

COMP9321 2019 T1 Group Project: Heart Disease

Weight: 20 Marks (Bonus: 4 Marks)

Due: **23:59:59 Sunday, 14th Apr, 2019**

Introduction

In this assignment, you will be asked to form a group to build a web application which contain several tasks. As the data cleansing is pretty hard, so we did it for you.

Form Group

4 students per group (less than 4 is accepted). Register your group on WebCMS3 with group name.

Data Set

This data-set contains 76 attributes, but all published experiments refer to using a subset of 14 of them which comes from University of California Irvine. Attribute information (Although unnecessary, you can find more details in the footnote) ¹:

1. age
2. sex
3. chest pain type (4 values)
4. resting blood pressure
5. serum cholestoral in mg/dl
6. fasting blood sugar > 120 mg/dl
7. resting electrocardiographic results (values 0,1,2)
8. maximum heart rate achieved
9. exercise induced angina
10. oldpeak = ST depression induced by exercise relative to rest
11. the slope of the peak exercise ST segment
12. number of major vessels (0-3) colored by flourosopy
13. thal(Thalassemia): 3 = normal; 6 = fixed defect; 7 = reversable defect
14. target: will not be used in this assignment.

Pre-Requirement

- You can use any technology/algorithm covered in this course, it's your choice as long as your team members are agree with it.

¹<https://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/heart-disease.names>

- You can use Tensorflow or PyTorch to implement your Neural Network if you want one, Scikit for your machine learning algorithm.
- For frontend you can use any framework (e.g.React, Angular, Vue ...) but not compulsory. Back-end have to be python.
- If you don't want to split the frontend and backend you can use flask or other python package. It's all your choice.

Project Requirement

Here is the **key** requirements (this is an open project, and you can implement it freely as you see fit.):

- You need to visualize the statistics for basic information(Attribute) No. 3 - 13 by groups of age and sex with appropriate charts, and display on your web app;
- Your web app should be able to shows the potential important factors(attributes) related to heart diseases(see information/attribute No.13) by choosing suitable algorithm introduced in this course. Any display on your web app in an appropriate way.
- Your web apps shall contain a page where accept user inputs and predict the thal(Thalassemia) and display properly. (e.g., you can show the accuracy graph based on number of iterations in this page after click the predict button)

Bonus

Bonus marks will be awarded based on the grounds of:

- Possible advance features used for clustering in your scenario(even your features may not be significantly improving your performance).

Project Track

Please make a GitHub repository(Any source code control platform is accepted) for your group project which is easy for the collaboration. (Note: Make sure it's private)

Contribution

We will marking the contribution based on the git commit.

Marking Criteria

Teamwork: 5 Marks

Demonstration: 5 Marks

Proper visualization and display: 5 Marks

Prediction accurately: 5 Marks; we will rank your accuracy with all submissions, and your marks depend on the ranking; Top 5 teams will get full mark.

Bonus: 4 marks.

Demonstration

Your demonstration should present and explain your web application, how you figure out the potential important factors(attributes) and your methodology for the prediction. And if applicable, the bonus part as well. Contributions of each group member shall also be claimed.

Submission

Only one submission per group is allowed. Include all your source codes into a3.tar. If you used neural network, include the trained model and the training file.

Please include a “README” file to describe your design and a “requirements.txt” to state all the packages needed for your submission.

Commend on CSE machine:

```
give cs9321 assn3 a3.tar
```

Plagiarism

0 for COMP9321 2019 Term One.