# CECS 551 Advanced Artificial Intelligence

Fall 2022 - Mid-term

**Assignment**

For this assignment you are encouraged to work in a group. Please identify your team members [here](#_Team_members_in). Each 2 groups will be working on same problem statement; however, it is advised not to communicate your results with another group.

Please use Python as implementation language and [Google Colab](https://colab.research.google.com/) as a platform.

* The problem statements are from [UCI Machine Learning Repository](https://archive.ics.uci.edu/ml/index.php) and it is strongly recommended to explore the corresponding UCI dataset page for understanding the problem statement and the data description.
* Each group can access the **dataset** [here](https://csulb-my.sharepoint.com/:f:/g/personal/rahuldeo_vishwakarma01_student_csulb_edu/Et1Gwi_u1BVKjcmSfXzbj0QBe5dXWun8S4GsAHUxivUqvw?e=TKx9GZ) (requires CSULB SSO login for OneDrive).

**Deliverables**

1. Source code as Google Colab url (submission [example](https://colab.research.google.com/drive/1JkUjhTjR8MyLxwarJjqnh836BICfocTu))
2. Presentation slides for the results (upload in Dropbox)

**Submission date**: October **24**, 2022 (Monday)

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Teams | **Presentation in Class** | Q&A |
| 25-Oct (Tue) | Group 01 to Group 05 | 10 mins | 3 mins |
| 27-Oct (Thru) | Group 06 to Group 10 | 10 mins | 3 mins |

## Group 1 and Group 2

**Problem statement**

1. Perform exploratory data analysis and feature selection
2. Develop a machine learning model (at least one deep learning model) to predict the productivity performance of the employees and compare the performance metrics of the algorithms.

**Group 1**: Please consider the problem statement as **Regression**

**Group 2**: Please consider the problem statement as **Classification**

1. Recommend two approaches for increasing the productivity of the employees. You can make a few assumptions to support the hypothesis.

**Dataset**

<https://archive.ics.uci.edu/ml/datasets/Productivity+Prediction+of+Garment+Employees>

**Hints**

Feature selection: <https://github.com/AutoViML/featurewiz>

You may use Pandas Profiling (pip install pandas-profiling), but you are free to use any other library. <https://pandas-profiling.ydata.ai/docs/master/index.html>

Choose a right estimator (<https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html>) to develop a machine learning model

## Group 3 and Group 4

**Problem statement**

1. Forecast the sales of products based on stores in near future. Initially, you may want to start with few products in selected stores and then scale out the solution.
2. Compare the forecasting accuracy for next 10 weeks and 2 months.
3. Compare the efficiency of ARIMA, SARIMA, and Neural Network models for the prediction.

**Dataset**

<https://archive.ics.uci.edu/ml/datasets/Demand+Forecasting+for+a+store>

## Group 5 and Group 6

**Problem statement**

The task is how to aggregate stock units into groups so that the resulting inventory policies are sufficiently close to the policies that would have been generated if every unit was treated differently.

1. Perform cluster analysis using k-means and DBSCAN and validate using silhouette index
2. Compare the efficiency of model after performing PCA

**Dataset**

<https://archive.ics.uci.edu/ml/datasets/Stock+keeping+units>

## Group 7 and Group 8

**Problem statement**

1. Create a Recency, Frequency, and Monetary model (RFM) based customer segmentation. The main purpose of this task is to help the business better understand its customers and therefore conduct customer-centric marketing more effectively. Identify the main characteristics of the consumers in each segment.
2. Provide a customer centric recommendation to increase overall profit.
3. Predict customer profitability over time based on time series.

**Dataset**

<https://archive.ics.uci.edu/ml/datasets/Online+Retail+II>

Sample recommendation for problem statement 2 is shown below.

|  |  |  |  |
| --- | --- | --- | --- |
| Cluster | Customer type | RMF Characteristics | Action |
| 0 | Best customer | Frequent shoppers | Potential to be target customers for launch of new product |
| 1 | New customer | ? | ? |
| 2 | Risk of churn | ? | ? |
| 3 | Lost customer | ? | ? |

## Group 9 and Group 10

**Problem statement**

The dataset has five independent failure modes TWF, HDF, PWF, OSF, and RNF. If one of the five mode fails, then the process fails, and the machine is marked as failed. However, the end user or administrator is never aware of the feature which lead to failure of the process. The task is to design a model which address explainability for the failure.

