BACHELOR THESIS

ARTIFICIAL INTELLIGENCE

Radboud University

Identifying Patients at Risk for Suicidal Ideation and Key Factors Responsible

by Means of a Self-Explaining Neural Network

Author

Lukas L. Deis
s4588827
L.Deis@student.ru.nl
DeisLukas@gmail.com

Supervisor
Radboud University

Dr. Pim Haselager
Department of Artificial Intelligence
& Donders Institute
w.haselager@donders.ru.nl

Supervisors
Radboudumc

Dr. Rose Collard
Department of Psychiatry
rose.collard@radboudumc.nl

Dr. Peter Mulders
Department of Psychiatry
& Donders Institute
peter.cr.mulders@radboudumc.nl

Radboudumc
university medical center

10 February 2021

Abstract

Suicide is a major cause of death in all of Europe, and it is on the rise. Worldwide, suicide is the second common cause of death in the age group 10 to 30 years (Bachman, 2018). A self explaining neural network (SENN) was trained to predict if a person suffers from suicidal ideation and state which factors were important in that prediction. The explainability the SENN offers is a major benefit to clinical applications, as not only being able to detect suicidal ideation but also giving insight into the underlying issues is key to better treatment (WHO, 2019b). For this research data from the MIND-SET study was used, which is a study by the Radboudumc. It includes 705 participants, of which 574 suffer from common psychiatric disorders and 131 are healthy controls. While the dataset contains more, only demographics and answers to questions designed to evaluate different mental health issues were included. The best performing model had an accuracy of 85.3% on the test set with a sensitivity of 79.1% and a specificity of 89.1%; the PPV was approximately 81.538% and the NPV was 87.5%. The most important risk factors are from two questionnaires. One is the Outcome Questionnaire, designed to capture quality of life. Its total score as well as subscores for interpersonal relations and symptomatic distress are consistently relevant predictors of suicidal ideation. The second is a high Inventory of Depressive Symptomatology total score. Some factors seem to significantly reduce the risk, too, such as SF_PG, a score describing a good mental health. Surprisingly most other relevant risk reducing factors stem from the AQ, a questionnaire measuring autism characteristics.

Contents

1 Related work

The high performance that can be achieved with neural networks could be an important step to providing an automated solution to this issue. Yet neural networks are rarely used in this field (Shatte et al., 2019).

However, suicidal ideation has successfully been predicted on the basis of twitter posts and text mining before. In that study a sensitivity and specificity around 80% were achieved, but the important factors could not be identified (Roy et al., 2020) and the applicability beyond twitter is unclear.

There is another investigation into how precisely a number of mental health issues can be predicted on the basis of textual medical history at a psychiatrist. Again the important factors were not identified (Tran & Kavuluru, 2017).

To identify people that are at risk of suicidal ideation, especially those that currently would not get treatment, it is necessary to create a system that works with more general inputs than twitter-posts or medical history. That is only possible with insight into the underlying concepts.

Those underlying concepts have been investigated before. Using Bayesian networks, risk factors for suicidal ideation have been identified in a dataset of depressed people (Galiatsatos et al., 2015). Within that dataset the researchers achieved an accuracy of 83.51%. The identified, most relevant factors were mood depression, loss of interest or pleasure, unworthiness or guilt, living in a city and concentration in thoughts.

While Bayesian networks offer good explainability their performance is often sub-par to neural networks and deep learning approaches. Also, using a bigger dataset, closer to the general society, more insight could be gained into the important factors. If the potentially higher performance of the neural network translates to the performance of the explanations better explanations might be found.

2 Introduction

2.1 Motivation

Suicide is a major cause of death all over the world (WHO, 2019a), being the second common cause of death in the age group of 10 to 30 years (Bachman, 2018). That issue is even more pronounced in the European Union (Eurostat, 2020a).

Being able to detect suicidal ideation and understanding important factors is key in treating it better (WHO, 2019b). While some risk factors for suicidal ideation are known, they vary between countries (Nock et al., 2008) and diagnostic tools based on known risk factors are not sufficiently accurate (Runeson et al., 2017).

2.2 Goal

This project aims to expand our understanding of how suicidal ideation can be predicted by identifying risk-factors and taking the first steps to an automated solution. The general application of automated solutions, specifically machine learning, has been shown to carry a wide range of benefits to diagnosis, treatment and research of mental health (Shatte et al., 2019).

Suicide experiences a treatment gap of around 50%, meaning that only half the people that commit suicide get treated before they do (Bruffaerts et al., 2011). Thus, in this case specifically an automated solution could be used to identify people in the general population that may be at risk for suicide.

2.3 Explainability and Efficacy

While methods from the field of explainable AI are rarely used, in this case they are crucial.

For one, explainability provides a major benefit to clinical applications. Because suicidal ideation can not be treated directly, the root causes and underlying disorders have to be identified such that they can be treated instead. The explanations could provide the necessary insight.

For another, the reasons that the network expects a person to suffer from suicidal ideation could give insight into general risk factors for suicide, helping to identify and reach people at risk.

Naturally, it is desirable for the found factors to be applicable to a broad range of people. That is why a dataset including healthy controls has been chosen. Furthermore, mostly consisting of answers to questions designed to evaluate different mental health issues has been chosen.

2.4 Approach

The following research question has been formulated:

How accurately can a neural network that was trained on data from common psychological tests identify subjects that suffer from suicidal ideation, and what are the most important factors ?

In order to answer this question, 10 neural networks are trained to predict suicidal ideation from questionnaires. From the trained networks the most accurate

two are selected. Their explanations are analyzed to find out what the most important predictors are and the performance of those networks is noted.

A specific type of neural network, called SENN is used to provide explanations for the predictions made (see: section 4.2 Self Explaining Neural Networks).

2.5 Expectations

The explanations given by the SENNs are expected to be focussed on quality of life, general mental health as well as age and gender, because they have been identified as risk factors in the Netherlands, where the data was collected (GGD, 2017). Predicting the model's accuracy turned out to be difficult, but it is expected to perform above the baseline model which is explained in section 5.3 Baseline Model.

2.6 Overview

In the following chapters it will become clear how a SENN was used to identify people that experience suicidal ideation as well as factors relevant to that prediction. First, some recurring abbreviations will be explained, followed by preliminary knowledge which this research builds upon. Second an explanation of what data was used and how it was treated will be given. That is followed by an explanation of used measurements, including the baseline model, and a description of the utilized model; there is mention of the stability and ethical implications of the model. Afterwards the results are listed and conclusions are drawn before the discussion. The paper ends with a look at limitations, future research and possible applications.

3 Used Abbreviations

3.1 Included Questionnaires

OQ	Quality of Life, meant to measure the general state of one's mental health and life
IDS	Inventory of Depressive Symptomatology, meant to assess the severity of depression
SD	Demographics, meant to capture general demographics and basic background
ASI	Addiction Severity Index, meant to help assess the severity of drug use
CAARS	Conners' Adult ADHD Rating Scales, meant to assess the presence and severity of symptoms related to ADHD
AQ50	Autism spectrum Quotient, meant to quantify the expression of ASD (autism spectrum disorder) related traits
SF20	20 item Short Form, a general, short questionnaire to get a fast and broad idea of overall health
PID	Personal Inventory DSM-5, meant to assess different personality traits
NEMESIS	Netherlands Mental Health Survey and Incidence Study, designed to get an idea about mental health disorders over time
MATE	Measurements in the Addictions for Triage and Evaluation, meant to get insight into substance abuse

3.2 Important Questions

OQ Total	quality of life
IDS Total	severity of depression
SF20 question 19	being in especially good health

IDS question 2	having issues sleeping through
IDS question 6	being irritable
IDS question 10	feeling sad
CAARS question 2	being always busy as if driven by a motor
CAARS question 8	(still) throwing temper tantrums
OQ question 14	working or studying too much
OQ question 18	being lonely
OQ question 34	having muscle ache
OQ question 36	being nervous often
OQ IR	issues maintaining relationships with others
OQ IR Av	issues maintaining relationships with others, in relation to the average
OQ SD	general symptomatic distress
OQ SD Av	general symptomatic distress in relation to the average
NEMESIS question 10a1	having experienced sexual trauma
SF20 PG	psychiatric health
SF20 PG Scale	psychiatric health point scale
SF20 question 8	having had physical pain recently
SF20 question 16	feeling extremely sad
SF20 question 18	feeling as healthy as everybody else
AQ50 question 5	noticing sounds that others do not
AQ50 question 18	it is hard for others to throw in a word while you talk
AQ50 question 25	being able to deal with a broken routine
AQ50 question 35	usually being the last one to get a joke
AQ50 question 40	having enjoyed games that involve pretending as a kid
AQ50 question 45	having trouble to understand the goals of others
AQ50 question 50	finds it easy to play games that involve pretending with kids

4 Preliminaries

4.1 Neural Networks

4.2 Self Explaining Neural Networks

Self explaining neural networks (SENNs) propose a general way of utilizing deep learning while maintaining a high level of explainability. They achieve that by expanding on the concept of a linear regression model (LRM). LRMs are known to be quite easily interpretable as they simplify data by assuming linear separability. Their output can be visualized as a straight line that divides data points into two classes. However, their performance is limited as very few relations are truly linear.

