

Configuration Manual

MSc Research Project
Programme Name

Kannan Kumaran
Student ID: x20195061

School of Computing
National College of Ireland

Supervisor: Dr.Pramod Pathak & Dr.Paul Stynes

National College of Ireland
Project Submission Sheet
School of Computing



Student Name:	Kannan Kumaran
Student ID:	x20195061
Programme:	Programme Name
Year:	2022
Module:	MSc Research Project
Supervisor:	Dr.Pramod Pathak & Dr.Paul Stynes
Submission Due Date:	31/01/2022
Project Title:	Configuration Manual
Word Count:	677
Page Count:	5

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	30th January 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission , to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project , both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Kannan Kumaran
x20195061

1 Overview

Research on a Machine and Deep Learning Framework to Retain Customers based on their Lifetime Value was implemented utilizing Python language. The system requirements and how to perform this project is provided as instructions in this manual.

2 System requirements

2.1 Hardware requirements

Operating System Type: Windows 10, 64-bit operating system (x64-based processor)
RAM: 8GB

These are the hardware requirements which satisfies the implementation.

2.2 Software requirements

Google Colab is necessary to execute the Python Code. The other tools employed are Overleaf and Microsoft Excel. Browser required is Google Chrome/Firefox and to use Colab Pro for mounting drive, a Gmail account is needed.

3 Implementation

Dataset collection: The online retail sales data set is extracted from the data source:<https://archive.ics.uci.edu/ml/datasets/Online+Retail+II>. There will be a file entitled (online_retail_II.xlsx) that may be downloaded in one click

Storing of data in colab: The dataset should be maintained in any location on Google Drive in order to access it during code execution.

The newly generated colab notebook is shown in Figure 1. To mount the drive using Google Colab, trigger the google drive symbol (third icon) on the left panel of the top three icons.

On clicking drive icon, notebook will prompt you for authentication. A code will be emailed to your Gmail account, which you may copy and paste under area to enter authorization code.

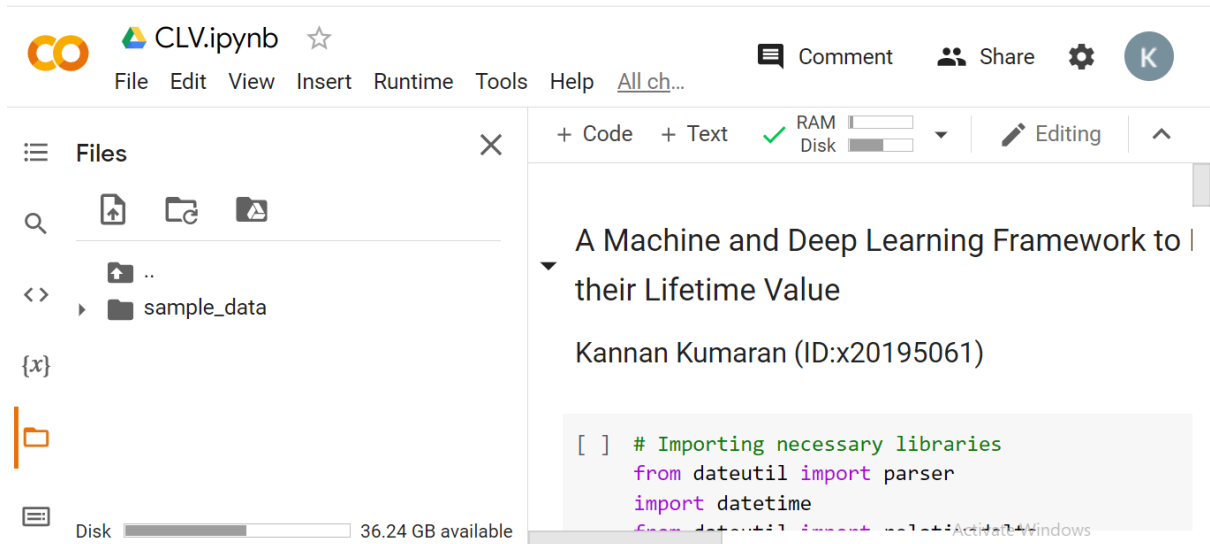


Figure 1: Data Panel of Google Colab

Figure 2 portrays the drive files in the left pane after the google drive was successfully mounted.

