**what is/was the most challenging machine learning problem you did, how you tackled it and what was the result?**

During my machine learning intern project at McKesson, one of the most challenging problems I encountered was developing a predictive maintenance model for medical equipment. The goal was to reduce unplanned downtime, optimize maintenance scheduling, improve equipment reliability, and achieve cost savings.

To tackle this challenge, I followed a systematic approach. First, I collected comprehensive historical data on medical equipment, including maintenance logs, sensor readings, and error codes. This data served as the foundation for developing the predictive maintenance model.

Next, I focused on data preprocessing to ensure data quality and completeness. I cleaned the collected data, handled missing values, and performed necessary transformations such as normalization and outlier removal. This step was crucial in preparing the data for subsequent analysis and model development. Here, I was tasked with managing a dataset of over 2 million records. For this task, I have Cleaned, pre-processed, and normalized a dataset of over 2 million records, improving the quality of training data thus improving model accuracy by 25%

Feature selection played a significant role in building an effective predictive maintenance model. I utilized techniques like correlation analysis, feature importance, and domain knowledge to identify the most significant features contributing to maintenance needs and failures. This step helped in reducing dimensionality and focusing on the most relevant information.

For model development, I experimented with various machine learning algorithms suitable for predictive maintenance, including random forest, support vector machines, and recurrent neural networks. I trained these models on the preprocessed data, optimizing hyperparameters to achieve the best performance.

To evaluate the model's effectiveness, I used appropriate evaluation metrics such as accuracy, precision, recall, and F1 score. I also performed cross-validation to ensure the model's generalizability and robustness.

Upon obtaining a well-performing model, I proceeded with the deployment and integration phase. I deployed the trained model in a production environment where it could receive real-time data from medical equipment. Integration with existing systems enabled continuous monitoring and maintenance scheduling based on predicted needs.

Throughout the project, I focused on monitoring the model's performance, collecting feedback from maintenance personnel, and iteratively refining the model. This iterative process ensured that the model remained accurate and reliable, adapting to changing equipment conditions over time.

As a result of this project, we were able to achieve significant outcomes. Unplanned downtime was reduced, minimizing disruptions to healthcare services and patient care. Maintenance scheduling was optimized, allowing efficient resource allocation. Equipment reliability and performance improved, leading to fewer failures and extended equipment lifespan. The implementation of predictive maintenance also resulted in cost savings through the avoidance of emergency repairs and operational efficiency gains.

Overall, the project successfully demonstrated the value of predictive maintenance using machine learning at McKesson, delivering tangible benefits such as reduced downtime, improved reliability, and cost savings. By leveraging historical data and advanced machine learning techniques, we were able to transform maintenance practices and enhance the organization's operational efficiency.