To tackle that, they modify the way the LRM learns: The weights of a SENN are allowed to depend on the input, allowing the model to picture relations that are not linear without changing the way the model can be viewed and visualized, assuming a certain input. While there are many ways to determine the optimal weights for every input, deep learning is chosen as it will allow us to train the whole model on data without further intervention or expert knowledge.

SENNs, contrary to most other self explaining models, do not suffer significant performance impacts compared to a regular deep neural network (Melis and Jaakkola, 2018). It was considered that SENNs have recently been criticized for offering unreliable explanations (ZhengIn et al., 2019). That critique is based on the way that important concepts are determined in this technique. It is also unclear how well

understandable, even if perfectly accurate, the concepts would be in this case. That, in combination with the previously mentioned critique led to the decision to avoid that part of the proposed model. This part of the SENN will be replaced with an identity mapping of the input such that every parameter is treated as a separate concept.

4.3 MIND-SET

The data used in this research were already collected by the Radboudumcs Department of Psychiatry and the Donders Institute. It is made available for research purposes as MIND-SET (Measuring Integrated Novel Dimensions in Neurodevelopmental and Stress-related Mental Disorders). A list of other publications that utilize the same dataset is included as 13.1 MIND-SET Publications.

The purpose of MIND-SET is understanding underlying common and unique mechanisms of psychiatric comorbidity (Collard, 2021). The data-collection for MIND-SET was planned to run from June 2016 to the end of 2020. As this research was begun in September 2020 the collection was not completed yet. Instead of the planned total of 800 participants, the utilized dataset contains a total of 705 participants (aged between 18 and 76).

There are two groups of participants. Patients that are diagnosed with one or more of the following disorders: ADHD, autism spectrum disorder, mood disorder, anxiety disorder and addiction disorder (Radboudumc, n.d.). Next to the patients, a psychiatrically healthy control group is included (n=131).

While the dataset includes a variety of information, for this project only two types of data will be used: Demographics and neuropsychological data, including diagnoses of the disorders mentioned above, but mostly common psychological tests used at psychological treatment facilities. The utilized tests are: OQ, SD, ASI, IDS, CAARS, AQ, SF, PID, NEMESIS, MATE; their purpose is briefly explained in section 3 Used Abbreviations.

By limiting the utilized information to this set, all non-structured data is excluded. That enabled the uniform implementation of the beneficial adaptations to the model, explained in section 5.4.2 Specific adaptations to the model.

Within those tests subjects evaluate themselves how they feel and how much they agree with statements about themselves as well. The precise format of the questions varies per questionnaire. Additionally, results of simple tasks that measure characteristics like attentiveness are included.

4.4 Utilized measurements

All utilized measurements are based on the same, four variables that are collected while testing the model.

TP: The number of samples that were correctly identified as positive cases.

TN: The number of samples that were correctly identified as negative cases.

FP: The number of samples that were falsely identified as positive cases.

FN: The number of samples that were falsely identified as negative cases.

4.4.1 Accuracy

Accuracy gives an idea about the performance of a model on all aspects, encoding how many of all made predictions were correct. It is calculated with the following formula: TP + TN / (TP + FP + TN + FN)

4.4.2 Sensitivity and Specificity

The sensitivity gives an idea of how well the model identifies positive cases. It is the ratio of correctly identified positive cases to all positive cases, which can be expressed in the formula: TP / (TP + FN).

The specificity, in turn, gives an idea of how well the model identifies negative cases. It is the ratio of correctly identified negative cases to all negative cases, which can be expressed in the formula: TN / (TN + FP).

4.4.3 Positive and Negative Predictive Value

In the medical field, models are often judged by their positive and negative predictive values (PPV and NPV respectively). Those two measurements essentially encode the trust that one should put into a positive or negative diagnosis by the model. For easy comparison within the medical field, the PPV and NPV of the most performant model will be included.

The PPV or “precision” gives an idea of the chance that a disease is present given a positive prediction. Put differently, how many of the cases that are identified as positive are indeed positive cases, using the formula: TP / (TP + FP).

The NPV or “recall” gives an idea of the chance that a disease is not present given a negative prediction. Put differently, how many of the cases that are identified as negative are indeed negative cases, using the formula: TN / (TN + FN).

4.5 Suicidal Ideation

There are different definitions of suicidal ideation (Harmer et al., 2020) and unifying them is beyond the scope of this research. What can be agreed upon is, that while the act of suicide refers to actively ending ones own life, not everyone who considers doing so goes through with it. Furthermore in the majority of cases persons think of suicide and death before going as far as committing it (Harmer et al., 2020).

For the purposes of this research the following was decided:

The thought of ending ones life, which may or may not go as far as planning the suicide, is referred to as suicidal ideation. Suicidal ideation is treated as a symptom of numerous psychiatric disorders, not a separate one.

5 Method

5.1 The Data

5.1.1 Participants

This research utilizes data from the ongoing MIND-SET study (Measuring Integrated Novel Dimensions in Neurodevelopmental and Stress-related Mental Disorders) which is executed at the Radboudumcs Department of Psychiatry and the Donders Institute. The inclusion criteria to that study are explained in more

detail in section: 4.3 MIND-SET. Generally, it can be said that most participants in the study have a psychiatric condition. Also, a psychiatrically healthy control group, consisting of 131 is included.

The used dataset contains 705 participants, of which 619 filled in the target question for training. The intuitive decision was that participants who did not answer the target question could not be classified and their data could not be used. Thus, the plan was to ignore all data collected from those participants.

However, those patients were not actually removed but counted as not experiencing suicidal ideation. While there is no hard evidence that these 86 participants actually did not experience suicidal ideation, they were judged to have a low risk to do so by the researchers collecting the data, otherwise they would have been asked to fill in the question. Possible effects of this are assessed in section 8 Discussion. The upside of this is, that the size of the dataset was not reduced by 12.2%.

5.1.2 Variables

The participants filled in a maximum of 451 items concerning mental health as well as some demographics, for details on this please refer to section 4.3 MIND-SET.

5.1.2.1 The Target Question

The risk for suicidal ideation will be extracted from question eight of the OQ, asking for it directly. The question (translated from Dutch) reads:

“I am considering to end my life”; the possible answers (also translated) are: “Never”, “Rarely”, “Sometimes”, “Often”, “Almost all the time”.
Any answer other than “Never” was counted as experiencing suicidal ideation.

A similar question was asked in the IDS (question 18), where the participants have to say which statement they agree with the most. The options are as follows (translated from Dutch): “I do not think about suicide or death.”, “I feel like my life is empty and question if it is still worth the effort”, “Several times a week I think about suicide or death” and “Several times a day I think about suicide or death; or I made plans to commit suicide”.

The question from the OQ was chosen over the question from the IDS for its stricter focus on suicidal ideation; as thinking about suicide and death is not necessarily the same as considering it.

For this research question 8 from the OQ was chosen as it is more specific. The answers to IDS question 18 were entirely removed from the used data because they are still very similar to the target question, making it too easy for the model to predict the data, removing the need to learn the other present relations.

5.1.3 Data augmentation

Due to the relatively small size of the dataset, the performance of the network is expected to be suboptimal. Although that was planned, due to time constraints the data was not augmented to enlarge the dataset.

5.1.4 Splitting the Data

To make sure that the model was not trained and tested on the same data, it was split three ways:

- 67.5% was used for training → 475 participants
- 07.5% was used for validation → 53 participants
- 25.0% was used for testing → 177 participants

5.2 Assessments and Measures

Using accuracy as the main measurement for the performance of the network seems the most understandable. It is the one singular measurement that gives an idea about the performance of the model on all aspects, which is why it is focussed on in the research question. The sigmoid output of the model (range (0,1)) is binarized into positive predictions (experiences SI) and negative predictions (does not experience SI). The threshold for that process is automatically determined such that the accuracy is maximized. As the dataset is not balanced, a relatively high accuracy of 64.4% can be achieved by bluntly predicting the same value all the time (see: section 5.3 Baseline Model). To prevent relying on a possibly deceptively high accuracy, the sensitivity and specificity have been included as measurements as well. In a medical context, measurements often called precision and recall in an AI context, are usually called positive predictive value (PPV) and negative predictive value (NPV) respectively. People with a medical background are used to judging theperformance of a model by

the PPV and NPV. Therefore, those two measurements have been included for the most accurate model to enable a quick comparison for people with that background.

5.3 Baseline Model

To put the accuracy of the model into perspective it is important to have a baseline. Only that way it can even be said if the model is better than chance or not. Given that the prevalence of SI in this dataset is 35.6% a model could always predict that a patient is not suffering from suicidal ideation and be 64.4% accurate.
While that model would have a specificity of 100% it would have a sensitivity of 0%, which is something that should be avoided for the created model.

5.4 Model

To identify the relevant factors any prediction the network makes needs to be backed up by reasoning. Such an output is also advantageous for any application in treatment of suicidal ideation, as suicidal ideation is not treated itself usually. Rather the root causes have to be identified such that they can be treated. The reasons that the network expects a person to suffer from suicidal ideation could give insight into these underlying issues, simplifying the treatment. This limits the choice of model significantly as common, performant networks are often hard to understand due to their black-box nature.

5.4.1 Utilizing a Self-Explaining Neural Network (SENN)

SENNs are a special kind of neural network designed by Melis and Jaakkola (2018). While the widespread perception is that explainability and performance of neural networks oppose each other, it was shown by this network that this is not the case.

