

A woman with long blonde hair, seen from behind, stands in a modern interior looking out a large glass window. She is wearing a black, backless dress. Outside the window is a lush tropical garden with green lawns, various plants, and a thatched roof structure. The scene is bright and sunny.

HOTEL RESERVATIONS

Scarlett
Valentin

08/01/24

TABLE OF CONTENTS

1

**BUSINESS
PROBLEM**

2

**DATA
UNDERSTANDING**

3

**DATA
PREPARATION**

4

MODELING

5

EVALUATION

6

NEXT STEPS

TABLE OF CONTENTS

1

**BUSINESS
PROBLEM**

2

**DATA
UNDERSTANDING**

3

**DATA
PREPARATION**

4

MODELING

5

EVALUATION

6

NEXT STEPS



BUSINESS PROBLEM

Goal: Combat hotel cancellations and no-shows.

How:

- Create a model to predict the number of hotel booking cancellations.
- Evaluate the model to assess the efficacy in using it to predict cancellations.
- Determine the 4 most important factors that contribute to a cancellation.

TABLE OF CONTENTS

1

**BUSINESS
PROBLEM**

2

**DATA
UNDERSTANDING**

3

**DATA
PREPARATION**

4

MODELING

5

EVALUATION

6

NEXT STEPS

DATA UNDERSTANDING

Hotel Reservations Dataset:

- 36,275 entries of unique bookings ranging from 2017 to 2018
- Includes 17 attributes of customer reservation details
- Target column: Booking Status

