IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, February 2026
Impact Factor: 8.2

Application of Linear Algebra in Machine Learning

Sonali Venunath Dighe[1] and Shweta Sachin Bibave[1*]
Department of Mathematics.
S. N. Arts, D. J. Malpani commerce and B. N. Sarda Science college (Autonomous), Sangamner, Dist. Ahilyanagar (M.S), India.
Affiliated to Savitribai Phule Pune University
shwetabibave@sangamnercollege.edu.in

Abstract: *Data representation is a critical foundation of machine learning, as it determines how raw information is structured, stored, and transformed into mathematical forms that algorithms can process. The effectiveness of representation directly influences the accuracy, efficiency, and interpretability of models. This paper investigates the role of data representation using a student dataset as a case study. The dataset includes both numerical features, such as hours studied, attendance percentage, and exam scores, as well as categorical features, including gender and extracurricular participation. These diverse attributes provide an ideal context for demonstrating multiple representation techniques, including vectors, matrices, tensors, one-hot encoding, and dimensionality reduction.*

The study applies regression, classification, and neural network models to the dataset, highlighting how proper representation improves predictive performance. For example, logistic regression achieved significantly higher accuracy when categorical variables were encoded using one-hot representation compared to raw categorical labels. Principal Component Analysis (PCA) reduced dimensionality while retaining over 95% of the variance, thereby simplifying training and improving computational efficiency. Neural networks using dense embeddings further enhanced prediction accuracy, demonstrating the importance of advanced representation methods..

Keywords: Data Representation, Machine Learning, Vectors and Matrices, Tensors, Machine learning applications

DOI: 10.48175/IJARSCT-31207