School of Computing 

COMP5200M Scoping and Planning Document

|  |  |
| --- | --- |
| **Student Name:**  Wenjie Zhang | |
| **Programme of Study:**  Msc Advanced Computer Scienc**e** | |
| **Provisional Title of Project:**  Machine learning classification of aneurysms using imaging and simulation data | |
| **Name of External Company** (if any)**:**  None | |
| **Supervisor Name:**  Toni Lassie | |
| **Type of Project:**  Empirical Investigation | |
| ***NOTE to student****: ensure you have discussed the content with the supervisor before submitting this document to Minerva. Submit an* ***electronic version*** *of this report in pdf via the appropriate link in Minerva; with filename of the format <surname><year>-SP ( e.g. SMITH15-SP.pdf).* | |
| **Signature of Student:** | **Date:** |

Contents

[1. Background Research for the project 1](#_Toc429570687)

[1.1 Context 1](#_Toc429570688)

[1.2 Problem statement 1](#_Toc429570689)

[1.3 Possible solution 1](#_Toc429570690)

[1.4 How to demonstrate the quality of the solution 1](#_Toc429570691)

[2. Scope for this project 1](#_Toc429570692)

[2.1 Aim 1](#_Toc429570693)

[2.2 Objectives 2](#_Toc429570694)

[2.3 Deliverables 2](#_Toc429570695)

[3. Project schedule 2](#_Toc429570696)

[3.1 Methodology 2](#_Toc429570697)

[3.2 Tasks, milestones and timeline 2](#_Toc429570698)

[3.3 Risk assessment (if appropriate) 2](#_Toc429570699)

[References 3](#_Toc429570700)

[Appendix A. How ethical issues are addressed 3](#_Toc429570701)

## 1. Background Research for the project

### 1.1 Context

Intracranial aneurysms(IA) is a cerebrovascular disease that mainly occurs in the cerebral artery, it is distinguished by pathological dilation of blood vessels[3]-[6]. And IA has been common nowadays, there are approximately three out of every hundred people got IA, and each patients’ IA got chance to rupture, which will leading to death[2]. So this motivate me to do a project as a computer scientist, that would help doctors and researchers who are recently working on curing IA; furthermore, to solve some recent known problem in this project.

### 1.2 Problem statement

Clarification of the problem – for example, refinement of the problem that was given as a project idea, or emerged research hypothesis from some background study.

By known the high risk of death cause by the rupture IA, and look into some past researchs on the rupture risk of IA[8][9], these invent me to work on a solution of predicting rupture risk of IA using the known computer science methods and algorithm. Through some literature research, I find out some recent researchs about predicting the IA focus on machine learning[1][7], they train the machine learning model use the parameters of shape of the IA and patient status to predict the rupture risk of the IA; there also some other researches about using deep learning model such as common DNN and CNN, feed the model with images that been taked from different angles of 3D modelled IA and predict the rupture or risk of the IA.

So this motivate me to work on this project, which give a proper way to predict the possibility of rupture for patients’ IA, so the doctor or hospital can give these patients some appropriate treatment plans and surgeries.

### 1.3 Possible solution

Initial thoughts on possible solution(s) to be tackled in the project - what may or may not be pursued depending on speed of progress, challenges and how to approach the development of a potential solution, which modules or computing topics will be the building blocks for the solution.

Should cite background reading or systematic research conducted so far to provide justification of your initial thoughts.

Collect the relative data from different website, and preporocess these data and input them in to RNN, CNN, Random Tree and Logistic regression model to form a model that will give a prediction on the IA cure possibility.

### 1.4 How to demonstrate the quality of the solution

Calculate the p-value and accuracy for each model, present the ROC curve and AUC, compare them and conclude the result. Then compare my result with the models develop by others previously, conclude the adventsge and improvement of my models compare to theirs

## 2. Scope for this project

This section specifies what the project will deliver. It should be written in a concise manner, to be used as a basis for assessment.

### 2.1 Aim

The aim of the project is the overall top-level goal. It might be helpful to consider this in conjunction with the project title.

* Design and create Machine learning and Deep learning classifier model that classify cerebral aneurysms, base on the input data (morphological features of different cerebral aneurysms and simulation-derived features), give out a predicted succeed probability of the treatment by endovascular flow diversion.
* Compare these model, and improve the model that gives the best accuracy, then train the model again, try to improve the trained model to give a better prediction accuracy.

### 2.2 Objectiv**e**s

List up to five objectives. When you phrase an objective, think about how you can demonstrate its achievement.

