SVKM's NMIMS
Mukesh Patel School of Technology Management & Engineering
Program: B.Tech

Course: Machine Learning
Experiment No.01

SAP ID: 70572300052	Name: Arushi Pandey
Class : 4th sem	Batch : Cseds
Date of Submission: 13-01-25	

B.1 TASK 1

Paper 1

i. **Paper Title:** An overview of machine learning applications for smart buildings

ii. **ML application :** Applications include building energy management, HVAC (Heating, Ventilation, and Air Conditioning) control, demand response, anomaly detection, and energy optimization in smart buildings.

iii. **Category of ML application:** Supervised Learning

iv. **Reasoning for category of ML application:** As supervised learning is used in scenarios like energy consumption prediction and fault detection where historical data with labels is available for training.

v. **Algorithms used:** Linear regression, Random Forest (RF), Artificial Neural Networks (ANNs), and Deep Learning.

vi. **Key concepts:**
The paper explores how ML is transforming building intelligence with applications in optimizing energy use, improving occupant comfort, and responding to dynamic environmental conditions. Emphasis is placed on autonomous AI agents, particularly reinforcement learning systems, which adapt based on feedback and optimize processes like HVAC and lighting. Digital twins and data-driven models are highlighted as critical tools for training and deploying these AI systems.

Paper 2

i. **Paper Title:** Automatic dataset builder for Machine Learning applications to satellite imagery

ii. **ML application :** The tool assists in creating datasets for Earth Observation (EO) and Remote Sensing (RS) using Machine Learning (ML). Applications include environmental monitoring, resource mapping, and geospatial analysis.

iii. **Category of ML application:** Supervised Learning

iv. **Reasoning for category of ML application:** The tool supports dataset preparation and preprocessing, which forms the foundation for supervised ML by generating labeled datasets (e.g., Sentinel-1 and Sentinel-2 data). Preprocessing techniques such as normalization and patch extraction align the data for ML model training.

v. **Algorithms used:** Min-Max Normalization, Standardization, Preprocessing steps compatible with convolutional neural networks (CNNs) for geospatial and environmental applications.

vi. **Key concepts:**
The paper presents a Python-based tool for automating dataset creation tailored for ML applications in satellite imagery. It enables users to prepare data from Sentinel-1 and Sentinel-2 satellites with features such as normalization, cloud masking, patch extraction, and dataset cleaning. This automation streamlines the tedious and time-consuming tasks associated with preparing geospatial data, making it a valuable resource for researchers in Earth Observation.

Paper 3

i. **Paper Title:** Novel application of unsupervised machine learning for characterization of subsurface seismicity, tectonic dynamics, and stress distribution

ii. **ML application :** The study applies unsupervised machine learning to analyze subsurface seismicity, delineate tectonic zones, and characterize stress distribution in seismic-prone areas, with a case study focused on the Makran Subduction Zone.

iii. **Category of ML application:** Unsupervised Learning

iv. **Reasoning for category of ML application:** Unsupervised learning techniques, particularly clustering, are utilized to identify patterns and categorize seismic events without predefined labels. This method effectively handles the inherent uncertainty in subsurface data and reveals hidden structures.

v. **Algorithms used:** Fuzzy C-Means (FCM), K-means Clustering and Hierarchical Clustering

vi. **Key concepts:**
This study employs Fuzzy C-Means (FCM) clustering to map and analyze seismic activity across the Makran Subduction Zone. It identifies six distinct clusters associated with tectonic and subsurface features. The findings enhance understanding of seismicity, enabling better risk mitigation and disaster preparedness. The research underscores AI's potential to optimize seismic hazard assessment and infrastructure design while providing actionable insights into tectonic processes.

B.2 Task 2

B.4 Conclusion:

This experiment highlighted the differences between supervised and unsupervised learning, emphasizing how labelled data drives predictions in supervised learning, while unsupervised learning identifies hidden patterns in unlabelled data. Key skills like reading various data formats and extracting metadata were discussed. Tools like Numpy and Pandas proved essential for handling and analysing structured data efficiently, demonstrating the importance of foundational techniques in machine learning workflows.