**Developer’s guide for ML Face Tagging Project.**

The below are the steps which we have planned and started working on for our Face tagging project.

* **Project Review**: Youth Exploring Science (YES) program is created to help underserved students in St. Louis area. Students participate in YES program to gain skills in Science, Technology, Engineering, Arts and Mathematics fields. High school students will gain real-world skills, career readiness training through YES program. Students in the YES program also promote the importance of STEAM-related careers. The SLSC Photo Archive Face project is to Identify and Tag people from the images client or user uploads, next once the detected faces tagging is completed we need to store the organized result in a database readily available to access, the client with the help of photo tagging interface would like to sort photos with respect to student which helps them to keep track of each student’s progress.
* **Installation and Setup**: We have decided to use python as our main programming language. We have done our initial python flask project setup and created CI/CD pipeline for continuous integration of the code. Our CI/CD pipeline includes automated testing of python scripts.
* **Data Preparation**: Our data contains images of individuals with their associated names. We need to identify and Tag people from the images to the correct name. We have connected with the saint Louis center to obtain the dataset. The Youth Exploring Science program has thousands of photos of students. As of now, the photos are organized by event and date, our project helps them to sort photos by student with their corresponding names. This would help them top track students’ progress and makes them easier to collect the information related to each student.
* **Selecting and Training an ML model:** Complete data is splitted into training, validation and testing data in the ratio 70:15:15. We have selected few ML models namely CNN, Multi cascade CNN, recurrent neural networks for face detection in our project. We need to train the data on multiple models and choose the one with high accuracy.
* **Testing:** Testing the predictions made by the model with testing data which we have splitted initially and evaluating the accuracy by performance meaures namely mean square error.
* **Building a User Interface:** Building a simple user interface where user can upload an image and on click of search button, there will be an API call happening which retrieves the data of that particular image and displays it on UI.