

Summary week 5 and 6

Week 5: Model evaluation and supervised learning

The fundamentals of supervised learning, such as its types and important algorithms, were the main topic of discussion this week. It explored structural risk reduction, Occam's razor, and model complexity. Numerous assessment measures were covered, including regression metrics (MSE, RMSE, R2) and classification metrics (precision, recall, F1-score). Finding the optimal hyperparameters, resolving unbalanced classes, and dividing datasets into training and testing sets were all practical procedures. Furthermore, regression models were presented, with examples in Python using scikit-learn, and focus was placed on subjects including multivariate regression, error analysis, and model assessment.

Week 6: Logistic and Linear Regression

The theory and real-world uses of logistic and linear regression were covered this week. Covariance and significance between features, the creation of linear regression, and classification using linear models were the first steps. After a review of generalization and model complexity, logistic regression formulation and training were conducted. Examples showed how to utilize regularized linear models to prevent overfitting in logistic regression in Python. Regression-based feature selection was presented, emphasizing the significance of dimensionality reduction and interpretability. Applied coding exercises at the end of the week reinforced real-world application.

Reading List Summary

- **External Resources:** Research papers on regression and classification methods, especially regularisation techniques.
- **Websites:** Documentation of scikit-learn for regression, classification, and hyperparameter tuning.
- **Book Chapters:** Standard ML texts covering supervised learning, regression analysis, and logistic regression theory.
- **Code Libraries:** Python's scikit-learn (for regression and classification), numpy, and pandas for data preprocessing, matplotlib and seaborn for model performance visualisation.

Reflection on Knowledge Gained

Through Weeks 5 and 6, I developed a strong understanding of supervised learning, particularly how models are trained, validated, and tested. I learned how model complexity affects performance and how evaluation metrics guide model improvement. I also understood the practical application of regression models, both linear and logistic, and the role of regularisation in controlling overfitting. Importantly, I now appreciate the importance of feature relevance, hyperparameter optimisation, and class imbalance handling in building reliable machine learning models. This knowledge provides a solid foundation for applying regression techniques to real-world datasets.