

# Project Proposal

sepsis prediction AI model

by Radoslav Radev

ICT & AI semester 4

14/03/2023

ver. 2.0

## Introduction:

This project aims to develop a model that can predict sepsis.

Sepsis is a life-threatening condition caused by the body's response to an infection. Early recognition and timely intervention are essential to improve patient outcomes.

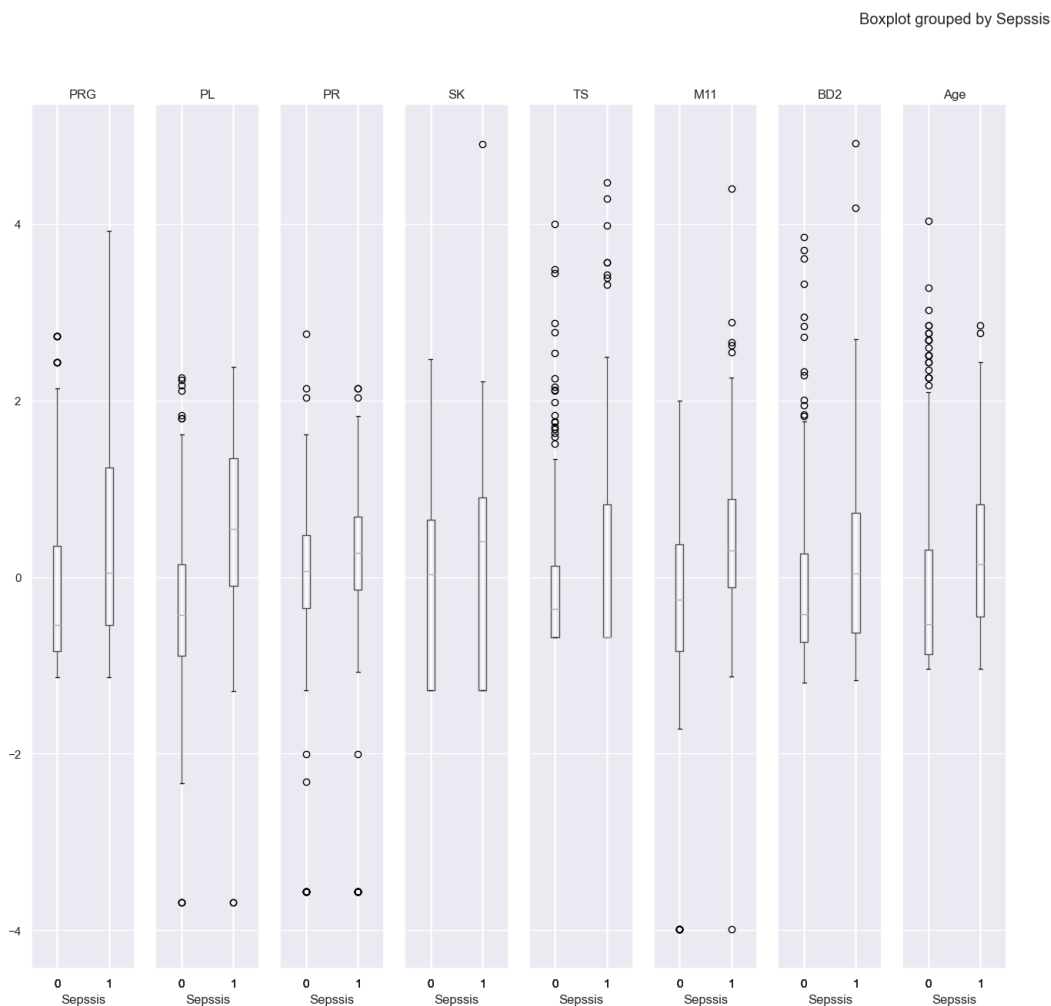
## Domain Understanding:

Please take a look at the Domain Understanding document where I have provided outpoints from my exploratory research, co-reflection with a fellow student, and an interview with an expert

