

Danmarks  
Tekniske  
Universitet



---

# Project Plan: Logic for Neural Networks assisted by Neuroevolution

---

## AUTHORS

Abhista Partal Balasubramaniam - s210246

October 1, 2022

# 1 Basic Project Abstract

Neuro-evolution is a genetic approach to train and develop neural networks and LEN (Logic Explained Networks) are neural networks that provide logical explanations to the decisions/outputs given by neural networks after training has been completed. The aim of this project is to study the theory behind the two unique ideas and find a method to combine them to provide a more robust neural network and explanations.

LENs tend to trade accuracy for the quality of logical explanations. To achieve these good readable explanations, LENs in the end go through a step called pruning which removes small/weak connections between layers to keep the output logical explanations readable.

Neuro-evolution allows layers and nodes to be added and removed from the neural network during the training process. Thus, allowing the neural network to grow only if the current neural network architecture cannot solve the problem at hand.

Therefore, combining the above two ideas will provide a way to develop neural networks which may have the optimal architecture, which can then be used to provide the logical explanations without using the pruning step in the creation of LENs. Since, LENs require human understandable features as inputs the above algorithm can be tested using old arcade video games for performance.

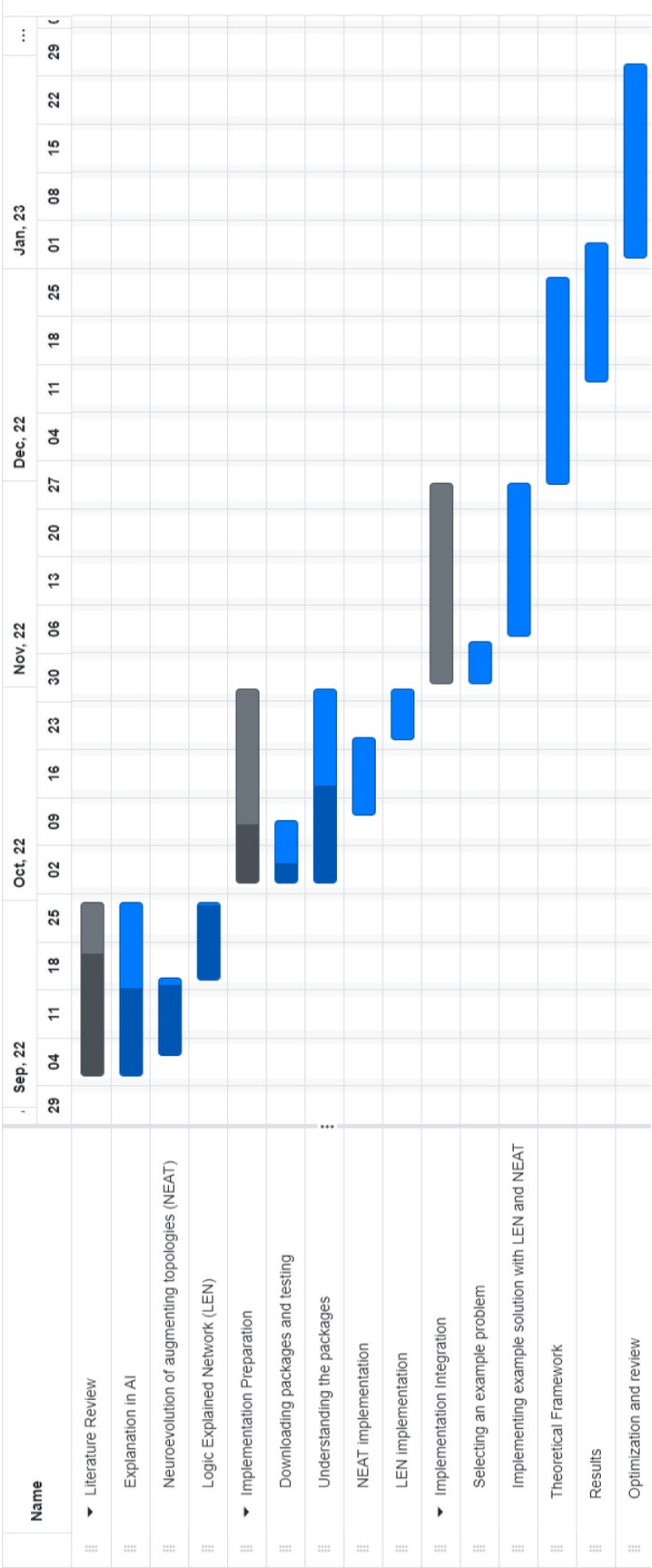


Figure 1: Project Gantt Chart