

Readify - Intelligent Assistant to Improve Reading and Comprehension Skills in English Language

Project ID: 24-25J-27

LOG BOOK

Sooriyaarachchi M.D.A – IT21173790

BSc. (Hons) Degree in Information Technology

Specializing in Software Engineering

Department of Computer Science and Software Engineering


Sri Lanka Institute of Information Technology Sri Lanka

Table of Contents

DECLARATION	2
1. Monthly Progress	3
2. Individual Task Allocation Process	11
2.1 Assigned Tasks.....	11

DECLARATION

I declare that this is my own work and this Logbook does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
Sooriyaarachchi M.D.A	IT21173790	

Signature of the Supervisor
(Dr. Dasuni Nawinna)

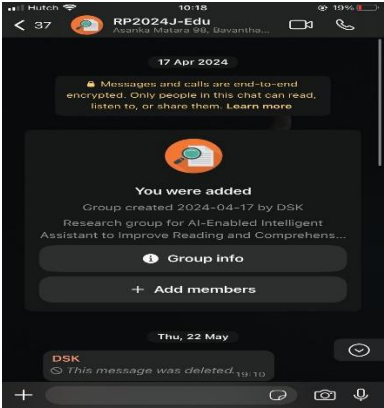
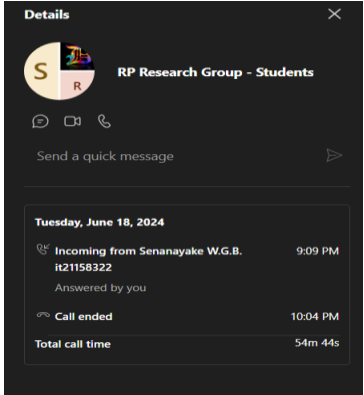
Date

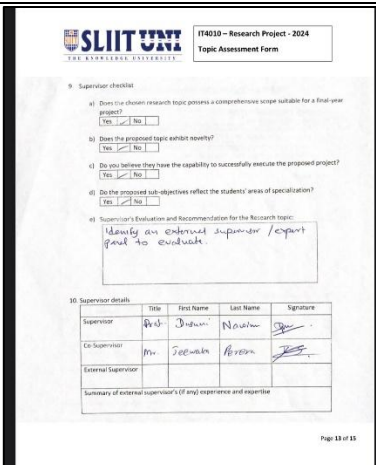
.....

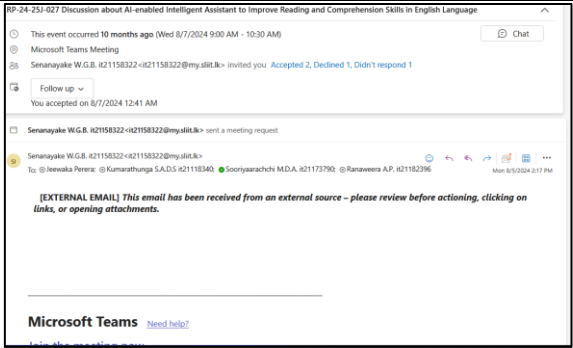
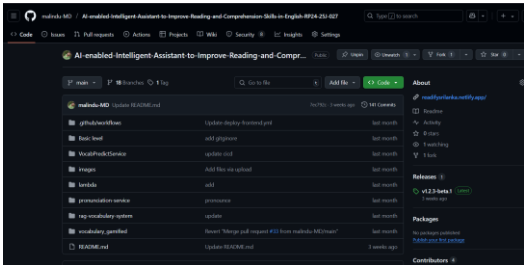
.....

