COMP4026 Project Plan

Student: Tzu-Chun Yu

Student ID: 20700115

Course: MSc Projects

1. Project Title

Towards a Knowledge-Augmented Chatbot for Personalised Breastfeeding Support

2. Statement of the Research Problem

Breastfeeding mothers often face challenges when using breast pumps, such as discomfort, uncertainty about optimal suction settings, and lack of timely support. Existing apps provide basic tracking but rarely offer personalised, context-aware guidance. Chatbots built on general breastfeeding knowledge deliver uniform responses that fail to address individual needs. To address this gap, this project explores how behavioural data from breast pump usage—such as suction levels, session patterns, and user feedback—can be integrated with expert knowledge to build a chatbot that delivers personalised, real-time support. By enabling AI systems to adapt to individual conditions, we aim to improve user experience and offer meaningful, scalable support to breastfeeding mothers.

3. Related Work

Breastfeeding support technologies have advanced in recent years, particularly through mobile health (mHealth) apps and AI chatbots. For example, **LactApp** uses decision trees to deliver personalised answers [4][3]. Similarly, chatbots like **Lhia** apply NLP to answer common breastfeeding questions [5]. These systems demonstrate high usability and value in guiding mothers through common challenges.

However, most current solutions lack integration with real-world behavioural data—particularly from breast pump usage—limiting their ability to adapt dynamically to each user's situation. Studies show that comfort and pain are major factors affecting continued pump use [2], and research in HCI highlights the value of personalisation, real-time tracking, and adjustable suction patterns for improving user experience [6].

Moreover, while parents generally support breastfeeding, they often face emotional, physical, and environmental challenges that are not easily addressed by generic advice [1]. This suggests a need for systems that can learn from users' lived experiences and usage data.

To address this gap, this project proposes a knowledge-augmented chatbot that combines trusted external sources with personalised insights from breast pump usage patterns. Using EDA and rule-based methods, the system will generate user-specific profiles, identify behavioural trends, and deliver tailored responses via Retrieval-Augmented Generation (RAG). This approach aims to enhance user comfort, optimise pump settings, and support mothers with context-aware guidance grounded in both data and empathy.

4. Methodology

This dataset is provided by the industry partner, Biamo Inc., and consists of 796 simulated entries derived from breast pumps. While synthetic, the data is designed to mirror real-world distributions and will be processed with anonymisation procedures in the final dissertation.

This project is divided into five main phases combining data understanding and EDA, user profile extraction and rule-based behaviour modelling, knowledge constructuion, personalised chatbot development and integration, and demo page implementation and testing. First, exploratory data analysis (EDA) will be used to identify global usage patterns to recommend baseline pump settings. Second, rule-based methods will be applied to extract personalised usage profiles—such as preferred suction settings or short sessions. In parallel, a knowledge base of trusted breastfeeding resources (e.g., WHO, NHS) will be constructed and embedded using vector-based retrieval techniques. A Retrieval-Augmented Generation (RAG) framework will then be developed to allow the chatbot to generate personalised responses by combining user behaviour summaries with relevant external content.

The system will be implemented on the AWS platform, utilising services such as AWS SageMaker, AWS S3, AWS Lambda, and AWS EC2. For chatbot integration, the project will consider using HuggingFace Models. A lightweight web interface will be built to demonstrate the chatbot and display behavioural insights. To evaluate the solution, usability testing will be conducted through an online questionnaire focusing on user satisfaction and perceived support.

This approach integrates behavioural data and expert knowledge in a user-centred way to deliver practical, personalised breastfeeding support.

5. Programme of Work

- Work Package 1 (WP1) – Literature Review and Project Planning

  Conduct a focused literature review on breastfeeding technologies, personalised health tools, and chatbot applications. Identify research gaps and define the project scope, objectives, and work plan.

  Milestone 1 (M1) – Literature review and project proposal document finalised.

- Work Package 2 (WP2) – Data Understanding and EDA

  Explore breast pump usage datasets to understand behavioural trends such as session duration, suction levels, and frequency. Tasks include cleaning and structuring session logs, identifying global usage patterns for initial setup while using breast pumps by conducting EDA.

