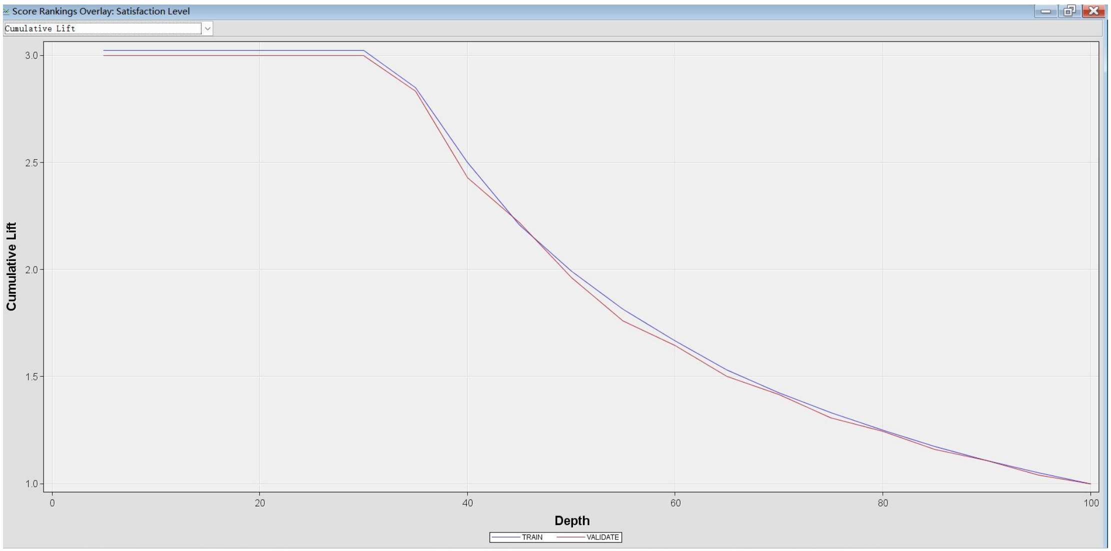
Results and Analysis:

The following results analysis and learning outcomes are supplements to AA1

Decision tree



The close overlap of the two curves indicates that the model's predictions are very similar to the actual situation. This indicates that the model has good predictive power and is able to identify the target variable accurately

The proximity of the proportion of "unsatisfied" customers in the training and validation sets suggests that the model has some generalisation ability at these nodes. For Index 1 and Index 2, the inconsistency between the training and validation sets may require further investigation.

