Machine Learning Pre-requisites

Introductory Notes:

Please practice python coding. It's not mandatory but it makes life easier. Following are some questions which shall help get prepped for the course. You are free to practice, if you prefer, these questions using other coding languages.

If you find these assignments tough, its recommended to practice some coding by the time the course starts. You may refer to following websites for help if you are new to python.

- https://www.guru99.com/python-tutorials.html
- https://www.w3schools.com/python/

For python users, you may use Google Colab (https://colab.research.google.com/) for practice. Alternately, you may install Anaconda (https://www.anaconda.com/) for coding on your local system.

For programmers in other languages, you have to identify & choose your own IDE's.

Additional websites which provide lots of help & material are:

- https://www.kaggle.com/ -- create an account here. This provided lots of projects and data to learn.
- https://www.datacamp.com/
- https://www.hackerrank.com/
- https://towardsdatascience.com/ contains lots of learning materials.

Additionally, it's a good idea to refresh your knowledge on basics of matrices, statistics and probability. Refresher for these would be provided in the class as well.

Questions

Q1. Create a vector, V1, (length = 100) of random number. Sort them in increasing order.

Q2. Find out what happens when perform V1*3. Multiply V1 with 3 to scale each value 3 times.

Q3. Find the mean and standard deviation for the vector.

Q4. Make a matrix of zeros with 4 rows and 3 columns. Fill the matrix with random numbers. Convert the matrix into a single dimensional array.

Q5. Take a string {S1 = "I am a great learner. I am going to have an awesome life."}. Search for the substring "Am" and find the occurrence count.

Q6. S2 = "I work hard and shall be rewarded well". Add both the strings to make a single string S3.

Q6. Taking the above string S3, split the string into words by using "white space" and "period(.)" as splitters. Put the words into an array. Find the length of the array.

Q8. Remove words "I", "Am", "to" & "and" from the array. Also, remove words containing more than 6 characters. Find the length of the array.

Q9. Consider the date "01-JUN-2021". Split the date and find the values for date, month and year separately. Find the numerical value for the month using calendar sequence (Jan = 1, Dec = 12).

Q10. Create an excel file with 3 columns. The 3 columns contain the names of city, state and PIN code. Make 10 row entries for 10 cities. (Refer table below.) Load this excel file into a data table. Make a fourth column that contains a string with the city & state joined with a comma separation. Write the table back into an excel file.

City	State	PIN Code
BENGALURU	KA	560001
CHENNAI	TN	600001
MUMBAI	MH	400001
MYSURU	KA	570001
PATNA	BH	800001
JAMMU	JK	180001
GANDHI NAGAR	GJ	382001
HYDERABAD	TS	500001
ERNAKULAM	KL	682001
AMARAVATI	AP	522001

Q11. Sort the vector V1 in increasing order and plot it. Observe the color of the plot & change the color of the plot to red.

Q12. Create another vector V2 containing the square of the number present in V1. Make a plot of both V1 and V2.