  Technologies used: Python (Pandas, Matplotlib, Seaborn), Jupyter Notebook via AWS SageMaker, AWS S3

  Milestone 2 (M2) – Summary report of usage patterns and EDA results completed.

- Work Package 3 (WP3) – Behavioural Profiling and Rule-Based Modelling

  Develop individual user profiles based on pump usage data by identifying key indicators such as frequent suction adjustments or short sessions. Extract each user's preferred configurations and store personalised summaries in a structured database.

  Technologies used: DynamoDB

  Milestone 3 (M3) – Personalised user profiles and rule-based logic implemented.

- Work Package 4 (WP4) – Knowledge Base and RAG Integration

  Collect authoritative breastfeeding resources (e.g., WHO, NHS), process them into retrievable formats, and embed them for semantic search. Integrate the knowledge base into a Retrieval-Augmented Generation (RAG) framework.

Technologies used: HuggingFace Models (all-MiniLM-L6-v2), FAISS or OpenSearch on EC2

Milestone 4 (M4) – Functional knowledge base with working vector search integrated into the system.

- Work Package 5 (WP5) – Personalised Chatbot Development

  Design and implement a chatbot system that combines individual user data with retrieved knowledge to generate personalised responses. This involves prompt design, API integration, and response handling.

  *Technologies used: HuggingFace Models (*microsoft/DialoGPT-medium*)*

  Milestone 5 (M5) – Working prototype of the chatbot capable of delivering tailored responses.

- Work Package 6 (WP6) – Demo Interface and Usability Testing

  Create a simple web-based interface that connects to the chatbot and visualises behavioural summaries. Conduct usability testing through an online questionnaire to assess user satisfaction and relevance of responses.

  Technologies used: Streamlit and Flask or AWS Lambda, online survey tool (e.g., Google Forms or Microsoft Forms)

  Milestone 6 (M6) – Completed usability evaluation with analysed feedback and demo interface deployed.

- Work Package 7 (WP7) – Dissertation Writing

  Prepare the final report by summarising results, methodology, evaluation, and key reflections. Include time for feedback and proofreading.

  Milestone 7 (M7) – Final dissertation submitted.

- Work Package 8 (WP8) – Presentation and Demo Preparation

  Prepare slides, demo, and summary to present the project's goals, methods, and results. Rehearse and ensure the chatbot runs smoothly.

  Milestone 8 (M8) – Final project presentation delivered.

The diagram is shown in Figure 1, and the simulated web pages are shown in Figure 2.