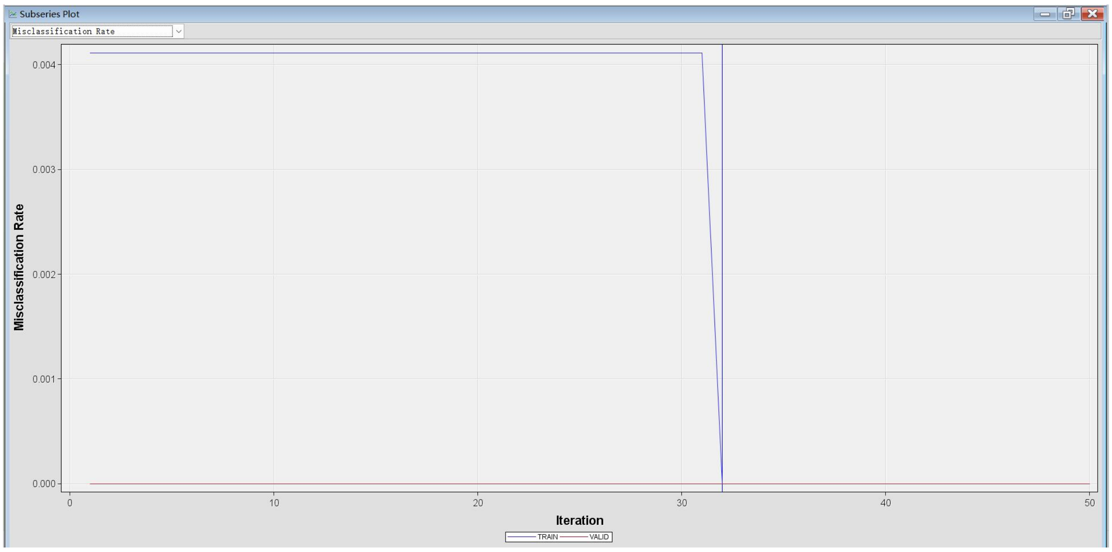
Assessment Score Rankings

Data Role=TRAIN Target Variable=Satisfaction_Level Target Label=Satisfaction Level

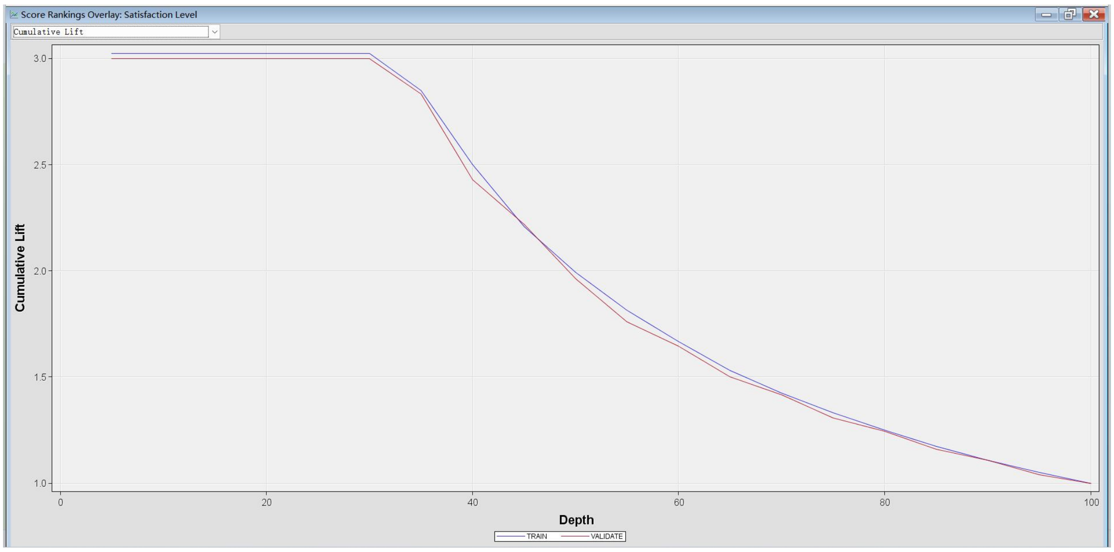
Depth	Gain	Lift	Cumulative Lift	% Response	Cumulative % Response	Number of Observations	Mean Posterior Probability
5	202.469	3.02469	3.02469	100.000	100.000	13	1.00000
10	202.469	3.02469	3.02469	100.000	100.000	12	1.00000
15	202.469	3.02469	3.02469	100.000	100.000	12	1.00000
20	202.469	3.02469	3.02469	100.000	100.000	12	1.00000
25	202.469	3.02469	3.02469	100.000	100.000	13	1.00000
30	202.469	3.02469	3.02469	100.000	100.000	12	1.00000
35	184.884	1.76440	2.84884	58.333	94.186	12	0.58333
40	150.000	0.00000	2.50000	0.000	82.653	12	0.00000
45	120.721	0.00000	2.20721	0.000	72.973	13	0.00000
50	99.187	0.00000	1.99187	0.000	65.854	12	0.00000
55	81.481	0.00000	1.81481	0.000	60.000	12	0.00000
60	66.667	0.00000	1.66667	0.000	55.102	12	0.00000
65	53.125	0.00000	1.53125	0.000	50.625	13	0.00000
70	42.442	0.00000	1.42442	0.000	47.093	12	0.00000
75	33.152	0.00000	1.33152	0.000	44.022	12	0.00000
80	25.000	0.00000	1.25000	0.000	41.327	12	0.00000
85	17.225	0.00000	1.17225	0.000	38.756	13	0.00000
90	10.860	0.00000	1.10860	0.000	36.652	12	0.00000
95	5.150	0.00000	1.05150	0.000	34.764	12	0.00000
100	0.000	0.00000	1.00000	0.000	33.061	12	0.00000

