The group members are Kunj Dedhia and Prateek Agarwal.

The two of us will be splitting the following work: Prepare the dataset and set up the pipeline using Kaggle's own time series API. Clean data and compress if necessary. Evaluate different machine learning frameworks for the use case and decide on architectures. Train architectures, Hyperparameter optimization, final testing and evaluation, infographic generation as well as insights.

Our plant is to follow along and compete in the currently active Riiid AIEd Challenge Kaggle Competition. The challenge is to investigate and track students' knowledge by prediction whether students are able to answer their next questions correctly.

This question is interesting because its success can serve as a prediction to how powerful the next generation of education applications can be in helping students on an individual basis. It is impossible for students nowadays to get personalized education, especially for online education like Khan academy, and using a neural network to predict student performance and produce individual insights can be promising for the future of MOOC education.

We will be using a time series dataset provided by the host company which tracks each student's performance on successive questions, and provides question metadata as well as a lecture metadata to focus on the question at hand.

We propose to use a RNN and CNN architecture concurrently to factor in the spatio-temporal features of the dataset. There are no existing implementations since this is a new dataset. We will do readings pertaining to test taking/making which can help us in feature extraction/selection as well as RNN architectures to extract the time series data.

We will evaluate our results on the provided test set and we will try to maximize our numerical accuracy.

Link : https://www.kaggle.com/c/riiid-test-answer-prediction