Current Progress:

We have made significant progress in our project, "Cultivating Tomorrow: AI Revolution in Greenhouse Agriculture." Our journey began with a comprehensive exploration of the Indian landscape of greenhouse farming and automation powered by AI. Despite challenges in finding Indian-specific datasets, we managed to obtain valuable data from a machine learning competition organized by a research institution. This dataset pertains to cherry tomato production and originates from five distinct commercial parties operating in the Netherlands. Although not Indian-centric, this dataset serves as a robust foundation for our automation project.

The dataset comprises approximately 45 files, each containing detailed information. It initiates with weather data, followed by greenhouse climate parameters, crop parameters, resource utilization statistics, Grodan sensor data, and concludes with production quality and laboratory analysis of the crop. In total, we have access to several lakhs of data points, making it an extensive resource for training our AI models. With over 100 parameters to analyze, we delved into understanding the intricate relationships between them.

One of the primary challenges we encountered was handling missing values, which were prevalent throughout the dataset. To address this issue, we employed various strategies such as data splitting into training and testing sets, where the testing set retained the missing values. Through training on non-null values, we were able to predict and impute the missing values effectively. Additionally, we utilized techniques like forward fill, backward fill, and simple imputation methods to handle missing values in certain columns.

Once the data was cleaned and preprocessed, we embarked on visualizing it to gain deeper insights into the underlying patterns and relationships. We conducted correlation matrix analyses and multivariate analyses to uncover hidden trends and dependencies among the myriad parameters. These visualizations proved instrumental in refining our understanding of the dataset and guiding subsequent modeling efforts.

Overall, our progress signifies a crucial step forward in leveraging AI for greenhouse agriculture. By harnessing the power of data-driven insights and advanced machine learning techniques, we are poised to revolutionize greenhouse farming practices and pave the way for a more sustainable and efficient future.