DECISION TREE
EXPERIMENTAL ANALYSIS REPORT

CSE 318 Offline 4

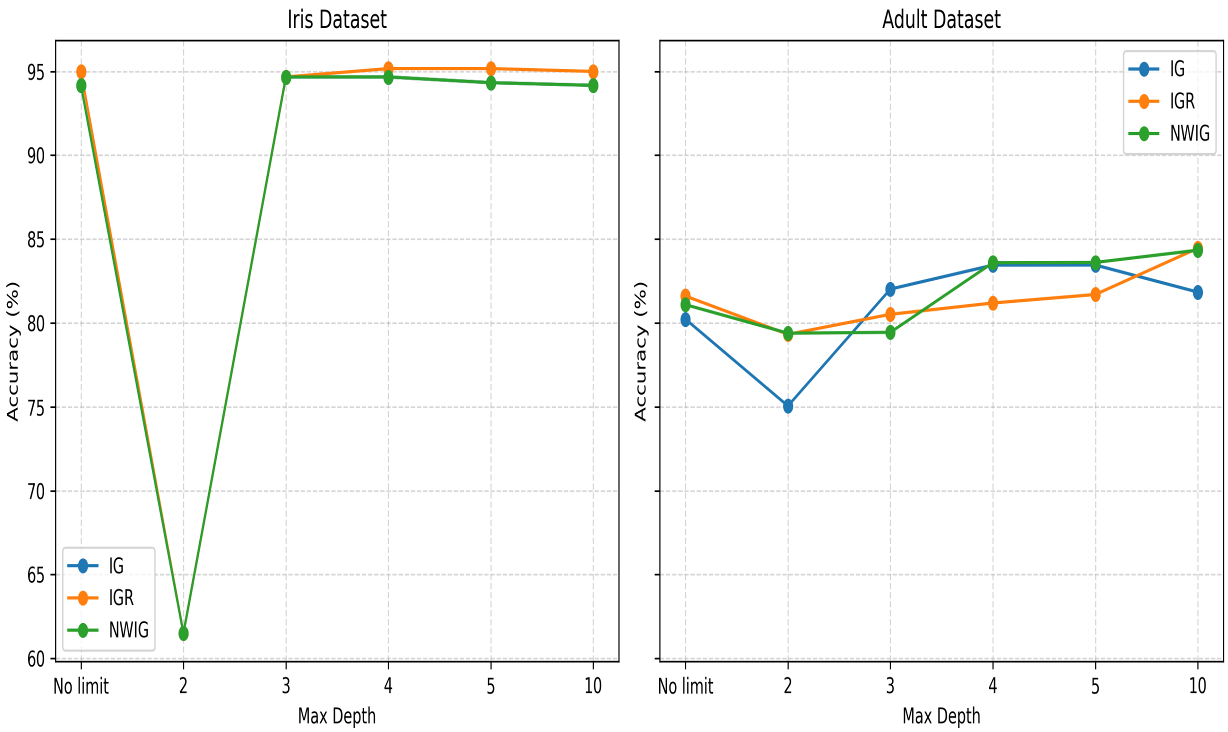
Shemanty Mahjabin
ID: 2105091

Iris Dataset Results Summary

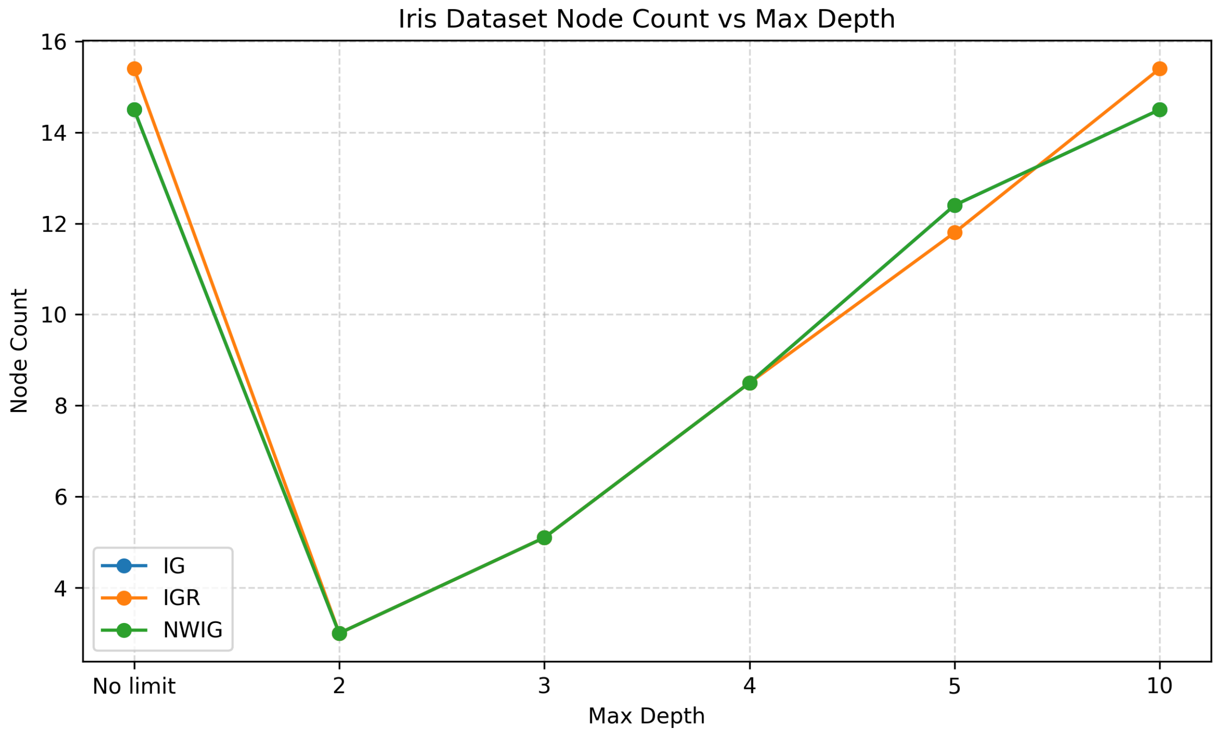
Criterion	Max Depth	Accuracy	Std Dev	Nodes	Avg Depth
IG	No limit	94.17	4.44	14.5	6.0
IG	2	61.50	5.87	3.0	2.0
IG	3	94.67	3.96	5.1	3.0
IG	4	94.67	3.96	8.5	4.0
IG	5	94.33	4.47	12.4	5.0
IG	10	94.17	4.44	14.5	6.0
IGR	No limit	95.00	3.67	15.4	6.3
IGR	2	61.50	5.87	3.0	2.0
IGR	3	94.67	3.96	5.1	3.0
IGR	4	95.17	3.82	8.5	4.0
IGR	5	95.17	3.50	11.8	5.0
IGR	10	95.00	3.67	15.4	6.3
NWIG	No limit	94.17	4.44	14.5	6.0
NWIG	2	61.50	5.87	3.0	2.0
NWIG	3	94.67	3.96	5.1	3.0
NWIG	4	94.67	3.96	8.5	4.0
NWIG	5	94.33	4.47	12.4	5.0
NWIG	10	94.17	4.44	14.5	6.0