## Data Sourcing:

The dataset I will be using for this individual challenge is provided by Johns Hopkins University, which is a private research university in Baltimore, Maryland U.S

for more detailed information please refer to the notebook of iteration1



after scaling the data and looking at the boxplots I have decided to use the following features for my model: Blood Work Result-1, patients age, Body mass index, and Blood Pressure I chose these features because they have the most significant difference between the positive and

negative cases in iteration 0, now I added the Blood Work Result-3, since after cleaning the data it looks like a good candidate to increase our positive accuracy a little bit

### Analytical Approach:

- **the type of problem**

this is a classification problem (we need to classify if the patient's parameters look like sepsis)

- **choice of models**

please take a look at the separate notebook I created, for choosing the best model for my case

- **evaluation**

iteration 1:

Accuracy: 0.7563025210084033

	precision	recall	f1-score	support
Negative	0.80	0.85	0.83	81
Positive	0.64	0.55	0.59	38
accuracy			0.76	119
macro avg	0.72	0.70	0.71	119
weighted avg	0.75	0.76	0.75	119

iteration 1 showed an improvement in performance over iteration 0, with an accuracy of above 75%. The precision and recall of negative cases indicating that the algorithm had a good ability to identify them. The precision and recall for positive cases have been increased by around 20% from iteration 0, the algorithm can still benefit from improvement in identifying positive cases.

iteration 2.1:

Accuracy: 0.7590361445783133

	precision	recall	f1-score	support
Negative	0.78	0.78	0.78	45
Positive	0.74	0.74	0.74	38
accuracy			0.76	83
macro avg	0.76	0.76	0.76	83
weighted avg	0.76	0.76	0.76	83

Cross validation results [0.74698795 0.71084337 0.69879518 0.75609756 0.76829268]

iteration2.1:

after playing with the hyperparameters of LogisticRegression() explicited in the code cell above I have determent that they don't have any significant effect on the accuracy of the model and left them on default.

However, balancing the dataset did have a significant effect on the accuracy of the model.

From the cross-validation accuracy scores we can see that he model is not overfitted or underfitted.

in our case both Recall (the proportion of true positives that were correctly identified by the model among all actual positive instances.) and Precision (the proportion of true positives (correctly classified positive instances) among all instances predicted as positive by the model) are important , and i am pretty happy with the results i managed to produce from this dataset

## Possible Use Cases:

there are 2 possible use cases:

**Use Case 1:** the system is implemented as standalone:

- medical professionals enter the patients' data, and receive a prediction for the likelihood of sepsis diagnoses
- medical professionals can check back with the system on later time and be provided with a suggestion list of patients who might need extra attention (are likely to have developed sepsis)

**Use Case 2:** the system is integrated as part of the existing hospital infrastructure and automatically process the available patients' data:

- medical professionals can be notified for patients that might need extra attention (are likely to have developed sepsis)

For this project's demonstration Use Case 1 will be developed

### *User story 1*

Title	Priority	Difficulty
Entering Data	MUST	8
<b>As a medical professional, I want to Enter my patient's data. So I can receive a prediction on the likelihood of my patient having sepsis.</b>		
<b>Acceptance Criteria</b>		
<b>Given the patient's information, When the Submit button, in the EnterNewPatient component of the website, is pressed The system:</b> <ul style="list-style-type: none"><li>- Stores the patient's information</li><li>- provides a prediction of how likely is this patient to have sepsis</li></ul>		
<b>No/Insufficient information entered:</b> <ul style="list-style-type: none"><li>- error message displaying which data is field incorrectly filled</li></ul>		

## *User story II*

Tittle	Priority	Difficulty
Checking for patients	MUST	8
<b>As a medical professional, I want to know if any of my patients deserve extra attention. So I can take the necessary extra steps to insure my patients` well-being.</b>		
<b>Acceptance Criteria</b>		
<b>When navigated to PatientsThatDeserveExtraAttention</b> <b>The system:</b> <ul style="list-style-type: none"><li>- displays set patients</li></ul>		