1. Monthly Progress

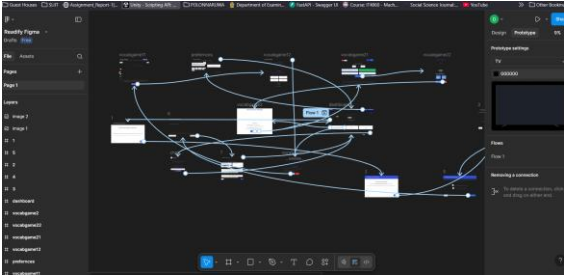
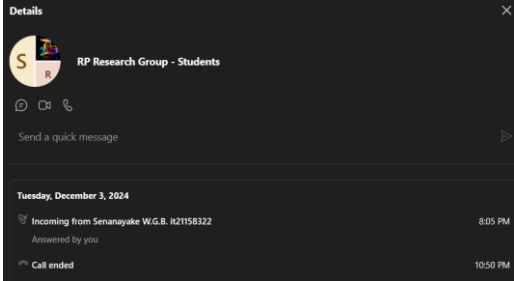
Month	Details and notes of work carried out, problems encountered and how solved etc., sketches and dimensions to be given wherever possible.
2024 -April	<ul style="list-style-type: none"> Reaching out to eligible supervisors to request their acceptance of our research proposal. (17.04.2024) <div data-bbox="610 573 1398 829"> </div> <div data-bbox="610 869 1398 1218"> </div> <div data-bbox="600 1268 1409 1675"> </div> <ul style="list-style-type: none"> Initial Discussion on Research Topic Functions via Teams with the supervisor (19.04.2024) Dr. Dasuni Nawinna Accepted our Proposal (19.04.2024)

<p>2024- May</p>	<ul style="list-style-type: none"> Supervisor Assigned Mr. Jeewaka Perera as a Co- Supervisor and Created WhatsApp Group.  <ul style="list-style-type: none"> Discussion on Research Topic Functions with our Co Supervisor at University (29.05.2024)
<p>2024 – June</p>	<ul style="list-style-type: none"> Created the TAF Document (15.06.2024 – 16.06.2024)  <ul style="list-style-type: none"> Discussion on Research Topic Functions with our Co Supervisor at University (20.06.2024) Signed and verified the TAF document from Supervisor, Co- Supervisor (20.05.2024 -21.05.2024)

	 <ul style="list-style-type: none"> Submitted the TAF Document (24.06.2024)
2024 - July	<ul style="list-style-type: none"> Topic Assessment evaluation. (05.07.2024) Accepted the Topic with minor changes (05.07.2024) Analyzed traditional vocabulary learning approaches. Researched AI techniques in educational apps. Conducted initial study on AI and vocabulary learning. Compared tools like Duolingo, Quizlet, and Memorize. Researched CEFR framework and gamification principles. Identified key technologies and tools for implementation. Studied CEFR proficiency levels and their implications. Explored possible model architectures (ML vs LLM).
2024 -August	<ul style="list-style-type: none"> Preparing for the Proposal presentation with the co-supervisor (02.08.2024) Prepared the Proposal presentation slides (03.08.2024) Preparing for the Proposal presentation with the Team (04.08.2024) Participated in the proposal presentation and accepted (Minor changes) with the research component (05.08.2024) Discussion about changes in the research project component with our co-supervisor on MS Teams (07.08.2024) Studied RAG, Lang Chain, Gemini, and Pinecone. Studied about Machine Learning (Supervised Learning)

	 <ul style="list-style-type: none"> Created proposal report (20.08.2024 – 23.08.2024) Meeting with the supervisor to obtain the signature for the proposal report (24.08.2024). Proposal report submitted to cloud folder (28-08-2024)
2024 – September	<ul style="list-style-type: none"> Discussion about Future Plans with our co-supervisor on Phone Call (10.09.2024) Collected CEFR-labeled learner text datasets. Created structured vocabulary dataset with metadata. Cleaned and normalized all text data. Applied POS tagging using spaCy. Integrated Named Entity Recognition for semantic depth. Analyzed class distribution across CEFR levels. Prepared training/test split (CSV format). Visualized word frequency and readability levels. Explored sentence embedding approaches. Team Discussion on Campus (22.09.2024) Found an external supervisor who is an English teacher and had a discussion with her. (30.09.2024)
2024 – October	<ul style="list-style-type: none"> Created GitHub Repository (02.10.2024) 