Adult Dataset Results Summary

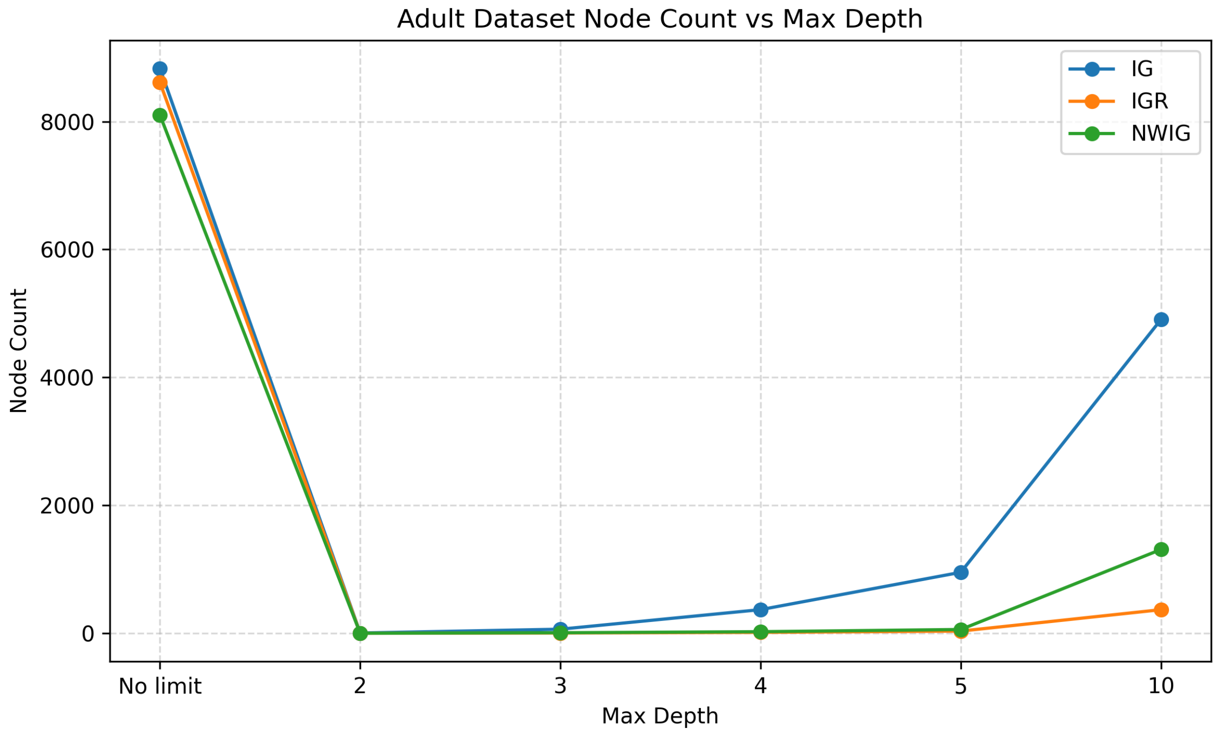
Criterion	Max Depth	Accuracy	Std Dev	Nodes	Avg Depth
IG	No limit	80.23	0.57	8827.2	32.8
IG	2	75.05	0.35	7.0	2.0
IG	3	82.02	0.26	66.4	3.0
IG	4	83.45	0.34	372.1	4.0
IG	5	83.46	0.32	955.6	5.0
IG	10	81.84	0.51	4907.0	10.0
IGR	No limit	81.61	0.53	8609.9	55.8
IGR	2	79.32	0.34	3.0	2.0
IGR	3	80.52	0.50	7.0	3.0
IGR	4	81.19	0.41	16.8	4.0
