

Reinforcement Learning Assigned Project Instructions

1. Getting started:

Install Python (version 2.7) and necessary packages at least including: [collections](#), [numpy](#), [pandas](#) and [mdptoolbox](#).

1. Download the dataset: MDP_Original_data.csv

2. Keep columns 1-6 untouched.

In the dataset, the columns 1-6 (index from 1) are static information about the data listed as follows: [student](#), [currProb](#), [course](#), [session](#), [priorTutorAction](#) and [reward](#).

[priorTutorAction](#) is the action that tutor takes for the corresponding problem in the row.

[Reward](#) is the reward of jumping from the previous state to current one taking corresponding action.

student	currProb	priorTutorAction	reward	state
0006-F14	1.0.1.0	PS	0	S1
0006-F14	1.0.2.0	WE	0	S2
0006-F14	1.0.3.0	WE	0	S1
0006-F14	1.0.4.0	PS	-94.07894737	S2
0006-F14	2.1.1.0	PS	0	S2
0006-F14	2.1.2.0	WE	0	S1
0006-F14	2.1.3.0	PS	161.8100398	S1

For example, in above sample table, we have two states S1 and S2. The starting state is S1. Based on the first two rows, we can see that S1 jumps to S2 by taking action WE and obtain reward 0. Based on row 2 and 3, we can conclude that S2 jumps to S1 by taking action WE and obtain reward 0. Finally, based on row 3 and 4, we can conclude that S1 jumps to S2 by taking action PS and obtain reward -94.07.

3. Generate or select up to 8 features from the remaining columns.

Your task is to extract or to select best feature set to model the learning environment. To complete this task, you need to carry out the following two basic tasks in any order or combined way:

- ❖ Feature Extraction/Selection: Implement your own feature selection (FS) or feature extraction (FE) algorithms to select or to extract features based on the remaining columns **MDP_Original_data.csv**. The features will be combined with the original columns (1-6) as **Training_data.csv**. **Please note the total number of selected/extracted features is capped at 8 strictly.**
- ❖ Discretizing features: Our MDP package here can only take discrete features. So you need also explore different ways of discretizing the continuous features.

4. MDP package

Run the provided python file Get MDP_policy.py to evaluation your project. The file will read a training dataset named as '**Training_data.csv**' and print out **the policy and the ECR value**.

- ❖ The format of the training data file should include up to 14 columns. The first 6 columns should be exactly the same (headers and content) as ones in the original dataset. From column 7-14 you need to place your selected/extracted features. If the features are in the list of features of original dataset, keep their original headers. If the features are abstract, just name them as f1, f2....
- ❖ Output file should be saved as output.pdf (screen shot is OK) and it should looks like something:
- ❖ You can execute python code using the following command:
python MDP_policy.py -input sample_output.csv.

The sample output is:

```
Policy:
state -> action, value-function
1:0 -> WE, 40.3113658941
2:1 -> PS, 57.3961960569
3:1 -> PS, 32.0464227157
4:1 -> PS, 39.4900063005
5:1 -> PS, 21.3578762853
6:1 -> PS, 25.2606056612
3:-1 -> WE, 30.5071197976
2:-1 -> WE, 25.4789814423
4:-1 -> WE, 17.8312569342
5:-1 -> WE, 17.5345286344
6:-1 -> WE, 0.0
ECR value: 40.3113658941
```

- ❖ The code will print out the policy and ECR value.
- ❖ You can use 'calculate_ECR' function as part of your own code, if you need to use 'calculate_ECR' as a step during your whole feature selection process.

5. Evaluation.

Using ECR to select the best policies you have explored.

6. Submission.

You should submit ONE .zip file and name it using the following name schema: A_ProjectGroupID.zip. The zip file should include the following files:

- ❖ ReadMe file: in this file you should inform the instructors how to run your code and all the necessary packages.
- ❖ **Training_data.csv**: submit the training data named **Training_data.csv for your final submitted policy.** .
- ❖ Your codes for feature discretization and feature selection/extraction
- ❖ Report your final “best” policy and the ECR_value in **Output.pdf**.