been reached, the encrypted CSV file is decrypted with a private key (a previously generated 24 character string). This ensures that the original health data can only be viewed by the application user. It must be noted, however, that Optimood does not yet require a password to login. For an even more optimal security, users should be prompted for their login credentials when they open Optimood. Unfortunately, this solution could also decrease the amount of anonymity.

The final category of requirements is accuracy: insights that are shown should be reliable and correct and the user should be informed about the statistical significance of generated information. These requirements are fulfilled in two ways: firstly, only moderate and strong correlations are shown; weak correlations are immediately rejected. For the machine learning part of the application, only the features that have the greatest influence on mood are delivered as an insight to the user. Secondly, Optimood always shows the confidence of its predictions. The importances of features are expressed in percentages and the strength of correlations are expressed in words (i.e. moderate or strong). These precautions are taken to maximise the application's accuracy.

Usability test question	Minimal score	Usability test score
How much effort did it take to fill in your data?	≥4	4.4
How easy was it to interpret and understand your insights?	≥4	3.9
How attractive was the design of the user interface?	≥4	4.6
To what extent was the user interface intuitive?	≥4	4.4
What is your overall score for Optimood?	≥4	4.1
Can you identify with the insights that were generated by Optimood?	≥4	3.7
Do the insights deliver any new or valuable information?	≥4	3.2

Table 6.1: Comparing the minimal scores as defined in the user requirements (section 4.2) with the actual usability test scores

7 Conclusion

In this research, an eHealth application was proposed that registers and analyses a wide variety of variables related to the user's life. The goal of this application was to automatically find patterns in the user's data and deliver this information in a concrete and insightful way. The most important variable in this regard was mood, as the primary purpose of Optimood was to find out which variables were associated with the best or worst mood. These influential factors were then delivered to the user in the form of insights.

A literature review was conducted to identify the most essential components of a mood analysis application. For the first component, in which data is registered, variables such as social activity, loneliness, stress and exercise were found to be determining factors for predicting mood. For the second component, in which data is primarily analysed, correlation coefficients and machine learning – specifically decision trees – were found to be effective.

The market research that followed after this literature review showed that there is a demand for Optimood. There are numerous applications and technologies that already collect health data, but none of these seem to perform any kind of advanced analysis. In recent years, computer data storage and computation power have increased tremendously. Data science and machine learning have also become very popular. By combining these developments with the storage of (health) data, the idea of Optimood came into existence. In the future, Optimood could even be used as a tool for psychotherapists to track the mental wellbeing of their patients.

The user requirements were then specified, partially by first discussing the ethical objections towards health data registration and analysis. The application certainly has the capability to positively influence the lives of users, but caution is advised as there is a trade-off between data storage and privacy. In some cases, generated insights may also rely on weak correlations or predictive models. After this ethical dissection, the requirements were grouped into three categories: usability, security and accuracy. Several user requirements were then further formalised by adding minimum thresholds for the usability test (as defined in table 4.2).

The application consists of a back and front end, respectively written in Python and HTML, CSS and JavaScript. The front end allows users to register data and view any insights that are delivered by the back end. The back end analyses data using correlation coefficients and a random forests regressor. There are also a number of hard-coded insights that are based on scientific research and recommendations from the World Health Organization.

To measure the effectiveness of the developed concept, a usability test was conducted. The application scored high on areas such as attractiveness and intuitive design, but the insights that were generated were somewhat mediocre and not very in-depth. To solve this problem, the amount of registered data should not only be increased; additional research and development is also needed to to build in more advanced insights and visualisations.

The goal of this thesis was to answer the research question: “to what extent can health data be used to provide users with practical advice and mood insights?” All in all, the Optimood prototype is an application that predominantly generates useful advice and mood insights, even though the methods of visualisation are not as effective as expected. Most users could identify with the insights that were generated, but they had expected more valuable information from it. The results were to a large extent dependent on the user's motivation to collect data. If more variables are collected for a longer duration, valuable insights could presumably be generated more easily.

7.1 Importance

The results of this research are important for several reasons. Firstly, depression is on the rise, especially among teenagers. By finding out the specific factors that influence mood for a user, a unique and individual solution can be designed. Optimood makes users more aware of the effects certain lifestyle factors can have, not only for users that have suffered from depression, but also for the general population. As market research suggested, a majority of people have suffered from periods of low mood and many of them were interested in the service Optimood provides.