IGR	5	81.70	0.33	36.7	5.0
IGR	10	84.46	0.26	371.4	10.0
NWIG	No limit	81.09	0.42	8100.9	37.2
NWIG	2	79.39	0.33	3.0	2.0
NWIG	3	79.44	0.33	11.0	3.0
NWIG	4	83.59	0.21	27.4	4.0
NWIG	5	83.61	0.22	61.8	5.0
NWIG	10	84.34	0.36	1312.4	10.0



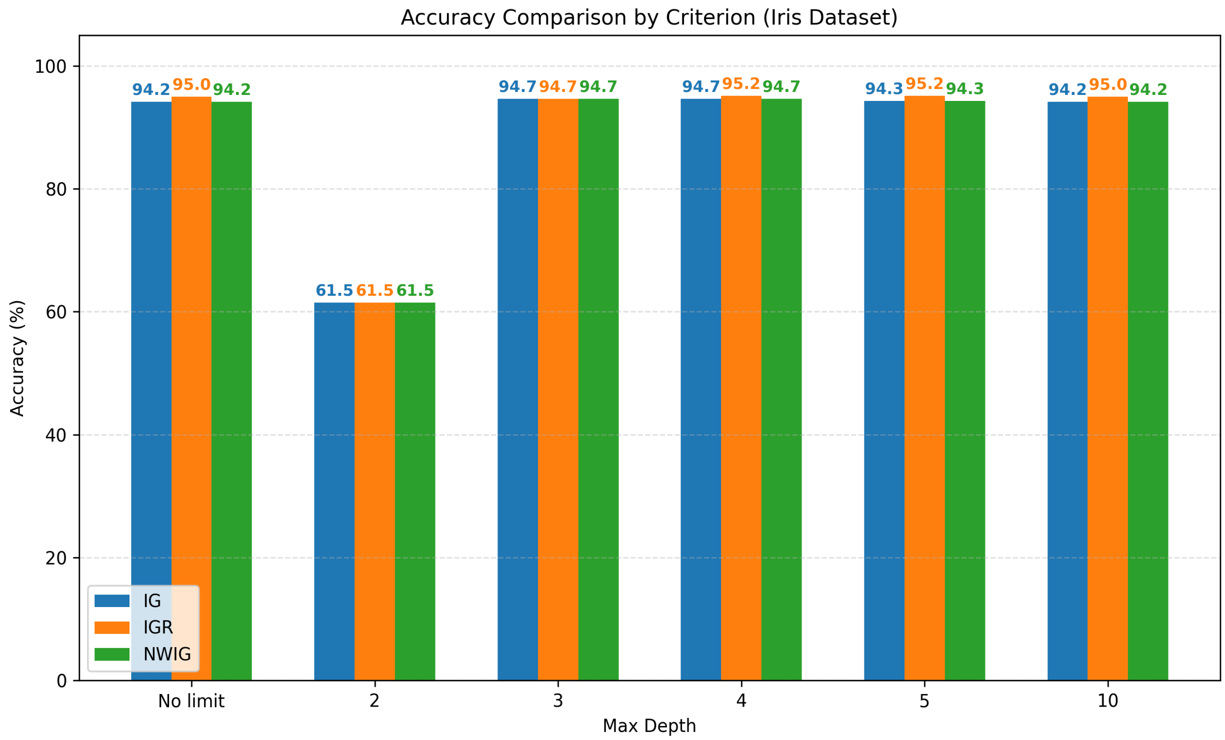
Iris Dataset
Adult Dataset
Accuracy (%)
Max Depth
IG
IGR
NWIG
No limit
2
3
4
5
10
95
90
85
80
75
70
65
60