## End Product:

I developed a RestApi Backend for my model that could be easily integrated in the already existing medical infrastructure

the endpoint /predict will return an immediate evaluation from the provided medical data and the system will save the record in the database

The screenshot shows a REST client interface for the POST /predict endpoint. The 'Parameters' section is empty. The 'Request body' section is marked as 'required' and has a dropdown menu set to 'application/json'. Below this, there is a 'Schema' tab showing a JSON schema for the request body. The schema defines the following fields: 'patientID' (string), 'PL' (integer), 'Age' (integer), 'M11' (integer), 'PR' (integer), and 'TS' (integer). The 'Example Value' tab is also visible, showing a sample JSON object.

```
{
  "patientID": "string",
  "PL": 0,
  "Age": 0,
  "M11": 0,
  "PR": 0,
  "TS": 0
}
```

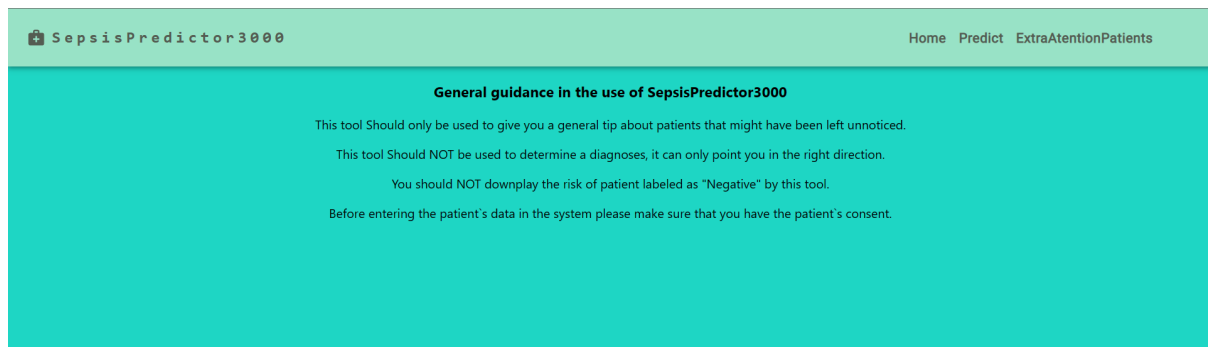
the endpoint /ExtraAttentionPatients will today entered positive patients ordered by certainty

The screenshot shows a REST client interface for the GET localhost:8000/ExtraAttentionPatients endpoint. The 'Send' button has been clicked, resulting in a '200 OK' status. The response is displayed in the 'Preview' tab, showing a JSON array of two patient records. The first record has a 'patientID' of '' and a 'certainty' of 1.0. The second record has a 'patientID' of 'M2435' and a 'certainty' of 1.0. The response also includes a 'date' field with a timestamp.

```
1 [
2   {
3     "patientID": "",
4     "sepsisPrediction": 1,
5     "certainty": 1.0,
6     "date": "2023-03-24T17:02:58.435570",
7     "PL": 148,
8     "Age": 50,
9     "M11": 33.0,
10    "PR": 72,
11    "TS": 0
12  },
13  {
14    "patientID": "M2435",
15    "sepsisPrediction": 1,
16    "certainty": 1.0,
17    "date": "2023-03-24T17:02:58.435570",
18    "PL": 85,
19    "Age": 25,
20    "M11": 26.6,
21    "PR": 66,
22    "TS": 94
23  }
24 ]
```

I also developed a webApplication (FrontEnd for the system) to be used standalone without any integration with other systems

the medical professional is greeted with the general guidance for use of the system



