Sensor-Based Human Activity Classification Challenge

Dataset: PIRvision_FoG_presence_detection

1 Dataset Overview

The PIRvision dataset contains occupancy detection data collected using a synchronized Low-Energy Electronically-chopped Passive Infra-Red (PIR) sensor node deployed in residential and office environments. Each observation corresponds to 4 seconds of recorded human activity within the sensor's Field-of-View (FoV).

1.1 Data Description

- **Tabular Features:**
  - **Date**: Observation date
  - **Time**: Observation time
  - **Label**: Activity label (target for classification)
  - **Temperature**: Ambient temperature (°F)
- **Time Series Features:**
  - **PIR values**: 55 analog sensor readings captured over 4 seconds.

1.2 Data Link

The dataset is available at:

`https://archive.ics.uci.edu/dataset/1101/pirvision_fog_presence_detection`

1.3 Team Details Submission

Team details submission link: **Click here to open**

2 Competition Task

Develop and rigorously evaluate a robust classification model, exclusively utilizing the provided dataset, to predict activity labels accurately. Your evaluation must employ 5-fold cross-validation. Clearly report your model's performance, including:

- **Mean Accuracy** across all folds.
- **Standard Deviation** of accuracy scores to demonstrate robustness and consistency.

Ensure your evaluation and reporting are clear, replicable, and methodologically sound.

3 Important Rules & Guidelines

- **Teams:** Exactly 2 students per team (mandatory).
- **No external datasets** are permitted. Only the provided dataset should be used.
- **Pretrained, Large Language Models (LLMs) and transfer learning** are strictly prohibited.
- **Originality required:** Cheating/plagiarism results in immediate zero marks for the project component.
- Deviations from instructions result in immediate zero marks.

4 Submission Requirements

Submissions must strictly adhere to the following format:

4.1 File Naming Convention

Submit one Jupyter Notebook (.ipynb) file:

`team_<team_number>.ipynb`

Checkpoint(s) files must also adhere strictly to the same naming convention provided.

4.2 Notebook Content

Clearly document every step with thorough explanations:

- Data loading, preprocessing, feature engineering
- Model building, training, testing, and evaluation
- Justification of chosen methods and hyperparameters

4.3 Visualizations & Plots

Mandatory clearly labeled visualizations:

- Training vs. Validation Loss curves
- Training vs. Validation Accuracy graphs

Additional meaningful visualizations beyond the mandatory requirements are encouraged and will be rewarded.

4.4 README File

Provide a detailed README file documenting:

- Libraries used
- Instructions on how to execute/run your code

4.5 Evaluation Function

Include the following Python function within your notebook to facilitate accurate evaluation:

```
def evaluate_model(datafile_location, checkpoint):
    # your code here
    return accuracy
```

Notes:

- Only one submission per team is accepted.
- Ensure all mandatory files (notebook, README, checkpoints) are included.
- Failure to submit correctly (even if one missing component) will lead to zero marks.
- Failure to reproduce your results using provided instructions or any deviation in training data will result in zero marks.
- Late submissions are strictly prohibited. No deadline extensions will be granted under any circumstances.

5 Evaluation & Presentation

Teams will participate in a scheduled evaluation after submission. The evaluation includes a brief **presentation (15 minutes)** clearly addressing:

- **Problem Understanding:** Demonstrate clear understanding of dataset and objectives.
- **Exploratory Analysis:** Insights derived from the data.
- **Feature Engineering & Selection:** Justify preprocessing and feature choices.
- **Model Selection & Validation:** Explain algorithm choices, hyperparameter tuning strategies, and validation methods.
- **Results & Insights:** Present and analyze final results.
- **Challenges & Future Improvements:** Discuss obstacles and propose improvements.

Follow-up Q&A evaluates your understanding and originality.

6 Project Expectations

- **Clear Problem Definition:** Demonstrate a solid understanding of the PIR sensor dataset and classification objectives.
- **Thoughtful Exploratory Data Analysis (EDA):** Visualize data distributions, correlations, anomalies, and insights effectively.
- **Robust Feature Engineering:** Show creativity in feature extraction and preprocessing to enhance performance.
- **Appropriate Algorithm Selection:** Provide clear justifications for your choice of model(s).

- **Rigorous Validation & Evaluation:** Accurately apply cross-validation, ensuring reproducibility.
- **Quality of Documentation and Code:** Clearly structured code with thorough commenting and easily reproducible steps.

7 Important Remarks

- Start early; no deadline extensions under any circumstances.
- Strict adherence to submission guidelines. Any missing component or deviation in results in zero marks.

Submission Deadline: 13-April-2025

Good luck! Showcase your creativity and analytical skills!