

Comparative Study of Supervised and Unsupervised Learning

Gaikwad Vaishnavi Tribhuvan¹ and Rohini S. Kapse²

M.Sc. Computer Science, K.T.H.M College, Nashik¹

Professor, Department Computer Science and Application, K.T.H.M College Nashik²

Abstract: *The field of Artificial Intelligence and Machine Learning is growing very fast and touching almost every industry today. This project work compares the two main approaches of machine learning – Supervised Learning and Unsupervised Learning. Supervised learning needs labeled data and gives very accurate predictions, so it is mostly used for classification and regression tasks. On the other side, unsupervised learning works on data without any labels and tries to find hidden patterns, natural groups or reduce dimensions using clustering and other methods.*

In this study, popular algorithms like Decision Tree, Support Vector Machine, Logistic Regression for supervised part and K-Means, Hierarchical Clustering and PCA for unsupervised part were implemented on standard benchmark datasets (Iris, Breast Cancer Wisconsin and Mall Customers dataset). The experiments were carried out in Python using scikit-learn library

Results clearly show that supervised models achieved high accuracy between 95–97%, whereas unsupervised models gave Silhouette Score in the range of 0.6–0.7 which is quite decent for pattern discovery. When labeled data is available, supervised learning is definitely better, but when labeling is costly or not possible, unsupervised methods become very useful. The study also suggests that future systems can get even better results by combining both approaches in hybrid or semi-supervised models.

Keywords: Supervised Learning, Unsupervised Learning, Classification, Clustering, K-Means, SVM, Decision Tree, PCA, Machine Learning Comparison

