COE3012 Computer System Engineering
Group Project Assignment

Project Title

Running a Lightweight Machine Learning Model on NodeMCU with Simulated Sensor Data

Objective

Design and implement an embedded system using the NodeMCU board that simulates sensor data and runs a lightweight machine learning (ML) model (e.g., decision tree, linear regression, or KNN) directly on the board. The system should demonstrate a practical application such as:

- Environmental monitoring and anomaly detection
- Smart health or fitness tracking
- Home automation with intelligent decision-making
- Predictive maintenance using simulated sensor data

Key Constraint

No real sensors are required. Students must use simulated data or publicly available datasets to train their models offline. The trained models should be converted into lightweight formats (e.g., decision rules or TensorFlow Lite models) and deployed on the NodeMCU for inference.

Project Requirements

- Use a NodeMCU board (ESP8266 or ESP32)
- Simulate at least two types of sensor data (e.g., temperature, motion, gas, acceleration)
- Train a lightweight ML model offline using Python with libraries such as `scikit-learn`, `TensorFlow`, or `Keras`
- Compress/export the trained model using appropriate formats (e.g., TensorFlow Lite, lookup tables, rule-based logic)
- Deploy the model to the NodeMCU using Arduino IDE or other compatible tools
- Demonstrate on-board prediction or classification using the simulated sensor data
- Optional: Transmit results to a web dashboard or mobile app via WiFi (MQTT or HTTP)

Deliverables

1. **Proposal Presentation (Week 3)**
    - Project idea and use case
    - ML algorithm and dataset to be used
    - High-level system diagram
    - Timeline of work
2. **Final Presentation (Week 12)**
    - Working demo on NodeMCU (using simulated inputs)
    - Explanation of model training and compression process
    - Results and accuracy metrics
    - Challenges and future improvements
3. **Final Report (6 pages, IEEE format)**
    - Abstract, Introduction, Literature Review
    - Dataset description and methodology
    - Model Training and Deployment
    - Results and Evaluation
    - Conclusion and References

Marking Scheme (Total: 100 Marks)

- Proposal Presentation: 15 marks
    - Clarity of idea and relevance: 5
    - Technical feasibility: 5
    - Planning and structure: 5
- Final Presentation: 30 marks
    - On-board model execution and results: 15
    - Technical explanation: 10
    - Creativity and innovation: 5
- Final Report: 35 marks
    - Organization and clarity: 5
    - Technical details: 10
    - Model training and deployment: 10
    - Evaluation and insights: 10
- Teamwork and Contribution (peer-reviewed): 10 marks
- Q&A during presentations: 10 marks

Group Composition

Each group must consist of 3–4 students. Roles should be clearly distributed (e.g., dataset researcher, model trainer, deployment programmer, documentation lead).

Submission Deadlines

- Proposal Presentation: **Week 3**
- Final Demonstration and Presentation: **Week 12**
- Final Report Submission: **Week 13**

Note

Plagiarism or reusing previous projects is strictly prohibited. Simulated data and open datasets must be properly cited. All work must be original and reproducible.