

# ConvLab2 PS

## **Introduction:**

ConvLab2 is an open source toolkit that allows researchers to build task-oriented dialogue systems with state-of-the-art models, perform an end-to-end evaluation, and diagnose the weakness of systems. It provides support for some large datasets and powerful models and is extremely flexible when it comes to customization. For more details on the full working and description of ConvLab2, refer to this [\[paper\]](#) or [\[repository\]](#). A lot of research by both academia and industries is focused around building dialog systems nowadays, hence this problem will help you inculcate strong fundamentals on this topic.

## **Task:**

ConvLab2 contains 4 main components, namely:

1. Natural Language Understanding (NLU)
2. Dialog State Tracking and Dialog Policy
3. Natural Language Generation (NLG)
4. Simulation and deployment

Each of the 4 components are a field of research themselves. The aim of the task can be reduced to the following three aspects:

1. Understand the codebase and play around with it.
2. Try to modify existing approaches
3. Deploy the model and submit a 2 page report of what you have done for evaluation
4. Make a 5-10 min presentation of your approach. (if you have done something novel).

You are not expected to come up with something novel (although doing so will guarantee you a prize ;) ), but to understand how dialog systems work so that in the course of time you are equipped with knowledge to pursue research.

Along with the above mentioned tasks, one more goal of this task is to initiate a discussion on this subject during ML talks. Everyone participating is expected to talk about the one or more components (both research based and code based) of the toolkit. The tentative schedule for the talks is as follows:

- Week 1: Setting up and doubts regarding the toolkit.
- Week 2: NLU
- Week 3: Dialog State Tracking and Dialog policy
- Week 4: Deployment

## **Rules:**

You may form teams with upto 2 members maximum.

Tentatively, by the beginning of February, we will start evaluating submissions.

## **Prizes:**

Would be decided based on the number of submissions, and the enthusiasm of participants.

There will be certificates for the best and the runner up submissions.