1. The main goal of term project is no more to find the minimum RMSE. The goal now is more job oriented where we will,

* Learn and implement a machine learning model which is currently in practice by IT companies.
* Design a static website which is professional and can be used to show the future employer during job interview.

2. In order to accomplish the goal, we have to work in following direction

* Look for minimum of 20-30 companies’ machine learning requirement for data scientist position and identify the models.
* Choose one of the machine learning model and master it in such a way that we can even teach the learned model like a professor.
* Setup an open source code repository for the project.
* Create the training dataset from the competitions’ (KDD cup 2010) training dataset or we can use the training dataset as is from the competition’s page. We should provide detailed discussion on how we choose the training data, validation data, and test data for the project.
* The website must include following three points.
* The problem description, which is not lengthy and must focus on machine learning keywords identified in the part 2(a). It must also have brief description of KDD cup 2010. Also show the kind of data used for the case study.
* Detailed but not lengthy description of the model used for the case-study. The description of machine learning method used should contain the mathematical formula with its explanation, text description of the method and illustrating examples to describe the method more clearly.
* To the point illustration of result which includes how we applied the machine learning method on the training data, validation data and test data.

3. TODO for progress report.

* Identify and pick minimum of one machine learning model which we will be learning and implementing. The description of machine learning method used should contain the mathematical formula with its explanation, text description of the method and illustrating examples to describe the method more clearly.
* Prepare the training dataset and give detailed description of: How we choose the training data, validation data, and test data for the project.
* Setup code repository.
* Come up with a professional team name (like a start-up).
* Design the website template.