Assignment: S2W8A2 - Stage 2 Deliverables

40 Points Possible

8/18/2024

Attempt 1

In Progress

NEXT UP: Submit Assignment

Add Comment

Unlimited Attempts Allowed

Details

Stage 2 Deliverables

Stage 2 focuses on preparing the data for analysis and modeling, ensuring its integrity and compatibility with the chosen analysis tools. Additionally, clustering analysis is performed to identify distinct customer segments based on key features.

Stage 2 Deliverables: Data Preparation and Modeling

2-1: Data Preparation Deliverables:

- A preprocessed dataset addressing missing data points and encoding categorical variables.

- The dataset is split into distinct training and testing sets to facilitate model validation.

- Appropriate scaling techniques are applied to normalise the data, enhancing model performance.

2-2: Clustering Analysis Deliverables:

- The optimal number of clusters is identified using the elbow method.

- A K-Means clustering model is trained on the dataset.

- Resulting clusters are visualised and labelled for interpretation and actionable insights.

By completing these deliverables, the project team ensures the data is appropriately prepared and analysed, setting the stage for effective modeling and predicting customer churn.

Submission Instructions

1. Repository Structure:

   - Create a new repository on GitHub/GitLab (or any other distributed version control) for the project submission.

   - Structure the repository with separate folders for each deliverable to maintain organisation and clarity.

2. Data Preparation Deliverables:

   - Create a folder named "Data_Preparation" to store all data preparation deliverables.

   - Within the "Data_Preparation" folder, include:

     - Preprocessed Dataset: Upload the preprocessed dataset file in a structured format.

- Training and Testing Sets: Provide separate files or folders for training and testing sets, including documentation on size and composition.

- Scaling Techniques Documentation: Include a document (e.g., PDF or Word) explaining the applied scaling techniques with relevant code snippets.

3. Clustering Analysis Deliverables:

- Create a folder named "Clustering_Analysis" to store all clustering analysis deliverables.

- Within the "Clustering_Analysis" folder, include:

- Optimal Number of Clusters: Upload documentation or analysis results indicating the optimal number of clusters and any supporting visualisations.

- Trained K-Means Model: Include the trained K-Means clustering model file and relevant code or scripts used for training.

- Visualisation and Labeling of Clusters: Provide visualisations of resulting clusters with labels for interpretation, explanations, or insights derived from the analysis.

4. Video Demonstration:

- Create a video (10 to 15 minutes long) demonstrating the key aspects of the project, including data preparation and clustering analysis. In the video, walk through the preprocessed dataset, training and testing sets, scaling techniques, the optimal number of clusters, the trained K-Means model, and the visualization of clusters. Provide explanations and insights throughout the demonstration, highlighting important findings and decisions made during the analysis process.

Submission Process:

- Submit a link to the GitHub/GitLab (or any other platform) repository as the project submission.

- Ensure that the repository is set to public or accessible to the project instructor for evaluation.

- Upload the video demonstration file.

> Note:
>
> - To submit your Stage 2 Deliverables, click on the ‘Submit Assignment’ button below.
> - Click on 'New Attempt' to make multiple submissions for different submission types.
> - The 'Project Manager' is responsible for submitting one copy of all the above documents. Individual members are NOT to submit any team documents at any time.

View Rubric

Stage 2 Deliverables (WD) Rubric

Criteria	Ratings		Pts
Structured Repository view longer description	**20 pts** **Full Marks** All elements as defined in the Assignment Instructions must be included to get full marks	**0 pts** **No Marks** No submission	/ 20 pts

Stage 2 Deliverables (WD) Rubric

Criteria	Ratings		Pts
Video Demonstration view longer description	**20 pts** **Full Marks** All elements as defined in the Assignment Instructions must be included to get full marks	**0 pts** **No Marks** No Submission	/ 20 pts

Total Points: 0

Keep in mind, this submission will count for everyone in your Project Groups group.

Choose a submission type.

Web URL | Media | Upload

I agree to the tool's **End-User License Agreement (https://acs.turnitin.com/integrations/cpf/eula/acs)**
I have read and understood ACS's Academic Misconduct Policy. I confirm that this submission is my original work and no part of it has been copied from any other source except where due acknowledgement is made. Furthermore, this work has not been previously submitted for this or any other course or unit. In submitting this assessment, I give ACS permission to retain, compare, copy, and archive for the purposes of detecting plagiarism and/or review by external examiners.

(https://acslearn.instructure.com/courses/1412/modules/items/135921)
Assignment
(https://acslearn.instructure.com/courses/1412/modules/items