

# Decision Trees Wrap-Up

Module 11 Conclusion



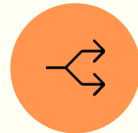
# What We Learned

Decision trees are powerful, interpretable models that recursively split data to make predictions. Let's recap the core concepts we covered in this module.



## Tree Structure

How decision trees work by creating a hierarchy of decisions based on feature values



## Split Selection

How trees choose the best splits using impurity measures to maximize information gain



## Key Metrics

Entropy, Gini impurity, and Information Gain for evaluating split quality



## Visualization

Reading and interpreting tree structures to understand model decisions



## Pruning

Controlling tree complexity to prevent overfitting and improve generalization

# Model Evaluation Recap

Evaluating model performance requires more than a single metric. We explored a comprehensive toolkit for assessing classification models.

O1

---

## Accuracy

Overall correctness—useful baseline but can be misleading with imbalanced data

O2

---

## Precision & Recall

Trade-offs between false positives and false negatives depending on application needs

O3

---

## F1 Score

Harmonic mean balancing precision and recall for a single performance measure

O4

---

## Confusion Matrix

Detailed breakdown of true/false positives and negatives revealing error patterns

O5

---

## ROC & AUC

Threshold-independent evaluation showing model discrimination across all cutoffs

# Key Takeaways

## Intuitive & Explainable

Trees mirror human decision-making, making them easy to interpret and explain to non-technical stakeholders

## Unstable & Sensitive

Small changes in data can lead to completely different tree structures, reducing reliability

## Overfitting Risk

Without constraints, trees memorize training data rather than learning generalizable patterns

## Pruning is Essential

Controlling tree depth and complexity is critical for building models that perform well on unseen data

## Beyond Accuracy

Choose evaluation metrics that align with your problem—precision, recall, or F1 may matter more than raw accuracy

# Moving Forward

Individual decision trees are just the beginning. Next, we'll explore how combining multiple trees creates more robust, accurate models.



## Decision Trees

Individual models—interpretable but unstable



## Ensemble Methods

Combining trees for superior performance

## What's Next

**Random Forests** aggregate predictions from many trees built on random subsets of data and features, reducing variance and improving stability.

**Gradient Boosting** builds trees sequentially, each correcting errors from the previous one, creating highly accurate predictive models.

These ensemble techniques deliver stronger, more stable, and more accurate predictions than any single tree.