1. Design a decision tree based explainable model and provide an explanatory interface
2. Compute the fairness of the model
3. Perform a what-if analysis using CeterisParibus on the given dataset

**Dataset**

<https://archive.ics.uci.edu/ml/datasets/AI4I+2020+Predictive+Maintenance+Dataset>

**Hints**

Library for eXplainable AI

<https://github.com/salesforce/OmniXAI>

<https://sfr-omnixai-demo.herokuapp.com/>

Fairness of the model

<https://dalex.drwhy.ai/python-dalex-fairness.html>

<https://dalex.drwhy.ai/python/>

CeterisParibus

<https://github.com/ModelOriented/pyCeterisParibus>

# Team members in each group

**Group 01**

|  |
| --- |
| [Aakanksha, .](javascript://) |
| [Bojja, Srikanth Reddy](javascript://) |
| [Bolem, Harsha Bhargav](javascript://) |
| [Brahmbhatt, Kaushal](javascript://) |
| [Chaudhary, Mudra Monghjibhai](javascript://) |

**Group 02**

|  |
| --- |
| [Chava, Rishikesh](javascript://) |
| [Chodavadiya, Misha Dipakbhai](javascript://) |
| [Christian, Almee](javascript://) |
| [Chunduri, Pavani](javascript://) |
| [Dabreo, Smith Elias](javascript://) |

**Group 03**

|  |
| --- |
| [Dandyala, Sree Guru Charan](javascript://) |
| [Deotare, Swapnil](javascript://) |
| [Ekbote, Darshan](javascript://) |
| [Gajera, Himanshu](javascript://) |
| [Goel, Devansh](javascript://) |

**Group 04**

|  |
| --- |
| [Gupta, Anshika](javascript://) |
| [Hiresadrahalli Dayananda, Arpitha](javascript://) |
| [Jariwala, Sarthak](javascript://) |
| [Joshi, Keval Dharmendra](javascript://) |
| [Katkamwar, Anish](javascript://) |

**Group 05**

|  |
| --- |
| [Keener, Jeremy Brent Hays](javascript://) |
| [Lokhande, Kashmira Anant](javascript://) |
| [Martinez, Anthony](javascript://) |
| [Mehta, Vruddhi Rakesh](javascript://) |
| [Mulakalapally, Soujanya](javascript://) |

**Group 06**

|  |
| --- |
| [Mullangi, Krishna](javascript://) |
| [Muni, Riya Rajesh](javascript://) |
| [Nathala, Varshini](javascript://) |
| [Panure, Anand](javascript://) |
| [Patel, Niti](javascript://) |

**Group 07**

|  |
| --- |
| [Patil, Vishal](javascript://) |
| [Popade, Aparna](javascript://) |
| [Pulla, Sai Lahari](javascript://) |
| [Ramrakhyani, Ankit](javascript://) |
| [Sagiraju, Jahnavi](javascript://) |

**Group 08**

|  |
| --- |
| [Saurabh, Sankalp](javascript://) |
| [Shah, Krimesh](javascript://) |
| [Shah, Suraj Nandkishor](javascript://) |
| [Shah, Varun](javascript://) |
| [Shaik, Waseem Farooq](javascript://) |
| [Sharma, Pragya](javascript://) |

**Group 09**

|  |
| --- |
| [Sharma, Siddharth](javascript://) |
| [Sondagar, Abhikumar](javascript://) |
| [Srivastava, Prakhar](javascript://) |
| [Tata, Atreya Pavan Kumar](javascript://) |
| [Tejani, Jonty Vijaybhai](javascript://) |
| [Thakur, Yash](javascript://) |

**Group 10**

|  |
| --- |
| [Thamma, Bhanu Sumanth Reddy](javascript://) |
| [Thube, Akshada](javascript://) |
| [Vanteddu, Santosh](javascript://) |
| [Venkatachalam, Shruthi](javascript://) |
| [Vora, Taher Murtuza](javascript://) |
| [Waghmare, Akshay Ramesh](javascript://) |