

## Assignment 3 – Maths for Machine Learning Part 1



**Objective:** Applying different linear algebra operations to Machine Learning dataset

**Step 1: Download** the **Red Wine Dataset** from <https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-red.csv>

The dataset description can be found in:

<https://archive.ics.uci.edu/ml/datasets/Wine+Quality>

**Step 2: Read** the dataset and store it as a 2D NumPy matrix (1599\*12) called XY.

**Step 3: Apply** the following matrix operations (independently to the source matrix of size 1599\*12) to the dataset matrix (Use built in functions using different libraries)

1. **Print** shape of the matrix XY.
2. **Flatten** the matrix XY into a 1D array
3. **Slice the matrix (column slicing):** Separate the features (1599\*11 matrix: all columns except the last one) of the dataset and target variable (1599\*1 matrix: last column) i.e. quality of wine.
4. **Slice the matrix (row slicing):** Divide the matrix XY into two equal sections of size 800\*12 each.
5. **Shuffle** the rows of the matrix XY
6. **Transpose** the matrix XY and print the shape
7. **Inverse** the matrix XY and print the shape
8. **Compute** the determinant of the matrix XY.

**Step 4: Create** a random matrix of size 1599\*12 named RXY and perform the following matrix level operations

1. **Add** XY and RXY
2. **Subtract** RXY from XY
3. **Outer product** between XY and RXY

**Step 5: Create** a random matrix of size 12\*1 named RXY2 and perform the **inner product** between XY and RXY2 (i.e. dot product between 1599\*12 and 12\*1). Print the matrix as well as its shape.

Step 6: Discover the real applications of the operations used in step 3,4 and 5.

**Suggested Package:** numpy. **Marking:** Marking is based on both **performance during the lab hours** as well as **complete submission**.