

Import required packages

Question 1

Setup a logistic regression model on the data at [adultUCI](#). Discuss the performance of your model using appropriate statistics. Use dummy variables to handle categorical variables.

1. Prepare the data. Create dummy variables for categorical variables. [See this](#)
2. Get feature matrix X, and target variable y (>50k or <50k)
3. Split data into training and testing
4. Normalize data using MinMaxScaler
5. Create a LogisticRegression object for modeling
6. Train the model with training data
7. Compare the precision, recall, and F1-score on the train and test data.
8. Improve the performance of the model on the test dataset.

Your code comes here

Question 2

Create the sparse COO representation ([ref](#)) of the adjacency matrix (no need to find the dense representation) of the following graph. Find its CSR representation using appropriate Python function. Visualize this matrix by using spy() function. Refer to the notebook of Lecture 10 live session.

Image source: [Mathworks](#)

Your code comes here

Question 3

The adjacency matrix of toy world-wide-web has been provided as a text file on the assignment page. Implement the page-ranking algorithm that displays the index of the 10 highest ranking web-pages. Also report on the time it takes to perform these calculations.