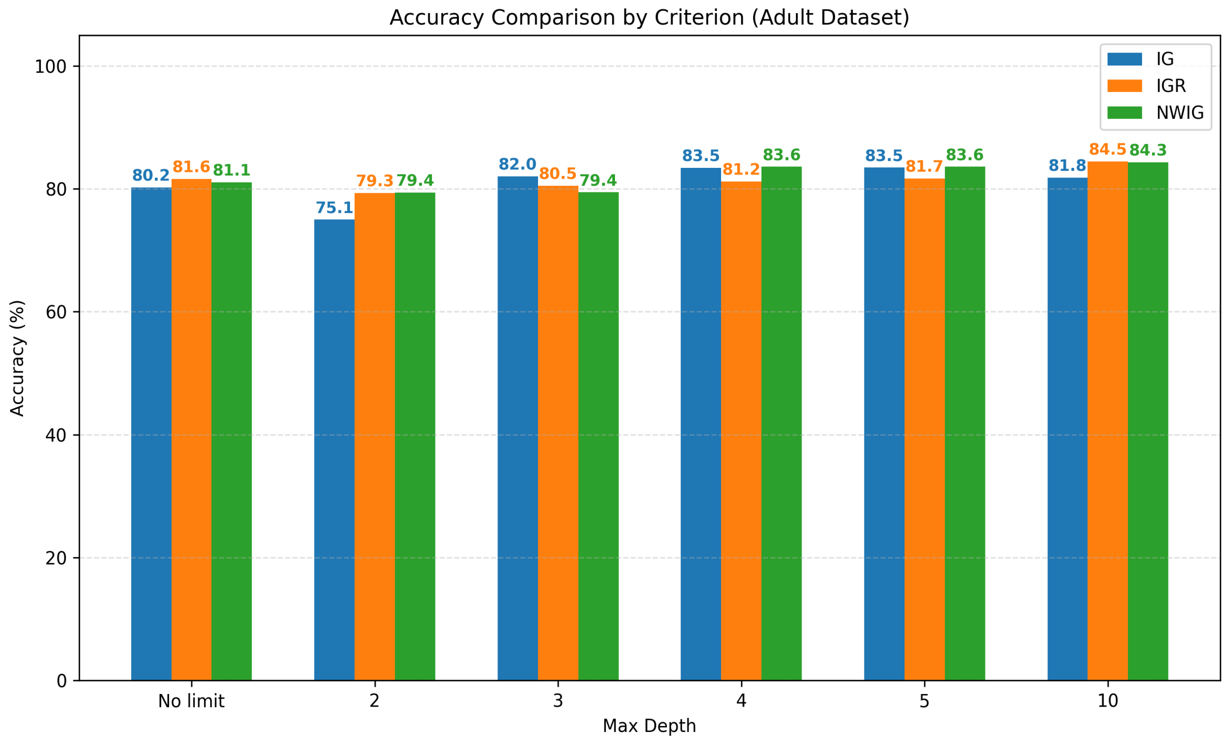
Iris Dataset Node Count vs Max Depth
Node Count
Max Depth
16
14
12
10
8
6
4
No limit
2
3
4
5
10
IG
IGR
NWIG



Adult Dataset Node Count vs Max Depth
IG
IGR
NWIG
Node Count
8000
6000
4000
2000
0
No limit
2
3
4
5
10
Max Depth



