7830 Machine Learning II

Final Project

COLUMBUS STATE
COMMUNITY COLLEGE

Final Project Overview

Design and present an end-to-end ML solution using a **public dataset** that aligns with topics from class:

- Neural Networks
- Deep Learning
- Recommender Systems
- Reinforcement Learning

Find a Public Dataset

Use open platforms:

Kaggle, UCI ML Repository, OpenML, Google Dataset Search, etc.

Verify the license before use!

🔓 Common Data Licenses:

- **CC BY** – Credit required
- **CC0** – Free to use with no restrictions
- **ODC-By / ODbL** – Attribution/Share-Alike for databases
- **GPL/Apache/MIT** – More for code, but sometimes seen in datasets

Organize Your Jupyter Notebook by AI Project Lifecycle

1. **Define the Business Problem**
   - What is the goal? Why does it matter?
   - State if it's: **Supervised**, **Unsupervised**, **Recommender**, or **Reinforcement and why you chose that method.**
2. **Data Acquisition & EDA**
   - Show key insights and visualizations
3. **Data Cleaning**
   - Handle missing values, transformations, scaling

COLUMBUS STATE
COMMUNITY COLLEGE

Organize Your Jupyter Notebook by AI Project Lifecycle

1

Modeling

Build one or more models aligned with your problem type

2

Model Evaluation

- Report metrics
- Show how you improved it
- Write a short paragraph with your **model assessment**

3

Deployment Plan

Briefly describe **how you could deploy** your model (e.g., API, web app, embedded system)