**User Interview Questions**

**Demographics**

1. What is your name?
2. What is the highest degree or level of school you have completed?
3. What is your current employment status?
4. If employed, in which domain do you work?
5. What is your job title?
6. What are your job responsibilities at work?
7. How many years of experience do you have?

**Previous experience**

1. What are your basic interests in the field of Machine learning?
2. Can you recall what Machine learning methods have you used uptil now?
3. What purpose did the method (Method 1, Method 2 etc.) serve?
4. Can you recall the Use case for the method (Method 1, Method 2 etc.)?
5. What kind of data have you worked on?
6. Is there any significant project/achievement, you would like to share with us as a Machine learning practitioner?
7. Do you have any experience in the research field? If yes, what is it?

**Use case**

1. Can you describe a project in which you have used hyper-parameter tuning?

**Data**

1. What datasets have you used in the project?
2. What were the features/attributes of the dataset (Dataset 1, Dataset 2 etc.)?
3. How many features/attributes were present in the dataset (Dataset 1, Dataset 2 etc.)?
4. What was the source of the dataset (Dataset 1, Dataset 2 etc.)?
5. Did you compile the data by yourself?
6. Was the data aggregated from multiple sources?
7. Was the data consistent?
8. Was the data accurate? Were there any outliers within the data?
9. Did you perform the data cleaning?
10. Did you ignore any attributes? If yes, why?
11. How much data was allocated to the training set?
12. How much data was allocated to the test set?
13. How much data was allocated to the validation set?
14. What steps did you follow for the data pre-processing?
15. Did you face any challenges in data pre-processing? What are they?
16. Was there any challenge due to data anomaly or inconsistency? If yes, what did you do to overcome it?

**Tasks**

1. What kinds of analysis you performed in the project using Machine learning? e.g., Classification, Regression etc.
2. How many tasks were there?
3. Was the task (Analysis 1, Analysis 2, etc.) supervised or unsupervised?
4. Did you use k-fold cross validation? What was the value of k?
5. How many runs were there?
6. What was/were the required/predicted outcome(s) of your analysis?
7. Did you get the required/predicted outcome? If not, what do you think was the problem?
8. Did you overcome the problem? If yes, what did you do to overcome the problem?
9. Were there any difficulties/challenges in performing the tasks?
10. How did you proceed to overcome the difficulties/challenges faced during the analysis?
11. Were there any limitations to your devised approach?
12. What were the pros of used approach? e.g., computational time, model efficiency etc.
13. Were there any other methods to perform the analysis apart from the one you used?
14. How much accuracy was achieved of the trained model?

**Hyper-parameter tuning**

1. What algorithm/method did you use for hyper-parameter tuning? e.g., Grid Search, Random Search etc.
2. Why did you use the specific algorithm/method?
3. Did you face any difficulty/challenge in implementing hyper-parameter tuning? If yes, what are they?
4. Is there any parameter tuning method you have designed, to address the difficulty/challenge you had? If yes, what are they and how does they work?
5. Did you utilize any machine learning library to perform hyper-parameter tuning?

**Grid search**

1. Did you use grid search as the hyper-parameter tuning algorithm in your project?
2. If no, why did not you use grid search?
3. Did you utilize any grid search library for it?
4. How did you design the grid e.g., search space etc.?
5. Did you face any difficulty while implementing grid search? e.g., time consumption, efficiency etc.
6. Did you face any difficulties working with grid search with large number of parameters e.g., n>100?
7. Do you recall the best score the model achieved?