Senior Design: Milestones, Timeline, and Matrix Effort

1) **Milestones**
    a) **Research different Decombilation LLM models methodologies.**
    b) **Generate source-pair dataset with malware samples**
    c) **Finetuned pretrained LLM**
    d) **Demo Ghidra plugin**
    e) **Integrating finetuned LLM with Ghidra**
    f) **Demo MalAI**

2) **Timeline**

Task	Start and Completion Dates
Research LLM models that could be utilized as an infrastructure and a starting point of the LLM models to be built.	10/9/2025 - 29/9/2025
Research possible decompilation methods.	22/10/2025 - 1/11/2025
Research how to utilize Ghidra as a plug-in.	1/10/2025 - 15/10/2025
Research and find datasets that will be utilized to train the LLM models.	8/10/2025 - 1/11/2025

Build a Ghidra plugin	**8/10/2025 - 29/10/2025**
Explored different variations of LLM4Decompile (Original Model, SK2)	**5/11/2025 - 12/11/2025**
Introduce malware samples into existing binary-source datasets (exebench, decompilebench, idioms	**12/11/2025 - 3/12/2025**
Explore possible tools to build datasets (Pairs of C/C++ and Binary) like CodableLLM	**12/11/2025 - 3/12/2025**
Integrate LLM finetuned model	**7/1/2026 - 28/1/2026**
Train LLM model to analyze executable files and decompile them into C-Code.	**10/12/2025 - 21/1/2026**

Test the models analysis of C code and Assembly code.	**17/12/2025 - 21/1/2026**
Document the analyzation performance of the built models in comparison to other recent models that target the same criteria.	**21/1/2026 - 28/1/2026**
Develop the analyzation page of the application that showcases the analyzation of both C and Assembly codes.	**28/1/2026 - 18/2/2026**
Develop the malware detection result page	**25/2/2026 - 11/3/2026**
Demo MalAI	**11/3/2026 - 25/3/2026**
Deploy Application (If feasible)	**1/3/2026 - 5/3/2026**
Develop a landing webpage for MalAI	**25/3/2026 - 30/3/2026**

3) Effort Martrix

Task	Start and Completion Dates	Long Nguyen (Hours of Effort)	Luqman Al Hasni (Hours of Effort)
Research LLM models that could be utilized as an infrastructure and a starting point of the LLM models to be built.	10/9/2025 - 29/9/2025	Primary 3 Hours	Secondary 2 Hours
Research possible decompilation methods.	22/10/2025 - 1/11/2025	Primary 4 Hours	Secondary 3 Hours
Research and find datasets that will be utilized to train the LLM models.	8/10/2025 - 1/11/2025	Primary 4 Hours	Secondary 2 Hours

Research how to utilize Ghidra as a plug-in.	**1/10/2025 - 15/1/2025**	**Secondary 2 Hours**	**Primary 4 Hours**
Build a Ghidra plugin	**8/10/2025 - 29/10/2025**	**Secondary 3 Hours**	**Primary 5 Hours**
Explored different variations of LLM4Decompile (Original Model, SK2)	**5/11/2025 - 12/11/2025**	**Primary 5 Hours**	**Secondary 4 Hours**
Introduce malware samples into existing binary-source datasets (exebench, decompilebench, idioms	**12/11/2025 - 3/12/2025**	**Primary 5 Hours**	**Secondary 3**

Explore possible tools to build datasets (Pairs of C/C++ and Binary) like CodableLLM	**12/11/2025 - 3/12/2025**	**Secondary 2 Hours**	**Primary 5 Hours**
Integrate LLM fintuned model	**7/1/2026 - 28/1/2026**	**Primary 6 Hours**	**Secondary 4 Hours**
Train LLM model to analyze executable files and decompile them into C-Code.	**10/12/2025 - 21/1/2026**	**Secondary 2 Hours**	**Primary 6 Hours**
Test the models analysis of C code and	**17/12/2025 - 21/1/2026**	**Primary 5 Hours**	**Secondary 2 Hours**

Assembly code.			
Document the analyzation performance of the built models in comparison to other recent models that target the same criteria.	**21/1/2026 - 28/1/2026**	**Secondary 2 Hour**	**Primary 5 Hours**
Develop the analyzation page that showcases the analysis of both C and Assembly codes.	**28/1/2026 - 18/2/2026**	**Secondary 2 Hours**	**Primary 3 Hours**
Develop the malware detection	**25/2/2026 - 11/3/2026**	**Secondary 4 Hours**	**Primary 6 Hours**

result page of the MalAI			
Demo MalAI	**11/3/2026 - 25/3/2026**	**Primary 4 Hours**	**Secondary 3 Hours**
Deploy Application (If feasible)	**1/3/2026 - 5/3/2026**	**Secondary 4 Hours**	**Primary 5 Hours**
Develop a landing webpage for MalAI	**25/3/2026 - 30/3/2026**	**Secondary 4 Hours**	**Primary 5 Hours**

According to the above Estimations the hours for each member are as below:

- **Luqman: 66 Hours**
- **Long: 62 Hours**