	<ul style="list-style-type: none"> • Designed overall system flow for retrieval and generation. • Studied semantic embedding with MiniLM model. • Created Pinecone vector database and schema. • Indexed vocabulary entries with semantic context. • Sketched RAG component-level diagrams. • Explored Lang Chain integration steps. • Mapped CEFR levels to vocabulary difficulty tiers. • Reviewed prompt engineering strategies. • Simulated vector searches using sample inputs. • Committed and pushed all implemented features to the 'vocab' branch in the Git repository. • External supervisor's physical meeting (26.10.2024) • Team discussion (26.10.2024)
2024 – November	<ul style="list-style-type: none"> • Team Discussion on Campus (05.11.2024) • Designed custom features: lexical richness, syntactic depth. • Calculated readability scores using Textstat. • Measured average sentence and word lengths. • Extracted token-level statistics (nouns, verbs, adjectives). • Developed parse tree depth algorithm. • Analyzed correlation between features and CEFR level. • Implemented named entity density measures. • Scaled features for model training. • Visualized feature importance using plots. • Finalized and stored feature vectors for training. • Trained baseline models (Logistic Regression, SVM, XGboost). • Implemented XGBoost for multiclass classification. • Evaluated model using accuracy, precision, recall, F1. • Visualized confusion matrix for performance analysis. • Saved trained model using pickle for deployment. • Committed and pushed all implemented features to the 'vocab' branch in the Git repository. • Created web application sketches, wireframes, and prototypes, along with discussions with team members (10.11.2024 – 20.11.2024).

	
<p>2024 – December</p>	<ul style="list-style-type: none"> • Team Discussion on Campus (01.12.2024) • Prepared the PP1 Presentation slides (02.12.2024) • Prepared for the PP1 Presentation with team (03.12.2024)  <ul style="list-style-type: none"> • Faced PP1 and got comments from the panel (04.12.2024) <ul style="list-style-type: none"> Panel Comment – Improve presentation skills • Submitted Checklist I (Git Repository) to the cloud location (CDAP) (08.12.2025) • Got a meeting with the co-supervisor to address issues found after PP1(16.12.2024) • Defined hyperparameter ranges for XGBoost tuning. • Used Bayesian optimization (BayesSearchCV). • Compared tuning strategies (grid vs Bayesian). • Improved validation accuracy with new parameters. • Retrained using best parameters and handcrafted features. • Designed game types: MCQs, fill-in-the-blank, synonym match. • Created prompt templates for Gemini LLM. • Developed game generation API logic. • Validated game output relevance and accuracy. • Developed chatbot prompt-response handling.

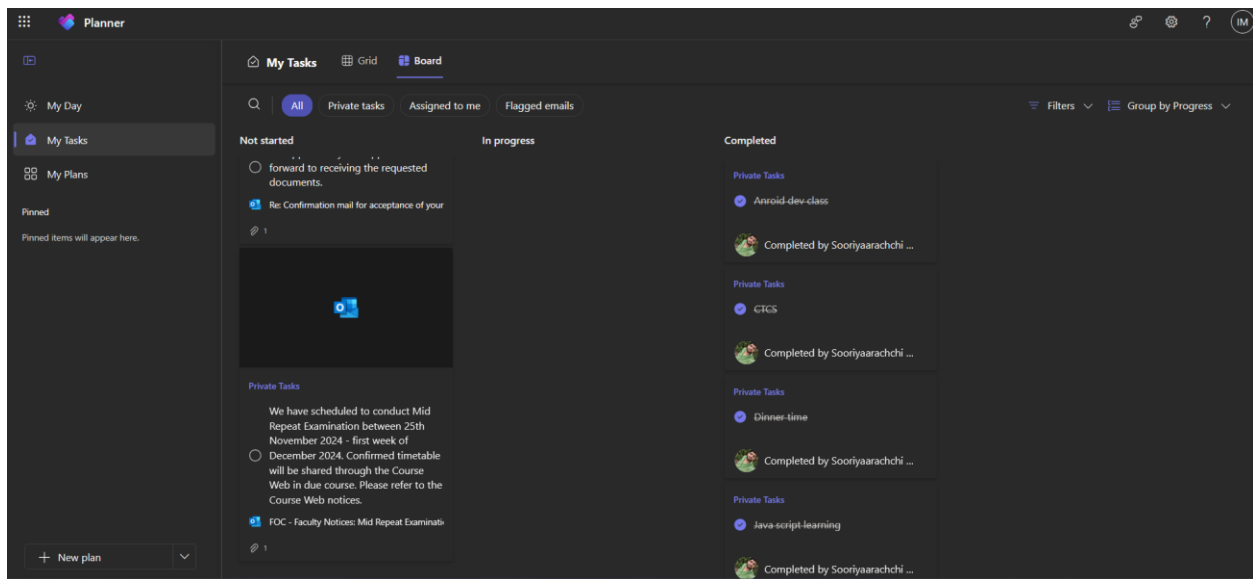
	<ul style="list-style-type: none"> • Personalized output based on user CEFR level and topic. • Finalized back-end logic for game generation and vocabulary level predict. • Committed and pushed all implemented features to the 'vocab' branch in the Git repository.
2025- January	<ul style="list-style-type: none"> • Team discussion (02.01.2024, 10.01.2024, 21.01.2024) • Initialized frontend project with Vite and React.js. (04.01.2024) • Designed CEFR prediction input and result display. • Created dynamic game interfaces for different levels. • Implemented chatbot UI for vocabulary support. • Styled components with Tailwind CSS. • Made UI responsive for mobile and desktop. • Improved UI elements based on user feedback • Committed and pushed all implemented features to the 'vocab' branch in the Git repository.
2025- February	<ul style="list-style-type: none"> • Merge other team members components through the GitHub (10.02.2025) • Attended an online session on "How to Write a Research Paper" conducted by our supervisor (11.02.2025, 18.02.2025) • Created research paper draft (09.02.2025 - 14.02.2025) • Submitted research paper draft (15.02.2025) • Integrated CEFR prediction and RAG game generation. • Configured CORS for frontend-backend communication. • Implemented error handling and retries. • Team discussion regarding PP2 (22.02.2025) • Arranged a meeting with the new external supervisor (27.02.2025) • Committed and pushed all implemented features to the 'vocab' branch in the Git repository.
2025-March	<ul style="list-style-type: none"> • Submitted final research paper to the cloud location (CDAP) (06.03.2025) • Created UI/UX Demo Video (15.03.2025) • Team discussion (07.03.2024, 11.03.2024) • Prepared PP2 Presentation slides (16.03.2025) • Prepared for PP2 presentation with the team (18.03.2025) • Submitted Checklist II to the cloud location (CDAP) (17.03.2025)