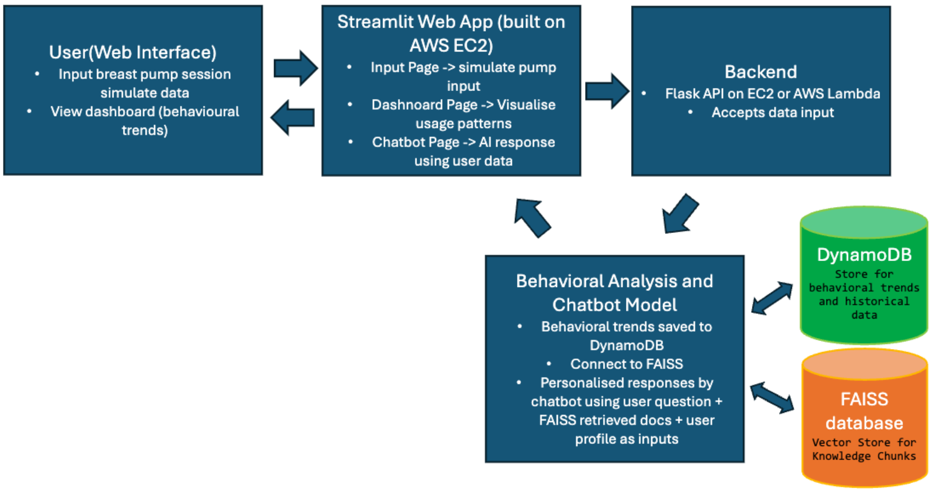
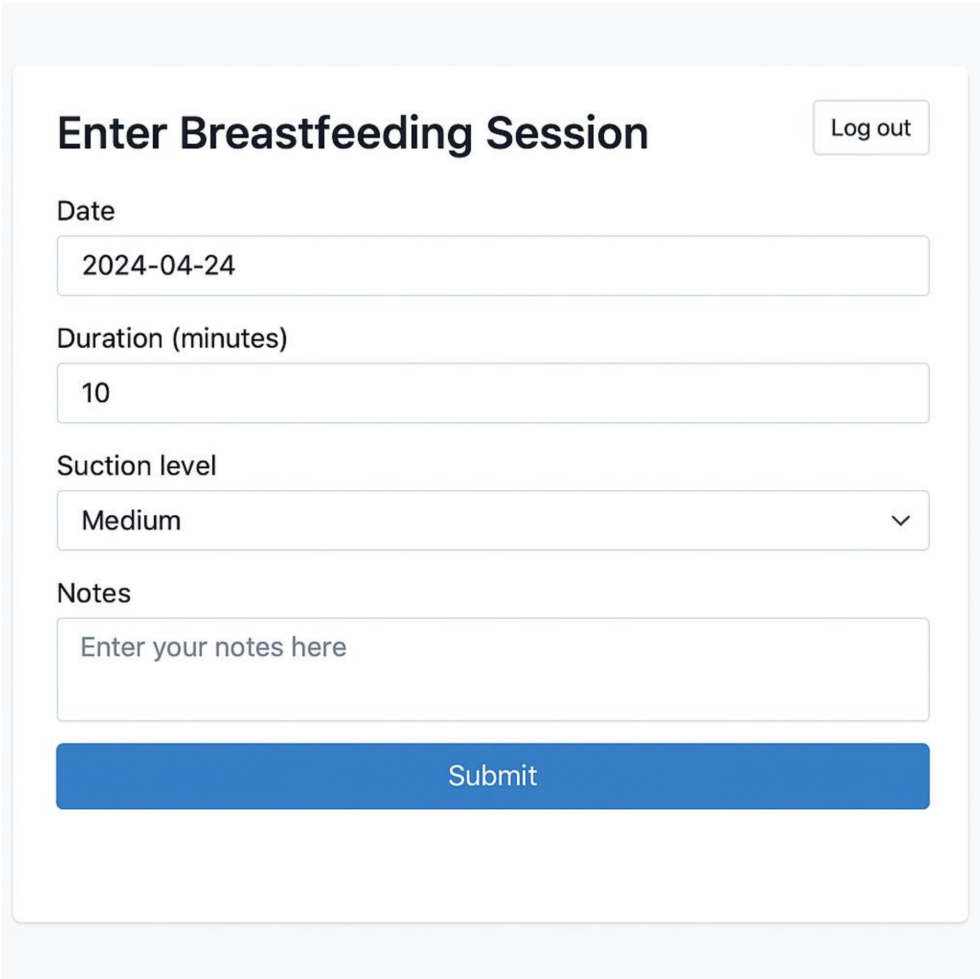
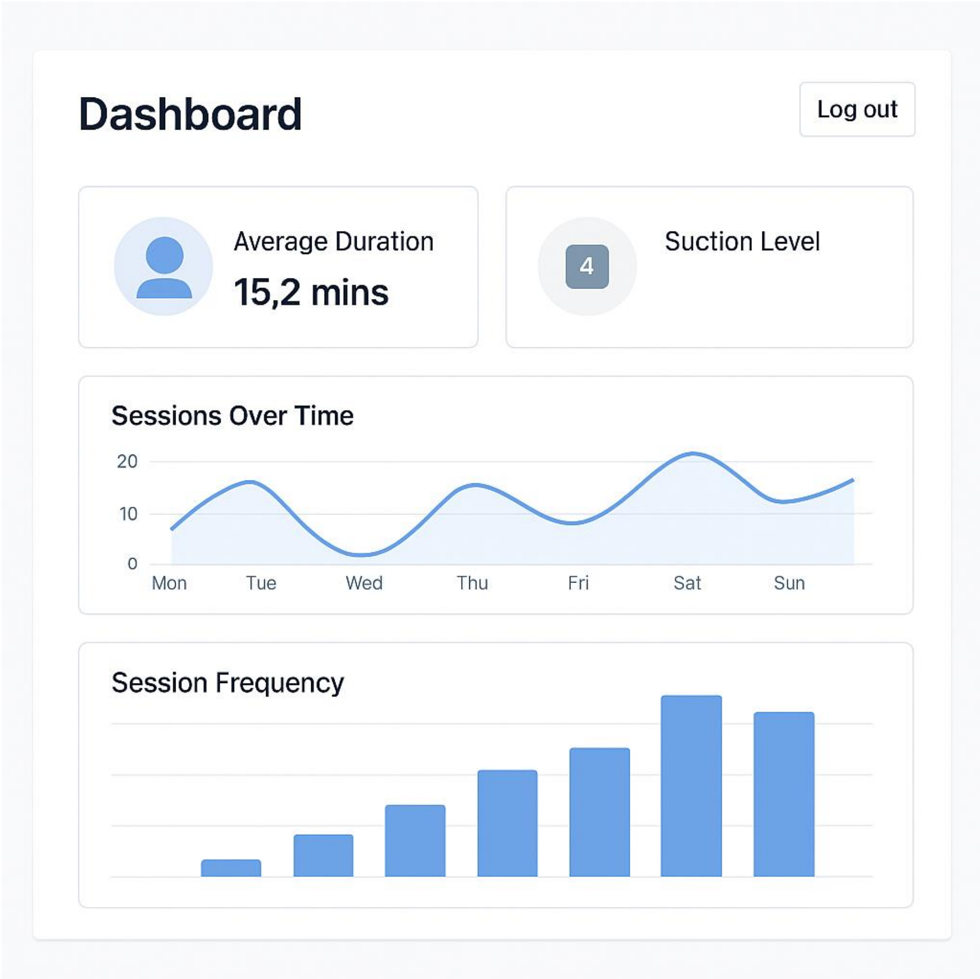


