Discussion 3. Project Topic

No unread replies.55 replies.

Please read the [Final Project Specification](https://uispringfield.instructure.com/courses/12540/files/1864593?wrap=1)[Download Final Project Specification](https://uispringfield.instructure.com/courses/12540/files/1864593/download?download_frd=1)carefully and give a one paragraph description of the topic that you choose for the final project.  **You can team up with another classmate to do the project.** In particular, you must answer the following questions in your description

1- What is the problem you are trying to address?

2- What  dataset(s) will you be using? how many variables do you have in the dataset ?  If you are using a public dataset, please provide a link to the dataset.

3- What type of model (classification, regression, clustering) will you be using to solve the problem  in question 1 ( Note: the type of the model depends on the type of your outcome variable) . If you are using a classification  or regression, explain what is the outcome variable that you are trying to predict/classify?

4- I**f you want to work in a team, please specify the name of the team members ( at most two). Please make only one submission per team and both the team members will receive  grade for the post.**

**Your post to this discussion forum is due by Sunday March 31st,  11:59 pm.**

[Learning from Imbalanced Insurance Data (kaggle.com)](https://www.kaggle.com/datasets/arashnic/imbalanced-data-practice)

Answer for the discussion board.

1- What is the problem you are trying to address?

The objective is to predict whether a customer will cancel or honor a hotel reservation based on various features.

2- What dataset(s) will you be using? how many variables do you have in the dataset ? If you are using a public dataset, please provide a link to the dataset.

Link - [Hotel Reservations Dataset (kaggle.com)](https://www.kaggle.com/datasets/ahsan81/hotel-reservations-classification-dataset/data)

There are 19 variables in total.

3- What type of model (classification, regression, clustering) will you be using to solve the problem in question 1 ( Note: the type of the model depends on the type of your outcome variable) . If you are using a classification or regression, explain what is the outcome variable that you are trying to predict/classify?

I will be using classification models and the outcome variable ‘booking\_status’. The DV will have 2 classes: Not Cancelled, Cancelled.

4- If you want to work in a team, please specify the name of the team members ( at most two). Please make only one submission per team and both the team members will receive grade for the post.