School of Computing

UNIVERSITY OF LEEDS

COMP5200M Scoping and Planning Document

Student Name: Wenjie Zhang
Programme of Study: Msc Advanced Computer Science
Provisional Title of Project: Machine learning classification of aneurysms using imaging and simulation data/ Machine learning and Deep learning method for predicting aneurysms rupture risk
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: **Wenjie zhang**	**Date:** **27/03/2024**

Contents

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]. IA has become common nowadays, there are approximately three out of every hundred people have IA, and each patient's IA has a risk of rupture, which will lead to death[2]. So this motivated me to do a project, that would help doctors and researchers who are recently working on curing IA; furthermore, to solve some recent known challenges in this project.

1.2 Problem statement

As the condition we stated in the last section, the IA of a patient has a risk of rupturing which will lead to death, which means an Unruptured intracranial aneurysm (UIA) become a life-threaten condition and treatment for it become more important nowadays. However, the UIA treatment which includes neurosurgical clipping and endovascular coiling has a risk of causing the IA rupture[11], so a proper measurement of the rupture risk for UIA is necessary before deciding the treatment plan for different patients.

When looking into some past research on the rupture risk of IA[8][9], we found out that the rupture risk of UIA depends on multiple factors of the IA and patient. As a computer science student, we believe there exist suitable computer science methods and algorithms to use these factors to measure the UIA rupture risk

All these made me work on this project, which aimed to give a proper way to predict the possibility of rupture for patients' IA using computer science algorithms and methods, so the doctor or hospital can give these patients some appropriate treatment plans and surgeries..

1.3 Possible solution

Through some literature research, we found some recent research about predicting the IA focus on machine learning[1][7], they train the machine learning model using the parameters of the shape of the IA and patient status to predict the rupture risk of the IA; there also some other researches about using deep learning models such as common Deep learning neural network (DNN) and convolutional neural network (CNN) [12][13], they feed the model with images that been taken from different angles of 3D modelled IA and predict the rupture or risk of the IA. Furthermore, recent research by H.Chao[14], uses the AneuX morphology database[15], which includes 750 IA samples to train a Point cloud-based Deep learning model called PointNet++ [17].

By looking through the research above, we conclude the possible way to predict the UIA rupture risk using the factors of IA patients is to use the machine learning and deep learning model, and the AneuX morphology database is a suitable database for training these models.

AneuX morphology database includes the 170 different morphometric parameters computed for the aneurysm models, which perfectly describe the shape of IAs and are suitable to feed to machine learning models.

As with some previous research [1][7], machine learning performs well with over 80% accuracy when fed with parameters of IAs and patients only, so some proper machine learning models can be trained using the AneuX morphology database and might give a good prediction. Moreover, a research by Detmer FJ. et al.[18] did a comparison for a previously developed IA rupture logistic regression model (LRM) with some machine learning classifiers, and concluded the performance for these models. Therefore, the classifier that has suitable performance in Detmer FJ. has been chosen in this project, which includes logistic regression, Lasso regression, Ridge regression and Random Tree, aimed to compare and conclude the best performance classifier or model when used on the AneuX database.

Research by Mirzat Turhon. et al.[12] has concluded a result the Deep learning model outperforms the traditional machine learning models when predicting UIA rupture risk using radiomics and morphological parameters. Concern this condition might also happen when fed with the AneuX morphology database, Deep learning models will also be trained using the AneuX database.

Also, we found out, that most research recently focuses on using the CNN deep learning model to predict UIA properties[1][12][13], it is suggested that there might be other Deep learning networks that give better predictions than CNN. One Deep learning model we found is Recurrent Neural Network (RNN), we will test this model in this project and compare the result with other machine learning and deep learning models in order to find the best solution model.

Moreover, because the AneuX database includes the 3D model of the IAs, it is also possible to feed these data to the Deep learning model that takes 3D models as input. There also some recent research [22][23] using PointNet[16], PointNet++[17], So-Net[19], SpiderCNN[20] and other 3D Point Cloud deep learning methods, training with the 3D model of the brain vessel and IAs gives a good performance on predicting some features of IA. So as for the next stage after the basic machine learning methods, CNN and RNN, some Point Cloud deep learning networks will be chosen and trained to compare with machine learning classifiers, CNN and RNN.

To sum up, we came up with a possible solution, which will use the AneuX database, preprocess these data and input them into RNN, CNN, Lasso regression, Random Tree and Logistic regression model to form a model that will give a prediction on the UIA rupture risk. If there is more time for another evaluation of the project, the PointNet and other models will be trained and compared with the models trained in the previous evaluation.

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 results with the models developed by others previously, and conclude the advantages and improvement of my models compared to theirs. Finally, look through the project, give the limitation of the work and possible future work.

2. Scope for this project

2.1 Aim

- Design and create some Machine learning and Deep learning classifier models that classify cerebral aneurysms, based on the input data (morphological features of different cerebral aneurysms and simulation-derived features), giving out a predicted rupture risk of a UIA.
- Compare these models, and try to improve the trained model to give a better prediction accuracy.
- Conclude the differences, pros and cons of the models
- Summarize the suitable prediction model for real-life conditionson

2.2 Objectives

- 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

- Final report
- Python Code
- Pre-trained Machine Learning Model
- trained deep learning or machine learning model

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.