Secondly, as eHealth technologies and data science are on the rise, it is only a matter of time before more data-driven applications will appear. Research is needed to further identify the possibilities and risks that are associated with this. Although the application certainly has potential, there are also several problems that have to be addressed. Section 7.2 will elaborate on these issues.

Thirdly, the underlying idea behind Optimood can be used in many other clinical and non-clinical domains. Machine learning has already been used to successfully predict problematic eating behaviour in women who suffer from eating disorders. Now, it appears data analysis may be just as effective at predicting what factors contribute to an optimal mood. The framework of Optimood could be applied to suit many different encouraging applications:

- Predicting psychotic episodes in schizophrenic patients by analysing different health variables; users could be made more aware of the factors that lead up to these episodes

- Optimising productivity for users by registering different lifestyle factors; the application could determine when the user is most productive
- Combining musical preferences from Spotify with mood data, so that users are presented with playlists that suit their emotional state

7.2 Recommendations

In this final chapter, a number of unanswered questions will be developed into research recommendations. In the first place, more research should be put into the insights generation component of Optimood. Several users stated that they had expected more from the insights that were generated. Further research should not only point out how to find more advanced and deeper patterns; more time should also be invested in finding better ways to visualise these patterns. The textual insights that were generated in the usability test were evidently confusing.

The usability test that was conducted for this research was an early-stage experiment. A larger usability test is needed in which users register more variables for a longer period of time (2-4 months preferably). In this way, Optimood will be able to find more interesting patterns with greater accuracy. If possible, the experiment should include participants who have suffered from depression. This may prove or disprove whether Optimood may be of any help to those that suffer from mood disorders. At the same time, support should be extended to mobile devices and automated data registration should be integrated. This will make it easier for users to register their data for a longer duration.

In conclusion, Optimood is a fairly promising eHealth technology that thrives on recent advances in the domain of data science. Although a working proof of concept has been realised, future research should focus on making insights more advanced and understandable.

8 Bibliography

Alpass, F. M., & Neville, S. (2003). Loneliness, health and depression in older males. *Aging & mental health, 7(3),* 212-216.

Belle, A., Thiagarajan, R., Soroushmehr, S. M., Navidi, F., Beard, D. A., & Najarian, K. (2015). Big data analytics in healthcare. *BioMed research international, Article ID 370194,* 2015.

Blumenthal, J. A., Babyak, M. A., Doraiswamy, P. M., Watkins, L., Hoffman, B. M., Barbour, K. A., ... & Hinderliter, A. (2007). Exercise and pharmacotherapy in the treatment of major depressive disorder. *Psychosomatic medicine, 69(7)*, *587*.

Centers for Disease Control and Prevention (CDC. (2011). Effect of short sleep duration on daily activities--United States, 2005-2008. MMWR. *Morbidity and mortality weekly report, 60(8),* 239.

Curtin, S. C., Warner, M., & Hedegaard, H. (2016). Increase in suicide in the United States, 1999-2014. *US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.*

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review, 65(3),* 122-135.

De Moor, M. H. M., Beem, A. L., Stubbe, J. H., Boomsma, D. I., & De Geus, E. J. C. (2006). Regular exercise, anxiety, depression and personality: a population-based study. *Preventive medicine, 42(4),* 273-279.

Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing.

Freed, E. X. (1978). Alcohol and mood: An updated review. *International Journal of the Addictions, 13(2),* 173-200.

Illari, P., & Russo, F. (2014). Causality: Philosophical theory meets scientific practice. OUP Oxford.

Institute for Quality and Efficiency in Health Care. (2017). Depression: How effective are antidepressants? Retrieved from https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0087089/.

Institute of Medicine. (2005). Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate. Retrieved from https://www.nap.edu/read/10925/chapter/1.

Kaneita, Y., Ohida, T., Uchiyama, M., Takemura, S., Kawahara, K., Yokoyama, E., ... & Fujita, T. (2006). The relationship between depression and sleep disturbances: a Japanese nationwide general population survey. *The Journal of clinical psychiatry, Feb 2006*.

Kegler, S. R. (2017). Trends in Suicide by Level of Urbanization—United States, 1999–2015. *MMWR. Morbidity and Mortality Weekly Report, 66*.

