

Dr. Ravinder Bhatt

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2021

B.Tech V Semester

COURSE CODE: 20B1WCI531

MAX. MARKS: 35

COURSE NAME: Foundation for Data Science and Visualization

COURSE CREDITS: 2

MAX. TIME: 2 Hours

Note: All questions are compulsory.

1. [2.5 + 2.5 + 2]
 - a) What is the difference between unsupervised and supervised learning?
 - b) What is clustering? Describe an example algorithm that performs clustering. How can we know whether it produced decent clusters on our data set?
 - c) Prove that Euclidean distance is in fact a metric.
2. [2.5 + 2.5 + 2]
 - a) What do we mean when we talk about the bias-variance trade-off?
 - b) Prove that the eigenvalues of A and A^T are identical.
 - c) Prove that matrix multiplication is associative.
3. [2.5 + 2.5 + 2]
 - a) How do we handle categorical variables in Logistic Regression?
 - b) What assumptions are required for linear regression?
 - c) How to build a Regression Model in Python?
4. [2.5 + 2.5 + 2]
 - a) The number of patients who come daily to the emergency room (E.R.) of a certain hospital has a Poisson distribution with mean 10. What is the probability that, during a normal day, the number of patients admitted in the emergency room of the hospital will be less than or equal to 3?
 - b) A health sensor produces a stream of twenty different values, including blood pressure, heart rate, and body temperature. Describe two or more techniques you could use to check whether the stream of data coming from the sensor is valid.
 - c) We often say that correlation does not imply causation. What does this mean?
5. [2.5 + 2.5 + 2]
 - a) Find the least squares straight line fit to the four points $(0, 1)$, $(1, 3)$, $(2, 4)$, and $(3, 4)$
 - b) Find the number of parameters in the general solution of $Ax = 0$ if A is a 5×7 matrix of rank 3.
 - c) Show that $u = (-2, 3, 1, 4)$ and $v = (1, 2, 0, -1)$ are orthogonal vectors in R^4 .