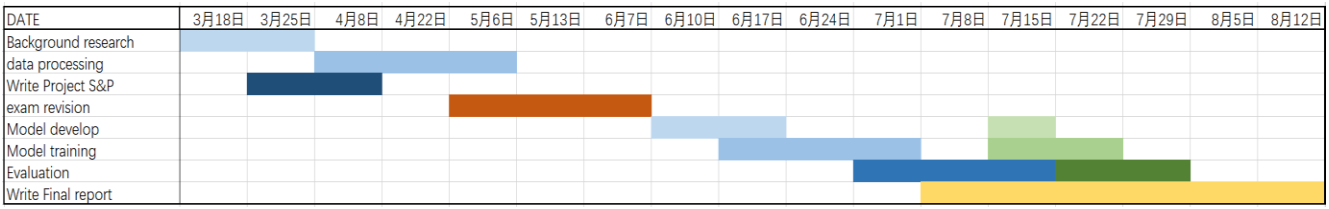
3.1 Methodology

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 looking into the dataset we observe that the data includes 4 kinds of cut types for every sample of IA, cut1, cut2, dome and ninja. As some samples are missing data for cut2, we will split the data into 3 different training sets and validation sets based on the cut types cut1, dome and ninja.
- Use the morphology values, and patient information(sex, age, location of IA) for 3 different kinds of cut type 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 needs large dataset).
- Give Evaluation for the 21 models using prediction accuracy, p-value and AUC.
- If there is more time, do the Point cloud deep learning andevaluate the models

3.2 Tasks, milestones and timeline

Project plan Gantt chart.



3.3 Risk assessment

- The Aneux database might be too small for deep learning model training
- The result might be used by others with out licences

References

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

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

[1] Jaehyuk Heo. et al. Prediction of Intracranial Aneurysm Risk using Machine Learning. Nature, Scientific reports (2020)

[2] Peiying Li, Yongchang Liu. et al. A deep-learning method for the end-to-end prediction of intracranial aneurysm rupture risk. Patterns 4, 100709 (2023)

[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] Kim HC, Rhim JK, Ahn JH, Park JJ, Moon JU, Hong EP, Kim MR, Kim SG, Lee SH, Jeong JH, et al. Machine Learning Application for Rupture Risk Assessment in Small-Sized Intracranial Aneurysm. *Journal of Clinical Medicine*. (2019);

[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).

[10] Hackenberg KAM, Hänggi D, Etminan N. Unruptured intracranial aneurysms. Stroke 49(9):2268–2275 (2018)

[11] Algra AM, Lindgren A, Vergouwen MDI .et al. Procedural clinical complications, case-fatality risks, and risk factors in endo vascular and neurosurgical treatment of unruptured intracranial aneurysms: a systematic review and meta-analysis. JAMA Neurol 76(3):282–293(2019)

[12] Mirzat Turhon .et al. Development and validation of a deep learning model for prediction of intracranial aneurysm rupture risk based on multi-omics factor. European Society of Radiolog (2023)

[13] Yang H, Cho K- C, Kim J- J, et al. Rupture risk prediction of cerebral aneurysms using a novel convolutional neural network- based deep learning model. J NeuroIntervent Surg 2023;15:200–204. (2022)

[14] Cao H, Zeng H, Lv L, Wang Q, Ouyang H, Gui L, Hua P and Yang S, Assessment of intracranial aneurysm rupture risk using a point cloud-based deep learning model. Front. Physiol. 15:1293380. (2024)

[15] Juchler, N., Bijlenga, P., & Hirsch, S. AneuX morphology database (v1.0) (2022). [Data set]. Zenodo. https://doi.org/10.5281/zenodo.6678442

[16] R. Q. Charles, H. Su, M. Kaichun and L. J. Guibas, "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation," *2017 IEEE Conference on Computer*

Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 77-85, doi: 10.1109/CVPR.2017.16.
[17] Charles R. Qi, Li Yi, Hao Su, and Leonidas J. Guibas. PointNet++: deep hierarchical feature learning on point sets in a metric space. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 5105–5114.90. (2017)
[18] Detmer FJ, Lückehe D, Mut F, Slawski M, Hirsch S, Bijlenga P, von Voigt G, Cebral JR. Comparison of statistical learning approaches for cerebral aneurysm rupture assessment. Int J Comput Assist Radiol Surg. 2020 Jan;15(1):141-150. doi: 10.1007/s11548-019-02065-2. Epub 2019 Sep 4. PMID: 31485987.
[19] Jiaxin Li, Ben M. Chen, Gim Hee Lee; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9397-9406 (2018)
[20] Yifan Xu, Tianqi Fan, Mingye Xu, Long Zeng, and Yu Qiao. 2018. SpiderCNN: Deep Learning on Point Sets with Parameterized Convolutional Filters. In Computer Vision – ECCV 2018: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VIII. Springer-Verlag, Berlin, Heidelberg, 90–105. (2018) https://doi.org/10.1007/978-3-030-01237-3_6
[21] Wang J, Liu J, Xu Z, et al. N-PointNet: A multi-layer embedded deep learning model for 3D intracranial aneurysm classification and segmentation. Research Square; 2023. DOI: 10.21203/rs.3.rs-2760399/v1. (2023)
[22] D. Shao, X. Lu and X. Liu, "3D Intracranial Aneurysm Classification and Segmentation via Unsupervised Dual-Branch Learning," in *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 4, pp. 1770-1779, April 2023, doi: 10.1109/JBHI.2022.3180326. (2023)
[23] X. Yang, D. Xia, T. Kin and T. Igarashi, "IntrA: 3D Intracranial Aneurysm Dataset for Deep Learning," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2020, pp. 2653-2663, doi: 10.1109/CVPR42600.2020.00273.(2020)

Appendix A. How ethical issues are addressed

- The machine learning and deep learning model used might have copyright and usage licence issues, more attention needs to be paid to these when using models
- The AneuX database used in this project is public but still has its own copyright and usage license issues, these have to be concerned when through the project