TABLE OF CONTENTS

1

**BUSINESS
PROBLEM**

2

**DATA
UNDERSTANDING**

3

**DATA
PREPARATION**

4

MODELING

5

EVALUATION

6

NEXT STEPS

DATA PREPARATION

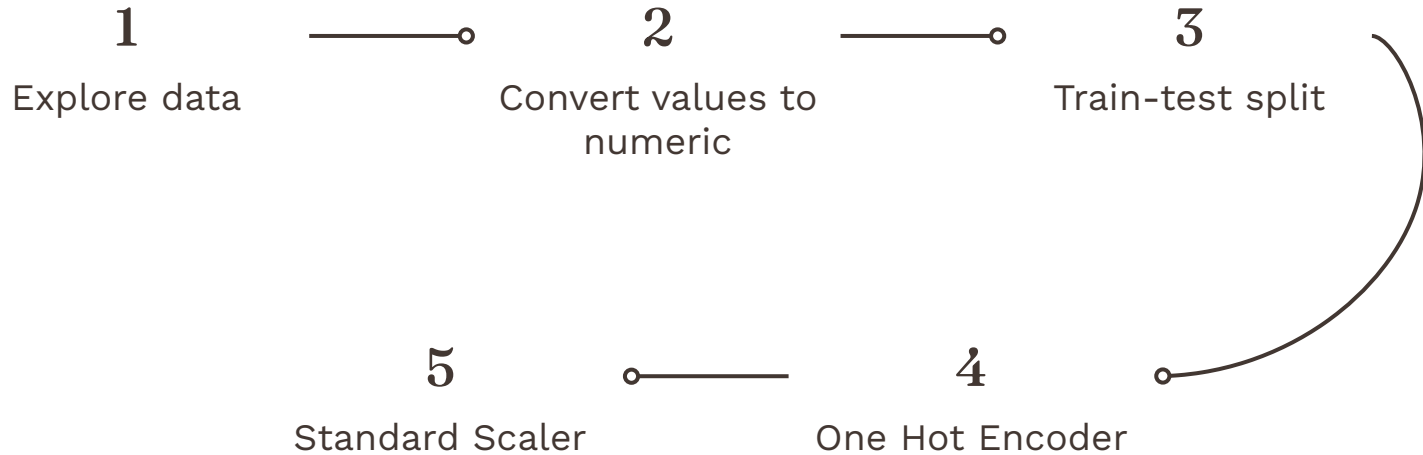


TABLE OF CONTENTS

1

**BUSINESS
PROBLEM**

2

**DATA
UNDERSTANDING**

3

**DATA
PREPARATION**

4

MODELING

5

EVALUATION

6

NEXT STEPS

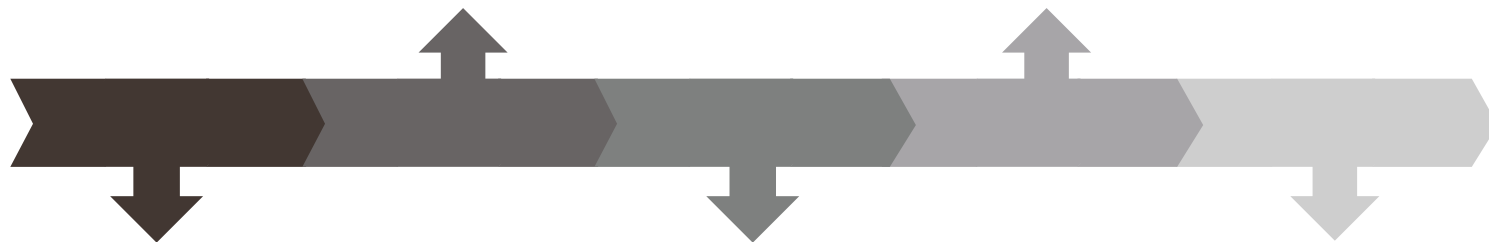
MODELING

Significant Features

Removed features
with $p\text{-value} > 0.05$

Final Model

`max_depth = 7`



Logistic Regression

Used for binary
outcomes

Decision Tree Classifier

Hyperparameter Tuning & Pruning:

- Maximum depth
- Minimum samples leaf
- Minimum leaf sample size
- Maximum features

Feature Importance

1. Lead time
2. Online booking
3. Average price per room
4. Number of special requests

TABLE OF CONTENTS

1

**BUSINESS
PROBLEM**

2

**DATA
UNDERSTANDING**

3

**DATA
PREPARATION**

4

MODELING

5

EVALUATION

6

NEXT STEPS

EVALUATION

85%

ACCURACY

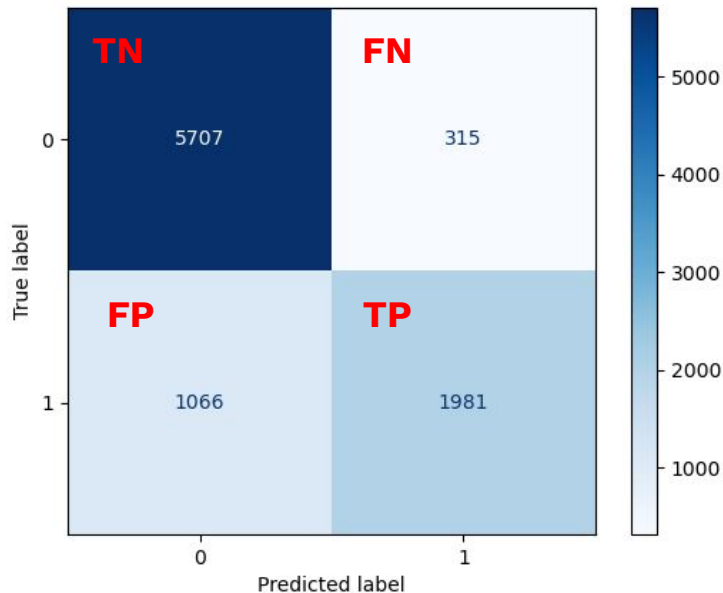
86%

PRECISION

0.8

AUC

CONFUSION MATRIX



315

retained
bookings that
were predicted
to cancel

TABLE OF CONTENTS

1

**BUSINESS
PROBLEM**

2

**DATA
UNDERSTANDING**

3

**DATA
PREPARATION**

4

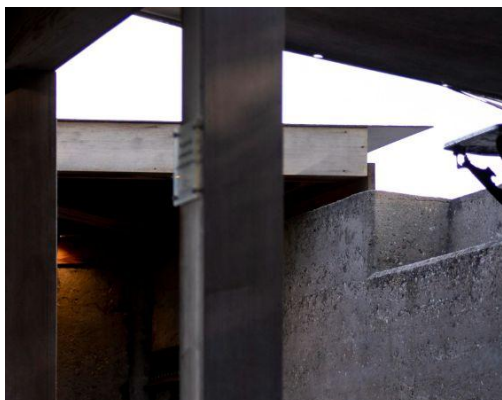
MODELING

5

EVALUATION

6

NEXT STEPS



NEXT STEPS

1. Experiment with different classifier models to **increase the precision score** of the predictions.
2. **Gather recent data** from our guest reservations to continue tuning and testing the model.
3. Further **investigate the top 4 factors** of hotel cancelations.



THANK YOU!

Do you have any questions?

Contact Information:

Email: scarlettmvalentin@gmail.com

GitHub: [@scarlettvalentin](https://github.com/scarlettvalentin)

LinkedIn: linkedin.com/in/scarlett-valentin/

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon** and infographics & images by **Freepik**