Applied Machine Learning Project Sheet

Course: CS652 Applied Machine Learning

Duration: about 13 weeks

Project Objectives:

1. Apply machine learning techniques to solve a real-world problem.
2. Gain hands-on experience in machine learning workflow, e.g. data preprocessing, model selection, training, and evaluation.
3. Develop skills in critical analysis and problem-solving within the context of machine learning.

Project Instructions:

1. Project Proposal (Week 1, by 31 Jan.):

- Submit a one-page proposal outlining:
    - Problem statement
    - Objectives
    - Dataset to be used (public datasets encouraged)
    - Proposed machine learning techniques
    - Expected challenges

2. Data Collection & Preprocessing (Weeks 2-3, by 14 Feb.):

- Acquire and preprocess data:
    - Data cleaning (e.g. handling missing values, outliers)
    - Data transformation (e.g. normalization, encoding)
    - Feature selection/extraction (if necessary)

3. Model Development (Weeks 4-6, by 7 Mar.):

- Choose appropriate machine learning algorithms.
- Train multiple models and compare their performance.
- Perform hyperparameter tuning (optional).

5. Model Evaluation (Week 7-8, by 21 Mar.):

- Evaluate models using relevant metrics (e.g. accuracy, precision, recall, F1-score).
- Perform cross-validation (optional).

6. Result Analysis & Improvement (Week 9-10, by 4 Apr.):

- Analyze the results and discuss strengths and weaknesses.
- Implement improvements or alternative approaches (if necessary).

7. Final Presentation (Week 11-12, by 25 Apr.):

- 15-minute presentation covering problem, methodology, results, and insights.
- Include visualizations and a live demo if possible.

8. Report (Week 13, by 2 May.):

- 8-10 pages including:
    - Introduction & problem statement
    - Literature review (related works)
    - Methodology
    - Experiments & results
    - Discussion & conclusions
    - References & Appendices (code snippets, additional results)

Expected Outcomes:

1. A functional machine learning model addressing the chosen problem.
2. A comprehensive report documenting the entire project process.
3. A clear and engaging presentation showcasing the project.
4. Demonstrated understanding of applied machine learning workflows.

Assessment Criteria: (total 50% of the course evaluation)

- Project Proposal: 5%
- Technical Implementation: 15%
- Model Performance & Evaluation: 10%
- Final Report: 10%
- Presentation: 10%

Additional Notes:

- Use any programming language or framework (e.g. Python with scikit-learn, TensorFlow, PyTorch).
- Ensure ethical considerations are addressed, especially in data usage.