At depths of 5% to 30% , the model achieves 100% cumulative lift and response rates, with all positive responses predicted at these depths. At depth 35%, the model's predictive performance begins to decline and the model begins to encounter unpredictable situations. Starting at 40% depth, predictive power is lost.

Gradient boosting



In most iterations, the misclassification rates for both the training and validation sets are very low, close to 0, indicating that the model has high classification accuracy on both datasets. There is a significant peak near iteration 30, after which the misclassification rate quickly returns to near 0. The reason for the sharp rise in misclassification rate after iteration 30 needs to be examined.



At the initial depth, the cumulative boost is higher than 1, which indicates that the model is very effective at predicting with high probability. In this interval, the model is significantly better at predicting positive classes than random predictions. As the depth increases, the cumulative lift value gradually decreases and tends to 1, which means that the model's ability for predicting positive classes with low probability is weakening and approaching the random level.

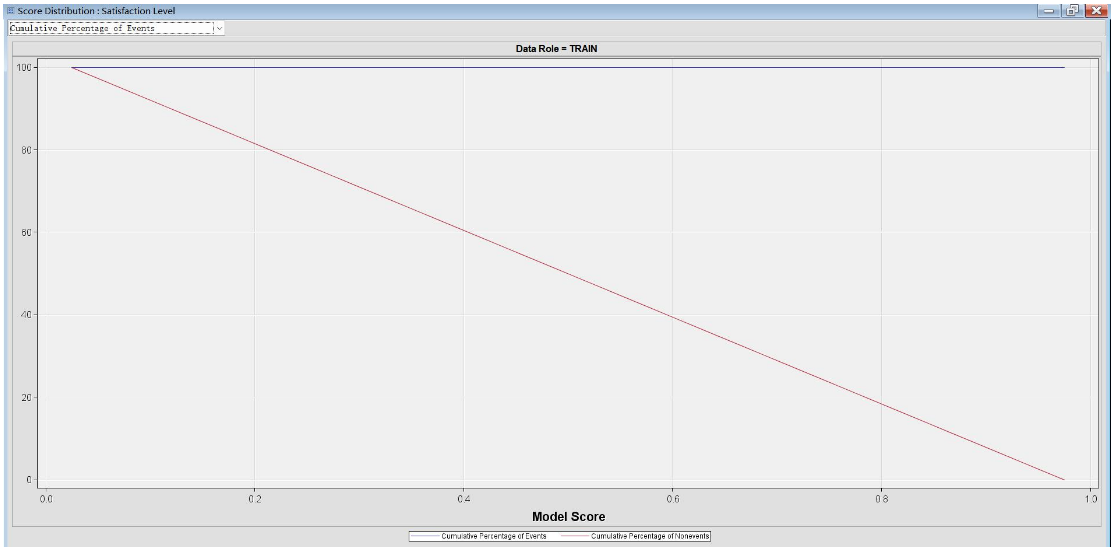
Assessment Score Rankings

Data Role=TRAIN Target Variable=Satisfaction_Level Target Label=Satisfaction Level