Knüppel, A., Shipley, M. J., Llewellyn, C. H., & Brunner, E. J. (2017). Sugar intake from sweet food and beverages, common mental disorder and depression: prospective findings from the Whitehall II study. *Scientific Reports, 7*.

Kross, E., Verduyn, P., Demiralp, E., Park, J., Lee, D. S., Lin, N., ... & Ybarra, O. (2013). Facebook use predicts declines in subjective well-being in young adults. *PloS one, 8(8),* e69841.

Lin, W., Orgun, M. A., & Williams, G. J. (2002). Mining temporal patterns from health care data. *Lecture notes in computer science, 222-231*.

Liu, J. J., Wei, Y. B., Forsell, Y., & Lavebratt, C. (2017). Stress, depressive status and telomere length: Does social interaction and coping strategy play a mediating role?. *Journal of Affective Disorders, 222,* 138-145.

Lupton, Deborah. (2012). The quantified self movement: some sociological perspectives. Retrieved from https://simplysociology.wordpress.com/2012/11/04/the-quantitative-self-movement-some-sociological-perspectives/.

Greer, Mark. (2004). Strengthen your brain by resting it. Retrieved from http://www.apa.org/monitor/julaug04/strengthen.aspx.

Mojtabai, R., Olfson, M., & Han, B. (2016). National trends in the prevalence and treatment of depression in adolescents and young adults. *Pediatrics*, e20161878.

Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24(3),* 69-71.

Pantic, I. (2014). Online social networking and mental health. *Cyberpsychology, Behavior, and Social Networking, 17(10),* 652-657.

Pross, N., Demazières, A., Girard, N., Barnouin, R., Metzger, D., Klein, A., ... & Guelinckx, I. (2014). Effects of changes in water intake on mood of high and low drinkers. *PLoS One, 9(4),* e94754.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems, 2(1),* 3.

Rohrer, J. E., Pierce, J. R., & Blackburn, C. (2005). *Lifestyle and mental health. Preventive medicine, 40(4),* 438-443.

Simon, E. B., Oren, N., Sharon, H., Kirschner, A., Goldway, N., Okon-Singer, H., ... & Hendler, T. (2015). Losing neutrality: the neural basis of impaired emotional control without sleep. *Journal of Neuroscience, 35(38),* 13194-13205.

Spanakis, G., Weiss, G., Boh, B., Lemmens, L., & Roefs, A. (2017). Machine learning techniques in eating behavior e-coaching. *Personal and Ubiquitous Computing, 1-15*.

Thomée, S., Härenstam, A., & Hagberg, M. (2012). Computer use and stress, sleep disturbances, and symptoms of depression among young adults–a prospective cohort study. *BMC psychiatry, 12(1),* 176.

Tudor-Locke, C., Craig, C. L., Brown, W. J., Clemes, S. A., De Cocker, K., Giles-Corti, B., … Blair, S. N. (2011). How many steps/day are enough? for adults. *The International Journal of Behavioral Nutrition and Physical Activity, 8, 79.*

Walsh, R. (2011). Lifestyle and mental health. *American Psychologist, 66(7),* 579.

Wang, X., Cai, L., Qian, J., & Peng, J. (2014). Social support moderates stress effects on depression. International journal of mental health systems, 8(1), 41.

World Health Organization. (2001). Mental Health: A Call for Action by World Health Ministers. Retrieved from
http://www.who.int/mental_health/advocacy/en/Call_for_Action_MoH_Intro.pdf.

World Health Organization. (2010). Global recommendations on Physical Activity for health. World Health Organization.

Yuenyongchaiwat, K. (2016). Effects of 10,000 steps a day on physical and mental health in overweight participants in a community setting: a preliminary study. *Brazilian journal of physical therapy, (AHEAD), 0-0.*

9 Appendix

Appendix A: Market research results

A: Gender

B: Age

C: Have you ever suffered from extended periods of low mood or depression?

D: Do you currently use any health apps?

E: How valuable do you believe an application such as Optimood could be?

F: Would you be interested in knowing more about the factors that affect your health and happiness?

