Data Science Internship Tasks

Task 1: Titanic Survival Prediction

\* Feelings: I found this task engaging and a great introduction to data science. The dataset was readily available, and the problem was well-defined. I enjoyed exploring the data, cleaning it, and building a simple model to predict survival. It was rewarding to see the model's performance and gain insights into the factors that influenced survival.

\* Conclusion: This task provided a strong foundation in data analysis, model building, and evaluation. It highlighted the importance of understanding the data and choosing appropriate techniques for the problem at hand.

Task 2: Movie Rating Prediction with Python

\* Feelings: This task was more challenging than the first, requiring me to think about feature engineering and regression techniques. I spent some time researching different approaches and experimenting with different models. It was a bit frustrating at times, but ultimately rewarding to see the model's performance improve.

\* Conclusion: This task deepened my understanding of regression analysis and the importance of feature selection. It also reinforced the iterative nature of data science, where experimentation and fine-tuning are often necessary.

Task 3: Iris Flower Classification

\* Feelings: This task was a good exercise in applying machine learning to a classic classification problem. The dataset was simple and easy to understand, which allowed me to focus on the modeling process. I enjoyed exploring different classifiers and comparing their performance.

\* Conclusion: This task demonstrated the power of machine learning for classification problems. It also highlighted the importance of choosing the right classifier for a given dataset and evaluating the model's performance using appropriate metrics.

Overall Conclusions

\* Skills Development: Through these tasks, I gained practical experience in data cleaning, exploratory data analysis, feature engineering, model building, and evaluation. I also became more familiar with common data science tools and libraries like Python, pandas, and scikit-learn.