The advantage of a SENN over a regular deep neural network is, that it will explain decisions based on training data. For example, it could not only say that a patient is likely to experience suicidal ideation but also what that specific combination of answers was relevant to that decision. This allows experts to reason about a diagnosis and formulate specific questions to look into the SENNs judgement. As the weights within the SENN are determined by deep learning, the performance can be expected to be comparable or relatively close to other deep learning approaches (Melis and Jaakkola, 2018). For more information on SENNs in general, please refer to section 4.2 Self Explaining Neural Networks.

5.4.2 Specific adaptations to the model

One critical aspect of the proposed model of SENNs was altered: the way in which the concepts are learned. As the data is categorical, the whole part of the model that is responsible to learn the categories was replaced by the preprocessed input data. This has been done before without stating the reasons explicitly (Hussain et al., 2020), but there are good arguments to follow this example. For one there is critique as to how reliable the concepts are if they are determined by a neural

network (ZhengIn et al., 2019). This is explained in more detail in section 4.2 Self Explaining Neural Networks.

For another, any complex relationship between the different input variables would be hard to interpret. Now a relevance score is calculated for all the singular inputs, allowing us to identify singular, general, risk-factors. If instead the relevance values would be calculated for specific combinations of these factors, those prototypes would have to be distinguishable. That is feasible in certain cases such as optical character recognition, where those 451 parameters would fit into a picture of 21 x 22 pixels. However, finding which parameters differ significantly between different prototypes, each of which is represented by a list of more or less differing values is not trivial.

To replicate the results of such a study, one would have to have access to all the input represented in such a prototype. Now, the singular identified variables can be tested and applied individually and immediately.

5.4.3 Implementation

5.4.3.1 General Implementation

The model is implemented in Python, using TensorFlow 2.0. The functional API was used to maintain as much overview as possible while providing the necessary flexibility. The full code with explanations can be found in the appendix.

5.4.3.2 Designed Output

The output of the network is a value between 0 and 1 as well as a relevance score for every input feature between -1 and 1. A negative value means that the variable is lowering the risk-score, while a positively valued parameter increases it; a value of 0 would mean that the variable is completely irrelevant.

To provide a classification, the output is binarized using a threshold of 0.5.

5.4.3.3 Structure

The SENN can be split into 3 sub-models:

1) Conceptizer
2) Parameterizer
3) Aggregator

In the following sections their separate Implementations will be explained.

5.4.3.4 Implementation Conceptizer

As discussed in section 5.4.2 Specific adaptations to the model, the Conceptizer is replaced by an identity mapping of the input.

5.4.3.5 Implementation Parameterizer

The original model does not specify what kind of deep learning is used for the Parameterizer. In this case the following structure was chosen:

The following block was repeated 4 times:

1) A fully connected layer with linear activation
2) A dropout layer
3) A fully connected layer with leaky ReLu activation

That block was followed by another fully connected layer with linear activation and a dropout layer.

The dropout layers were configured with a dropout rate of 10% and only active during training. They prevent the model from overfitting, by randomly (p=10%) setting any factor in the layer to 0. While the layers with linear activation are a simple way to construct a neural network, more complex relations in the data have to be captured by non-linearly activated units. Therefore, every other trainable layer instead utilizes a leaky ReLu activation function, where leaky ReLu was chosen over regular ReLu to prevent running into issues with vanishing gradients.

Because this is a binary classification problem, binary cross-entropy was selected as the loss-function to train this part of the neural network. Binary cross-entropy is the standard loss function in this case for several reasons, relating to it being a maximum likelihood estimator which comes with a number of benefits (Goodfellow et al., 2016).

While no other efforts have been taken to utilize that here, future research utilizing this model may profit from an additional property of binary cross-entropy: It usually provides good statistical calibration, enabling the models output to be interpreted as risk-scores instead of just predictions (oW_♦, 2019).

The result of this part of the network is used to generate the final classification, but also directly outputted. The vector of weights encodes the relevance of every concept, so in this specific case: every parameter.

5.4.3.6 Implementation Aggregator

The Aggregator multiplies every concept from the Conceptizer with the corresponding weight from the Parameterizer. The resulting values are then summed together before being passed through a sigmoid function. That keeps the networks' output readable by ensuring it is between 0 and 1. While using binary cross-entropy as the loss-function should provide a fairly good statistical calibration of the Parameterizers output and the value now is between 0 and 1, it was not tested if the model as a whole is properly statistically calibrated. Therefore, this value will not be treated as a probability.

5.4.3.7 Combination

In combination those different components look like this:

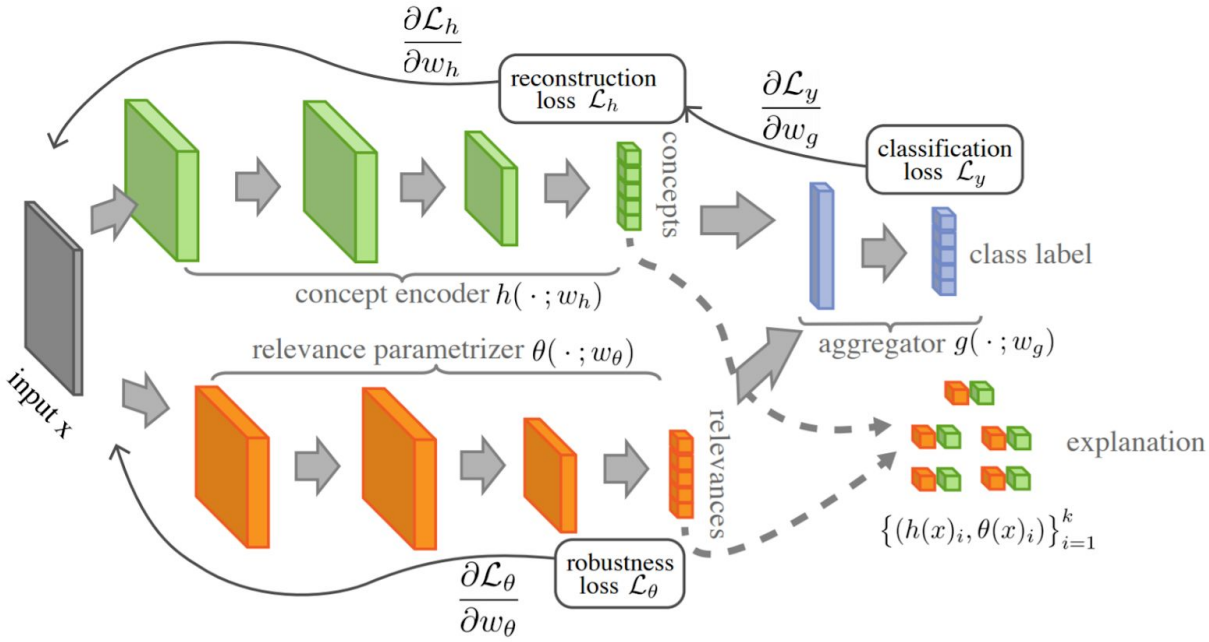

Figure 1: Graphical representation of the SENN model by Melis and Jaakkola, 2018.

The implemented version looks as follows:

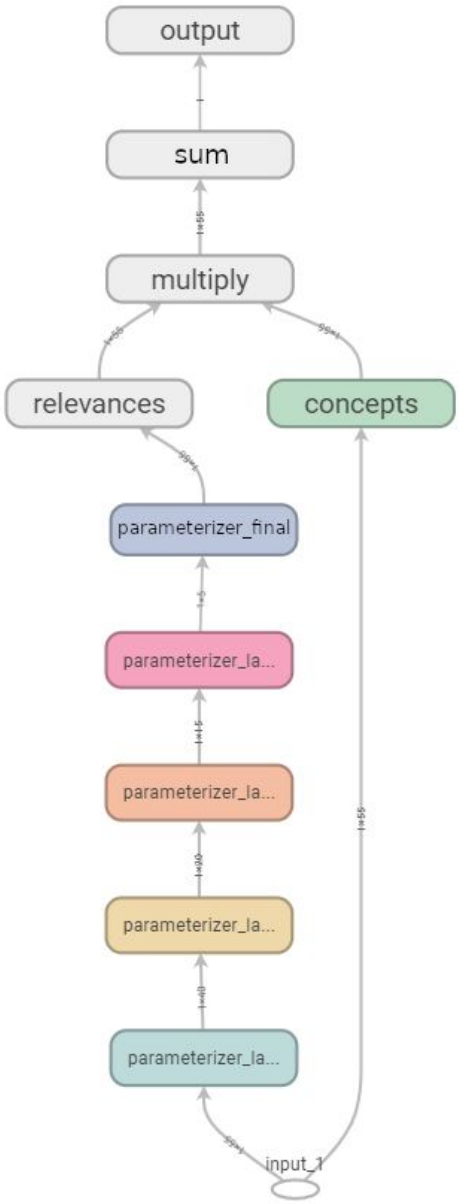

Figure 2: Graphical representation of the implemented model
The input at the bottom is passed through the 4 parameterizer-layers, which each include a linear, dropout and leaky-ReLu activated layer before being passed through the Parameterizers final layer, consisting of a linearly activated and a dropout layer. The resulting values are multiplied with the concepts, which are equal to the input, before being summed and passed through a sigmoid layer that produces the output.

5.4.3.8 Preprocessing

Before the data from the dataset can be worked with, it must be converted to numerical values. As that is a task that will be necessary for any possible follow-up research too it was decided that this should happen within the network.