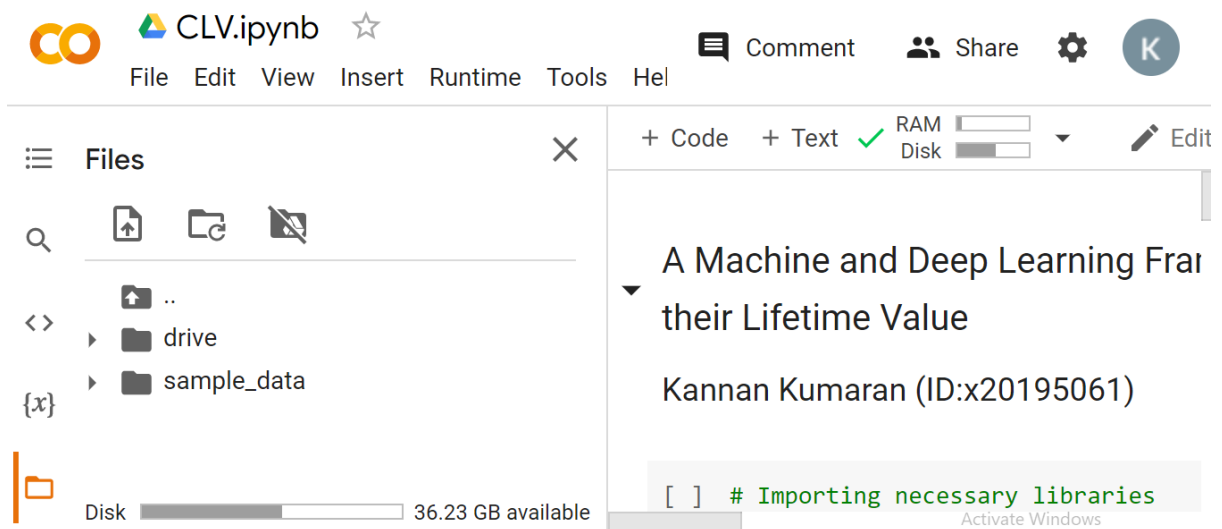


Figure 2: Visibility of mounted drive

To begin, some of the common libraries necessary for developing a Model have been downloaded and installed. These tools and libraries have all been updated to their most recent versions. Figure 3 displays the libraries that must be imported before the data can be loaded.

```
# Importing necessary libraries
from dateutil import parser
import datetime
from dateutil import relativedelta
!pip install lifetimes
!pip install openpyxl
!pip install squarify
!pip install KMeans
from sqlalchemy import create_engine
import datetime as dt
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from lifetimes import BetaGeoFitter
from lifetimes import GammaGammaFitter
from lifetimes.plotting import plot_period_transactions
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import StandardScaler
import squarify # treemap
import warnings
from lifetimes.utils import summary_data_from_transaction_data, calibration_and_holdout_data
```

Figure 3: Importing the required libraries

The dataset directory within the drive in the left panel has to be copied by right clicking and an option appears as copy path. The copied path should then be inserted inside the brackets of read excel() function as illustrated figure 4.

```
[ ] # Reading the excel format data set by using "Year 2010-2011" sheet named.

df= pd.read_excel("drive/My Drive/Online Retail/online_retail_II.xlsx", sheet_name = "Year 2010-2011")
df.head()
```

Figure 4: Setting data drive directory

The cell will now execute when you hit Shift+Enter, and the dataset will be successfully loaded. Also, Run all Cells could be selected for one time execution. The implementation stages can then be carried out in the sequence specified in the Python code.

Required Packages & libraries: The Data science functions and plotting methods can be accessed by importing the corresponding libraries.

These are the libraries that must be imported into the colab notebook in order to create models for customer lifetime value prediction, as illustrated in Figure 5.