G: If Optimood were available today, how likely would you be to try it?

market-research.csv

A	B	C	D	E	F	G
M	22	Yes	Yes	5	Yes	5
F	20	No	No	3	Yes	3
F	21	Yes	No	5	Yes	2
M	23	Yes	No	4	Yes	4
M	22	Yes	No	5	Yes	5
F	19	Yes	No	3	Yes	3
M	21	No	No	3	Maybe	2
M	23	Yes	No	4	Maybe	1
M	22	Yes	No	4	Maybe	4
M	21	Yes	No	3	Yes	4
F	26	Yes	No	4	Maybe	2
M	54	No	No	3	Yes	1

F	52	No	No	4	Yes	3
F	20	No	No	1	No	1
F	18	Yes	No	4	Yes	5
M	21	Yes	No	5	Yes	4
F	46	Yes	No	3	Yes	4
M	22	Yes	No	3	No	2
F	31	Yes	No	4	Yes	4
M	29	Yes	No	4	Yes	4
F	23	Yes	No	4	Yes	4
F	27	Yes	No	4	Maybe	4
F	31	No	Yes	4	Yes	5
F	29	No	Yes	4	Yes	4
F	37	Yes	No	4	Maybe	4
M	34		Yes	3	Yes	3
F	33	Yes	Yes	3	Yes	3
F	33	Yes	No	3	Maybe	3
M	65	Yes	No	3	No	1
M	20	Yes	No	4	Yes	4
M	24	Yes	No	4	Yes	4
M	22	Yes	No	4	Maybe	3
M	31		No	4	Yes	4
F	43	No	Yes	5	Yes	5
M	24	Yes	No	3	Yes	3
M	22	Yes	No	4	Yes	4

F	23	Yes	Yes	3	No	2
M	23	Yes	No	4	Maybe	2
F	25	Yes	No	2	Yes	3
F	17	Yes	No	3	Maybe	4
F	17	No	No	1	Maybe	2
M	22	No	No	4	Yes	2
M	21	Yes	No	5	Yes	5
M	23	Yes	Yes	3	Yes	3
M	19	No	No	2	Yes	2
M	20	Yes	No	4	Yes	5
M	21	Yes	No	2	Maybe	3
F	19	Yes	No	4	Yes	3
F	28	Yes	Yes	4	Yes	4
M	19	Yes	No	4	Yes	4
M	40	No	No	3	Maybe	5
F	29	Yes	Yes	2	Yes	2
M	19	Yes	No	3	Yes	3
M	21	Yes	Yes	5	Yes	5
M	26		Yes	3	Yes	4
M	18	Yes	No	2	Maybe	2
M	19	No	No	4	Yes	3
M	20	Yes	No	3	Maybe	2
M	21	Yes	No	4	Yes	4
F	22	Yes	No	3	Yes	3

F	19	Yes	No	4	No	2
F	18	Yes	No	3	Yes	3
F	21	Yes	No	1	Maybe	1
M	23	Yes	No	2	Yes	2
F	22	Yes	Yes	4	Yes	2
F	18	Yes	No	4	Yes	4
F	21	Yes	No	4	Yes	2
F	19	Yes	No	3	Maybe	4
F	19	Yes	No	5	Yes	4
F	20	No	No	3	No	1

Appendix B: Datasets for usability testing

A: Mood

B: Stress

C: Water

D: Sugar

E: Fruit

F: Sleep

G: Steps

H: Exercise

I: Loneliness

J: Alcohol

K: Social media

L: Social activity

1.csv

A	B	C	D	E	F	G	H	I	J	K	L
8	3	1,5	3	1	5	7000	0	1	0	0	1
7	3	2	2	1	5	7500	0	2	0	0	0
7	3	1,5	1	1	4,5	6000	0	1	0	0	1
8	2	1,5	2	1	5,5	5500	0	1	0	0	1
8	2	1,75	2	1	5	6000	0	1	0	0	1
7	3	1,5	2	0	5,5	4500	0	1	0	0	1
7	4	1,5	3	1	7	5000	0	1	0	0	1
8	4	2	0	1	4,5	5500	0	1	0	0	1
8	4	1,5	2	1	6	5500	0	1	0	0	1
7	3	1,5	3	1	5	5500	0	1	0	0	1
7	3	1,5	2	1	4,5	5500	0	1	0	0	1

