# 7CS997 Independent Studies Progress Report

## Student/supervisor details

**Submission deadline for signed and completed progress report:**

* **Progress report 1: 16/02/2024 12:00 Midday**

Project Title: **XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX**

Name: **XXXXXXXXXXXXXX**

Student Id: **XXXXXXXXX**

MSc: Big Data

Supervisor: **Dr Oluwarotimi Samuel**

**Progress Report: 1**

## 1. Record of meeting

On the 9th of February 2024, a meeting was held to discuss the progress of my project. I provided an update in line with the Gantt chart, emphasizing my current focus on exploring pertinent literatures. My introductory chapter was presented and reviewed, also the choice of data for my project was discussed and approved by my supervisor .

Our discussion encompassed various algorithms, including Random Forest, Support Vector Machine, Decision Tree, Naive Bayes, Logistic Regression, and the Ensemble Method. Within this conversation, my supervisor clarified the distinction between Random Forest and the broader concept of the Ensemble Method.

In the discourse on literature review, I presented the reviewed literatures and their relevance to my study. My supervisor provided guidance on effective literature review techniques tailored to the objectives of my project. The critical importance of conducting a thorough literature review to obtain insights for my project was also emphasized by my supervisor.

The meeting concluded with the assignment for me to continue and complete the reading of relevant literatures according to the Gantt chart. The next meeting is scheduled for 19/02/2024.

## Evaluation of progress

Drawing from our February 9th meeting and my project outline, my engagement with crucial literature aligns well with the project's initial stage, aiming to create a robust foundation for my research. The approval of my choice of data and algorithms signifies a key development in the research's methodological approach. The valuable guidance and insights from my supervisor have been essential in refining my approach, especially in deciding the most appropriate machine learning models for my project. Looking ahead, with the literature review in progress and our next meeting scheduled for February 19th, my path towards meeting the project's goals is clearly marked.

This progress coupled with the completion of the introductory chapter, not only demonstrates engagement with my supervisor but also a commitment to the strategic execution of the project, charting a direct path for the upcoming comparative analysis and subsequent phases of my project.

## Actions/targets for next meeting

For my upcoming meeting with the supervisor, the primary agenda will involve discussing key progress points in my research project. By this point, the literature review is anticipated to be completed. I will shift my focus to emerging aspects of the project and initiate discussions on follow-up steps to be presented for feedback.

**Literature Review Recap**: Provide a concise summary of the literature review, highlighting the key findings, insights, and gaps identified in the existing research related to machine learning applications in breast cancer prediction and prognosis.

**Data Collection**: In this phase, I will present to my supervisor the advancements on the data collection and preprocessing journey on the acquired datasets for breast cancer. I have already shared with my supervisor the dataset obtained from Kaggle.

**Methodological Framework**: Develop a robust methodological framework that outlines the overall approach to be taken in my research. I will refine the specific research objectives based on insights gained from the literature review to ensure that the objectives align with the goal of enhancing the accuracy, reliability, and utility of machine learning models in predicting and prognosing breast cancer.

**Dataset Preparation and Exploration**: Continue with data preprocessing by thoroughly preparing the breast cancer datasets. Explore and gain a deeper understanding of the dataset characteristics, identifying potential challenges or biases that may need addressing during the analysis. This not only ensures the integrity of subsequent analyses but also positions the research to uncover meaningful insights that contribute to the advancement of breast cancer prediction and prognosis.

**Model Selection Decision**: Based on the characteristics and insights gained from the dataset, I will proceed to select appropriate machine learning models for breast cancer prediction and prognosis.

This comprehensive agenda highlights a holistic approach to advancing the research, from literature review synthesis to methodological refinement, data preparation, and model selection. Every aspect collaboratively works towards the main objective of improving the accuracy, reliability, and usefulness of machine learning models in predicting and understanding breast cancer.

## 4.Record of meetings held

*The table below should contain the meeting schedule that you planned with your supervisor at the start of the project. In the “Confirmation” column on the right, you should confirm the meeting has been held by including a “Y” or “N” in the relevant box. If there has been a need to change the original meeting schedule planned, you should update this table.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Project week | Meeting date | Meeting time | Venue | Confirmation |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

XXXXXXXXXXX XXXXXXXXXXXXXXXXXXX

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Student Name Supervisor Name

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Student Signature Supervisor Signature

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date Date