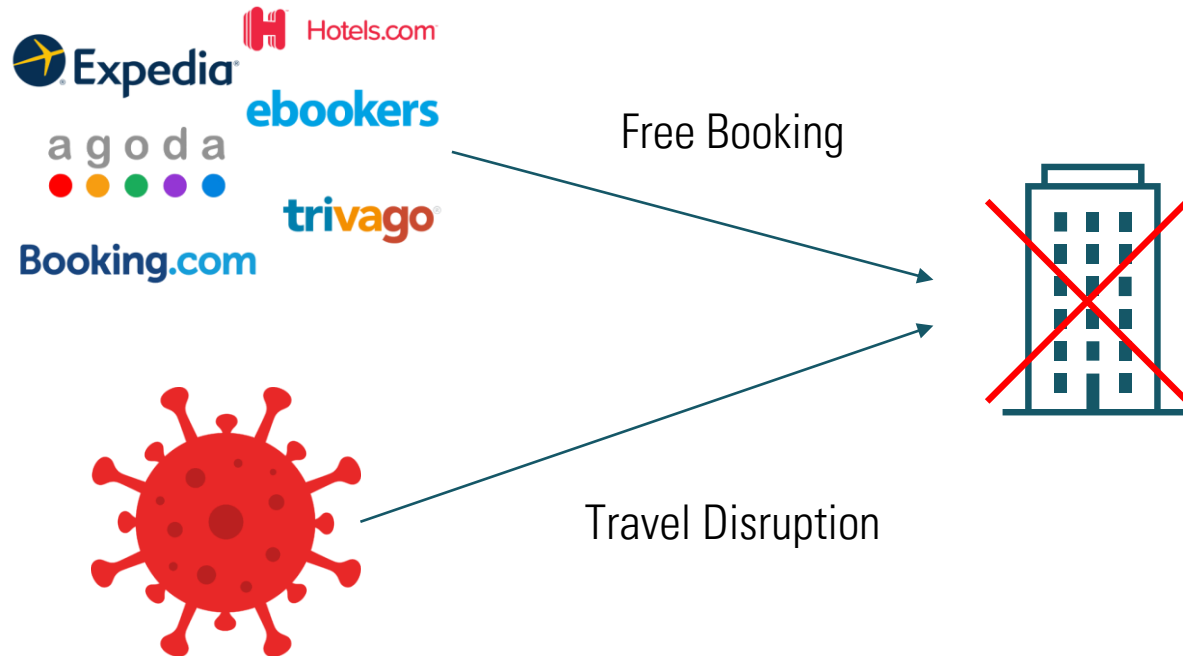
A photograph of a modern glass skyscraper with a grid-like facade, partially obscured by a white diagonal shape that serves as a background for the text.

# HOTEL BOOKING CANCELLATION PREDICTION APP

“Plan for what it is difficult while it is easy, do what is great while it is small.”

— Sun Tzu, The Art of War

# MOTIVATION



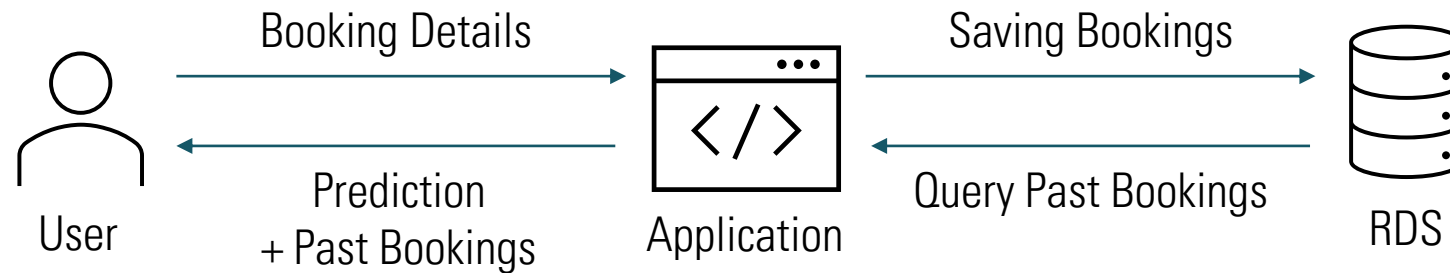
- 35.9% average booking cancellation rate in 2020 and 31.9% average booking cancellation rate in 2021 hurt hotels' revenue by a lot, lowering their revenue and profits.
- By carrying out booking cancellation predictions, the hotels can refine their pricing strategy and resources allocation.