The input gets passed into TensorFlows Preprocessing layers.
They are trained separately but on the same training data as the general network afterwards. There are different preprocessing layers for different types of input.

Dates are converted to the time that passed since that date, converting birth-dates to ages. All numerical inputs are normalized to make sure that something is not seen as more important just because it was measured on a bigger scale. Textual answers are translated by building up dictionaries of possible inputs (the choices of answers for questions) and outputting one-hot vectors encoding the answers; usually a question has 5 possible answers to capture the levels of severity.

5.4.4 Tackling Class Imbalance

In the data, the prevalence of SI is 35.6%. That class-imbalance could have led to a low sensitivity of the model. To avoid this, the minority class was oversampled so that there is an equal number of positive and negative cases in the training data. Due to that, some participants in the training data will be duplicates.

5.5 Stability

Splitting the data into separate sets for training, validating and testing as well as upsampling the minority class includes several random variables that can not easily be controlled. By design, the different datasets change randomly with every split. During upsampling, from all given samples of the minority class, samples are drawn at random. Given the varying inputs, all outputs such as performance and relevance scores have to be expected to vary between runs. The procedure will be repeated 10 times to get some indication of how stable the model is. Specifically, 5 different splits for the training sets will be utilized, each will be used to train 2 models. As a higher accuracy implies that a model captures the structure of the data better, it is expected that the best performing networks will also have the most accurate explanations. Thus, for further evaluation the two models with the highest accuracy will be chosen.

5.6 Ethical implications

Automated solutions are often seen as a way to counter human biases. However, human biases can be and usually are ingrained into datasets and designs made by humans. That way machine learning may learn and reinforce human bias. Potentially, that can lead to unfair treatment, which is a concern especially if that bias concerns protected groups[1](Schönberger, 2019). Especially in the medical field, bias and discrimination must be avoided. By validating the reasons the network gives for the prediction, any bias could be identified, but caution is still recommended. While

[1] Such as race, age, sexual orientation, ability, or belief.

the model should reveal any bias, the human user still needs to manually identify it as such, which can be less trivial than it sounds.

Consider the following:

The model could consistently predict people with a specific sexual orientation to suffer from suicidal ideation. Is that the result of discrimination against people of that group or a bias?

Without more data that explicitly contains information about to which protected group a person belongs it can not actually be tested.

No matter how many samples the network is trained on, if only relatively few of those samples belong to a specific group that may lead to issues. The factors relevant to that specific group may not be learned as accurately as the factors that are relevant to the majority of samples. If such factors exist that are specifically and only relevant to such a group, they would not be learned as well. While measures to tackle such imbalances exist and have been implemented in this research, to raise the sensitivity of the model in general (see section 5.4.4 Tackling Class Imbalance) they were not taken for any other case.

6 Results

6.1 The accuracy

All models trained on data from common psychological tests when identifying subjects that suffer from suicidal ideation lies between 75.1% and 85.3%.

The baseline model (which is explained in detail in section 5.3 Baseline Model) would have an accuracy of 64.4% while the best performing model had an accuracy of 85.3%; next to a sensitivity of 79.1% and a specificity of 89.1%. The sensitivity was expected to be lower than the specificity given the prevalence of 37.9%. That accuracy is not only higher than the baseline models but also what was achieved in the only comparable study that was found, 83.51% (Galiatsatos et al., 2015). The second best model (No.7) had an accuracy of 83.1%, providing a sensitivity of 79.4% and a specificity of 85.3%.

The model with the lowest accuracy was trained on the same train-test split as the best performing model. It provided an accuracy of 75.1% being on par with the best models specificity of 87%, but providing only 59.7% sensitivity. That test set had a prevalence of 38.4%.

6.2 Most important factors

6.2.1 Underlying structure

To look into what the most important factors are, an analysis was performed within two clusters. The clusters were based on if the network predicted a participant to experience suicidal ideation or not. The relevance scores associated with the parameters were compared, assuming that especially informative factors would be more relevant to one group than the other. However, it seems that while there are differences, they are very small; within a few percent points of the relevance score.

(Data not shown)

As explained in section 5.4.3.2 Designed output, how influential a parameter was is measured in a normalized relevance score. So the score with the highest influence is seen as 100%.
In the appendix all questions mentioned in the upcoming graphs are explained.

Usually, the questionnaires are evaluated by combining the points for different answers. The data included the raw, combined, point scores as well as scaled scores (corresponding to the points). Effectively, the same information was given twice in different formats. For simplicity parameters like this will be seen as equal in further analysis.

6.2.2 Risk raising factors

6.2.2.1 Best Performing model - No.9

For both groups, a high OQ total score seems to be the most predictive of high risk. That score is associated with a generally low quality of life.

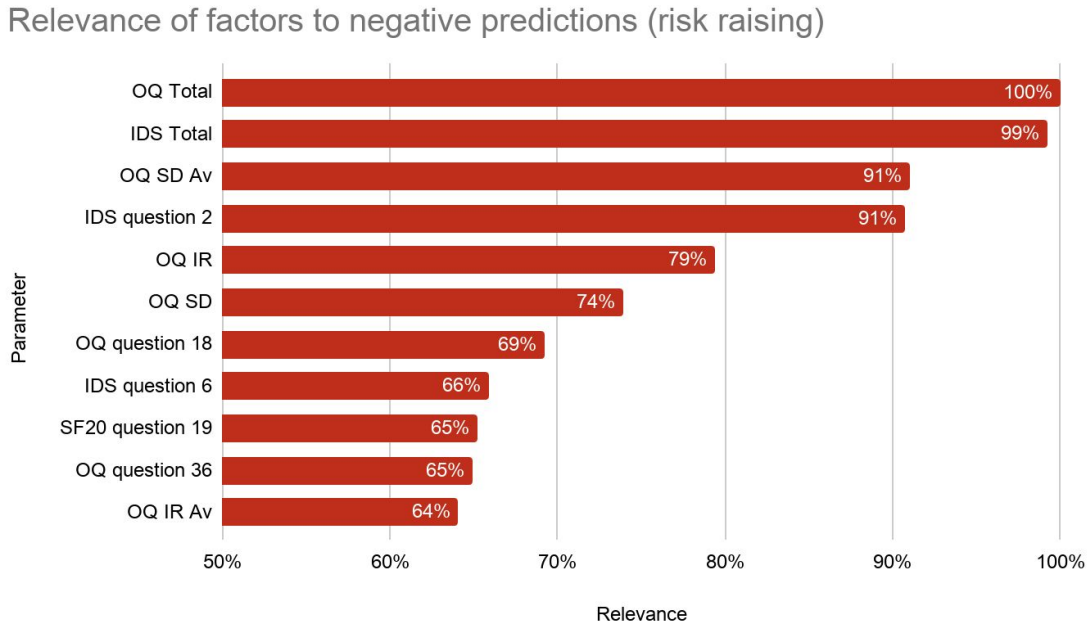

Figure 3: Important factors that raise the risk to negative predictions in model 9. They can be interpreted as (in that order): quality of life (100%), severity of depression (99%), general symptomatic distress in relation to the average (91%), having issues sleeping through (91%), issues maintaining relationships with others (79%), general symptomatic distress (74%), being lonely (69%), being irritable (66%), feeling like one is in especially good health (65%), being nervous often (65%), issues maintaining relationships with others in relation to the average (64%)

In generally high-scoring samples, some of the same parameters were used, but they were seen as more important. Some parameters however seem to only be considered for high-risk patients.

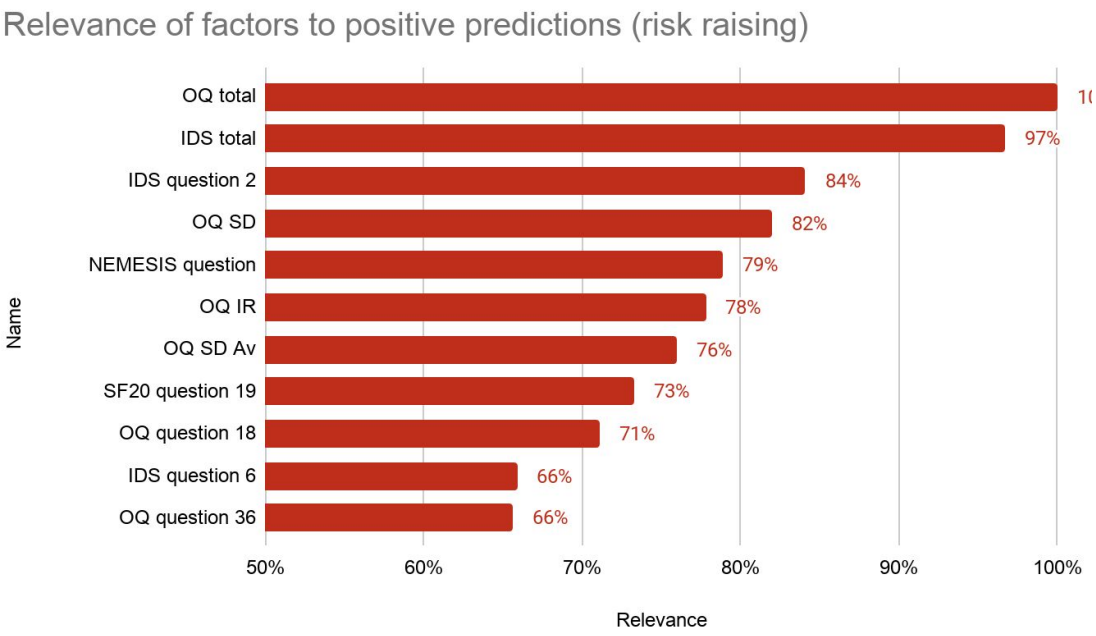

Figure 4: Important factors that raise the risk to positive predictions in model 9. They can be interpreted as (in that order): quality of life (100%), severity of depression (97%), having issues sleeping through (84%), general symptomatic distress (82%), having experienced sexual trauma (79%), issues maintaining relationships with others (78%), general symptomatic distress in relation to the average (76%), feeling like one is in especially good health (73%), being lonely (71%), being irritable (66%), being nervous often (66%)

Some of those most important parameters are from the OQ, which is commonly used in cases related to suicidal ideation. Its total score is, although mildly, influenced by a direct question for suicidal ideation that is removed from the input and used as the target. However, the choice was made to leave the OQ total score in the input data, as it is one of the most commonly used tests for this purpose and the final score is influenced by many factors. While parameters from the OQ dominate the picture for the next most important parameters there are some that are not part of it.