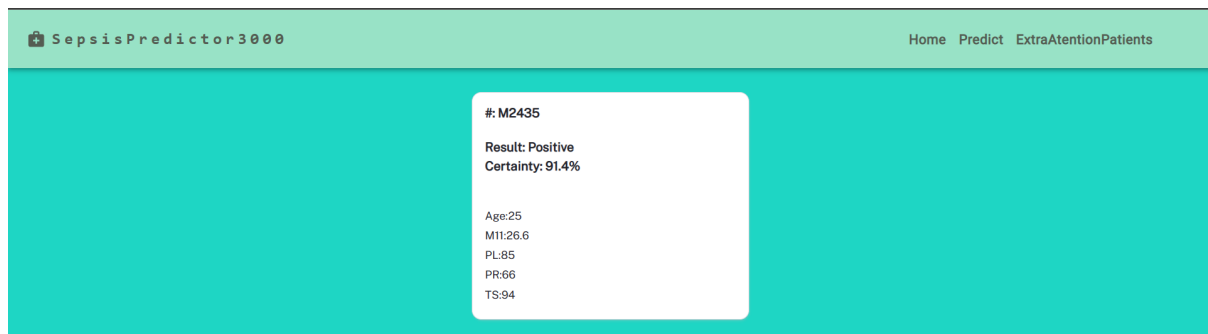
the medical professional can enter a new result in the system from the **Predict** page

The screenshot shows the 'Predict' page of the SepsisPredictor3000 application. The header is identical to the Home page. The main content area contains a form with the following fields:

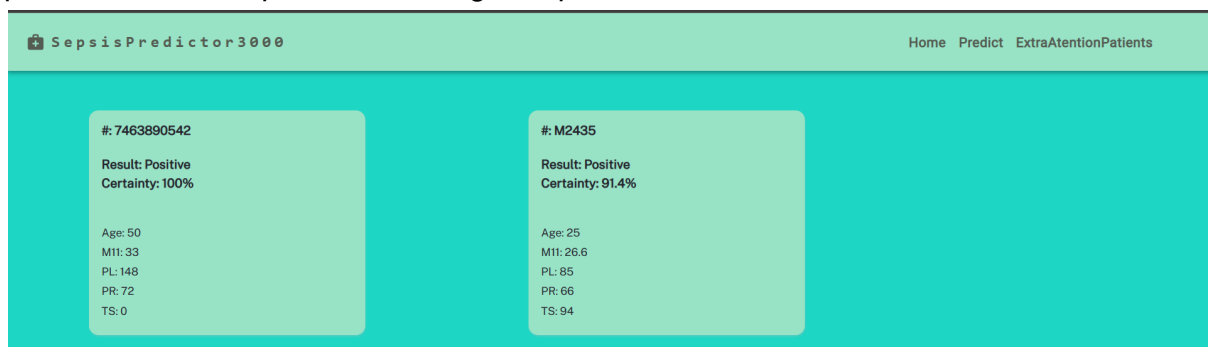
- Patient ID: M2435
- Age: 25
- M11: Body mass index: 26.6
- PL: Blood Work Result-1 (mu U/ml): 85
- PR: Blood Pressure (mm Hg): 66
- TS: Blood Work Result-3 (mu U/ml): 94

A 'Submit' button is located below the input fields.

after that the medical professional is instantaneously provided with the result



if the medical professional navigates to the ExtraAttentionPatients page he will be presented with set patients that might require a second look



The end Product is fully developed and available at:

<https://github.com/raga70/SepisiPredictionModel>



## **General guidance in the use of the system:**

**This tool Should only be used to give you a general tip about patients that might have been left unnoticed.**

**This tool Should NOT be used to determine a diagnosis, it can only point you in the right direction.**

**You should NOT downplay the risk of patient labeled as "Negative" by this tool.**

**Before entering the patient`s data in the system please make sure that you have the patient`s consent.**

## **Consulting the stakeholders:**

after presenting and discussing the solution with my medical contact, she was optimistic about it and stated that it could definitely be integrated as an extra precaution (while the current procedures for diagnosing stay the same and the tool can be used for giving them an early tip that a sepsis diagnosis might be possible), and agreed that integrating it with existing infrastructure would benefit medical professionals by not wasting their valuable time in the use of the system.