

Mid.

* Lec - 1 (2nd 8)

* Lec 1

→ Theory

→ AI, ML, DL
definition (sq)

→ Traditional & ML programming
difference.

→ Types of ML definition, necessary

* Lec - 2

→ K-NN implementation

→ why K-NN is called a lazy learner?
→ How to find optimal value
of k in K-NN? (in detail)
algo method.

→ k এর value বড় বা খুচ (ছোট)
কি হবে? / what will be effect
of taking large or small value
of k ?
optimal

→ K-NN distance formula 3ট, 1

→ Pros & cons of K-NN.

→ 10 to 15

o Algo implementation.

o Algo theke short

ques → 5 marks

o concept short q &
performance u ques

Programming VS

→ implt from Mahesh huddar youtube.

→ logistic vs linear Regression

sigmoid.

→ logistic R এর implt.

Lec: 5 → Model performance → page 2 এর ২ টা

→ Dataset কে কত ভাগে ভাগ করা যাবে? 3 types.
(train test & split এর জোড়ি, কখন use করা হবে আর
কেন use হবে?)

→ k fold cross validation কখন use করা হবে
image বা table করে,

→ confusion Matrix implt in a dataset table

→ precision বের করা এই type এর 2
confusion matrix table করে
then ans.

→ performance measure evaluate করা,

→ Drawbacks of accuracy.

Mahesh
Huddan

→ Know the implementation & the performance measure →

→ supervised & unsupervised ML Algo (class & regression)

* Lec 3 → univariate linear Reg → multivariate, multiple, Uni definition

→ Implementation practice with data page 23. (hypothesis & cost function)

→ LR pros & cons.

→ PR definition, imple

→ what is the modification of LR? ans.

* Lec 4 → sub function of gradient Descent imple with data

→ How to choose step size / of alpha? Learning Rate / α

→ How to choose α

→ ~~α value~~ affect to J

→ α value must be small
(how to choose it?)

Lec : 6 → Bias & variance for ~~fitting~~ ^{fitting} over ~~fitting~~ ^{under} fitting
→ formula द्वारा बात बताते हैं,

Lec : 7 → unsupervised → k means clustering.
→ pros & cons.
→ implementation.

Lec : 8 → supervised learning → Naive Bayes

→ Naive Bayes ^{does} implt.

→ why Naive Bayes called Naive ?

CT 2 (A)

CSE 427: Topics of Current Interest

Time: 25 minutes

26 Aug 2

Name:

Student ID:

1. Given the following dataset:

X	Y	Z
2	4	3
3	5	4
9	8	6
4	6	7
10	9	8

Consider two centroids: $C_1 = (3, 4, 1)$ and $C_2 = (9, 9, 9)$. Using a clustering algorithm, **select** the given data points to their respective clusters based on the nearest centroid. Use two iterations only.

CT 2 (B)

CSE 427: Topics of Current Interest

Time: 20 minutes

18 Aug

Name:

Student ID:

1. Given the following dataset:

X	Y
2	4
3	5
9	7
4	3
10	9

Consider two centroids: $C_1 = (3, 4)$ and $C_2 = (9, 7)$. Using a clustering algorithm, **select** the given data points to their respective clusters based on the nearest centroid. Use two iterations only.

CT 1 (C)

CSE 427: Topics of Current Interest

Time: 20 minutes

28 July 2020

Name:

Student ID:

1. Based on the following dataset, use the k-Nearest Neighbors (k-NN) algorithm with $k = 3$ to predict whether a patient with a Glucose Level of 130 mg/dL, BMI of 25.0, and Age of 35 has diabetes (1) or not (0).

Glucose Level (mg/dL)	BMI	Age	Diabetes (1) / No Diabetes (0)
110	22.5	26	0
140	28.0	45	1
125	24.5	34	0
160	30.2	50	1
115	23.8	29	0
145	27.5	42	1
170	32.0	55	1
105	21.0	22	0
135	26.0	45	1
155	29.5	44	0

There are Three Questions. Answer all of them. Part marks are shown in the margins.

1. a. Describe the concept of gradient descent in the context of optimizing a machine learning model.
- b. Given the following scenarios, identify whether the issue is likely due to high variance, high bias, or both:
 - I. A model performs exceptionally well on the training set but poorly on the test set.
 - II. A model performs poorly on both the training and test sets.
 - III. A model performs decently on the training set but its performance fluctuates significantly across different test sets.
2. a. Consider the following dataset where x and y represent the two input variables corresponding to coordinate points, while "Class" serves as the dependent variable.

x	y	Class
-2	1	B
-1	0	A
0	2	B
1	1	A
2	3	B
3	0	A
3	2	B

Classify the new data point (1,3) using Euclidean distance with 3-NN and analyze the impact of increasing the value of k to a larger number, relative to the size of the dataset, on the model's prediction.

3. a. Explain the differences between linear regression and multivariate linear regression in modeling relationships between independent and dependent variables. Use two different dataset examples: one where linear regression is applied and another where multivariate linear regression is applied.

Mid Semester Examination, Fall 24
Program: B.Sc. in CSE
4th year 1st Semester

Course Title: Topics of Current Interest
Time: 1 hour.

Credit Hour: 3.00

Course Code: CSE 427
Full Marks: 20

There are Two Questions. Answer all of them. Part marks are shown in the margins.