7	3	1,5	2,5	1	5	5500	0	1	0	0	1
7	2	1,5	2,5	1	8	6500	0	2	0	0	0
6	2	1,5	3	1	8	6000	0	2,5	0	0	0
6	2	1,5	3	1	5	8000	0	2	0	0	1
7,5	2	1,5	2,5	1	6	5500	0	2	0	0	0
8	2,5	1,5	2,5	1	5,5	5500	0	1	0	0	1
8	2	1	3	1	4,5	5000	0	1	0	0	1
7	3	1,5	3,5	1	5	5500	0	1	0	0	1
8	2	1,5	4	1	6,5	5500	0	1	0	0	0
7,5	2	1,5	4	1	7	5000	0	1	0	0	1
8	2	1,5	3	1	5	5500	0	1	0	0	1
8	2	1,5	3	1	7	5000	0	1	0	0	0
7,5	2,5	1,5	3	1	4,5	5000	0	1	0	0	1
7	3,5	1,5	3	1	4,5	5500	0	2	0	0	1
8	2,5	1,5	3,5	1	4	6000	0	1,5	0	0	1

2.csv

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
7	2	1	1	1	8	6000	1	2	0	1	1
6	2	1	1	1	8	3000	0	3	0	1	0
7	2	1	1	1	8	6000	1	2	0	1	1
6	2	1	1	1	8	3000	1	3	0	1	0
7,5	2	1	1	1	8	6000	1	1	1	1	1
7	2	1	1	1	8	3000	0	3	0	0	1
6	2	1	1	1	8	3000	0	3	0	1	0

7,5	2	1	1	1	8	3000	0	3	1	1	1
8	2	1	1	1	8	7000	0	2	0	1	1
8	2	1	1	1	8,5	7000	1	2	0	1	1
6	2	1	1	1	7,5	2000	0	3	0	1	0
8	2	1	1	1	8	6000	1	2	0	1	1
6	2	1	1	1	8	3000	0	3	0	1	0
7	2	1	1	1	8,5	4000	0	3	0	1	0
8	2	1	1	1	8	8000	1	2	0	1	1
7	2	1	1	1	7,5	6000	0	2	0	1	0
8	2	1	1	1	8	6000	1	2	0	1	1
6	2	1	1	1	8	4000	0	3	0	1	0
8	2	1	1	1	7,5	6000	1	2	0	1	1
8	2	1	1	1	8	4000	0	3	0	1	1
7	2	1	1	1	8,5	6000	0	3	0	1	0
8	2	1	1	1	8	6000	1	2	0	1	1
7,5	2	1	1	1	8	4000	0	3	0	1	1
8	2	1	1	1	8	6000	1	2	0	1	1
9	2	1	1	1	8	14000	0	1	0	1	0

3.csv

A	B	C	D	E	F	G	H	I	J	K	L
7	2	2	3	1	7	0	0	2	0	0	0
6	3	2	2	1	8	0	0	2	0	1	0
6	2	2	3	1	8	0	1	1	0	0	1
7	2	3	1	1	8	0	0	1	1	1	1

8	1	2	2	1	7	0	0	1	0	1	1
7	2	2	2	1	9	0	0	2	0	1	0
6	2	1	3	1	8	0	1	1	1	1	1
8	3	2	2	1	7	0	1	1	0	0	1
7	2	2	3	1	9	0	0	1	0	1	0
7	3	2	2	1	9	0	0	2	0	1	0
7	2	1	1	1	8	0	0	1	0	1	1
6	2	2	1	1	8	0	0	1	1	0	1
8	1	2	2	1	7	0	0	2	1	1	1
7	2	2	2	1	8	0	1	1	0	1	1
8	1	1	2	1	8	0	1	1	0	1	1

4.csv

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
7,5	1,5	0,5	2	1	7	3698	0	1	0	1	1
5	2	0,2	2	0	8	6809	0	2	0	1	0
8	3	0,7	2	0	7	11912	1	1	0	1	1
7	2	0,7	2	0	8	4658	0	1	0	1	1
6	4	0,3	4	0	7,5	7570	0	1	0	1	1
8	3	0,2	2	0	8,5	5500	0	1	0	1	0
7	2	0,5	3	1	7	6852	0	1	0	1	1
6	4	0,5	2	1	7	4372	1	1	0	1	1
8	1	0,4	2	0	7,5	5585	0	1	0	1	1
9	1	0,3	2	0	8	9787	1	1	0	1	1

