1 Introduction

Decision trees are a powerful machine learning tool used for classification and regression tasks, representing decisions in a flowchart-like structure. This tutorial provides a comprehensive guide to building a decision tree using the ID3 (Iterative Dichotomiser 3) algorithm, focusing on classification. We'll use the "Play Tennis" dataset to demonstrate the process, including calculations for Entropy and Information Gain, and show how to construct and use the tree for predictions. This guide is beginner-friendly yet detailed, with a complete example dataset.

2 Why Decision Trees?

Decision trees are popular because:

- They are intuitive and easy to visualize.
- They handle both categorical and numerical data.
- No data normalization is required.
- They capture non-linear relationships.

Applications include spam detection, customer churn prediction, and medical diagnosis. We'll use the ID3 algorithm, which selects splits based on Information Gain.