



## Assignment - 1

**Course Name:** Advanced Machine Learning Lab

**Course Code:** PMDS602P

**Instructor's Name:** Dr. Gayatri S. Panicker

**Slot:** L7 + L8

Fall Semester 2025 - 26

VIT - Vellore

---

**Due Date: August 31, 2025**

---

### Basic Instructions:

- Write each program in a new script file.
- In your assignment, write your code in full.
- Also attach screenshots of the output obtained on running your program.
- The uploaded file on VTOP *must be a PDF file* (Word files will *not* be evaluated).

Write programs in Python to evaluate the following problems.

1. Generate a set of multivariate three-dimensional data following normal distribution with mean  $\mu = [4, 6, 8]^T$  and covariance matrix

$\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}$  having size 1000. Plot the generated points in a 3d plot.

Further, generate a dataset of two-dimensional points following normal distribution, forming a circle with radius 5 centred around  $[1, 0]^T$ . Plot the generated points.

2. Identify (with justification) the outlier(s) amongst the data  $A = \{[1, 0]^T, [0, 1]^T, [1, 1]^T, [1, 1.5]^T, [1.5, 1]^T, [1.5, 1.5]^T, [2, 1.5]^T, [1.5, 2]^T, [1, 2]^T, [10000, 15000]^T, [20000, 15000]^T, [1, 15000]^T\}$ . Consider the level of significance to be 5%. If you modify the level of significance to be 10% how would your answer change? If it were 15%?