Accuracy Comparison by Criterion (Iris Dataset)
Accuracy (%)
Max Depth
0
20
40
60
80
100
No limit
2
3
4
5
10
94.2
95.0
94.2
61.5
61.5
61.5
94.7
94.7
94.7
94.7
95.2
94.7
94.3
95.2
94.3
94.2
95.0
94.2
IG
IGR
NWIG


Accuracy Comparison by Criterion (Adult Dataset)
IG
IGR
NWIG
Accuracy (%)
Max Depth
No limit
2
3
4
5
10
0
20
40
60
80
100
80.2
81.6
81.1
75.1
79.3
79.4
82.0
80.5
79.4
83.5
81.2
83.6
83.5
81.7
83.6
81.8
84.5
84.3

Analysis and Observations

DATASET CHARACTERISTICS

Iris Dataset:

- 150 samples, 4 features, 3 classes
- Well-separated, linearly separable data
- Classic benchmark dataset for classification

Adult Dataset:

- 30,162 samples, 14 features, 2 classes
- Mixed categorical and numerical features
- Real-world income prediction problem

CRITERION PERFORMANCE ANALYSIS

- **IG**: Simple, intuitive, strong baseline. Stable on Iris, strong on Adult.
- **IGR**: Corrects for attribute bias. Best on Iris, strong on Adult at deeper trees.
- **NWIG**: Balances gain and split quality. Consistent, best at shallow depths.

DEPTH ANALYSIS AND OVERFITTING

Pruning Effects:

- Limiting tree depth (pruning) generally improves generalization, especially for Adult.
- For both datasets, best accuracy is often at moderate depths (3-5).
- Pruning reduces model complexity and risk of overfitting.

Overfitting Indicators:

- For Adult, node count explodes at depth=0 (no limit) but accuracy does not improve.
- For Iris, accuracy is stable across depths, indicating low overfitting risk.
- Standard deviation is higher for shallow trees, especially on Iris.

TREE COMPLEXITY ANALYSIS

Node Count vs Accuracy Trade-off:

- Shallow trees (depth 2): very low node count, but lower accuracy (especially for IG).
- Moderate depths (3-5): best trade-off between accuracy and complexity for all criteria.
- No limit: node count increases rapidly, but accuracy plateaus or even drops (overfitting).

Complexity Observations:

- Adult dataset: node count grows rapidly with depth, especially for IG and IGR.
- NWIG and IGR tend to produce more compact trees at moderate depths.
- All criteria show diminishing returns in accuracy beyond depth 4-5.

CONSISTENCY AND ROBUSTNESS

- IGR and NWIG are most robust to overfitting and perform consistently across depths.
- IG is a reliable baseline but can be outperformed by IGR/NWIG at most depths.
- Standard deviation is generally low for Adult, higher for Iris due to smaller sample size.

DATASET-SPECIFIC FINDINGS

Iris Dataset:

- All criteria reach >94% accuracy at moderate depths (3-5).
- IGR and NWIG slightly outperform IG at most depths.
- Overfitting is not a major concern; accuracy is stable.

Adult Dataset:

- Pruning is essential: node count explodes at depth=0, but accuracy does not improve.
- NWIG and IGR outperform IG at most depths.
- Best accuracy is achieved at moderate depths (4-5), with much lower complexity.

RECOMMENDATIONS

For Iris-like Datasets (small, well-separated):

- Use IGR or NWIG at depth 3-5 for optimal accuracy and simplicity.
- IG is a solid baseline, but IGR/NWIG may offer slight improvements.
- Pruning is less critical, but still helps reduce complexity.

For Adult-like Datasets (large, mixed features):

- Always prune (limit depth to 4-5) to avoid overfitting and excessive complexity.
- Prefer NWIG or IGR for best accuracy and compact trees.
- IG is acceptable but may yield larger trees and slightly lower accuracy.

UNEXPECTED PATTERNS

- For Iris, accuracy is nearly identical for IGR and NWIG at all depths.
- For Adult, node count for IG at depth=0 is extremely high, but accuracy does not improve.
- Standard deviation is surprisingly low for Adult, indicating stable results.

TRADE-OFFS OBSERVED

Accuracy vs Complexity:

- Depth 2: Simple trees, lower accuracy (especially for IG).
- Depth 3-5: Best balance for all criteria.
- No limit: Marginal accuracy gains, huge complexity cost.

Criterion Selection:

- IG: Simplicity, but can be outperformed by IGR/NWIG.
- IGR: Best for varied attribute cardinalities, robust to overfitting.
- NWIG: Consistent, compact trees, strong performance at shallow depths.