

Reflective Journal

The wine classification project deepened my understanding of supervised machine learning, particularly how feature selection directly impacts model performance. Beyond understanding the functions and purpose of logistic regression, this assignment helped me see how it behaves in practice when experimenting with different features and observing changes in accuracy. It became clear that not all variables contribute equally. Features like alcohol, color intensity, and proline significantly improved performance, which reinforced the idea that model success depends heavily on choosing informative predictors rather than simply increasing model complexity.

One of the main challenges I faced was understanding why certain feature combinations performed better than others. Initially, I assumed adding more features would automatically increase accuracy. However, through experimentation, I realized that irrelevant or redundant features can introduce noise and reduce generalization performance. To overcome this, I focused on analyzing the dataset's characteristics and considering how each feature might logically relate to wine classification. Concepts such as correlation, variance, and linear relationships became much more tangible. Logistic regression's decision boundary felt less abstract once I saw how numeric features could separate classes in a structured way. It also reinforced my understanding of overfitting and why simpler models often perform well on clean, well-structured datasets.

Several questions emerged during the project. For example, would a more complex model like Random Forest significantly outperform logistic regression, or would it risk overfitting? I also became curious about feature scaling and whether standardizing the variables would further improve accuracy.

Overall, this project changed my approach to machine learning problems. Instead of immediately thinking about more advanced algorithms, I now recognize the importance of understanding the data first. Thoughtful feature selection, proper evaluation, and model simplicity can often lead to strong performance. The experience emphasized that machine learning is not just about algorithms, but about reasoning, experimentation, and critical thinking.