TripleSec_PDF

true

2021-08-29

Introduction

In a day and age, where more data per patient is collected and has to be evaluated, health care providers are starting to rely on clinical decision support systems (CDSS). The CDSS support health care providers by sifting through the patient data and suggesting further steps of treatment depended upon evidence based clinical guidelines [1].

The Triple Sec Algorithm is a clinical support algorithm for risk assessment of Alzheimer's disease (AD).

AD is the most common cause of a decline in mental ability [2]. It is a highly complex disease that surfaces differently with each patient. After it becomes visible it is hard to treat, so an early diagnosis and early treatment increases the chances of the efficiency of the treatment [3]. According to current statistics [4] over 50 million people are living with a decline in mental abilities, dementia, today and a lot of them have not been diagnosed. Statistics estimate that someone in the world develops dementia every 3.2 seconds, leading to about 10 million new dementia cases each year [4]. Due to the number of cases there is an urgent need for early and reliable diagnosis.

The problem with neurological diseases is the incomplete knowledge of them and the lacking consistency between different clinical, laboratory and imaging findings. For the task of singling out affected individuals with confidence we propose the created Triple Seq algorithm, which identifies patients dependent on user given risk thresholds by linking several predictors to the diagnostic procedure which are optimal to determine the given risk level, therefore categorizing the patients into risk categories dependent on their neurological symptoms. These categories, high risk, low risk and indiscriminate, represent the risk of the patient progressing to Alzheimer's disease, dependent on the given biomarker values in the completed examinations. The categories are defined by cutoff values of the biomarkers. These cutoff values are calculated dependent on user defined risk-threshold and ratio of patients. The algorithm allows healthcare providers to justify more invasive or more expensive treatment options for high risk converters [5].

Alzheimer

Alzheimer's disease (AD) is a neurodegenerative disease and cause of dementia [2,6].

The term dementia describes a group of neurodegenerative symptoms concerning memory, language usage and orientation. There are several different forms of dementia and several different diseases that cause it, though AD is by far the most common accounting for 60-80% of the cases [2]. The course of progression of AD is usually categorized in three stages, early or mild, middle or moderate, and late or severe, with a continually decrease in cognitive and functional capabilities ending in the patients death [6]. Even though the disease affects so many people and death is certain, the cause is still unclear. There are various hypotheses ranging from genetic mutations to chronic inflammations to protein abnormalities [6]. The biggest risk factor is age, which makes the diagnosis that much harder as the first signs of AD, such as short term memory loss and subtle problems with attentiveness, planning, abstract thinking and flexibility, are often mistaken for normal aging [7].

Diagnosis

Triple Sec

The Triple Sec algorithm is a clinical support algorithm for the risk assessment of patients to have Alzheimer's disease.

The cause of AD is unknown which complicates finding precise methods for the identification of the disease. Usually the way clinicians identify that a patient has Alzheimer' disease is by doing different neurological and imagining tests in order to find out if the patient has AD or a different neurological disease, seldom are clinicians able to do costly or invasive procedure since the earlier tests do not correspond with each other therefore never creating a good base for them to justify their need for invasive tests. The Triple Sec algorithm allows clinicians to link the different test together and receive a risk assessment therefore giving them a base to make decisions about further treatments or tests.

