



QUIZ (2)

Data Mining
IS – IT – BIO – SW
2025

Quiz (1)

- **Real Estate Pricing Prediction:**

- Dataset: House prices based on size (sq. ft.) and bedrooms.
- Use the following Classifiers:
 - KNN (k=3)
 - Naïve Bayes
- Note: use 80% of the data as a training.
- Compute The Accuracy
- New House to Classify:
 - Size: 2300 sq. ft.
 - Bedrooms: 3

Size (sq. ft.)	Bedrooms	Price (Category)
1500	2	Affordable
2000	3	Affordable
2200	3	Moderate
3000	4	Expensive
2700	4	Moderate
1800	3	Affordable
2500	3	Moderate
3200	5	Expensive

Quiz (2)

- **Simple Classification:**

- Dataset: Animal Features & Classification.
- Use the following Classifiers:
 - Decision Tree
 - Naïve Bayes
- Note: use 80% of the data as a training.
- Compute The Accuracy
- New Object to Classify:
 - Has Fur: No
 - Can Fly: Yes
 - Lays Eggs: Yes

Animal	Has Fur	Can Fly	Lays Eggs	Class
Dog	Yes	No	No	Mammal
Eagle	No	Yes	Yes	Bird
Dolphin	No	No	No	Mammal
Penguin	No	No	Yes	Bird
Bat	Yes	Yes	No	Mammal
Sparrow	No	Yes	Yes	Bird
Whale	No	No	No	Mammal
Chicken	No	No	Yes	Bird

Quiz (3)

- **Penguin Classification:**

- Dataset: Penguin species (Adelie, Gentoo, Chinstrap) using bill length (mm) and flipper length (mm).
- Use the following Classifiers:
 - KNN (k=3)
 - Decision Tree
- Note: use 80% of the data as a training.
- Compute The Accuracy
- New Penguin to Classify:
 - Bill Length: 42.0 mm
 - Flipper Length: 190 mm

Bill Length (mm)	Flipper Length (mm)	Species
39.1	181	Adelie
45.4	195	Gentoo
40.3	185	Adelie
50.1	205	Gentoo
47.2	194	Chinstrap
38.9	179	Adelie
49.5	216	Gentoo

Transaction ID	Items Purchased
1	Milk, Bread, Eggs
2	Bread, Butter
3	Milk, Bread, Butter
4	Milk, Eggs
5	Bread, Butter, Eggs

Quiz (4)

- **Market Basket Analysis:**
 - Dataset: Grocery Store Transactions.
 - Use the following techniques:
 - Apriori
 - FP-Growth
 - Note: 50% Support – 70% Confidence.
 - List the Association Rules.

Quiz (5)

- **Electronics Store Purchase:**
 - Dataset: Electronics Store Transactions.
 - Use the following techniques:
 - Apriori
 - FP-Growth
 - Note: 50% Support – 70% Confidence.
 - List the Association Rules.

Items:

- L = Laptop
- M = Mouse
- K = Keyboard
- H = Headphones
- C = Charger

Transaction ID	Items Purchased
1	L, M, K
2	M, K
3	L, M, H
4	M, K, H
5	L, C
6	L, M, K, C

Quiz (6)



- **Stock Market Data:**

- Dataset: Simple Stock Prices.
- Use the following Classifiers:
 - Linear Regression
- Compute The Accuracy (r2_score function)
- New House to Classify:
 - Opening Price (\$): 107.
 - Trading Volume: 13000

Day	Opening Price (\$)	Trading Volume	Closing Price (\$)
1	100	10000	102
2	102	12000	105
3	105	15000	108
4	108	9000	106
5	106	11000	109

- **Market Basket Analysis:**

- Dataset: Grocery Store Transactions.
- Using Apriori, find Frequent Itemsets (Support $\geq 50\%$).

Transaction	Items
1	milk, bread
2	bread, eggs
3	milk, eggs
4	milk, bread, eggs
5	bread, cookies