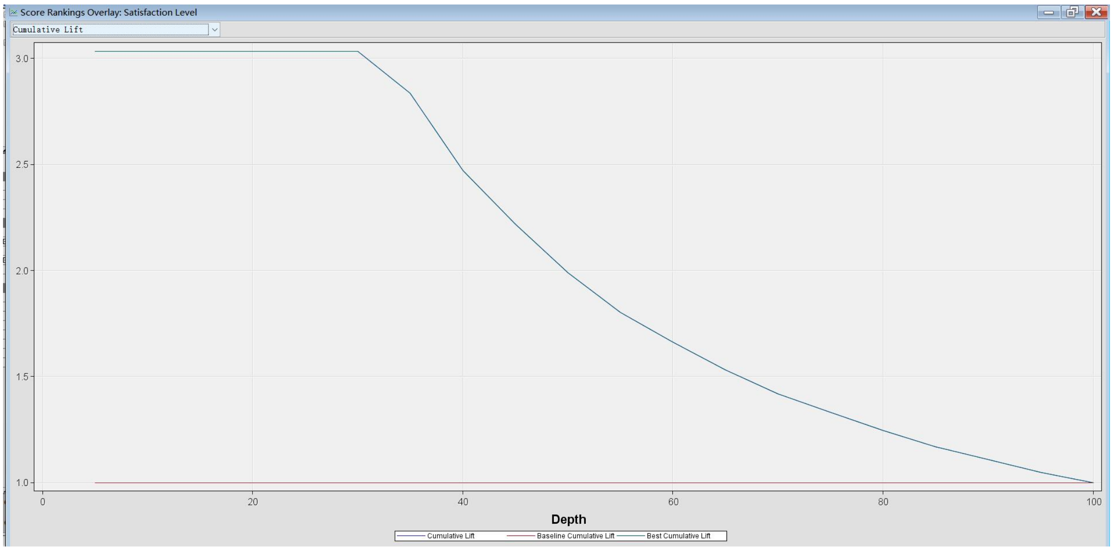
Depth	Gain	Lift	Cumulative Lift	% Response	Cumulative % Response	Number of Observations	Mean Posterior Probability
5	202.469	3.02469	3.02469	100.000	100.000	13	0.40296
10	202.469	3.02469	3.02469	100.000	100.000	12	0.40296
15	202.469	3.02469	3.02469	100.000	100.000	12	0.40296
20	202.469	3.02469	3.02469	100.000	100.000	12	0.40296
25	202.469	3.02469	3.02469	100.000	100.000	13	0.40296
30	202.469	3.02469	3.02469	100.000	100.000	12	0.40208
35	184.884	1.76440	2.84884	58.333	94.186	12	0.35952
40	150.000	0.00000	2.50000	0.000	82.653	12	0.30082
45	120.721	0.00000	2.20721	0.000	72.973	13	0.30082
50	99.187	0.00000	1.99187	0.000	65.854	12	0.30082
55	81.481	0.00000	1.81481	0.000	60.000	12	0.30063
60	66.667	0.00000	1.66667	0.000	55.102	12	0.29969
65	53.125	0.00000	1.53125	0.000	50.625	13	0.29951
70	42.442	0.00000	1.42442	0.000	47.093	12	0.29852
75	33.152	0.00000	1.33152	0.000	44.022	12	0.29852
80	25.000	0.00000	1.25000	0.000	41.327	12	0.29852
85	17.225	0.00000	1.17225	0.000	38.756	13	0.29852
90	10.860	0.00000	1.10860	0.000	36.652	12	0.29852
95	5.150	0.00000	1.05150	0.000	34.764	12	0.29852
100	0.000	0.00000	1.00000	0.000	33.061	12	0.28974

The model was successful in predicting a 100% positive response within 5% to 30% of the depth. However, the performance of the model decreased significantly with increasing depth.

Random forest



This figure shows that as the model score increases, the percentage of cumulative events decreases linearly, while the percentage of cumulative non-events increases linearly.
This linear relationship indicates that the model does not show significant predictive power for distinguishing between events and non-events, i.e., the model's score is not particularly effective at distinguishing between satisfied and dissatisfied customers.