QUESTION 1 [10 MARKS]

- a. Demonstrate the differences between Classification and Regression with example. [5]
- b. Consider the following dataset: [5]

X	Y
2	4
3	5
9	8
4	6
10	9

Consider two centroids: $C_1 = (3, 4)$ and $C_2 = (9, 9)$. Using a clustering algorithm, select the given data points to their respective clusters based on the nearest centroid. Use Manhattan distance and two iterations only.

QUESTION 2 [10 MARKS]

- a. A mobile app company wants to predict whether a user will renew their subscription (1) or cancel it (0) based on their average daily app usage (minutes), number of notifications received per day, and number of months subscribed. Predict whether a user with 35 minutes of daily usage, receives 4 notifications per day, and has been subscribed for 9 months will renew or cancel utilizing the k-NN algorithm ($k = 3$).

Avg Daily Usage (mins)	Notifications per Day	Months Subscribed	Renew (1) / Cancel (0)
45	5	12	1
20	2	4	0
50	6	15	1
10	1	2	0
40	4	10	1
25	3	6	0
55	7	18	1
15	1	3	0
48	5	14	1
30	3	8	0

Given a dataset containing 12 instances with their actual (y) and predicted class (\hat{y}) labels for a binary classification task, construct the confusion matrix and measure Precision, Recall, and F_1 -Score. Assume the positive class is labeled as 1 and the negative class as 0.

Instance	1	2	3	4	5	6	7	8	9	10	11	12
y	1	0	1	0	1	0	1	0	1	1	0	0
\hat{y}	1	0	0	1	1	0	1	0	1	0	0	1

Alright, let's go step by step.

a. Gradient Descent Concept

Gradient Descent is an **optimization algorithm** used to minimize the loss (or cost) function of a machine learning model.

- The idea:
Imagine the loss function as a curve (or surface in higher dimensions). Each point represents how well the model is performing for a given set of parameters (weights, biases).
- Gradient descent starts with random parameters and **iteratively updates them** in the direction that reduces the loss.


- Update rule:

$$\theta := \theta - \alpha \cdot \nabla J(\theta)$$
 - θ = model parameters
 - α = learning rate (step size)
 - $\nabla J(\theta)$ = gradient (slope of loss function)


This continues until convergence (minimum loss) or until a stopping condition is met.

b. Bias–Variance Identification


1. A model performs exceptionally well on the training set but poorly on the test set.

- This means the model memorized the training data but failed to generalize.
 **High Variance (Overfitting)**
-

2. A model performs poorly on both the training and test sets.

- The model cannot even capture patterns in the training data.
 **High Bias (Underfitting)**
-

3. A model performs decently on the training set but its performance fluctuates significantly across different test sets.

- The model has learned some patterns but is unstable when generalizing.
 - This inconsistency indicates **sensitivity to different data samples**.
 **High Variance**
-

Final Answers:

- 1 → High Variance
- 2 → High Bias
- 3 → High Variance

Do you want me to also make a **comparison table** for bias vs variance with symptoms (training/test performance), so you can use it for quick exam revision?

Got it 👍 Let's break it down clearly.

a. Difference Between Linear Regression and Multivariate Linear Regression

1. Linear Regression (Simple Linear Regression)

- Models the relationship between **one independent variable (X)** and **one dependent variable (Y)**.
- Equation:
$$Y = \beta_0 + \beta_1 X + \epsilon$$
- Here, β_0 is the intercept, β_1 is the slope, and ϵ is the error term.
- Goal: Find the best-fit line through the data points.

2. Multivariate Linear Regression (Multiple Linear Regression)

- Models the relationship between **two or more independent variables (X1, X2, ..., Xn)** and **one dependent variable (Y)**.
 - Equation:
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon$$
 - Instead of a line, the model fits a **hyperplane** in higher dimensions.
 - Goal: Capture how multiple factors together affect the outcome.
-

Examples

1. Linear Regression Example (Single Variable)

- **Dataset:** Predicting a student's exam score (**Y**) based only on the number of study hours (**X**).
 - **Model:**
$$\text{Score} = \beta_0 + \beta_1(\text{Hours of Study}) + \epsilon$$
$$\text{Score} = \beta_0 + \beta_1(\text{Hours of Study}) + \epsilon$$
 - **Interpretation:** Each additional study hour increases the predicted score by β_1 .
-

2. Multivariate Linear Regression Example (Multiple Variables)

- **Dataset:** Predicting a house price (**Y**) based on multiple features:
 - X_1 : Size of the house (sq. ft)
 - X_2 : Number of bedrooms
 - X_3 : Distance from city center
 - **Model:**
$$\text{Price} = \beta_0 + \beta_1(\text{Size}) + \beta_2(\text{Bedrooms}) + \beta_3(\text{Distance}) + \epsilon$$
$$\text{Price} = \beta_0 + \beta_1(\text{Size}) + \beta_2(\text{Bedrooms}) + \beta_3(\text{Distance}) + \epsilon$$
 - **Interpretation:**
 - β_1 shows how much price changes per unit increase in size.
 - β_2 shows effect of extra bedrooms.
 - β_3 shows effect of distance from city center.
-

✓ Summary in One Line:

- **Linear regression** → one predictor, one response.
 - **Multivariate linear regression** → multiple predictors, one response.
-

Would you like me to also create a **side-by-side table** comparing the two (with equations, visualization, and examples) so you can revise quickly?