How to Explain Your Project in an Interview
Here is a script you can practice.

(S) Situation: The Problem
"In today's world, many of us use ride-sharing apps like Uber and Ola. However, their dynamic pricing can often feel like a 'black box.' The fare for the same route can change dramatically based on the time of day or demand, which can be frustrating for customers. I was interested in demystifying this process and understanding the key factors that determine a ride's cost. To do this, I decided to build a fare prediction model using a large, publicly available dataset of New York City taxi rides, which contained over 55 million records."

(T) Task: The Goal
"My main goal was to engineer a machine learning model that could accurately predict the fare of a taxi ride based on various inputs. My objectives were to:
Clean and process a massive dataset to make it usable.
Engineer new, insightful features that would improve the model's predictive power.
Train and compare different regression models to find the best performer.
Ultimately, identify which factors (like distance, time, or location) have the most significant impact on the final fare."

(A) Action: The Process
"I took a systematic, step-by-step approach using Python and its data science libraries.
Data Processing: First, I used the Pandas library to load and clean the 55-million-row dataset. This involved handling missing values, removing impossible trip distances, and filtering out outlier fares that would have skewed the results.
Feature Engineering: This was the most critical part. I created several new features from the raw data. I calculated the direct trip distance using the pickup and dropoff coordinates. I also extracted the hour of the day, day of the week, and month from the timestamp, as I hypothesized that rush hour and weekends would heavily influence the price.
Modeling: Using Scikit-learn, I experimented with two different models. I started with a baseline Linear Regression model. Then, I trained a more complex Decision Tree Regressor, which I believed would better capture the non-linear relationships in the data. I split my data into training, validation, and test sets to ensure my model's performance was reliable.
Deployment: To make my project interactive, I even built a simple frontend with HTML, CSS, and JavaScript. This web page allowed a user to input their trip details and get an instant fare prediction from my model."

(R) Result: The Outcome
"The project was a success. The Decision Tree model significantly outperformed the baseline, achieving a Mean Absolute Error (MAE) of $1.85. This means, on average, my model's prediction was only off by less than two dollars from the actual fare—a very strong result.
Most importantly, the model's analysis confirmed that while trip distance is the primary factor, the time of day is the second most influential variable, causing prices to spike during morning and evening rush hours. This project not only allowed me to apply my Python and machine learning skills to a real-world problem but also successfully provided a data-driven explanation for the 'black box' of dynamic pricing, giving clarity to the concerns customers have."