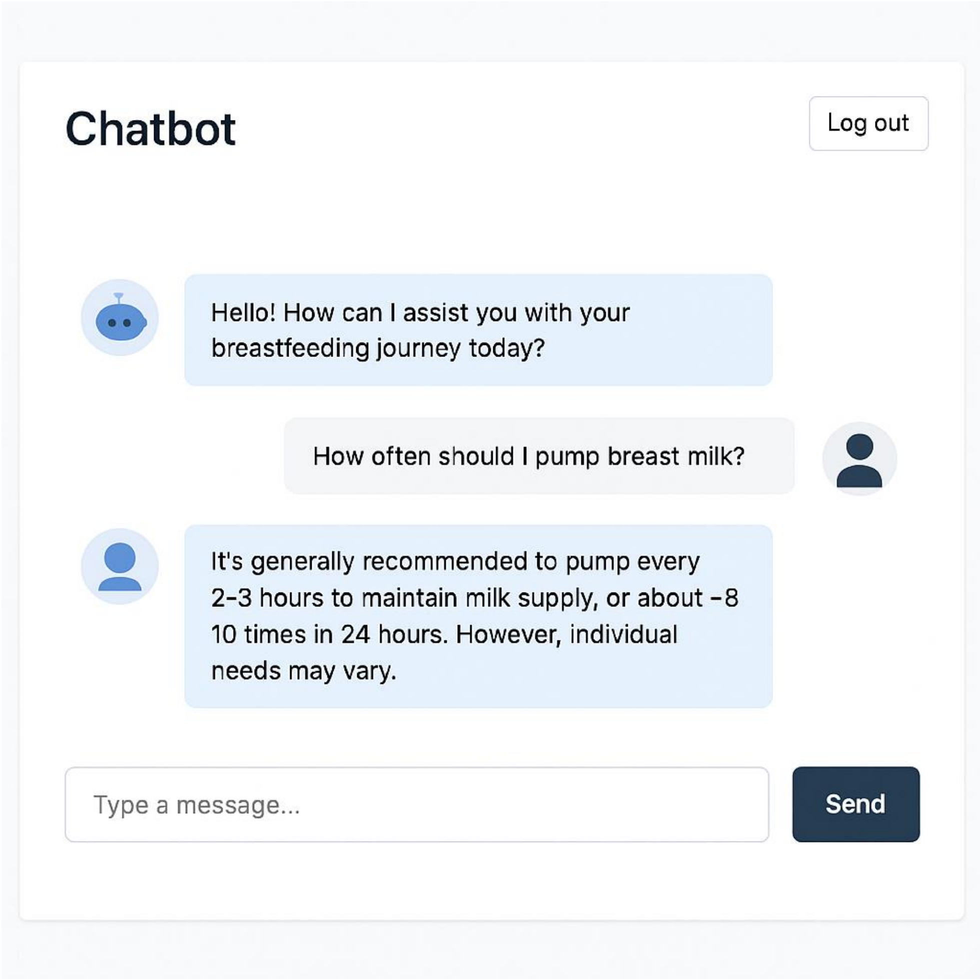
Figure 1. Personalised Breastfeeding Chatbot Pipeline



