What?

The goal is to evaluate the role humans play in implementing a deep learning model. We will evaluate the performance metric of a deep learning model when trained by different people with different knowledge background. Different knowledge background refers to the experience and prior knowledge of a person in training a deep learning model.

- **RQ:**

  How it will affect the final performance of a deep learning model when it is trained by different people?

- **HP1:**

  The final performance metrics will vary based on the hyperparameters chosen by the people. There will be variance.

- **HP2:**

  The final performance metrics will be dependent on the background knowledge of people training or deciding the hyperparameters values.

- **HP3:**

  The above research can be generalized for other deep learning models.

- **HP4:**

  The above research can be generalized for other types of problems solved using deep learning models.

Why?

There are many research papers published in deep learning trying to solve a problem using deep learning models and claiming to achieve certain performance metrics in solving the problem in hand. It is a common occurrence that these papers do not mention the complete list of hyperparameters and their specific values used to train the deep learning model.

We couldn't find any prior research to quantify the variation in the final performance metrics due to the absence of these hyperparameters if any. This will help us estimate the true range of deep learning model's metrics and the impact of hyperparameters on a deep learning model.

How?

1. Choose a paper explaining a deep learning model used to solve a certain problem.
    - Should I choose the paper that specifies all the hyperparameters and its values used for training the model?
        - It will make it easy to verify that even when all hyperparameters value is defined, results may vary.
2. Implement the deep learning model as mentioned in the paper.
3. Verify the final performance metrics of the model.
    - Check for an initial variance after this step.
4. Implement a pipeline and interface for the users enabling them to choose the values for the hyperparameters and store their final model results.
    - It includes collecting data about user's prior knowledge in deep learning
5. Analyze the data collected for the variance in the final performance of the model and correlation between the user's choices and final metric results obtained.
6. Document the results and look at ways to collect more data.

    Optional?

    - Extend this thesis for other models and other tasks solved by deep learning.
    - Try including more and more students/researchers/staff members to gain more insights.

Challenges:

1. If we get minimal variances for the research paper that specifies all the hyperparameters and its values, then research we are doing is more about finding the impact of hyperparameters in training a model. That could be more easily done by eliminating humans and experimenting with hyperparameter values. Why do we need humans?
2. Choice of the model.
    - The training time of a model is the key parameter. It should be optimal as it directly impacts the resources like CPU time, costs for running those machines and the number of participants that can be accommodated.