# LINK TO APP

<http://msia423-361876428.us-east-1.elb.amazonaws.com/>

# DATA & RDS

- The data is extracted from Hotel Booking Demand Datasets written by Nuno Antonio, Ana de Almeida and Luis Nunes. The dataset comprises of bookings for a resort hotel and a city hotel, including the bookings that effectively arrived and were cancelled.



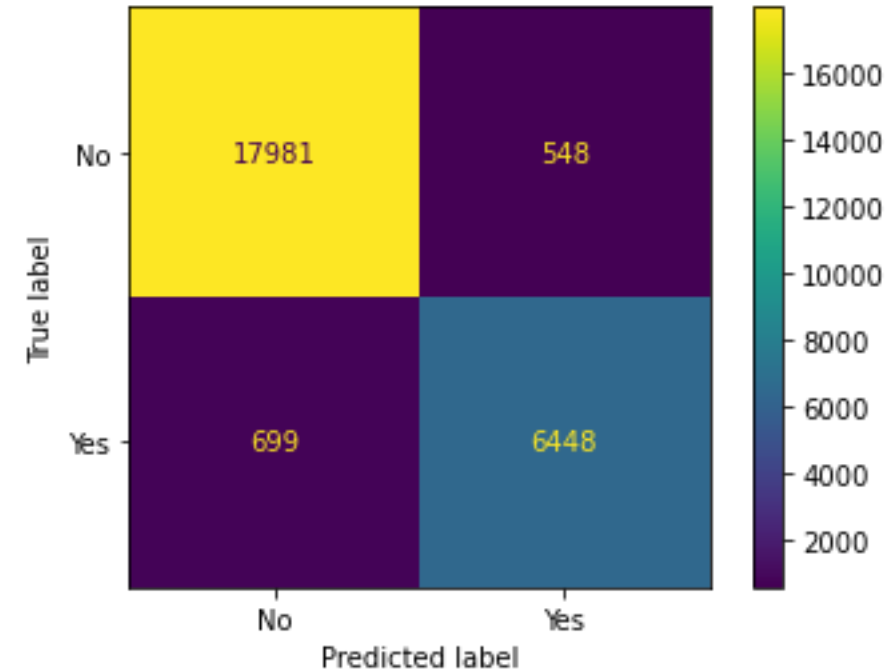
Source: Nuno Antonio, Ana de Almeida, Luis Nunes, Hotel booking demand datasets, Data in Brief, Volume 22, 2019, Pages 41-49, ISSN 2352-3409, <https://doi.org/10.1016/j.dib.2018.11.126>. (<https://www.sciencedirect.com/science/article/pii/S2352340918315191>)



# MODEL & SUCCESS CRITERIA

Label	Precision	Recall	F1-score	Support
0	0.96	0.97	0.97	18529
1	0.92	0.90	0.91	7147
Accuracy			0.95	25676
Macro Avg	0.94	0.94	<b>0.94</b>	25676
Weighed Avg	0.95	0.95	<b>0.95</b>	25676

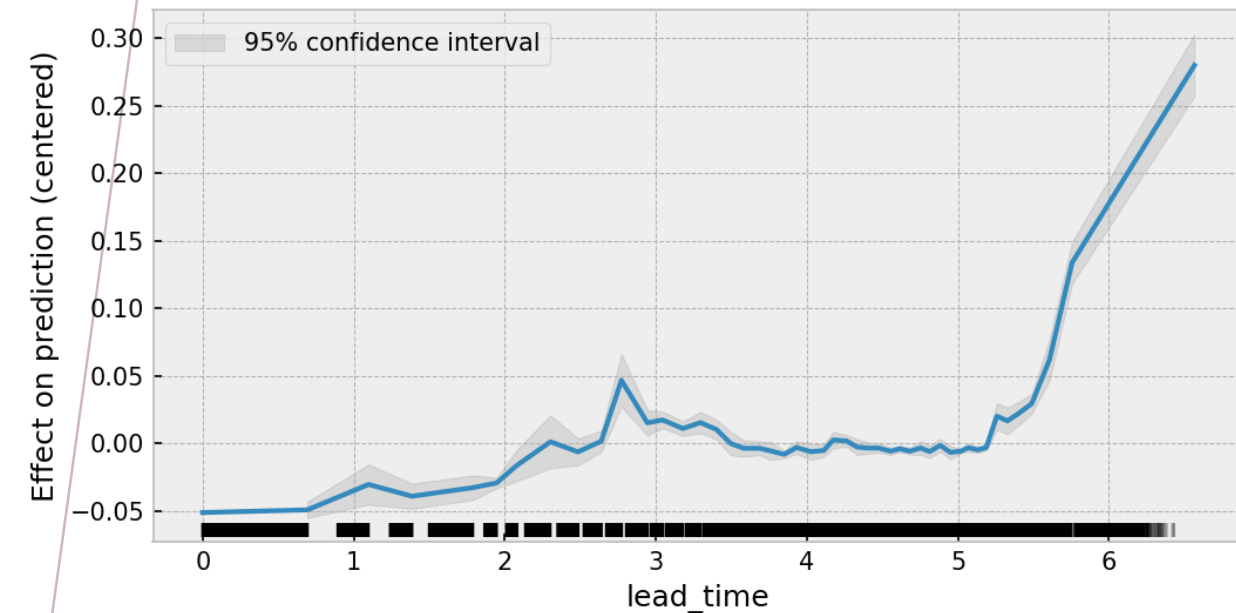
The **decision tree model** achieved **macro average F1-score of 0.94** and **weighted average F1-score of 0.95**, which fulfils the **predefined success criteria of F1-score greater than 0.9**. The model is deployed **online**, and can make prediction on the fly based on user's input.