6.2.2.2 Second best performing model - No.7

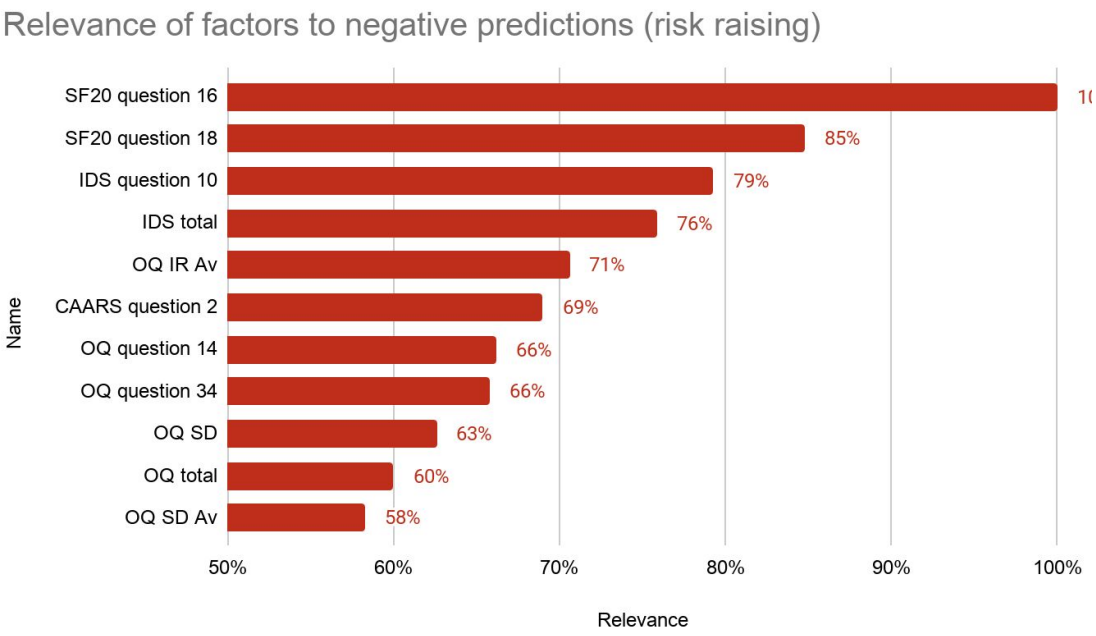

Figure 5: Important factors that raise the risk to negative predictions in model 7. They can be interpreted as (in that order): feeling extremely sad (100%), feeling as healthy as everybody else (85%), feeling sad (79%), severity of depression (76%), issues maintaining relationships with others in relation to the average (71%), being always busy as if driven by a motor (69%), working or studying too much (66%), having muscle ache (66%), general symptomatic distress (63%), quality of life (60%), general symptomatic distress in relation to the average (58%)

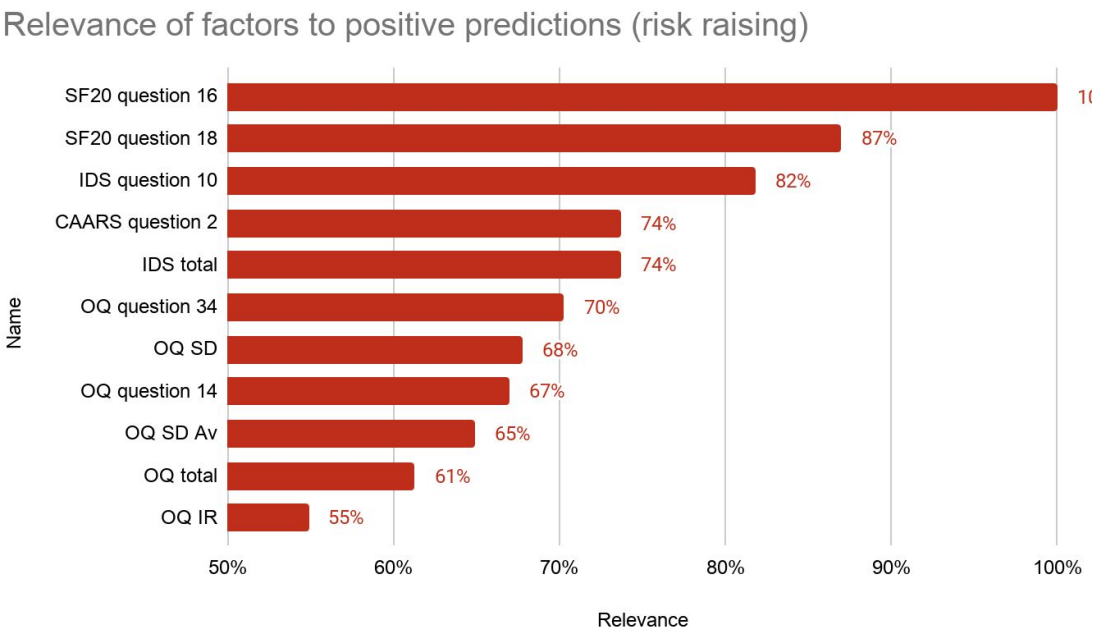

Figure 6: Important factors that raise the risk to positive predictions in model 7. They can be interpreted as (in that order): feeling extremely sad (100%), feeling as healthy as everybody else (87%), feeling sad (82%), being always busy as if driven by a motor (74%), severity of depression (74%), having muscle ache (70%), general symptomatic distress (68%), working or studying too much (67%), general symptomatic distress in relation to the average (65%), quality of life (61%), issues maintaining relationships with others (55%)

6.2.3 Risk reducing factors

Note: The values are negative which only indicates that they adversely affect the risk.

While the order and relevance scores do differ a bit, the parameters seem to be more or less the same again for both clusters in both models.

In the best performing model, within both the low-scoring and high-scoring cluster question 25 of the AQ50 seemed to be the factor reducing the risk most with a relevance score of -86.8% and -87.7% respectively. That question asks about how well the participant can deal with a broken routine, i.e. their flexibility.

6.2.3.1 Best Performing model - No.9

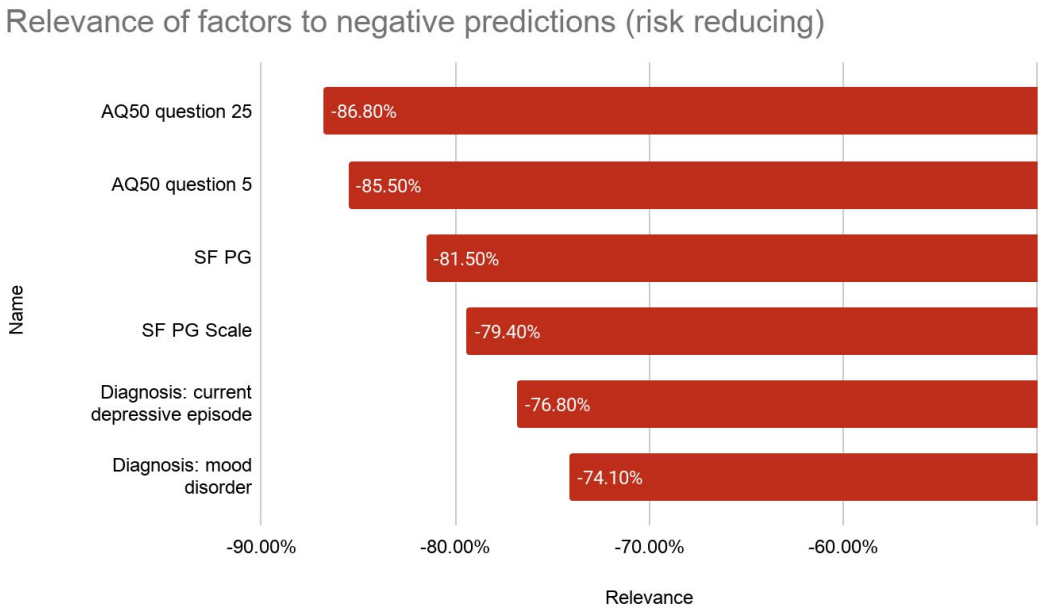

Figure 7: Important factors that reduce the risk to negative predictions in model 9. They can be interpreted as (in that order): being able to deal with a broken routine (-86.80%), noticing sounds that others do not (-85.50%), psychiatric health (-81.50%), psychiatric health point scale (-79.40%), currently experiencing a depressive episode (-76.80%), suffering from a mood disorder (-74.10%)