The algorithms goal is to categorize patients with neurological symptoms into three categories. These three categories, high risk, low risk and indiscriminate, represent the risk of the patient progressing to Alzheimer's disease, dependent on the given biomarker values. The categories are defined by cutoff values of the biomarkers. These cutoff values are calculated dependent on user defined risk-threshold and ratio of patients. The user defined risk-thresholds, pHigh and pLow, are given to the algorithm. They present the percentage of patients which are supposed to have a certain value $>=x/ <=x$ to determin the cutoff values.
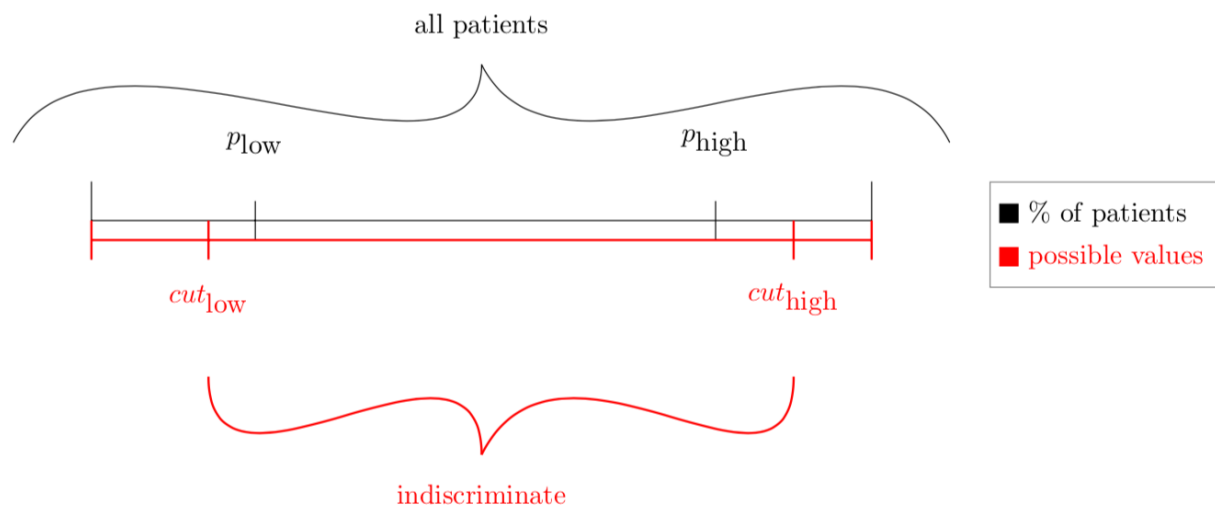
This calculation is repeated for each biomarker given from the user, filtering out the patients which have been categorized into high or low risk, therefor minimizing the amount of patients the algorithm categorizes as indiscriminate ("unknown risk").
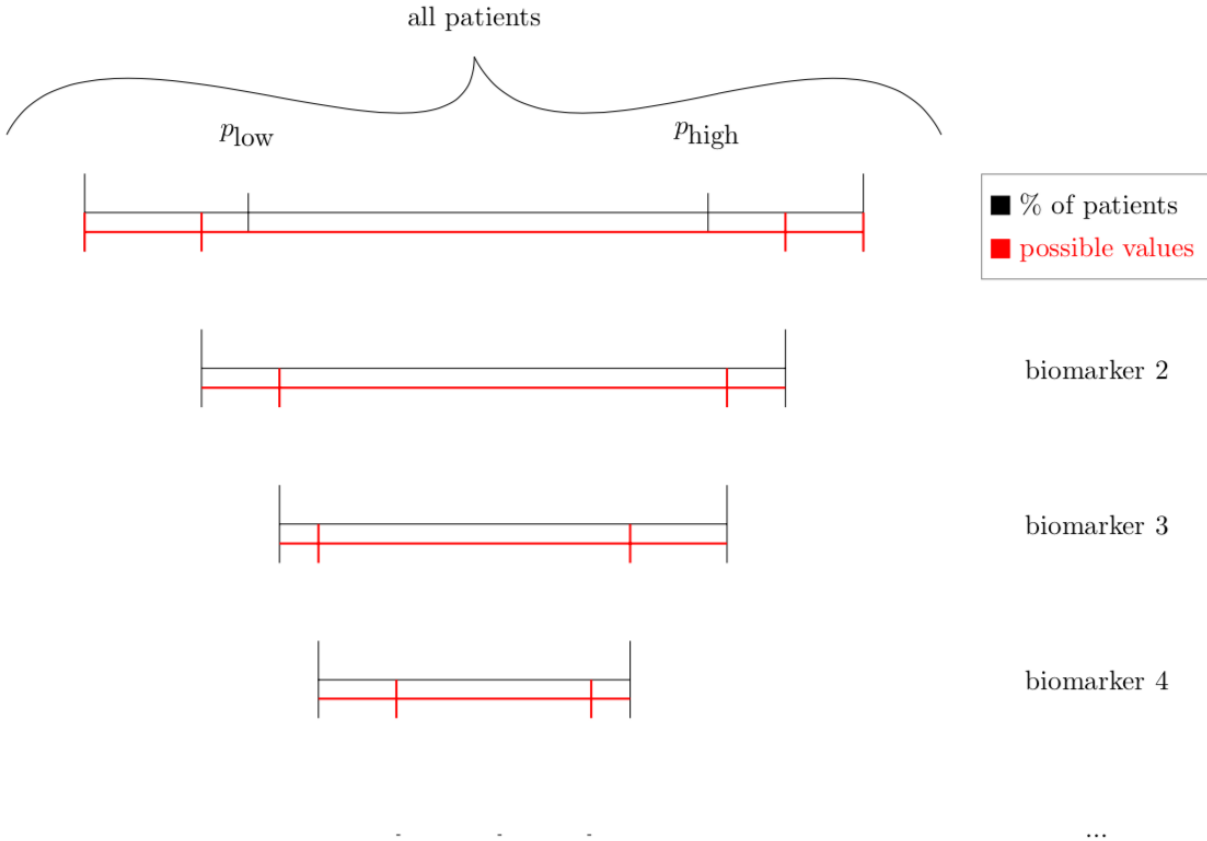
Preinstallation

TripleSec depends on one package. readr

Please ensure to install the readr package before using the TripleSec package.

