# Course Project: Bone Age Prediction from Hand Radiographs

Pattern Recognition And Machine Learning Course Project Specification

## 1. Problem Statement

The objective of this project is to estimate a patient's **bone age** (a continuous value measured in years) from a given **hand X-ray image**. Bone age reflects skeletal maturity and is widely used in pediatrics to assess growth disorders and endocrine abnormalities.

This task is primarily a **regression problem**, where the model must learn a mapping:

$$f : \text{Image} \rightarrow \text{Age (in years)}$$

Students must also build a **classification model** by categorizing bone age into discrete developmental stages (e.g., *Child*, *Adolescent*, *Adult*).

## 2. Dataset Description and CSV File Format

Students will use the **RSNA Bone Age Dataset**, which contains thousands of hand X-ray images and corresponding chronological ages.

Each X-ray image is accompanied by metadata stored in a CSV file. The general structure of the CSV file is as follows:

| **Column Name** | **Description** | **Example Value** |
|---|---|---|
| `id` | Unique image identifier (without extension) | 12345 |
| `bone_age` | Bone age of the patient in months (target for regression) | 156 |
| `sex` | Biological sex of the patient (M or F) | F |

**Image Files:** The dataset folder typically contains images named as: `<id>.png` or `<id>.jpg`, each corresponding to one record in the CSV file.

### Necessary initial steps

- Resize all images to a fixed size.
- Split the dataset into `train`, `validation`, and `test` sets (e.g., 70/15/15 split).

# 3. Expected Outputs and Deliverables

## A. Core Deliverables

1. **Code:** A clean, reproducible Python codebase or Jupyter Notebook implementing:
    - Data loading and preprocessing.
    - Feature extraction using classical methods (e.g., HOG, texture descriptors) or CNNs.
    - Regression model training and evaluation.
    - Appropriate bone age classification.
    - Scatter plots of predicted vs. actual age values.
    - Heatmaps (e.g., Grad-CAM) to visualize influential image regions.
    - Discussion of bias or error trends across sex or age groups.
2. **Model Explanation:**
    - Architecture or model design summary.
    - Justification for preprocessing and feature choices.
    - Visualization of learned representations.
    - Perform gender-wise performance and bias analysis.
3. **Result Report:**
    - Comparison between different approaches.
    - Plot of predicted vs. true ages on validation/test set.
    - Interpretation of errors and discussion of difficult samples.

### C. Final Submission Must Include

1. Well-documented code (Jupyter Notebook or Python script).
2. Trained model weights or reproducible training instructions.
3. Short report (3–5 pages) including:
   - Abstract, Methodology, Results, and Discussion.
   - Error analysis and insights from model behavior.
   - References to datasets and any pretrained models used.

## 4. Evaluation Metrics

### Regression Metrics

- **Mean Absolute Error (MAE):**

  $$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

  Measures the average absolute deviation between predicted and true ages.

- **Root Mean Squared Error (RMSE):**

  $$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

  Penalizes larger errors more heavily, providing a sense of overall prediction spread.

- **Coefficient of Determination ($R^2$):**

  $$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

  Indicates how well the regression model explains the variance in bone age.

## Classification Metrics

- **Accuracy:** Fraction of correctly predicted age-stage classes.
- **Precision, Recall, and F1-score:** Useful for imbalanced bins.
- **Confusion Matrix:** Visualization of true vs. predicted categories.
- **Quadratic Weighted Kappa (QWK):** Measures agreement between predicted and true ordinal classes while discounting random agreement.