National Institute of Technology Kurukshetra

Machine Learning and Data Analytics-CSIC 221 (Civil Engineering)

Mid-Term Examination-I, Odd Semester, 2025-26

Time: 50 Minutes Total Marks: 15

Note: Attempt all the questions. If require any missing data; then choose suitably.

Q1. A) Write any two differences and any two similarities between Binomial distribution and Poisson distribution.

[2 marks]

B) A student is applying for Masters Course in 8 US Universities and believes that she has in each of the eight universities a constant and independent 0.42 probability of getting selected. Answer the following questions: **(i)** what is the probability that she will get call from at least 3 universities? **(ii)** What is the probability that she will get calls from exactly 4 universities? **[1.5 +1.5= 3 marks]**

Q2. Height of ten men are taken from a population given as (unit in feet and inches) 5.8, 5.9, 5.2, 5.7, 5.6, 6.2, 5.7, 5.1, 6.3, 5.8. Transform this height attribute using i) Standardization ii) Normalization

[4 marks]

Q3. Marks of few students in Mathematics and English are given in the table below. Find the Spearman Rank correlation. **[3 marks]**

	Marks									
English	56	75	45	71	62	64	58	80	76	61
Maths	66	70	40	60	65	56	59	77	67	63

Q4. Consider the following sentences

S1: I love travelling to Shimla.

S2: The minister sanctioned funds for development.

S3: I love travelling on hills.

S4: I love to boost the immunity with zinc.

The list of key words in the vocabulary = [love, travelling, Shimla, minister, sanctioned, funds, development, hills, boost, immunity, zinc].

i) Convert the sentences into vectors based on keywords given, by neglecting the non-keyword words.

ii) Find cosine similarity between all possible pairs of sentences converted in vector form.

iii) Find the Manhattan distance between all possible pairs of sentences in vector form. **[3 marks]**