Decision Tree Evaluation Report

2105166

Overview

This report investigates the behavior of decision tree classifiers using three splitting criteria:

- **Information Gain (IG)**
- **Information Gain Ratio (IGR)**
- **Normalized Weighted Information Gain (NWIG)**

We evaluate their performance on:

- **Iris**: A small, clean dataset with numerical features.
- **Adult**: A large real-world dataset with both categorical and numerical attributes.

Tree depth is varied from 1 to 10. For each configuration, average accuracy and node count are computed over 20 experiments.

Visualizations

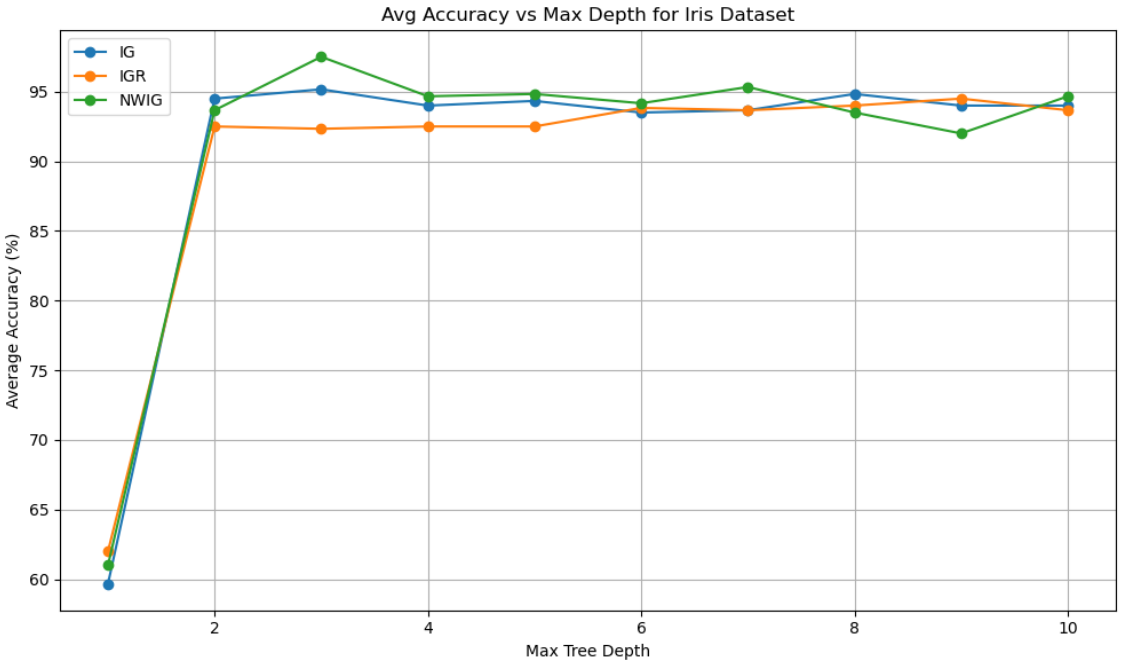

Figure 1: Iris Dataset: Accuracy vs Max Depth

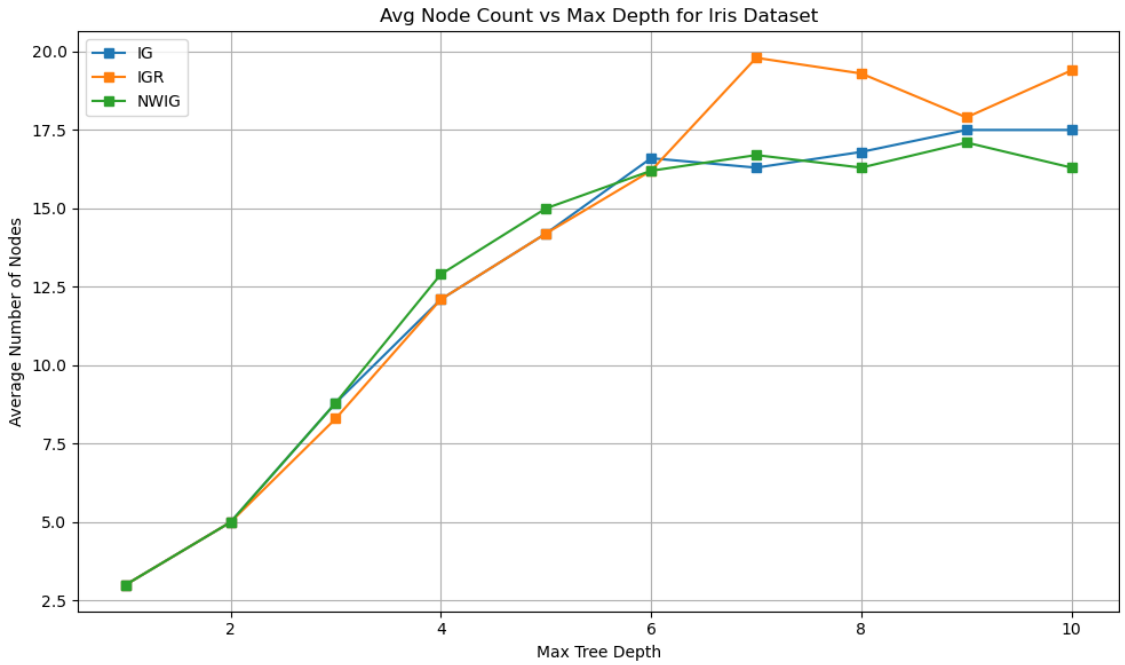

Figure 2: Iris Dataset: Node Count vs Max Depth

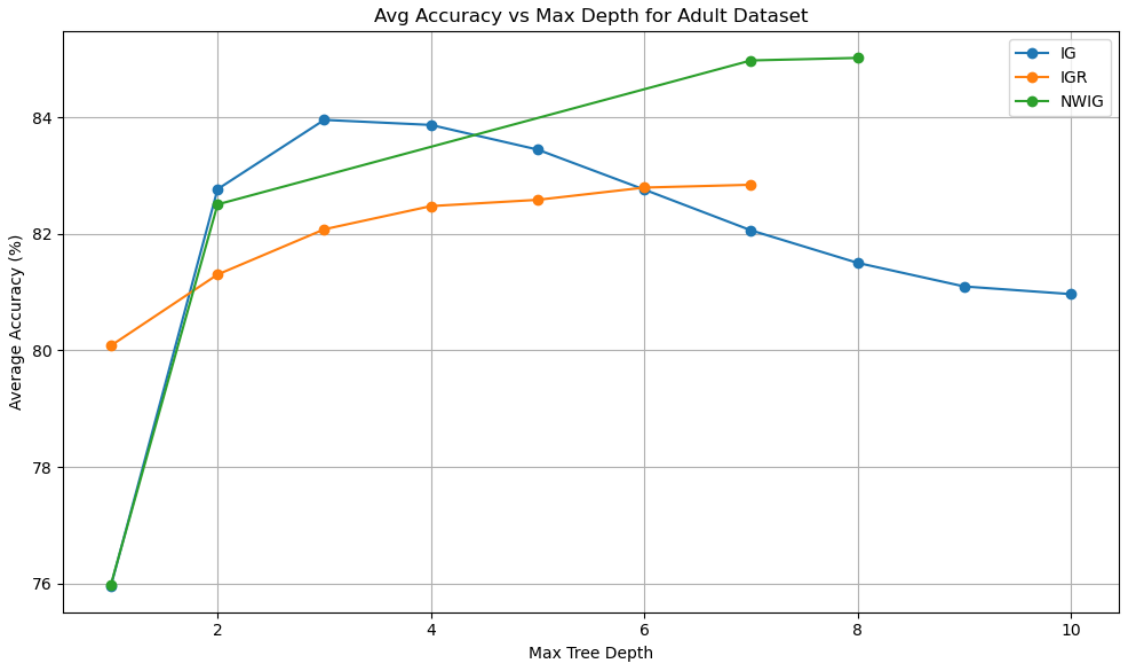

Figure 3: Adult Dataset: Accuracy vs Max Depth

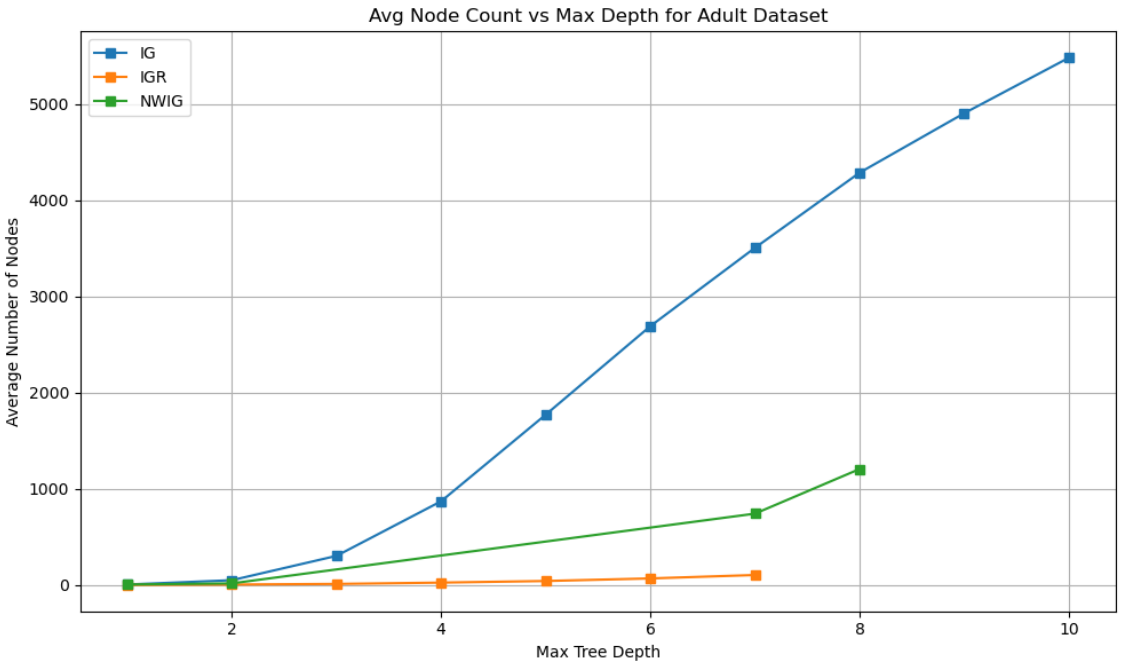

Figure 4: Adult Dataset: Node Count vs Max Depth

Result Tables

Iris Dataset

