

## Daily Diary: Machine Learning (Weekend Classes)

### November 4th (Saturday)

- **Topic:** Introduction to Machine Learning, Supervised Learning & Unsupervised Learning
  - **Notes:**
    - Defined Machine Learning as a field of study that gives computers the ability to learn without being explicitly programmed.
    - Discussed the key concepts: learning from data, making predictions, and improving performance over time.
    - Explored different types of machine learning: Supervised, Unsupervised, and Reinforcement Learning.
    - Focused on supervised learning, where the algorithm learns from labeled data.
    - Discussed regression and classification as two main types of supervised learning problems.
    - Examined examples of regression (predicting house prices) and classification (spam detection).
    - Explored unsupervised learning, where the algorithm learns from unlabeled data.
    - Discussed clustering (grouping similar data points) and dimensionality reduction (reducing the number of features) as key unsupervised learning techniques.
    - Studied examples of clustering (customer segmentation) and dimensionality reduction (principal component analysis).

### November 5th (Sunday)

- **Topic:** Reinforcement Learning
  - **Notes:**
    - Introduced reinforcement learning, where an agent learns to interact with an environment to maximize rewards.
    - Discussed the concept of an agent, environment, actions, and rewards in reinforcement learning.
    - Explored examples of reinforcement learning (game playing, robotics).

### November 11th (Saturday)

- **Topic:** Train & Test, Validation Split
  - **Notes:**
    - Emphasized the importance of splitting data into training and testing sets.
    - Explained the purpose of a validation set for model selection and hyperparameter tuning.
    - Discussed techniques for splitting data effectively (e.g., stratified sampling).

### November 12th (Sunday)

- **Topic:** Performance Metrics
  - **Notes:**
    - Explored various performance metrics for evaluating machine learning models:
      - **Regression:** Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R-squared
      - **Classification:** Accuracy, Precision, Recall, F1-score, AUC-ROC
    - Discussed the choice of appropriate metrics based on the problem and business goals.

### November 18th (Saturday)

- **Topic:** Overfitting & Underfitting
  - **Notes:**
    - Defined overfitting and underfitting as key challenges in machine learning.
    - Explained how overfitting occurs when a model performs well on training data but poorly on unseen data.
    - Discussed techniques to prevent overfitting (e.g., regularization, cross-validation).

### November 19th (Sunday)

- **Topic:** Linear Regression
  - **Notes:**
    - Introduced linear regression as a simple yet powerful supervised learning algorithm for predicting continuous values.
    - Discussed the underlying assumptions of linear regression (linearity, independence, normality, homoscedasticity).
    - Explored the concepts of R-squared and adjusted R-squared for model evaluation.

### December 1st (Saturday)

- **Topic:** Introduction to Scikit-learn
  - **Notes:**
    - Introduced Scikit-learn, a popular Python library for machine learning.
    - Explored the key features of Scikit-learn, including its comprehensive collection of algorithms, data preprocessing tools, and model evaluation metrics.
    - Practiced using Scikit-learn to implement simple machine learning models.

### December 2nd (Sunday)

- **Topic:** Logistic Regression
  - **Notes:**
    - Introduced logistic regression as a widely used algorithm for binary classification.

- Discussed the sigmoid function and its role in logistic regression.
- Explored the concepts of precision, recall, and the ROC curve for evaluating binary classifiers.

### **December 8th (Saturday)**

- **Topic:** F-Score, Decision Trees
  - **Notes:**
    - Defined the F1-score as a harmonic mean of precision and recall.
    - Introduced decision trees as a non-parametric supervised learning algorithm for both classification and regression.
    - Explored the concept of information gain and how it is used to build decision trees.

### **December 9th (Sunday)**

- **Topic:** Cross-Validation, Bias vs. Variance
  - **Notes:**
    - Discussed cross-validation techniques (k-fold, leave-one-out) for model evaluation and hyperparameter tuning.
    - Explored the bias-variance tradeoff, a fundamental concept in machine learning that helps understand model performance.

### **December 15th (Saturday)**

- **Topic:** Ensemble Approaches
  - **Notes:**
    - Introduced ensemble learning as a powerful technique that combines multiple models to improve performance.
    - Discussed bagging and boosting as two main types of ensemble methods.

### **December 16th (Sunday)**

- **Topic:** Bagging, Boosting
  - **Notes:**
    - Explored bagging (e.g., Random Forest) and boosting (e.g., AdaBoost, Gradient Boosting) in more detail.
    - Discussed the advantages and disadvantages of bagging and boosting.