



VII SEMESTER B.TECH. (COMPUTER SCIENCE ENGINEERING)

END-SEMESTER EXAMINATIONS, DEC-JAN 2022

SUBJECT: MACHINE LEARNING WITH BIG DATA [CRA 4057]

PART B

(24/12/2021)

Time: 75 Min

MAX. MARKS: 20

Instructions to students:

Students must write the answers in their own handwriting on a white sheet of paper. The details to be mandatorily written on the answer sheet are : **Name, Semester, Roll Number, Registration Number, Course name, Signature with date [in each sheet].**

On completion of answering the assignment, students need to scan/image (using device of your choice) all the answer sheet/s in sequence and **save it with file name whose file name convention should be Registration Number followed by "CRA4057" and "Date" in PDF form. Ex: 200905011_CRA4057_24DEC2021.pdf**

Upload the above pdf file containing the answers to MS Teams Assignment platform within the given schedule. **CLICK ON "ADD WORK" OPTION, ATTACH THE PDF AND THEN CLICK ON "TURN IN".**

- | | | |
|------------|---|----------|
| 1A. | Illustrate the steps involved in machine learning process. | 5 |
| 1B. | When does overfitting and underfitting occurs in classification and what are their results? How to avoid overfitting? | 3 |
| 1C. | Illustrate with examples how missing values are handled in machine learning? | 2 |
| 2A. | Explain Decision Tree algorithm in detail. What are Entropy and Information gain in Decision tree algorithm? | 4 |
| 2B. | What is Linear Regression? What Are the Drawbacks of the Linear Model? | 3 |
| 2C. | What is the difference between Regression and classification ML techniques? What are Recommender Systems? | 3 |