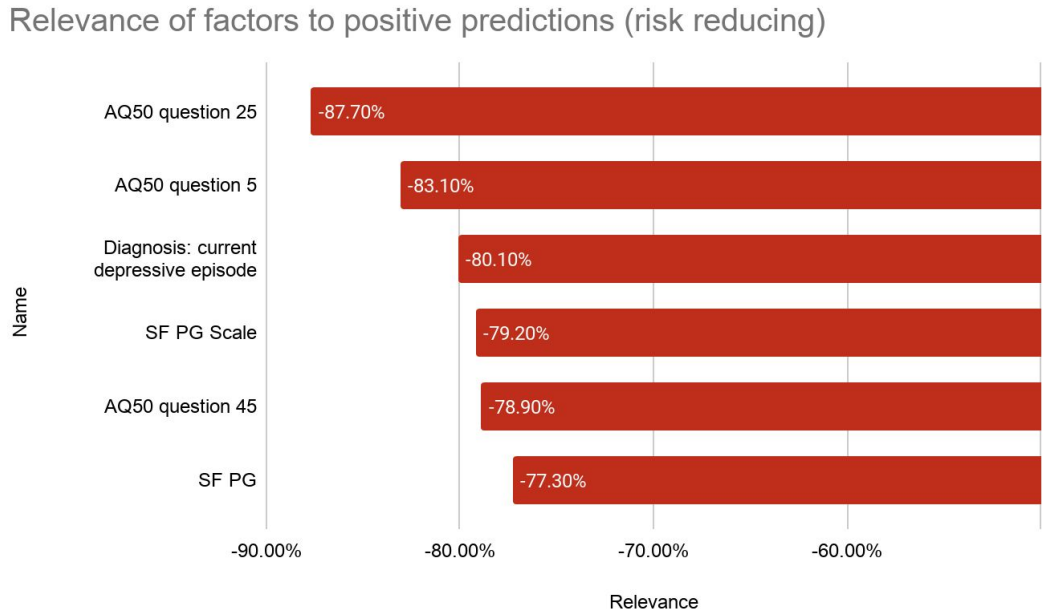

Figure 8: Important factors that reduce the risk to positive predictions in model 9. They can be interpreted as (in that order): being able to deal with a broken routine (-87.70%), noticing sounds that others do not (-83.10%), currently experiencing a depressive episode (-80.10%), psychiatric health point scale (-79.20%), having trouble to understand the goals of others (-78.90%), psychiatric health (-77.30%)

6.2.3.2 Second best performing model - No.7

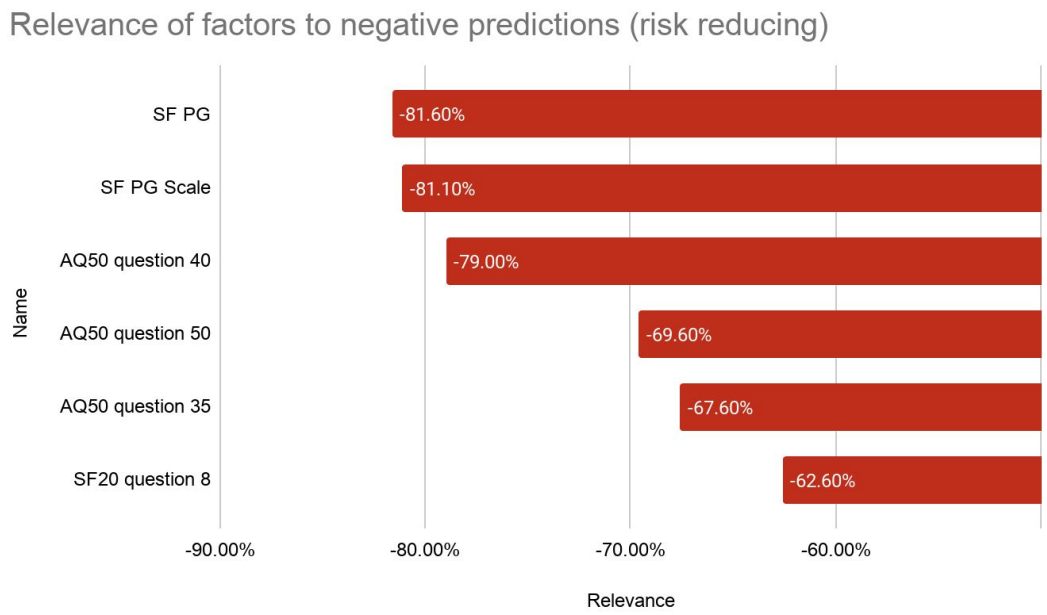

Figure 9: Important factors that reduce the risk to negative predictions in model 7. They can be interpreted as (in that order): psychiatric health (-81.60%), psychiatric health point scale (-81.10%), having enjoyed games that involve pretending as a kid (-79.00%), finds it easy to play games that involve pretending with kids (-69.60%), usually being the last one to get a joke (-67.60%), having had physical pain recently (-62.60%)

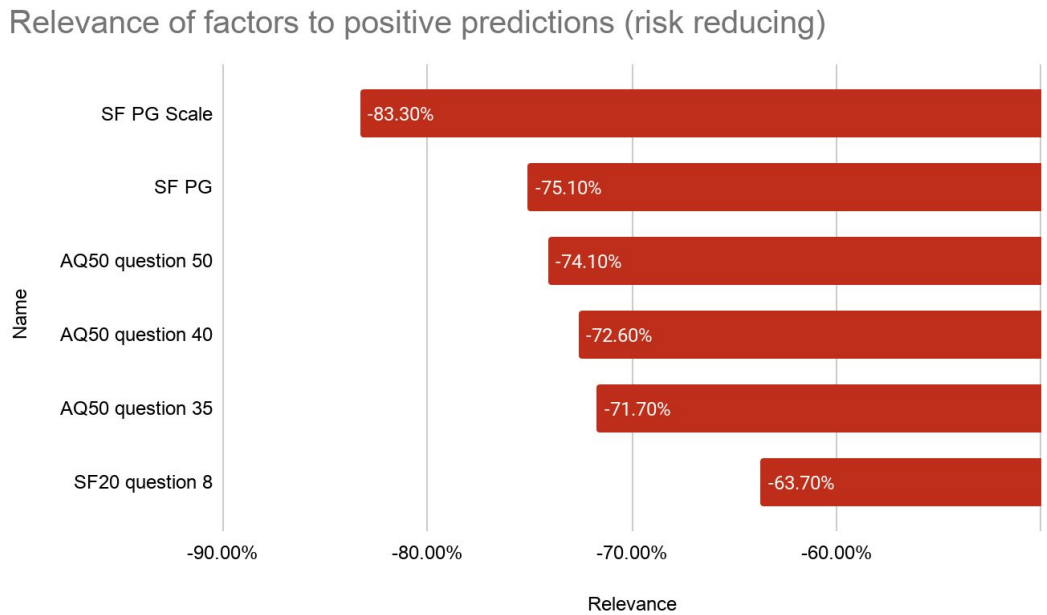

Figure 10: Important factors that reduce the risk to positive predictions in model 7. They can be interpreted as (in that order): psychiatric health point scale (-83.30%), psychiatric health (-75.10%), finds it easy to play games that involve pretending with kids (-74.10), having enjoyed games that involve pretending as a kid (-72.60%), usually being the last one to get a joke (-71.70%), having had physical pain recently (-63.70%)

6.3 Additional Model

The plan was to store all 10 created models for reasons of reproducibility. Due to a mistake in the code however, that did not happen. To provide at least one model along with this paper another model was trained. It reaches almost the same accuracy (84.18%) and should be interesting to consider.

Interestingly, while the performance is similar, it relies on different factors. While the two best performing models from the initial run heavily rely on combined scores such as the total scores of the OQ and IDS questionnaire, this model considers specific sub-questions of those scores much more relevant.

It seems to focus on what reduces risk, rather than what increases it, as the two most relevant parameters are ones that reduce risk. Specifically, question 18 of the AQ50 is considered the most relevant predictor. It states “While I am talking it is hard for others to throw in a word.” The second most important parameter is the general mental health (PG) score from the SF, reducing the risk with 83.8% relevance. Only then parameters that raise the risk follow. Interestingly the 3rd most relevant parameter is question 8 from CAARS (asking if one experiences temper tantrums), raising the risk with 76.5% relevance. According to this model a high level of education also reduces the risk with a relevance of 70.1%.

7 Conclusion

The goal of this research was to answer the following question:

How accurately can a neural network that was trained on data from common psychological tests identify subjects that suffer from suicidal ideation, and what are the most important factors ?

7.1 Performance

The best performing model had an accuracy of 85.3%, showing that this accuracy is achievable. While this is already better than the baseline model which would have an expected accuracy of 64.4%, it is probable that this can be improved by means outlined in the discussion. The positive predictive value of the best performing model was 81.538%, its negative predictive value was 87.5%.