The model showed good results in identifying potential "satisfied" customers, with the model's ability to identify satisfied customers decreasing with increasing depth.

Assessment Score Rankings

Data Role=TRAIN Target Variable=Satisfaction_Level Target Label=Satisfaction Level

Depth	Gain	Lift	Cumulative Lift	% Response	Cumulative % Response	Number of Observations	Mean Posterior Probability
5	203.509	3.03509	3.03509	100.000	100.000	9	1.00000
10	203.509	3.03509	3.03509	100.000	100.000	9	1.00000
15	203.509	3.03509	3.03509	100.000	100.000	8	1.00000
20	203.509	3.03509	3.03509	100.000	100.000	9	1.00000
25	203.509	3.03509	3.03509	100.000	100.000	9	1.00000
30	203.509	3.03509	3.03509	100.000	100.000	8	1.00000
35	183.607	1.68616	2.83607	55.556	93.443	9	0.55556
40	147.143	0.00000	2.47143	0.000	81.429	9	0.00000
45	121.795	0.00000	2.21795	0.000	73.077	8	0.00000
50	98.851	0.00000	1.98851	0.000	65.517	9	0.00000
55	80.208	0.00000	1.80208	0.000	59.375	9	0.00000
60	66.346	0.00000	1.66346	0.000	54.808	8	0.00000
65	53.097	0.00000	1.53097	0.000	50.442	9	0.00000
70	41.803	0.00000	1.41803	0.000	46.721	9	0.00000
75	33.077	0.00000	1.33077	0.000	43.846	8	0.00000
80	24.460	0.00000	1.24460	0.000	41.007	9	0.00000
85	16.892	0.00000	1.16892	0.000	38.514	9	0.00000
90	10.897	0.00000	1.10897	0.000	36.538	8	0.00000
95	4.848	0.00000	1.04848	0.000	34.545	9	0.00000
100	0.000	0.00000	1.00000	0.000	32.948	8	0.00000

Within depths of 5% to 30%, the gain, boost, and cumulative boost values remain consistent and the cumulative response rate is 100%. After 30% of the depth, the cumulative boost value starts to decrease and reaches 35% of the depth, the posterior probability decreases to 0.55556, and from 40% of the depth, the boost value decreases significantly and the response rate decreases to 0%, which indicates that the model is not able to effectively differentiate which customers are satisfied.

learning outcomes:

Model Understanding: Learned how to interpret decision tree outputs, including key metrics such as Gain, Lift, and Cumulative Lift, and how they reflect the model's ability to predict positive events at different depths.

Application of Performance Metrics: Understood the importance of metrics such as response rate and posterior probability in real-world business decisions, and how to make strategic adjustments based on these metrics.

Model Diagnostics: Learned how to diagnose model strengths and weaknesses by evaluating model performance at different depths, and how this information can guide model optimisation.

Importance of dataset size: Recognised the impact of dataset size on model performance and stability, and how to be aware of the risk of overfitting when data volumes are small.

Challenge faced:

Data limitations: The small size of the dataset limits the generalisation ability and stability of the model, making model performance assessment potentially less accurate.

Limitations of model performance: Model performance decreases as depth is increased, requiring a deeper understanding of the causes of performance degradation and exploration of ways to improve the model.

Risk of overfitting: For small datasets, the model may be too complex, capturing noise in the data rather than underlying patterns.
Feature engineering and model selection: There may be a need to explore more complex feature engineering or try different models to improve predictive performance, which requires additional statistical knowledge and machine learning skills.

How to overcome difficulties:

Data enrichment: Adding more data to train the model, using techniques such as data augmentation or external data sources to improve the robustness of the model.
Model Optimisation: Adjust model parameters to improve model performance.
Cross-validation: Perform cross-validation to evaluate the generalisation ability of the model.
Feature Selection: Perform feature selection and engineering to remove irrelevant features and introduce new information-rich features.