**Assignment - 01**

**Deadline: 11:55 PM, Tuesday, August 05, 2025**

**-----------------------------------------------------------------**

**List of Modifications:**

| **Sl No** | **Time of modification** | **Modification** |
| --- | --- | --- |
| **1** | **01:33, July 23, 2025** | **Specification uploaded** |

**Task Details:**

### Task Overview

In this assignment, you will explore a popular machine-learning problem hosted on [Kaggle](https://www.kaggle.com/competitions/titanic/overview), a platform renowned for machine-learning competitions.

The problem is based on the **Titanic disaster**, where your goal is to predict whether a passenger survived or not. This task will help you gain hands-on experience with basic classification models.

<https://www.kaggle.com/competitions/titanic/overview>

### Steps to Complete the Task:

**Download Data**

* Download the train.csv and test.csv files from the Kaggle Titanic competition page.

**Model Training and Prediction**

* Train your classification model using train.csv. Use *DecisionTreeClassifier* from the scikit-learn library. You have to make the decision on what to do with missing values and string values.
* Generate predictions for test.csv and save the results in a file named prediction.csv.

**File Submission Format**

* Follow the **Submission File Format** provided on Kaggle carefully. Errors often arise due to improper formatting. Refer to the sample [submission](https://drive.google.com/file/d/1Cw5ct_9UsmyOpfC1NW5yk0ThyzRySW85/view?usp=sharing).

**Report Preparation**

* Write a report that includes:
  + The features used in your model.
  + The accuracy achieved (attach a screenshot of the Kaggle evaluation).

**Submission Guidelines:**

Submit the following files:

1. **Report File**: A document in **PDF** or **DOCX** format containing your report.
2. **Source Code**: Your notebook file (**.ipynb**) containing the code used to solve the task.

#### Folder and File Naming

* Create a folder named after your **application number** (e.g., 1705001).
* Add both your report file and the notebook file into this folder.
* Compress the folder into a **ZIP file** (e.g., 1705001.zip). **Note**: Other formats like .rar or .7z will **not** be accepted.

#### Example Submission Steps

For application number **1705001**:

1. Create a folder named **1705001**.
2. Place the report file (PDF/DOCX) and source code (.ipynb) in the folder.
3. Compress the folder into a **ZIP file** named **1705001.zip**.
4. Submit the ZIP file on [**Google Classroom**](https://classroom.google.com/c/Nzg4ODQ4OTQ1MzU1?cjc=freokyht).

Find your application no from [here](https://docs.google.com/spreadsheets/d/1KhemM7YcJLv4rQoG7mwnVgsAuzWqQT57/edit?usp=sharing&ouid=102397830300864945659&rtpof=true&sd=true).

**Optional**:

You may try other models like ANN, Logistic Regression, SVM, etc., and report the accuracy of each model.