7.2 Important Factors

Age and gender have been identified as relevant factors in suicidal ideation by previous research (GGD, 2017) and were thus expected to be relevant to this model. However, they were not at all included in the relevant factors of either model.

7.2.1 Risk Factors

The questionnaires that contain almost all the most important parameters are the Outcome Questionnaire (OQ) and Inventory of Depressive Symptomatology (IDS). The most important parameter to an elevated risk are the OQ total score and two of

its subscales; one encoding general symptomatic distress (SD) and the other one reflecting issues with interpersonal relations (IR). The IDS total score is also considered very relevant by both models, with the individual questions 2 and 10 being notably more relevant than the others. Question 2 is about issues with sleeping through a night while question 10 is about being sad.

Furthermore, the questions 16, 18 and 19 of the Short Form survey (SF), seem relevant. While question 16 is about being “so sad that nothing can cheer one up”, questions 18 and 19 are concerned with being general health.

The only question that is relevant to an elevated risk from a different survey is question 2 of CAARS, a questionnaire designed to look into ADHD. That question is about being busy all the time as if ‘driven by an external motor’.

7.2.2 Protective Factors

While some factors differ between the models there are some things they have in common. The risk-reducing variables prominently feature the subscore of the SF concerned with mental health (PG). It is represented within the top 5 relevant protective factors by both, its scaled outcome and point scale in 3 out of 4 cases. Essentially meaning that this variable is counted double.

Most other questions featured in any model stem from the AQ, a questionnaire designed to look into autism. That questionnaire is not commonly used to diagnose or look into suicidal ideation. However, it looks like those questions do provide additional insight. Specifically, the following questions were featured: 5, 25, 35, 40,

50. Question 5 is about noticing sounds that others do not which could be interpreted as being attentive to one's surroundings. Question 25 is about dealing with an interrupted routine, while question 35 is about usually being the last one to get a joke. Question 40 asks if someone enjoyed games that involve pretending things when they were a kid and question 50 asks if the person finds it easy to play games that involve pretending things with kids. Questions 35, 40 and 50 can be summed up as ‘being imaginative’.

Three other factors have been identified within the top five factors of both models. Two of them are diagnoses: having a mood disorder and currently experiencing a depressive episode. For both it is undetermined why they seem to reduce the risk of experiencing suicidal ideation instead of raising it. The third is question 8 from the SF20 which concerns having had physical pain recently. This seems contradictory at first, as one would expect bodily pain to be related to symptomatic distress as described in the OQ (OQ SD), which was identified to be a risk-factor. However, the OQ SD captures a number of variables, and physical pain could still somehow reduce the risk of experiencing suicidal ideation when considered alone.

8 Discussion

While not all the models were equally accurate (75.1% - 85.3%), there were several models that performed similarly well; the two best models were only 2.1% points apart (85.3% and 83.1%). Still, their judgement seems to be based on different information. This might be due to overlap between questionnaires or for other reasons. Optimally these models would rely on different structures in the data, in which case ensemble methods would be a promising way to increase performance. Either way, both explanations for decisions should be seen as similarly justified.

At first glance, this accuracy is comparable to the previous research into the important factors to suicidal ideation by Galiatsatos et al. (2015). However, that research was done using Bayesian networks which were not tested on a separate training set. The accuracy measured there only reflects how well the networks trained in that research were able to classify the data that was used in their training. As that study focussed on finding important factors, that did not present an issue. When comparing that accuracy to the one achieved here however, it does. It has to be expected that the performance of that model would be worse when classifying samples from a different dataset. In contrast, within this research a part of the data was held back during training and exclusively used for testing later.

The identified important parameters have to be taken with some caution. While they do seem to be good predictors for the network, we do not understand why for specific cases they seem to work well. It is unreasonable to assume that all people that suffer from suicidal ideation do this for the same reasons and all factors are

always equally important. The analysis is not fine-grained enough to look at those nuances and differences between the underlying specific patterns. Even if that was possible, the learning process of a neural network is based solely on correlation, so found patterns do not necessarily imply causative relationships.

However, the identified risk-factors are related to mental health and seem to match what was expected. While the protective factors were much less expected, and should be further investigated.

The reported accuracy has been compared to a base-model, but the significance of the findings has not been verified by a permutation analysis or anything alike. That leaves the size of the possibility that the data in the test-set was skewed in favour of a higher classification accuracy unknown.

The measured accuracy might have also been influenced by a possible misclassification of 12.2% of the input samples. However, as explained in section 5.1.1 Participants, the likelihood of that is low. Even if one was to assume that they were, noisy data would most likely reduce the performance of the model, indicating that it would actually work better than was shown here.

9 Limitations

For one, clinicians might have trouble handling the output of such a predictive system properly. The output of the system will not always be accurate, and it would be dangerous if doctors started relying on the (statistics based) analysis of a neural network too much. It is also possible that the insight such a neural network gives is usually right but ignores certain special cases. Careful consideration is necessary to ensure that the quality of care does not worsen for people with atypical conditions.

It will be important to keep in mind that the dataset is not a representative sample of the society. Most people in the dataset have some sort of condition that might change the way they react to different influences. Despite the included control group it can not simply be assumed that the findings generalize to the general population.

It is still under debate to what extend people can accidentally commit suicide, for example by unintentionally going further than intended with non lethal self-harm. In such cases, detecting suicidal ideation may not be an effective way of preventing the suicide.

10 Future Research

One could investigate if a more regular deep neural network in combination with post-hoc explanation methods could reveal different information about the factors. However, it seems reasonable to first try to improve the performance of this model. This could be done by fine-tuning parameters such as the number of epochs during training or the size of the layers. It is also an option to combine several of these models in an ensemble approach or utilizing different data.

Either way, a bigger dataset could help improve the performance. If one was to find a label-invariant transformation that could be applied to the data, be it from the same dataset or another, data augmentation would be an option. The transformation could easily be integrated in the code utilized in this project.

With or without those adaptations, it would be interesting to see if and how the performance changes with altercations to which data is used. One could train the model on only combined scores of the questionnaires or remove those combined scores, leaving only singular questions. That would certainly change the explanations, possibly revealing new important factors.

Likewise, one could investigate how the model performs on a more limited subset of questions in general or on a dataset with people that represent the general population better. Even looking into different topics all together is possible utilizing this basic setup.

11 Possible Applications

The concept shown here can be used as it is when treating patients that suffer from suicidal ideation. Patients in psychiatric care settings are usually asked to fill in at least some questions contained in the dataset. If those answers were fed to such a model and the prediction is correct, the explanations could give insight into why the patient is experiencing suicidal ideation, aiding in treatment.

Also, it should be possible to find proxies for the identified risk-factors in other settings. Training a similar model on those may enable the implementation of early warning systems in schools and other settings where entrusted persons can be informed and try to reach out to the person at risk.

12 Acknowledgements

I would like to thank my supervisors Dr. Pim Haselager, Dr. Rose Collard and Dr. Peter Mulders for their invaluable feedback and insight. Without the combination of their perspectives this would not have been the same. Without the critical questions of Meilina Reksoprodjo however, probably nobody would have been able to understand what I made of that insight; thank you! Dr. Petra Muckel made sure that people are not only able to understand me, but also feel the comfort of proper formatting and style, which I am very grateful for. Had I lost motivation though none of this would have mattered though. I want to thank all my friends and family for their continued support and especially my boyfriend Julian for helping me to have the energy for this. Thank you!

12 References

Bruffaerts, R., Demyttenaere, K., Hwang, I., Chiu, W., Sampson, N., Kessler, R., . . . Nock, M. (2011). Treatment of suicidal people around the world. British Journal of Psychiatry, 199(1), 64-70. https://doi.org/10.1192/bjp.bp.110.084129

Eurostat (2020a, June 19). Causes of death - standardised death rate. Retrieved September 25, 2020, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Causes_of_death_%E2%80%94_standardised_death_rate,_2017_(per_100_000_inhabitants)_Health20.png

Eurostat (2020b, June 02). Causes of death - standardised death rate by residence. Retrieved from https://appsso.eurostat.ec.europa.eu/nui/show.do?query=BOOKMARK_DS-417853_QID_-1C267FA_UID_-3F171EB0&layout=SEX,L,X,0;GEO,L,Y,0;UNIT,L,Z,0;TIME,C,Z,1;AGE,L,Z,2;ICD10,L,Z,3;INDICATORS,C,Z,4;&zSelection=DS-417853AGE,TOTAL;DS-417853ICD10,X60-X84_Y870;DS-417853UNIT,RT;DS-417853INDICATORS,OBS_FLAG;DS-417853TIME,2017;&rankName1=ICD10_1_2_-1_2&rankName2=TIME_1_0_-1_2&rankName3=UNIT_1_2_-1_2&rankName4=AGE_1_2_-1_2&rankName5=INDICATORS_1_2_-1_2&rankName6=SEX_1_2_0_0&rankName7=GEO_1_2_0_1&rStp=&cStp=&rDCh=

&cDCh=&rDM=true&cDM=true&footnes=false&empty=false&wai=false&time_mode=NONE&time_most_recent=false&lang=EN&cfo=%23%23%23%2C%23%23%23.%23%23%23

Melis, D. A., & Jaakkola, T. (2018). Towards robust interpretability with self-explaining neural networks. In Advances in Neural Information Processing Systems (pp. 7775-7784).