Installation via GitHub

```
require(readr)
install_github("zaynabhammoud/TripleSec")
library(TripleSec)
```

Available Functions

pruneTree   calculate cutOff table for given data

This function divides the given input data frame into training and test data. All patients but one are used as training data and are used in the function calCutOff, which returns a matrix containing the calculated cut off values. Using this matrix and the test patient, pruneTree then calls predictTree. This function assesses the risk of conversion for the test patient.

The function pruneTree returns a variable result with a link to a table containing the average cutoff values for each feature and a data table with the patients and feature and their risk assessment, once it has iterated through all input data and using each patient as a test patient.

The function requires the following arguments:

- **inputData** - The prepared input data. The data can be prepared via the function `prepareData(data, firstFeature, featureList)`

- **pHigh** - The high risk threshold
- **pLow** - The low risk threshold

Example

```
inputData <- prepareData(data, firstFeature, featureList)
pHigh <- 0.8
pLow <- 0.2
result <- pruneTree(inputData, pHigh, pLow)
```

calCutOff   Sequential training of the algorithm

This function is used to calculate the cut off values for low-risk and high-risk groups.

The values are calculated by the proportion between converters with a value and all patients having this certain value corresponding with the given pHigh and pLow values. For each feature cutoff calculation all patients that have not been categorized into the high or low risk categories will be used.

The function calCutOff returns a variable fit, which contains links to the calculated cutOff table, filtered data and effectiveness of the cutOff values.

The function requires the following arguments:

- **inputData** - The prepared input data for training of the algorithm, which is provided by pruneTree.
- **pHigh** - The high risk threshold.
- **pLow** - The low risk threshold.

predictTree   Sequential prediction of a patient conversion risk

This function calls the recursive function predictTreeRec. The function returns a variable with a link to a patient data table containing their risk assessments and a link to the average cutoff value matrix.

The function predictTree returns a variable result with its links to the cutOff matrix and predicted patient information.

The function requires the following arguments:

- **cutoffValues** - The calculated cutOff values, which are provided by the function calCutOff.
- **inputData** - The inputData/the patient which is supposed to be predicted (provided by pruneTree).

predictTreeRec   Sequential prediction of a patient conversion risk

This recursive function calculates the cutoff value for each parameter dependent upon the patient ratio, it also verifies if the model is effective (cut_high > cut_low).

The function predictTreeRec returns a variable result with its links to the cutOff matrix and predicted patient information.

The function requires the following arguments:

- **cutoffValues** - The calculated cutOff values, which are provided by the function calCutOff.
- **inputData** - The inputData/the patient which is supposed to be predicted (provided by pruneTree).

prepareData   Filters and transforms given data

This function transforms the given data into a data frame containing only the columns necessary for the TripleSec algorithm, therefor for the risk assessment.

The function requires the following arguments:

- **data** - The data which is given by the user.
- **firstFeature** - The name of column containing the feature data which is most relevant.

- **featureList** - A list contaning all column names of the features which the algorithm is supposed to work with.

Testing

Results

References

1. CDC. How to Implement Clinical Decision Support Systems [Internet]. Centers for Disease Control and Prevention. 2021 [cited 12. August 2021]. available on: https://www.cdc.gov/dhdsp/pubs/guides/best-practices/clinical-decision-support.htm

2. Dementia vs. Alzheimer's Disease: What is the Difference? [Internet]. Alzheimer's Disease and Dementia. [cited 12. August 2021]. available on: https://alz.org/alzheimers-dementia/difference-between-dementia-and-alzheimer-s

3. Alzheimer verstehen [Internet]. [cited 12. August 2021]. available on: https://www.alzheimer.de/

4. ADI - Dementia statistics [Internet]. [cited 12. August 2021]. available on: https://www.alzint.org/about/dementia-facts-figures/dementia-statistics/

5. Early probability-driven prediction of disease progression: a new algorithm applied to Alzheimers disease. :13.

6. Alzheimer-Krankheit. In: Wikipedia [Internet]. 2021 [cited 19. August 2021]. available at: https://de.wikipedia.org/w/index.php?title=Alzheimer-Krankheit&oldid=214843556

7. What is Alzheimer's? [Internet]. Alzheimer's Disease and Dementia. [zitiert 19. August 2021]. Verfügbar unter: https://alz.org/alzheimers-dementia/what-is-alzheimers

8. Ren J. Automated Detection of Early-Stage Alzheimer's Disease. :94.

9. Frost C. Molecular Dynamics Simulation: Applications from Alzheimer's Disease to Photopharmacology. :208

10. Sturm D, Biesalski A-S, Höffken O, Herausgeber. Neurologische Pathophysiologie: Ursachen und Mechanismen neurologischer Erkrankungen [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2019 [cited 29. August 2021]. available on: http://link.springer.com/10.1007/978-3-662-56784-5

11. HealthITSecurity. Understanding the Basics of Clinical Decision Support Systems [Internet]. HealthIT-Analytics. 2017 [cited 29. August 2021]. available on: https://healthitanalytics.com/features/understanding-the-basics-of-clinical-decision-support-systems