Metric Term	Definition
Precision	: Percentage of correct prediction
Recall	: Fraction of positive that were correctly identified
F1 Score	: Weightage average between Precision and Recall
Support	: Number of occurrences

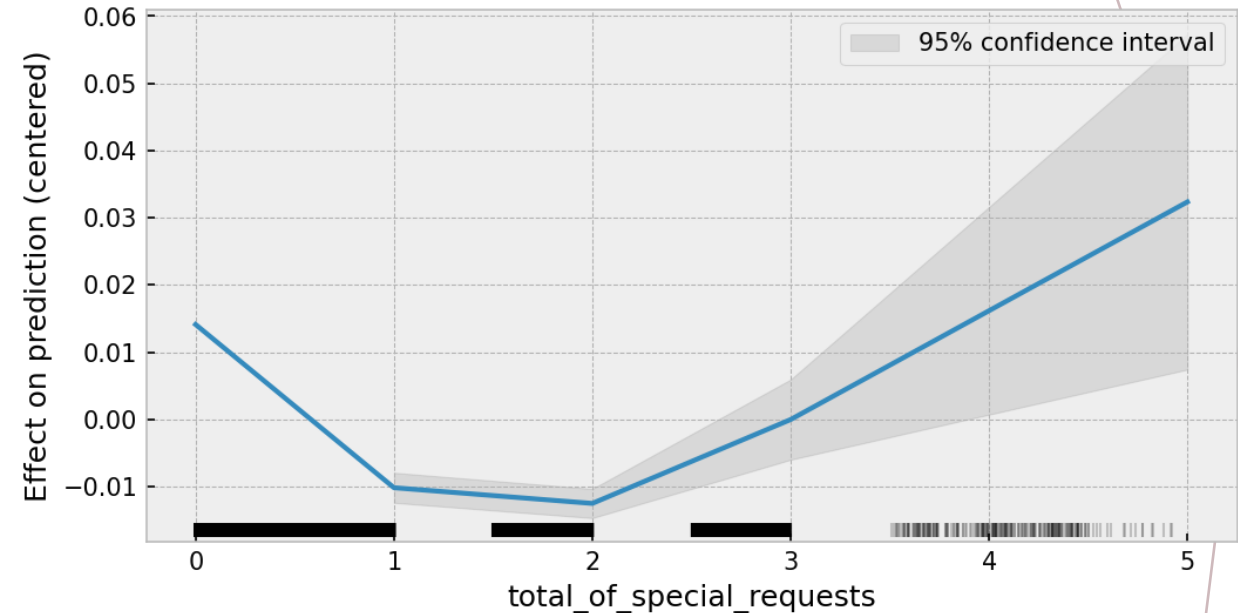
# INSIGHTS

1D ALE Plot - Continuous



Higher lead time will lead to higher predicted probability of cancellation.

1D ALE Plot - Continuous



Bookings without special request or with a high number of special requests will lead to higher predicted probability of cancellation.



# THANK YOU

Jia Hao Lo

[jialo2022@u.northwestern.edu](mailto:jialo2022@u.northwestern.edu)

<https://www.linkedin.com/in/jiahaolo/>