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 5, Issue 8, March 2025
Impact Factor: 7.67

Intelligent Cloud Resource Allocation Using AI: Optimizing Performance and Cost-Efficiency in Multi-Cloud Environments

Ravinder Ramidi
Birla Institute of Technology and Science - Pilani, India

Abstract: *Cloud computing has transformed organizational IT infrastructure, yet inefficient resource allocation persists as a critical challenge leading to excessive costs and suboptimal performance. This article introduces an AI-powered system that dynamically manages cloud resources across multi-cloud environments through intelligent automation. The solution employs advanced machine learning models for predictive workload analysis, enabling proactive rather than reactive resource management. Through provider-agnostic resource pooling, smart workload placement, and dynamic redistribution, the system optimizes across diverse cloud environments while accounting for multiple dimensions including performance requirements, costs, network considerations, and compliance constraints. Cost-aware resource selection incorporating instance optimization, spot instance usage, storage tiering, and commitment planning delivers substantial savings while maintaining performance. Additionally, the system enhances resilience through predictive failure detection, automated remediation, and continuous improvement mechanisms that learn from past incidents. These capabilities collectively address the growing complexity of cloud environments, helping organizations maximize their cloud investments while improving operational reliability.*

Keywords: Machine learning prediction, Multi-cloud orchestration, Cost optimization, Self-healing resilience, Resource utilization

DOI: 10.48175/IJARSCT-24533