(a) Breast pump data input session



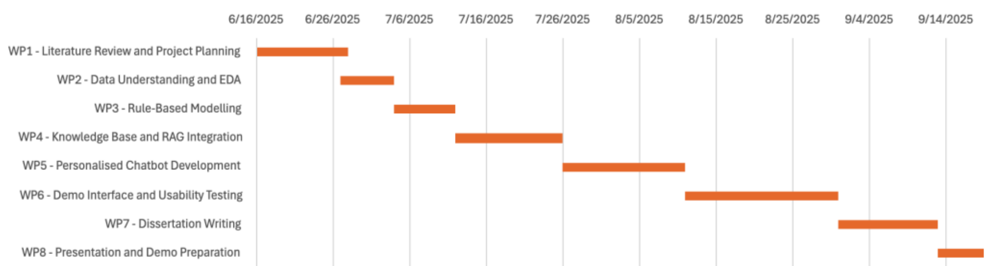
(b) personalised historical data



(c) Chatbot interface

Figure2. Web Application Conceptual Diagram

6. Time arrangement



Task Name	Start Date	End Date	Duration
WP1 – Literature Review and Project Planning	6/16/2025	6/27/2025	12
WP2 – Data Understanding and EDA	6/27/2025	7/3/2025	7
WP3 – Rule–Based Modelling	7/4/2025	7/11/2025	8
WP4 – Knowledge Base and RAG Integration	7/12/2025	7/25/2025	14
WP5 – Personalised Chatbot Development	7/26/2025	8/10/2025	16
WP6 – Demo Interface and Usability Testing	8/11/2025	8/30/2025	20
WP7 – Dissertation Writing	8/31/2025	9/12/2025	13
WP8 – Presentation and Demo Preparation	9/13/2025	9/18/2025	6

7. Reference

[1] K. Tang, K. Gerling, L. Geurts, and K. Spiel, ‘Understanding the role of technology to support breastfeeding’, in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–13.

[2] G. E. Becker, ‘Measuring mothers’ viewpoints of breast pump usage’, International Journal of Environmental Research and Public Health, vol. 18, no. 8, p. 3883, 2021.

[3] A. A. S. Almohanna, S. Meedya, E. Vlahu-Gjorgievska, and K. T. Win, ‘Exploring user experiences with a persuasive mHealth app for breastfeeding: an empirical investigation’, International Journal of Human–Computer Interaction, pp. 1–18, 2024.

[4] P. Quifer-Rada, L. Aguilar-Camprubí, A. Padró-Arocas, and D. Mena-Tudela, ‘Evaluation of the usability and utility of LactApp, a mHealth for breastfeeding support’, International Journal of Medical Informatics, vol. 180, p. 105240, 2023.

[5] S. Agudelo-Pérez, D. Botero-Rosas, L. Rodríguez-Alvarado, J. Espitia-Angel, and L. Raigoso-Díaz, ‘Artificial intelligence applied to the study of human milk and breastfeeding: a scoping review’, International Breastfeeding Journal, vol. 19, no. 1, pp. 1–15, 2024.

[6] C. D’Ignazio, A. Hope, B. Michelson, R. Churchill, and E. Zuckerman, ‘A Feminist HCI Approach to Designing Postpartum Technologies: “ When I first saw a breast pump I was wondering if it was a joke”’, in Proceedings of the 2016 CHI conference on human factors in computing systems, 2016, pp. 2612–2622.