Homework – Week 3: Precision Metrics and Model Evaluation

Course: Introduction to Machine Learning
Andres Mendez-Vazquez

Instructions

Answer the following questions using complete sentences. Be clear and concise. You may answer by hand or type your answers in a PDF document.

Part I: Precision vs Recall

1. Explain in your own words the difference between **precision** and **recall**.

2. For each of the following situations, indicate whether **precision** or **recall** is more important, and explain why:

   - (a) Email spam detection.
   - (b) Diagnosing a serious disease (e.g., cancer).
   - (c) Detecting fraud in financial transactions.
   - (d) Facial recognition for unlocking a phone.

Part II: Bias and Variance (Qualitative)

1. In your own words, describe what it means for a model to have:

   - (a) High bias
   - (b) High variance

2. Suppose you train a logistic regression model on a dataset and get the following results:

   - Training Accuracy: 75%
   - Testing Accuracy: 74%

Do you think this model suffers more from bias or from variance? Justify your answer briefly.

3. Suggest one way you could try to reduce the bias of a model. (You may give an idea even if we have not covered all models yet.)

Optional Challenge (Bonus)

Find a real-world example (news, science, health, industry, etc.) where it is especially important to maximize either precision or recall. Briefly describe the scenario and explain which metric is more critical.