Academic Year 2024-2025

Home
Announcements
Modules
Discussions
Quizzes
Assignments
Grades
Collaborations
People

Coursework 2 - Model Building and Validation

30 Points Possible

Due: Mon, 7 Apr 2025 15:30

Attempt 1

In Progress
NEXT UP: Submit assignment

Add comment

Unlimited Attempts Allowed

Details

F78DS Coursework 2

Due date: Monday 7th April 2025 - 3:30 pm (GMT)

Previous

Submit assignment

Next

Coursework 1 covered the process of conducting *descriptive analytics*, whereas the objective of this coursework (Coursework 2) is to conduct *predictive analytics*, through machine learning on a dataset using Python in the Jupyter Notebook environment. This coursework will test your ability to:

- Read and describe the data using basic statistics and exposure to some natural language processing** terms,
- Conduct some feature selection,
- Split the dataset into training and testing (for this coursework, it may sometimes be referred to as validation dataset),
  - And an opportunity (not required for this Coursework) to create multiple training and testing datasets, in order to conduct cross validation.
- Conduct multi-class classification using Naive Bayes (NB)**,
- Communicate the output of your analysis using the Quadratic Weighted Kappa (QWK) measure**,
- Experience independent model evaluation through submission for an in-class Kaggle competition.

** Not taught in this course, you are to explore and elaborate these in your report submission. This will be a mild introduction to life-long learning to learn by yourself.

The links (from lectures) to further reading for the Cohen Kappa Score:

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1), 159–174. https://doi.org/10.2307/2529310

If you like a shorter explanation: https://datatab.net/tutorial/cohens-kappa

Choose a submission type

Upload

More

rag a file here, or

Choose a file to upload

File permitted: PDF, IPYNB

or

Canvas File

I agree to the tool's End-User License Agreement
This assignment submission is my own, original work