The installation of the lifetime package is shown in Figure 6. This software must be loaded before the beta-geometric and gamma-gamma models can be used to forecast future purchases and expenditures.

```

from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import StandardScaler
import squarify # treemap
import warnings
from lifetimes.utils import summary_data_from_transaction_data, calibration_and_holdout_data

from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import r2_score

from sklearn.cluster import KMeans
from yellowbrick.cluster import KElbowVisualizer

from sklearn.metrics import silhouette_score
from scipy.cluster.hierarchy import linkage
from scipy.cluster.hierarchy import dendrogram
from scipy.cluster.hierarchy import cut_tree

```

Figure 5: Necessary libraries for models and evaluation methods

```

!pip install Lifetimes
Requirement already satisfied: Lifetimes in /usr/local/lib/python3.7/dist-packages (0.11.3)
Requirement already satisfied: scipy>=1.0.0 in /usr/local/lib/python3.7/dist-packages (from Lifetimes) (1.4.1)
Requirement already satisfied: pandas>=0.24.0 in /usr/local/lib/python3.7/dist-packages (from Lifetimes) (1.1.5)
Requirement already satisfied: dill>=0.2.6 in /usr/local/lib/python3.7/dist-packages (from Lifetimes) (0.3.4)
Requirement already satisfied: autograd>=1.2.0 in /usr/local/lib/python3.7/dist-packages (from Lifetimes) (1.3)
Requirement already satisfied: numpy>=1.10.0 in /usr/local/lib/python3.7/dist-packages (from Lifetimes) (1.19.5)
Requirement already satisfied: future>=0.15.2 in /usr/local/lib/python3.7/dist-packages (from autograd>=1.2.0->Lifetimes) (0.16.0)
Requirement already satisfied: pytz>=2017.2 in /usr/local/lib/python3.7/dist-packages (from pandas>=0.24.0->Lifetimes) (2018.9)
Requirement already satisfied: python-dateutil>=2.7.3 in /usr/local/lib/python3.7/dist-packages (from pandas>=0.24.0->Lifetimes) (2.8.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-packages (from python-dateutil>=2.7.3->pandas>=0.24.0->Lifetimes) (1.14.0)
Requirement already satisfied: openpyxl in /usr/local/lib/python3.7/dist-packages (2.5.9)
Requirement already satisfied: et-xmlfile in /usr/local/lib/python3.7/dist-packages (from openpyxl) (1.1.0)
Requirement already satisfied: jdcal in /usr/local/lib/python3.7/dist-packages (from openpyxl) (1.4.1)
Requirement already satisfied: squarify in /usr/local/lib/python3.7/dist-packages (0.4.3)
Requirement already satisfied: KMeans in /usr/local/lib/python3.7/dist-packages (1.0.2)

```

Figure 6: Installation of Lifetime Packages

Keras and Tensorflow packages are needed for model building of Deep Neural Networks along with other parameters which is presented in Figure 7

```
#Deep Neural Network
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from keras.callbacks import ModelCheckpoint, EarlyStopping
```

Figure 7: Libraries for DNN

Remaining implementation steps: The provided Python code may be run in a single click by selecting the runtime button from the main menu after importing the dataset into the colab notebook as described before.

After successfully importing the dataset, click Runtime, and then run all.

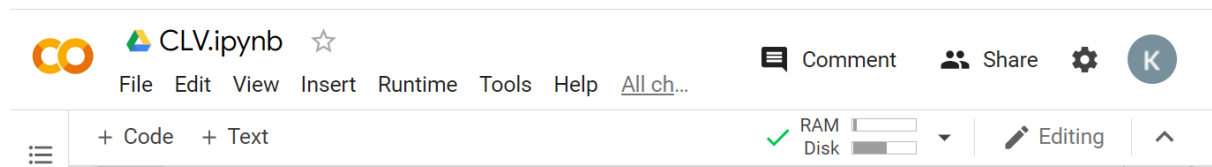


Figure 8: Colab Notebook Main Menu

4 Conclusion

To successfully execute the entire code in Google colab, follow the instructions listed above. The colab has built-in support for 12GB RAM, ensuring that the code runs smoothly and quickly. The Run time environment should be switched to GPU for faster execution and performance.