8	1	0,3	2	1	7	12920	1	1	0	1	0
7,5	2	0,3	4	0	8	10633	0	1	0	1	1
8	2	0,3	2	1	8	12925	1	1	0	1	1
9	2	0,3	2	0	8	4225	0	1	0	1	0
7	2,5	0,4	3	0	7	4663	0	1	0	1	1
7,5	1	0,7	2	1	9	8185	0	1	0	1	1
8	1	0,7	2	0	7	9046	0	1	0	1	1
7	2	0,7	3	1	8	7608	0	1	0	1	0
7,5	2,5	0,7	3	0	8,5	7608	0	1	0	0	1
8	1	0,4	3	0	8	7608	0	1	1	1	1
9	1	0,7	4	0	8	7608	0	1	0	1	1

5.csv

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
7	1	2	2	1	7,5	8000	1	1	0	1	1
7	2	2	2	1	8	10000	1	1	0	1	1
7,5	1	2	2	1	8	12000	1	1	0	1	1
6,5	1	2	3	1	8,5	14000	1	1	0	1	1
7	1	2	2	1	8	10000	1	1	0	1	1
7,5	1	3	4	1	7	8000	1	1	0	1	1
7	1	2	2	1	8,5	12000	1	1	0	1	1
7	1	2	2	1	7,5	12000	1	1	0	1	1
6,5	1	3	3	1	8	10000	1	1	0	1	1
6,5	1	2	2	1	8	10000	1	1	1	1	1

7	2	2	2	1	7,5	16000	1	1	0	1	1
2	3	4	1	0	2	1000	0	2	0	1	0
3	3	4	1	0	3	500	0	2	0	1	0
5	2	2	1	0	6	2000	0	1	0	1	0

6.csv

A	B	C	D	E	F	G	H	I	J	K	L
7	1	0,4	2	0	7,5	8338	1	1	0	1	1
7	1	0,5	1	0	7	6000	0	1	0	1	1
6,5	2	0,5	2	0	7,5	6777	0	1	1	1	1
6,5	1	1	2	0	7	6354	1	1	0	1	1
6	2	0,5	1	0	6	4050	0	1	1	1	1
7	2	0,5	1	0	7	3243	0	1	0	1	1
6	2	1	1	0	7	5336	1	1	0	1	1
7,5	2	1	2	0	7,5	11645	0	1	0	1	1
6,5	2	1	2	0	7	10350	0	2	0	1	1
7	1	0,5	1	0	7	7156	0	1	0	1	1
6	2	1	2	0	6,5	7011	1	2	0	1	1
7	1	1	2	0	7	8312	0	1	0	1	1
8	1	1,5	2	0	6,5	11410	0	1	1	1	1
7	1	1	1	0	7	10810	0	1	0	1	1
6,5	2	0,5	2	0	6,5	7030	1	1	0	1	1
6	3	1	3	0	6	8010	0	2	0	1	1
7	2	1	2	0	6,5	10320	0	1	0	1	1
7	1	1,5	3	0	6,5	9140	1	1	1	1	1

7	3	1	2	1	7	10112	0	1	0	1	1
6,75	2	0,5	1	0	6	9841	0	2	0	1	1
6,75	1	1	1	1	7,5	11310	0	1	0	1	1
6,75	2	0,5	1	0	6,5	9314	1	1	0	1	1
6,75	1	1	2	0	7	7810	0	1	1	1	1
6,75	1	1,5	2	0	7	8520	0	1	0	1	1
6,75	1	1	2	1	6,5	10250	0	1	1	1	1

7.csv

A	B	C	D	E	F	G	H	I	J	K	L
8,5	1,5	1,5	2	0	7	8167	0	1	0	0,3	1
8	2,5	1,5	2,5	1	6,5	8167	0	1,5	0	0,3	1
9	1,5	1,2	2	0	7	8167	1	1	0	0,2	1
9	1	1,5	2	0	7	8167	0	1	0	0	0
8	2	1	4	0	7	7000	0	1	1	0	1
7,5	2	0,8	2	0	7,5	7500	0	1	0	0,1	0
7	1,5	1	1,5	0	7	8167	0	1,5	0	0,1	1
8	1,5	1,5	2	0	6,5	8167	0	1	0	0,1	0
8	2,5	0,75	2,5	0	7	13000	0	1	1	0	1
8,5	2,5	0,75	1,5	0	7	11000	1	1	0	0	1
9,5	1,5	1	2	0	7	8167	0	1	1	0,25	0
9	1,5	1	2,5	1	7	8167	0	1	0	0,25	0
8,5	2	1,5	1	1	7,5	8167	0	1	0	0	0
8,5	2,5	1	1,5	1	7	8167	0	1,5	1	0,2	1
9	1	1,3	1	0	7	8167	0	1	0	0,25	0

