
Work Integrated Learning Programmes Division
Cluster Programme in Data Science and Engg.

# Assignment 2

## DSECL ZC416

## Mathematical Foundations for Data Science

---

**Instructions**

1. This is **NOT** a group activity. Each student should do the problems and submit individually.
2. Assignments have to be handwritten and uploaded as a single pdf file with name BITSID.pdf
3. Do **NOT** use the built-in library functions and use any programming language of your choice.
4. Write the necessary theory and then your observations in the answer-script. **NO** need to include the code. It is a must to include graphical outputs wherever applicable.
5. Submissions beyond 6th of September, 2023 23.59 hrs would not be graded
6. Assignments sent via email would not be accepted
7. Copying is strictly prohibited. Adoption of unfair means would lead to disciplinary action.

---

Answer all the questions

**Q1)** Write a code to implement the three versions of the gradient descent namely. vanilla gradient descent, mini batch gradient descent and stochastic gradient descent for a data set consisting of at least 1500 rows and 15 columns. Write the necessary theory and experiment on the convergence using the optimal batch size for the mini batch method. (4)

**Q2)** Implement the line search methods (bisection, golden section and Armijo rule) on the data set taken for **Q1)**. What do you observe on the time taken by the methods and the associated convergence? (3)

**Q3)** Given the function $f(x) = (x^2\cos(x) + \sin(x) - x)$, derive the variable step size ($\gamma$) that will ensure faster convergence of the gradient descent method and compare the results with (3)

i) constant step size

ii) decaying step size using various values of $\alpha_0$ and $k$

iii) Bold driver algorithm