Tinkerers' Lab IIT Hyderabad

Bootcamp and Mini-Project

Machine Learning

What is Machine Learning?

Machine Learning (ML) enables systems to learn from data, recognize patterns, and make intelligent decisions. From personalized recommendations to autonomous systems, ML is driving innovation across industries. It transforms data into actionable insights and enhances efficiency, accuracy, and decision-making processes.

Project Overview

You may choose any of the project ideas listed below or propose a project of your own, provided it aligns with the scope and objectives of this initiative. If you decide to work on a project outside the suggested topics, please ensure that it is well-defined, feasible within the given timeframe, and demonstrates meaningful application of relevant concepts and skills.

1. **Handwritten Digit Recognition (MNIST Dataset):** Train a simple classifier using the MNIST dataset to recognize handwritten digits (0-9).
2. **Spam Email Classifier:** Build a binary classifier to differentiate spam and non-spam emails. dataset
3. **Iris Flower Classification:** Use the Iris Flower Dataset to classify flowers into three species based on petal and sepal dimensions. Use different Machine Learning algorithms that you learnt on this data and report your observations. If you can build any of the models from scratch, do so. If not, make sure you explain how the algorithm works in brief in the report.

   Also include a comparative analysis of model performance based on standard evaluation metrics such as **accuracy, precision, recall, and F1-score**. Additionally, mention the key observations regarding model strengths, limitations, and potential improvements.

4. **Titanic Survival Prediction:** Predict passenger survival on the Titanic Dataset based on factors like age, gender, and ticket class. Use different Machine Learning algorithms that you learnt on this data and report your observations. If you can build any of the models from scratch, do so. If not, make sure you explain how the algorithm works in brief in the report.

   Also include a comparative analysis of model performance based on standard evaluation metrics such as **accuracy, precision, recall, and F1-score**. Additionally, key observations regarding model strengths, limitations, and potential improvements are discussed.

Mentors

- Saran Konala - ai23btech11023@iith.ac.in (9515186523)
- Adamya Varma - cs23btech11001@iith.ac.in (6393781295)
- Kaushal Morankar - cs23btech11037@iith.ac.in (7498583320)
- Sonit Patil - bm23btech11023@iith.ac.in (8788280330)

Materials & Resources

- Python and other basic Libraries:
  - Python Documentation
  - Numpy , pandas , matplotlib .
- Basic Mathematical Concepts:
  - Calculus: Study material
  - Linear Algebra: Video Lectures
  - Probability and Statistics: Study material
- Classical Machine Learning Algorithms:
  - Machine Learning Series by StatQuest
  - Andrew NG course
- Deep Learning:
  - Andrew NG course
  - 3B1B Series
  - Neural Network Playground
- Computer Vision:
  - Andrew NG course
  - CNN Explainer

- Natural Language Processing:
    - Andrew NG course
    - Embedding Visualization
    - Transformer Visualization
- Reinforcement Learning:
    - Video Lectures
    - Documentation
- Tensorflow:
    - TensorFlow Documentation
- PyTorch:
    - Video Lecture
    - Study Material

Submission

You may submit your work via GitHub, Google Drive, or any other suitable platform. Please ensure that your code is well-structured, properly documented, and follows best practices for readability and maintainability.

Documentation:
Along with your submission, include a **Report or README** that provides an overview of the project, including:

- **Why did you choose the project?**
- Your key learnings and insights from working on the project.
- Any challenges faced and how they were addressed.
- Instructions on how to set up and run the project.
- Dependencies or prerequisites required.

If there are any additional considerations (e.g., dataset usage, API keys, or deployment instructions), please include them in the documentation.

Note: Sessions will be conducted to explain the concepts.