9	1,5	1,5	1	1	7,5	3500	0	1	0	0,2	1
8,5	1,5	1	3	0	7,5	7000	0	1	1	0,25	1
9	1,5	1	2,5	0	8	8167	0	1	1	0,2	1

8.csv

A	B	C	D	E	F	G	H	I	J	K	L
5	4	1,5	4,5	0	10	2198	0	3	0	0,75	0
3	4	0,5	3,5	0	7,5	4896	0	3,5	0	0,5	0
7,5	2,5	1	3,5	0	7,5	4896	0	2	1	0,5	1
5,5	2	1,5	3	1	10,5	1234	0	2	0	0,5	0
4,5	3	0,8	3	0	10,5	556	0	2,5	0	0,5	0
6,5	2	1,5	2	1	9	5829	1	2	1	0,4	1
5	3	0,4	2	0	10	2495	0	2,5	0	0,3	0
6	3	0,5	2	1	9	5327	0	3	1	0,3	0
4,5	2,5	0,3	4	0	9,5	1865	0	3,5	0	0,2	0
3,5	3,5	0,2	4	0	8	4814	0	3	0	0,2	0
4,5	3	0,6	3	1	10,5	1522	0	3	0	0,2	0
6,5	2,5	0,4	2,5	0	8,5	3188	0	2,5	1	0,3	1
6	2	1	2	1	9,5	2627	0	2	0	0,2	1
5,5	2,5	0,4	2	0	9,5	3346	0	2,5	1	0,1	1
4	3	1	3	0	5	3191	0	2,5	0	0,2	0
5	2,5	0,6	2,5	1	7,5	2685	0	2	0	0,1	1
5,5	2,5	1,3	3	0	8	8032	0	2,5	0	0,3	0
4,5	3	1,2	2,5	1	10,5	2583	0	3,5	0	0,3	0
5,5	2,5	1,3	3,5	1	7,5	4405	0	2,5	1	0,2	1

4	3,5	0,8	3,5	1	9,5	1508	0	3,5	0	0,4	0
6,5	2,5	1,3	2,5	0	10	13558	1	3	0	0,2	0
5	3	0,8	3	1	8,5	2427	0	3	0	0,3	0
5,5	2,5	1	3,5	1	10,5	1573	0	3	0	0,3	0
5	3	0,6	3	1	8,5	1238	0	3	1	0,4	0
5	3,5	0,8	3	1	8,5	3343	0	3	0	0,3	0
4,5	3	0,8	3	0	10	3507	0	3	1	0,5	0
4	2,5	0,7	3	0	8,5	1109	0	3	0	0,4	0
5	3	0,8	3,5	1	6,5	3456	0	3	0	0,4	1
6,5	2,5	1,3	1,5	1	9,5	3901	0	3	0	0,3	0
5,5	3	1,2	2	1	10	2470	0	3	0	0,2	0
6	3	1,2	2	0	8	3954	0	3	0	0,4	0
4	3	0,7	2,5	0	10	1108	0	3,5	1	0,3	0
6	2,5	1,1	2	1	10,5	2466	0	3,5	0	0,3	1

9.csv

A	B	C	D	E	F	G	H	I	J	K	L
8	1	1,5	4	0	6	0	1	2	0	4	1
7	1	1,5	4	0	8,5	0	1	1	0	4	1
6	2	1,5	5	0	8	0	1	2	1	5	0
7,5	2	1,5	3	1	7	0	1	1	1	3	1
7,5	2	1,5	5	1	7	0	1	1	1	3	1
7,5	2	1,5	5	0	7	0	0	1	1	3	1
8	1	1	2	1	6	0	1	1	0	4	1
8	1	1	2	1	7	0	0	1	1	3	0

6	3	1	2	0	9	0	1	2	0	3	1
7	2	2	1	1	6	0	1	2	0	3	0
6	2	2	2	1	7	0	0	2	1	5	0
7	1	2	4	1	7	0	1	1	1	3	1
7	1	1,5	1	0	9	0	0	1	1	1	1
6	2	2	1	1	6	0	0	2	0	3	0
8	2	2	2	1	7	0	1	1	0	3	0
8	1	1	3	1	7	0	0	1	0	2	1
8	1	1	3	1	6,5	0	1	1	1	2	1
8	1	1,5	3	1	7	0	1	1	1	2	1
8	2	1	2	0	7	0	0	1	1	2	1
8	2	1	3	0	8	0	0	1	1	2	1
8	2	1	3	0	7	0	0	2	1	2	1

