Group Study

Rule Post Pruning

Introduction

In machine learning, particularly in decision tree learning, overfitting is a problem .

Common problem where a model fits the training data too well, capturing noise .

Rule post-pruning is an attempt to reduce its ability to generalize to unseen data .

The effective technique to address this is by simplifying the model after it has been simplified .

Rather than halting tree growth early (pre-pruning), rule post-pruning .

Allows the tree to grow fully and then trims it back, improving its generalization .

This method has been found to yield highly accurate hypotheses .

It is widely used in practical applications of decision trees .

Objective:

The objective of this study is to explore and understand the

concept of rule post-rule .

Pruning in decision tree learning, focusing on its methodology, effectiveness in decision tree learning

Reduce overfitting and its application in generating high-accuracy , and reducing overfitting .

This includes analysing the step-by-step generalizable and interpretable hypotheses .

The step process of converting a fully grown decision tree into a simplified rule-based decision tree is a step-by-step process of converting a fully grown decision tree into a simplified rule-based decision tree .

Model and evaluating its usefulness across various real-world domains.

Process:

Rule post-pruning is a successful method for finding high accuracy hypotheses

• Rule post-pruning involves the following steps:

• Infer the decision tree from the training set, growing the tree until the training set

Data is fit as well as possible and allowing overfitting to occur .

• Convert the learned tree into an equivalent set of rules by

creating a rule for each rule

Each path from the root node to a leaf node has a path from the root node to a leaf node .

• Prune(generalize) each rule by removing any preconditions that result in

Improved its estimated accuracy

• Sort the pruned rules by their estimated accuracy and consider them in this way

When classifying subsequent instances, the sequence is used to classify subsequent instances .

Converting a Decision Tree into Rules: Converting a Decision Tree into Rules

For example, consider the decision tree : the leftmost path of the tree in below .

The figure is translated into the rule .

IF (Outlook = Sunny) (Humidity = High)

THEN play tennis = No.

Given the rule above, rule post-pruning would consider removing the rule .

preconditions

(Outlook = Sunny) and (Humidity = High)

• It would select which of these pruning steps produced the greatest results

If there is improvement in estimated rule accuracy, then consider pruning the second one.

As a further pruning step, the precondition is a precondition .

• No pruning step is performed if it reduces the estimated rule accuracy

There are three main advantages by converting the decision tree to rules .

before pruning:

• Allows separate pruning decisions for the same attribute in different paths .

• Avoids complex restructuring of the tree when pruning high-level nodes .

• Improves readability and makes the model easier to understand

Application of Rule Post-Pruning Applications of Rule Post-Pruning

Medical Diagnosis: Helps simplify complex diagnostic rules for better diagnosis

Interpretability while maintaining accuracy is maintained .

Finance: Used in credit scoring and risk assessment to reduce overfitting while

Keeping decision criteria understandable is important.

Customer Behaviour Prediction: Applied in marketing and recommendation

Systems to extract general patterns from customer data .

Fraud Detection: Detects anomalous behaviours by refining rules

Conclusion:

Rule post-pruning enhances the generalization ability of decision tree models by enhancing rule post-pruning .

As illustrated, simplifying rules without compromising their predictive accuracy is illustrated .

In the example, the process involves evaluating each precondition in a rule .

By selectively removing those that do not contribute to or improve accuracy , we remove those that do not contribute to or improve accuracy .

pruning only when it results in a performance gain, the method ensures that

The final rules remain both effective and interpretable despite this selective simplification .

Helps avoid overfitting and leads to more robust models suitable for real-world use .

Decision-making tasks are

Presented by:

K. Swathi (22491A4722)

Khadharunnisa (22491A4746)

Y. Radha Krishna (22491A4756)