Introduction to AI – Fall 2023
Practice 2
Instructor: Dr. Shahab Nabavi
Deadline: 1402/08/15 23:59

In this problem set, You will work with Numpy and Pandas libraries in order to analyze NBA player salaries. So let's start:

STEP 1:

First of all, You must access the dataset. Kaggle has provided the dataset and you must download It from Kaggle. You can use this link, how to access the Kaggle dataset in Google Colab.

STEP 2:

After access to dataset, unzip it and read dataset using Pandas library.

STEP 3:

Now it is time to analyze the data:

1- Which player had the most salary? Plot Top 10 players.
2- Which player had the lowest salary?
3- How many team do we have in dataset?
4- Which team has paid the most salaries? Plot Top 10 Teams.
5- What was the average salary for players who were 30 or older? What about players who were less than 30?
6- Which position has been paid the most salaries? Plot average salaries in each Position.
7- Add a new column to the data frame named SPM that calculates the salaries per minute of players.
8- Calculate the Z-Score of the dataset with Numpy functions.
(Hint: You should remove unnecessary columns and convert categorical columns into one-hot vector.)

Note that You should handle the missing value first and then solve the above questions.

Do this problem set in the Google Colab and submit the .ipynb file with the Google Colab link in Courseware.

Name your file like HW2_StudentId.

GOOD LUCK