Criterion	Max Depth	Avg Accuracy (%)	Avg Nodes
IG	1	59.67	3
IG	2	94.50	5
IG	3	95.17	8.8
IG	4	94.00	12.1
IG	5	94.33	14.2
IG	6	93.50	16.6
IG	7	93.67	16.3
IG	8	94.83	16.8
IG	9	94.00	17.5
IG	10	94.00	17.5
IGR	1	62.00	3
IGR	2	92.50	5
IGR	3	92.33	8.3
IGR	4	92.50	12.1
IGR	5	92.50	14.2
IGR	6	93.83	16.2
IGR	7	93.67	19.8
IGR	8	94.00	19.3
IGR	9	94.50	17.9
IGR	10	93.67	19.4
NWIG	1	61.00	3
NWIG	2	93.67	5
NWIG	3	97.50	8.8
NWIG	4	94.67	12.9
NWIG	5	94.83	15.0
NWIG	6	94.17	16.2
NWIG	7	95.33	16.7
NWIG	8	93.50	16.3
NWIG	9	92.00	17.1
NWIG	10	94.67	16.3

Table 1: Iris Dataset Full Results

Adult Dataset

Criterion	Max Depth	Avg Accuracy (%)	Avg Nodes
IG	1	75.96	7
IG	2	82.77	52.05
IG	3	83.96	305
IG	4	83.87	872.45
IG	5	83.45	1772.65
IG	6	82.76	2692.15
IG	7	82.06	3506.15
IG	8	81.50	4286.75
IG	9	81.10	4904.65
IG	10	80.97	5480.95
IGR	1	80.08	3
IGR	2	81.30	7
IGR	3	82.08	13.3
IGR	4	82.48	27
IGR	5	82.59	43.65
IGR	6	82.80	70.4
IGR	7	82.84	105.95
NWIG	1	75.98	7
NWIG	2	82.51	19
NWIG	7	84.98	743.65
NWIG	8	85.02	1207.15

Table 2: Adult Dataset Full Results

Observations and Analysis

Iris Dataset

- NWIG achieved highest accuracy (97.5%) at depth 3 with only 8.8 nodes on average.
- IG had accuracy around 94–95% from depth 3 onwards but tree size grew unnecessarily.
- IGR performed well and offered better compactness with competitive accuracy.

Adult Dataset

- IG produced large trees (up to 5480 nodes) but showed limited accuracy improvement beyond depth 4–5.
- NWIG performed best at depth 8 with 85% accuracy and manageable tree size.
- IGR consistently achieved ¿82% accuracy with small trees, making it a good balance.

Conclusion

- **NWIG** is the most effective criterion in both datasets, offering superior accuracy and efficiency.
- **IGR** provides a solid trade-off between accuracy and complexity.

- **IG** tends to overfit and generate large trees especially in deeper configurations.
- Pruning or limiting depth is essential for real-world datasets like Adult to avoid overfitting and inefficiency.