Hussain A., Elbaghdadi, O., Bardarov I., Hoenes C. (2020, January 31). Self Explaining Neural Networks: A Review with Extensions. https://amanhussain.com/publication/self-explaining-neural-networks/

Radboudumc (n.d.). MIND-SET. Retrieved September 25, 2020, from https://www.radboudumc.nl/trials/mind-set

Nock, M., Borges, G., Bromet, E., Alonso, J., Angermeyer, M., Beautrais, A., . . . Williams, D. (2008). Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. British Journal of Psychiatry, 192(2), 98-105. https://doi.org/10.1192/bjp.bp.107.040113

GGD Brabant-Zuidoost (2017). Serieuze suïcidegedachten onder volwassenen. Retrieved January 23, 2021, from

https://www.ggdbzo.nl/ggdkompas/Documents/Infographic%20su%C3%AFcidegedachten%20volwassenen.pdf

Zheng, H., Fernandes, E., & Prakash, A. (2019). Analyzing the Interpretability Robustness of Self-Explaining Models. arXiv preprint arXiv:1905.12429.

WHO (2019a). Suicide. Retrieved January 23, 2021, from https://www.who.int/teams/mental-health-and-substance-use/suicide-data#:~:text=Close%20to%20800%20000%20people,prevent%20suicide%20and%20suicide%20attempts

WHO (2019b). Suicide. Retrieved January 23, 2021, from https://www.who.int/news-room/fact-sheets/detail/suicide

oW_♦(2019). What makes binary cross entropy a better choice for binary classification than other loss functions? Retrieved January 28, 2021, from https://datascience.stackexchange.com/questions/53400/what-makes-binary-cross-entropy-a-better-choice-for-binary-classification-than-o

Shatte, A., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. Psychological medicine, 49(9), 1426–1448. https://doi.org/10.1017/S0033291719000151

Roy, A., Nikolitch, K., McGinn, R., Jinah, S., Klement, W., & Kaminsky, Z. A. (2020). A machine learning approach predicts future risk to suicidal ideation from social media data. NPJ digital medicine, 3(1), 1-12. https://doi.org/10.1038/s41746-020-0287-6

Tran, T., & Kavuluru, R. (2017). Predicting mental conditions based on "history of present illness" in psychiatric notes with deep neural networks. Journal of biomedical informatics, 75, S138-S148. https://doi.org/10.1016/j.jbi.2017.06.010

Galiatsatos, D., Konstantopoulou, G., Anastassopoulos, G., Nerantzaki, M., Assimakopoulos, K., & Lymberopoulos, D. (2015, September). Classification of the most significant psychological symptoms in mental patients with depression using Bayesian network. In Proceedings of the 16th International Conference on Engineering Applications of Neural Networks (INNS) (pp. 1-8). https://doi.org/10.1145/2797143.2797159

Runeson, B., Odeberg, J., Pettersson, A., Edbom, T., Jildevik Adamsson, I., & Waern, M. (2017). Instruments for the assessment of suicide risk: a systematic review evaluating the certainty of the evidence. PLoS one, 12(7), e0180292. https://doi.org/10.1371/journal.pone.0180292

Schönberger, D. (2019). Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. International Journal of Law and Information Technology, 27(2), 171-203. https://doi.org/10.1093/ijlit/eaz004

Goodfellow,I., Bengio Y., & Courville A. (2016). Deep Learning. MIT Press, 172-175. http://www.deeplearningbook.org

Harmer, B., Lee, S., Duong, T., & Saadabadi, A. (2020). Suicidal Ideation. StatPearls.

Collard, R. (2021). Measuring Integrated Novel Dimensions in Neurodevelopmental and Stress-related Mental Disorders. Manuscript submitted for publication.

13 Appendix

13.1 MIND-SET Publications

A list of all publications that resulted from MIND-Set, kindly provided by Rose Collard.

No.	Title and authors	(Target) journal	Date first registration	Status (Date)	No.
1	~~Alexithymia mediates the relationship between childhood trauma and emotion regulation difficulties in psychiatric patients. Everaertd D., Schene, A.H., Schellekens, A.F.A., Collard, R.M., van Eijndhoven, P., van Oostrom, I., Vrijsen, J.N.~~	~~COTR~~	~~25-04-2018~~	~~In prep:~~ ~~30-06-2020~~	
2	Systematic Review of Affective Cognitive Biases in Autism Spectrum Disorder: Towards an Understanding of the Prevalent Comorbidity with Depression. Bergman, M.A., Schene, A.H., Constance Th.W.M., Vrijsen, J.N., Kan, C., van Oostrom, I.		4-5-2018	4-5-2018	

3	Affective attentional biases in Autism Spectrum Disorder and/or Major Depressive Disorder: an eye-tracking study. Bergman, M.A., Constance Th.W.M., Vrijsen, J.N., Rinck, M.M., Schene, A.H.		4-5-2018	4-5-2018	
4	Attentional biases in Neurodevelopmental & Mood Disorders: a Network approach Bergman, M.A., Constance Th.W.M., Vrijsen, J.N., Rinck, M.M., Brolsma, S.C.A. Schene, A.H.		4-5-2018	4-5-2018	
5	Continued stress from a network perspective (analysis on the resting state scan after stress induction, compared to the resting state scan after a neutral control condition). J. van Oort, I. Tendolkar, A. Schene, G. Fernandez, P. van Eijndhoven		04-06-2018	Accepted	1
6	Challenging the negative learning bias hypothesis of depression: reversal learning in a naturalistic psychiatric sample. S.C.A. Brolsma, J.N. Vrijsen, E. Vassena, M. Rostami Kandroodi, M.A. Bergman, P. van Eijndhoven,	Psychological Medicine	16-10-2018	Accepted	2

	R.M. Collard, H.E.M. den Ouden, A.H. Schene, R. Cools				
7	Negative learning bias in depression revisited: Enhanced neural response to surprising reward across psychiatric disorders. S.C.A. Brolsma*, E. Vassena*, J.N. Vrijsen, G. Sescousse, P. van Eijndhoven, R.M. Collard, A.H. Schene, R. Cools	Biological Psychiatry: Cognitive Neuroscience and Neuroimaging	16-10-2018	Accepted	4
8	BOLD activity in visual association areas is modulated by unexpected outcomes. S.C.A. Brolsma, J.N. Vrijsen, E. Vassena, A.H. Schene, R. Cools		16-10-2018	External research group will continue	
9	Relation between childhood adversity and the volume of the hippocampus and amygdala J. van Oort, I. Tendolkar, A.H. Schene, P. van Eijndhoven		13-11-2018	In prep	
10	Is an attentional deficit (ADHD) the "cure" for negative attentional bias in depressed patients? Schuthof, C., Tendolkar, I.,		01-06-2019	In prep: 08-12-2020	

	Collard, R.M., Eijndhoven, P., Schene, A.H., Vrijsen, J.N.				
11	The Importance of Perseverative Cognition for Both Mental and Somatic Disorders in a Naturalistic Psychiatric Patient Sample. Appel, J., Schene, A.H., Eijndhoven, P., Collard, R.M., Tendolkar, I., Vrijsen, J.N.		01-06-2019	Submitted: 08-12-2020	
12	Negative memory bias as a transdiagnostic cognitive marker for depression symptom severity. Duyser, F.A., Van Eijndhoven, P.F.P., Bergman, M.A., Collard, R.M., Schene, A.H., Tendolkar, I., Vrijsen, J.N.	Journal of Affective Disorders	24-07-2019	Accepted	3
13	Amygdala reactivity as neural correlate of negative memory bias in a naturalistic psychiatric patient sample. Duyser, F.A., Vrijsen, J.N., Van Oort, J., Collard, R.M., Schene, A.H., Tendolkar, I., Van Eijndhoven, P.F.	Biological Psychiatry: Cognitive Neuroscience and Neuroimaging	24-07-2019	In prep: 08-12-2020	
14	Anhedonia as a Transdiagnostic Symptom in symptom clusters of Depression, Anxiety Sensitivity, ADHD and Autistic Traits: A Network		30-06-2020	In prep: 08-12-2020	

	Approach. Guineau, M., Ikani, N., Rinck, M., Collard, R.M., Van Eindhoven, P.F.P., Schene, A.H., Becker, E., Vrijsen, J.N.				
15	Transdiagnostic brain-behavioral mapping using sparse multiple canonical correlational analysis (msCCA) regression Peter Mulders, Andre Marquand, Philip van Eijndhoven, Indira Tendolkar, Aart Schene		01-05-2020	Analysis started	
16	Neural correlates of repetitive negative thinking (relation of RNT (measured with the PTQ) with resting state functional connectivity and with stress induced changes in connectivity). J. van Oort, I. Tendolkar, R. Collard, D. Geurts, A.H. Schene, P. van Eijndhoven		30-06-2020	In prep	
17	Relation of psychiatric symptoms with resting state connectivity and stress induced changes in connectivity (a Linked ICA analysis). J. van Oort, I.		30-06-2020	Analysis started	

	Tendolkar, A.H. Schene, P. van Eijndhoven				
18	Pooled SRET project: an examination of the task to find optimal outcome variables/best predictors for psychopathology. Duyser, F.A., Van Eijndhoven, P.F., Schene, A.H., Tendolkar, I., Vrijsen, J.N.		30-06-2020	Cleaning up data started (November 2020)	

13.2 Full code

The full code, including results from all runs can be found here:

https://gitlab.com/deislukas/senn-identifies-si