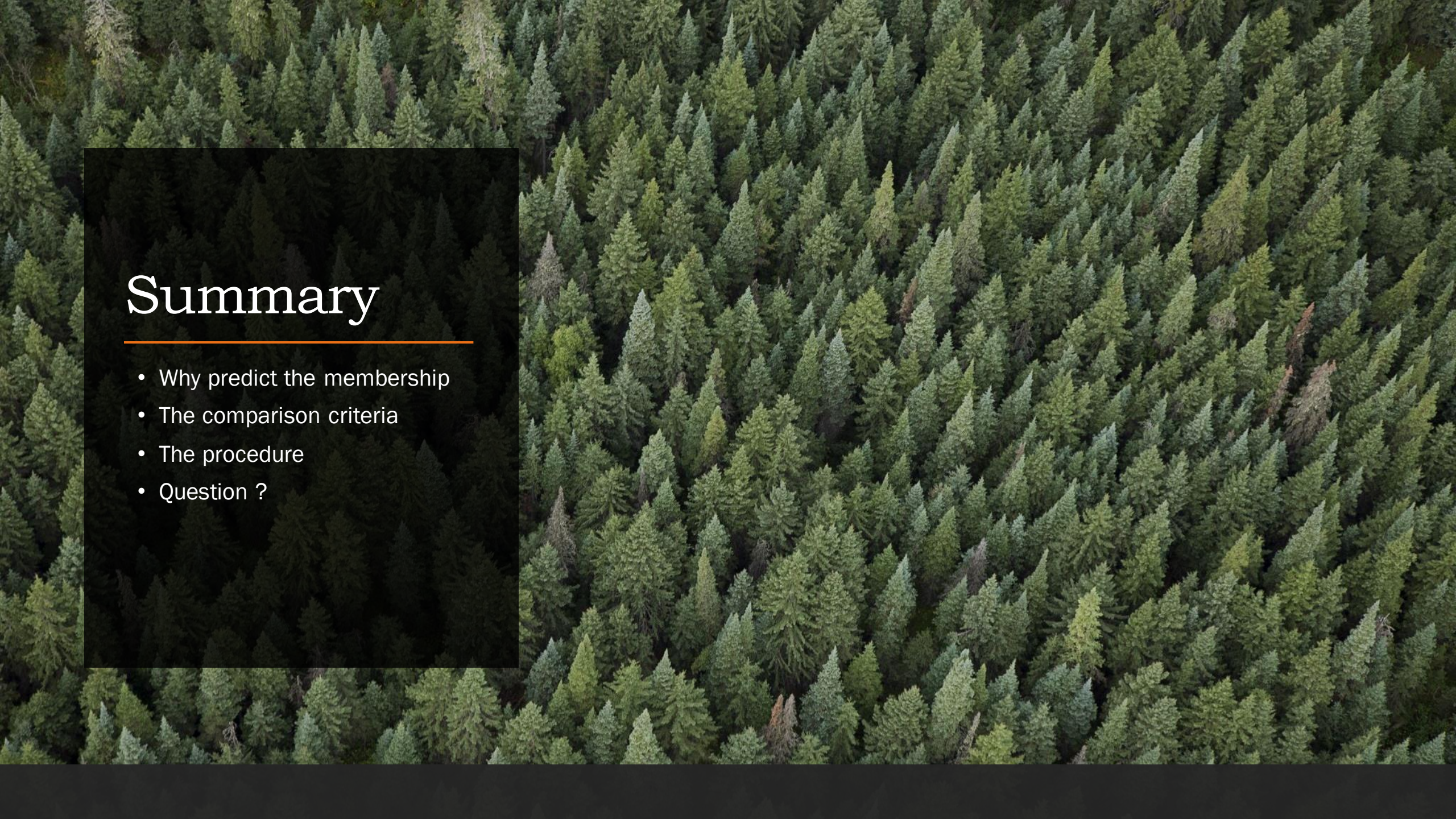
Prediction with Random Forest

MARTIN GUILLEMOT
STANEY SANON



Summary

- Why predict the membership
- The comparison criteria
- The procedure
- Question ?

Why predict the membership

Personalized Experience: Customize services to align with customer preferences.

Strategic Marketing: Optimize targeted marketing efforts based on predictive insights.

Behavioral Insights: Understand and anticipate customer loyalty patterns.

Membership Conversion: Identify non-members exhibiting member-like behavior to encourage membership enrollment.

The comparison criteria

Model Accuracy: Highlight the model's strong accuracy in predicting membership accurately. (Higher is better)

Classification Report:

- Precision: How correctly the model identifies members. (Higher is better)
- Recall: The model's ability to find all actual members. (Higher is better)
- F1-Score: A balance of precision and recall, indicating overall accuracy. (Higher is better)

Computational Feasibility:

- Resource Efficiency: The model's memory usage during operation. (Lower is better)
- Execution Time: How fast the model processes data and delivers results. (Lower is better)

The procedure

1

Initial Training with All Data:

- Training the model using the entire dataset, incorporating all features.
- Helps understanding the overall data structure and feature relevance.

2

Feature Importance Analysis:

- Identify the most important features.
- Determining which features significantly impact the model's predictions.

3

Refined Training with Key Features:

- Re-train the model, this time using only the identified important features.
- The model becomes more efficient and potentially more accurate.

4

Eliminating Redundant Data:

- Remove the 'Duration' feature, retaining only 'EnterTime' and 'LeaveTime'.
- This decision is based on avoiding redundancy.

Model Iteration

	All Features	Only Important Features	Delete Duration
Accuracy	0.82	0.86	0.87
Memory	17G	8G	7G
Speed	4 sec	4s	3s
Len(X)	29	5	4

Conclusion

Good Accuracy: Our model's accuracy is 0.87, which is very good for predicting membership.

Fast on My Computer: The model runs quickly on my computer, making it practical for everyday use.

Important Finding: 'EnterTime' and 'LeaveTime' are important features. If members have reserved times, the model is invalid.


Questions ?