Customer segmentation is a process to dividing customers into groups which possess common characteristics based on their age, gender, profession as well as interests. By doing so will enable the company will gain insights of customer's buying pattern or preferences, therefore the company will design a marketing strategy to target the most profitable segments.

An automobile company is considering entering new markets using their existing products. Prior entering the market, the analysts from the company conducted intensive market research and they found out that the behaviour of the new market is similar to their existing market.

In their existing market, the sales team has classified all customers into 4 segments (A, B, C, D) based on their customer's gender, age, profession, spending pattern, and etc. Segmented outreach as well as communication for different segment of customers (A,B,C,D) were conducted and this strategy has work well for the company. The company has been generating profits by changing the marking strategy to target the segmented customers. Therefore, the company plans to use the same strategy on the new markets. The sales team have identified 2627 new potential customers.

Hence, your job as data analysts is to develop a deep learning model to predict the right group of the new customers.

The criteria of the project are as follows:

1) Develop a deep learning model using TensorFlow which only comprises of Dense, Dropout, and Batch Normalization layers.
2) The accuracy of the model must be more than 80% with F1 more than 80%.
3) Display the training loss and accuracy on TensorBoard
4) Create modules (classes) for repeated functions to ease your training and testing process

Files to be submitted and uploaded to GitHub and LMS (submission link will be given on the assessment day):

1) Training, deployment scripts and classes (GitHub and LMS)
2) Dataset (.csv file) (GitHub and LMS)
3) Saved model in .h5 format and scalers (if any) in .pkl file format. (GitHub and LMS)
4) Training process plotted using Tensorboard can be snipped and saved as image file format (LMS) and use EarlyStopping callback to prevent overfitting.

5) The architecture of the model should be plotted using plot_model function and saved as .png file format. Include the image in README.md and also upload to LMS. (GitHub and LMS)
6) Performance of the model and the reports can be snipped and saved as image file to be included in the zip folder for LMS submission. (LMS and GitHub)
7) Include your GitHub URL directing to your assessment 2 in a text file then submit to LMS. (LMS)
8) **Don't forget to credit/cite the source of the data on your GitHub page**

   https://www.kaggle.com/datasets/abisheksudarshan/customer-segmentation

*Please zip all the required files into one folder then submit to LMS.

**Please save the dataset and model in 2 different folders to GitHub.

Complete the assessment and submit the files to LMS and GitHub by 5pm. Good Luck!!!

	100%	50%	0%
Task Completion (30%)	Scripts can be executed without any error on trainer's local machine.	-	Scripts fail to be executed on trainer's local machine.
Project requirements (30%)	Able to achieve the objectives of the project using relevant and appropriate approach.	Able to achieve the objectives of the project but using inappropriate approach such as brute forcing the solution.	Fail to achieve the objectives of the project.
Exploratory data analysis (30%)	Demonstrates strong understanding on the objectives of the project and performs relevant approach to process the data. Necessary data processing techniques such as, data loading, data cleaning, features selection and data preprocessing are performed and well justified.	Shows comprehensive understanding of the objectives of the project but uses incorrect or irrelevant approach to process the data. For example, removing NaN data when there is limited amount of samples in the dataset.	Shows limited understanding of the objectives of the project. Absence of data processing section in the code.
Code readability (5%)	Involves the usage of functions or methods for repeated tasks. Codes are easily readable and justified by including comments and description texts.	Minimal usage of functions or methods for repeated tasks. Available comments and descriptions but lack of details.	No usage of functions or methods for repeated tasks. Codes are difficult to read and understand. Missing descriptions and comments.
GitHub repo (4%)	Detailed and clear instructions of the project on README.md. Results such as graphs are also included in README.md as part of the project description.	Project successfully uploaded to GitHub repo but with incomplete README.md. Missing descriptions, instructions, and results.	Fails to upload project to GitHub repo and missing README.md
PEP8 compliance (1%)	Fully complies with PEP 8 Standard	Partially complies with PEP 8 Standard	Fails to comply with PEP 8 Standard
Total (100%)			