10.csv

A	B	C	D	E	F	G	H	I	J	K	L
8	1	1	2	1	10	5582	0	1	0	2	0
9	1	1	2	1	10	5688	1	1	0	2	0
6	4	2	3	1	7	220	0	2	0	2	1
7	2	1	2	1	10	2929	0	2	1	2	0
7	2	1	3	1	9	250	0	2	0	2	1
7	2	1	2	0	12	3748	0	2	1	2	0
6	3	1	3	1	7	5345	0	2	1	2	1
9	1	1	2	1	8	7061	0	1	1	3	1
10	1	1	1	1	10	7614	1	1	1	3	1

8	1	1	2	1	7	3972	1	1	0	2	0
6	3	1	1	1	12	2806	0	1	1	2	1
6	2	1	2	1	10	420	0	2	1	2	1
8	2	2	4	1	10	6158	1	2	1	2	1
8	1	2	4	1	10	14284	0	1	1	2	1
8	1	2	1	1	10	6937	1	3	1	2	0
6	2	2	3	1	10	298	0	4	0	2	0
8	2	2	1	1	10	3662	1	2	0	2	0
7	3	1	2	0	10	6769	0	2	1	2	0
8	2	2	4	0	10	4564	0	1	1	2	1
7	3	2	1	1	10	5650	0	3	1	2	1

Appendix C: Mood correlations overview

Variable	Variable	Correlation	Insight
mood	stress	positive	Higher stress levels seem to also accompany a better mood for you: you thrive under stress.
mood	stress	negative	Stress is negatively associated with the way you feel. When you are under a lot of stress, your mood suffers.
mood	water	positive	The more water you drink, the better you generally feel. Try to drink at least 2.7 litres of water per day.
mood	water	negative	Drinking less water seems to be related to a better mood. Try to drink at least 2.7 litres of water per day.
mood	sugar	positive	Eating more sugar is associated with a better mood for you. Eating too much sugar can be unhealthy, however.
mood	sugar	negative	A lower sugar intake seems to be related with an elevated mood. Try to cut down on sugar in your diet.
mood	fruit	positive	Eating more fruit seems to have a positive effect on your mood. Try to eat at least two servings of fruit per day.
mood	sleep	positive	More sleep seems to affect your mood in a positive way. In contrast, less sleep is associated with a worse mood.
mood	sleep	negative	You seem to feel the best on days when you get the least amount of sleep.
mood	steps	positive	The more steps you set per day, the better you seem to feel. For an optimal mood, aim for at least 7,000 steps.
mood	steps	negative	You generally feel better on days where you do not set a lot of steps.
mood	exercise	positive	For your health data, exercise seems to be associated with a positive mood.
mood	exercise	negative	You generally feel worse on days where you had more exercise.
mood	loneliness	negative	The lonelier you are, the worse you feel. In contrast, if your mood is better when you are not feeling lonely.
mood	alcohol	positive	For your health data, drinking more alcohol is related to a better mood.
mood	alcohol	negative	Drinking more alcohol seems to be associated with a more negative mood.
mood	social media	positive	You generally feel better when you spend more time on social media.
mood	social media	negative	Social media may have a negative effect on the way you feel.
mood	social activity	positive	The more people you meet, the better you seem to feel. Reach out to friends or family for an optimal mood.
mood	social activity	negative	For your health data, meeting more people is associated with a worse mood.

Appendix D: Usability test results

A: How much effort did it take to fill in your data?

B: How easy was it to interpret and understand your insights?

C: How attractive was the design of the user interface?

D: To what extent was the user interface intuitive?

E: What is your overall score for Optimood?

F: Can you identify with the insights that were generated by Optimood?

G: Do the insights deliver any new or valuable information?

usability-test.csv

A	B	C	D	E	F	G
5	3	4	5	4	4	3
5	3	3	3	3.5	4	4
4	5	5	5	4.5	5	4
4	2	5	3	3.5	2	2
5	5	5	5	4.5	4	4
4	3	5	5	3.5	4	3
5	3	5	5	4	2	2
5	5	5	4	5	5	4
3	5	5	4	4	3	3
4	5	4	5	4.5	4	3