	<ul style="list-style-type: none"> • Participated in PP2 presentation and viva. (19.03.2025) • Created detailed test plan and strategy. • Wrote unit tests for CEFR model pipeline. • Validated RAG output with test inputs. • Evaluated chatbot responses for correctness. • Committed and pushed all implemented features to the 'vocab' branch in the Git repository
2025-April	<ul style="list-style-type: none"> • Created Final report (01.04.2025 - 11.04.2025) • Submitted the final report to the cloud location (11.04.2025) • Enhanced the user interfaces and the user experience under the guidance of the external supervisor (17.04.2025) • Initiated the submission of the research paper to journals and conferences. (18.04.2025) • Team discussion (20.01.2024, 28.01.2024) • Committed and pushed all implemented features to the 'vocab' branch in the Git repository • Merge members' components through the git hub (30.04.2025)
2025-May	<ul style="list-style-type: none"> • Discussed testing methods for my component with the external supervisor (05.05.2025). • Deployed services using Docker containers. • Hosted APIs on aws Services. • Hosted Frontend on Netlify. • Created Deployment Report of Project (22.05.2025) • Submitted Checklist III to the cloud location (CDAP) (26.05.2025) • Created Research Project Website (25.05.2025) • Prepared the final presentation slides (24.05.2025) • Prepared for PP2 presentation with the team (25.05.2025) • Participated in the final presentation and viva (26.05.2025) • Participated in the Research Project Website Evaluation Viva (26.05.2025)
2025-June	<ul style="list-style-type: none"> • Got the RP- Acceptance notification (09.06.2025)

	<div> <div> 16th ICCCNT 2025 submission 7947 Inbox x </div> <div> 16th ICCCNT 2025 <16thiccn2025@easychair.org> to me Mon, Jun 9, 3:54 PM (11 days ago) </div> <div> <p>"Dear Authors, Paper ID: 7947 Title: Readyfy - Intelligent Assistant to Improve Reading and Comprehension Skills in English Language</p> <p>Congratulations! Your paper has been accepted. To ensure it is included in the presentation schedule, please make the necessary changes based on the below comments and update the same in the EasyChair login at the same Paper ID. The present Similarity/Plagiarism Index Including References in iThenticate of your paper: 9%</p> <p>Reviewer Comments:</p> </div> </div> <ul style="list-style-type: none"> Completed the RP – Registration process (19.06.2025)
--	---

2. Individual Task Allocation Process

2.1 Assigned Tasks



(MS Planner)