To summarise, characteristics of suitable objectives are:

* Deliverable -- you will hand them in!
* Measurable -- examiners are able to judge/quantify if you have done a good job.
* Appropriate -- they should solve a sufficiently difficult problem.
* Agreed – by your supervisor, assessor, other members of the School (where appropriate).
* Background research on previous similar projects to generate a basic idea
* Understanding the data by give research on the data
* Choose and collect suitable data for model training from a different website
* Background research on different Deep learning and machine learning model, aim to choose suitable classifiers and neural network structures
* Design input data types and preprocessing methods for the data
* Design prototypes of different Deep learning and machine learning model
* Use Python to create the preprocessor for the original data, so that data can be input to the model
* Use the Python libraries torch and sklearning to create the models
* Train and test the models with the preprocessed data, and give labels to different cerebral aneurysms.
* Conclude the result by observing the prediction accuracy to evaluate different models
* Change the structure of models and see if there will be some improvement in the prediction accuracy
* Evaluate the models again, repeat the processes a couple of times
* Write the final report that compares the different models and concludes the findings

### 2.3 Deliverables

These are items for assessment under ‘delivery’. These could be written up as sections in the final project report (e.g. comparison of algorithms, or feasibility assessment, or design documentation, etc.) or be handed in separately (e.g. code, user manual or installation guide etc.). It is important to have the agreement from the supervisor at this stage that these deliverables are suitable as delivery for the type of project.

To ensure that these deliverables are within the scope of the project, cross-referencing to the objectives may be a helpful check.

Some examples of deliverables: an analysis of current infrastructure, a comparative study of techniques or tools, a recommendation to the client, requirements specification, design documentation, algorithms, software functionality, a qualitative or quantitative evaluation study, and so on, as appropriate for the type of project.

* Final report
* Python Code
* Machine Learning Model prototype
* trained deep learning or machine learning model
* Excutebale software that visualize the algorith

## 3. Project schedule

The schedule for completion of the project should relate the activities (or tasks) to objectives or deliverables. A few milestones should be identified for self monitoring of progress.

### 3.1 Methodology

Outline the underpinning project approach that is appropriate for the chosen type of project. This should help to plan for the order of the activities /tasks.

* By look into the dataset I observe that the data include 4 kinds of cut type for every sample of IA, cut1, cut2, dome and ninja. As some sample is missing data for cut2, so I will split the data in to 3 different training set and vaildation set base on the cut type cut1, dome and ninja.
* Use the morphology values, patient information(sex, age, location of IA) for 3 different kinds of cuttype as training data and validation data feed to Lasso, Ridge, Logistic regression and Random forest model; Next feed the same data to a 3 Layer MLP, 3 layer CNN and 3 Layer RNN(this might cause the poor performance, because Deep learning need large dataset).
* Evalutation the 21 models use prediction accuracy, p-value and AUC.
* If there is more time, do the PointNet or SpiderNet, evaluate the model

### 3.2 Tasks, milestones and timeline

Any appropriate method of presentation is acceptable. A common method is the use of Gantt chart.

### 3.3 Risk assessment (if appropriate)

If there is any risk identified at this stage (e.g. availability of stakeholders, technical issues or suitable test data etc.), mitigating strategy should be discussed.

* If the model used required to be a pretrained model, it may cause some ethical issue.

## References

Further guidance from Skills@Library: [**http://library.leeds.ac.uk/skills-referencing**](http://library.leeds.ac.uk/skills-referencing)**.**

Also see Resources on ‘Writing Tips' for other guidance.

1. Prediction of Intracranial Aneurysm Risk using Machine Learning
2. A deep-learning method for the end-to-end prediction of intracranial aneurysm rupture risk
3. Kim, J. Y. et al. Executive Summary of Stroke Statistics in Korea 2018: A Report from the Epidemiology Research Council of the Korean Stroke Society. J Stroke 21, 42–59, https://doi.org/10.5853/jos.2018.03125 (2019).
4. Kim, T. et al. Incidence and risk factors of intracranial aneurysm: A national cohort study in Korea. Int J Stroke 11, 917–927, https:// doi.org/10.1177/1747493016660096 (2016).
5. Lee, E. J. et al. Rupture rate for patients with untreated unruptured intracranial aneurysms in South Korea during 2006-2009. J Neurosurg 117, 53–59, https://doi.org/10.3171/2012.3.JNS111221 (2012).
6. Kim, T., Kwon, O. K., Ban, S. P., Kim, Y. D. & Won, Y. D. A Phantom Menace to Medical Personnel During Endovascular Treatment of Cerebral Aneurysms: Real-Time Measurement of Radiation Exposure During Procedures. World Neurosurg, https://doi. org/10.1016/j.wneu.2019.01.063 (2019).
7. Machine Learning Application for Rupture Risk Assessment in Small-Sized Intracranial Aneurysm
8. Asari, S. & Ohmoto, T. Natural history and risk factors of unruptured cerebral aneurysms. Clin Neurol Neurosurg 95, 205–214 (1993).
9. Ronkainen, A. et al. Risk of harboring an unruptured intracranial aneurysm. Stroke 29, 359–362 (1998).

## Appendix A. How ethical issues are addressed

This is a University requirement. See Resources on 'Ethics relevant to computing projects' for guidance and discuss it with your supervisor. If no ethical issue is involved, a sentence to that effect will suffice.