Project Plan Felicia:

v1-2: Literature reading, establishing data structures / data + organisation of necessary existing code

v3-8: Practical data analysis, interpretation, survival analysis

v9: Documentation of code and report writing

v10: Report writing and preparation of presentation

Breast cancer is a major global health concern, affecting millions of women each year. Among the different subtypes of breast cancer, triple-negative breast cancer (TNBC) is the most aggressive, with poor prognosis and a high risk of relapse. Recent studies have highlighted the role of PD-L1 and anti-PD-L1 in breast cancer, with immunotherapy showing promising results in patients with PD-L1-positive tumours. However, accurately scoring PD-L1 expression can be challenging due to inter- and intraobserver variability in pathologist scoring.

In this 10-week project, the student will investigate the utility of a deep learning (DL) model for PD-L1 scoring in TNBC. Specifically, the student will test a DL model trained on external data to assess its performance on our SCAN-B TNBC dataset. The student will compare the DL scores generated by the model with pathologist scores and automated cell counts, with the goal of identifying the most accurate and reliable approach for PD-L1 scoring in TNBC.

In addition, the student will perform survival analysis on PD-L1 scores to determine whether PD-L1 expression is associated with patient outcomes in our cohort. The ultimate goal of this project is to improve our understanding of PD-L1 expression in TNBC and develop more reliable methods for PD-L1 scoring that can be applied in the clinic