1 Demonstrate following pre-processing operations using R/Python 40

a) Deleting missing values
b) Replacing missing values
c) Imputing missing values
d) Work with categorial variables
e) Work with outliers

Use Dataset: titanic_train.csv

2 Journal 05

3 Viva 05

1 Demonstrate Simple Linear Regression model using R/Python 40

a) Define Problem Statement
b) Define Null Hypothesis
c) Perform Pre-processing operations on dataset
d) Prepare Model
e) Use Model for prediction
f) Evaluate Model

Use Dataset: Use any suitable dataset

2 Journal 05

3 Viva 05

1 Demonstrate Multiple Linear Regression using R/Python [40]

a) Define Problem Statement
b) Define Null Hypothesis
c) Perform Pre-processing operations on dataset
d) Prepare Model
e) Use Model for prediction
f) Evaluate Model

Use Dataset: Use any suitable dataset

2 Journal [5]

3 Viva [5]

1 Perform following tasks [40]

a) Create any R markdown document implementing any machine learning algorithm of your choice.
b) Upload it in your RStudio account

2 Journal [5]

3 Viva [5]

1 Demonstrate Logistic Regression using R/Python [40]

a) Define Problem Statement
b) Define Null Hypothesis
c) Is it classification or prediction problem. Explain.
d) Perform Pre-processing operations on dataset
e) Prepare Model
f) Use Model for prediction
g) Evaluate Model

Use Dataset: Use any suitable dataset

2 Journal [5]

3 Viva [5]

1 Perform following Hypothesis testing methods using R/Python [40]

a) One sample t-test
b) Two sampled t-test
c) Paired sampled t-test
d) ANOVA (F-TEST)

Use Dataset: Use any suitable dataset

2 Journal [5]

3 Viva [5]

1 Implement Decision Tree Algorithm (Classifier) using R/Python [40]

a) Define Problem
b) Implement Decision Tree Algorithm on suitable dataset.
c) How to evaluate the above algorithm?

Use Dataset: Use any suitable dataset

2 Journal [5]

3 Viva [5]

1 Implement PCA using R/Python [40]

a) What is Dimension reduction?
b) What are different methods for dimension reduction?
c) Why Dimension reduction is important?
d) Implement PCA Algorithm on suitable dataset.
e) How to evaluate the above algorithm?

Use Dataset: Use any suitable dataset

2 Journal [5]

3 Viva [5]

1 Perform following task using MongoDB [40]

a) Create suitable database in MongoDB.
b) Create suitable collection in database.

c) Insert 3 documents in above collection.
d) Perform CRUD operation on documents inserted in collection.
e) What is use(s) of MongoDB database?

2 Journal [5]
3 Viva [5]

1 Implement K-means clustering using R/Python [40]
a) What is clustering?
b) Write steps of K-means clustering algorithm.
c) How to determine best value of k?
d) Implement K-means clustering on suitable dataset.
e) How to evaluate the above algorithm?
Use Dataset: Use any suitable dataset
2 Journal [5]
3 Viva [5]

1 Implement Time-series forecasting using R/Python [40]
a) What is time-series data? Give example.
b) Define the problem.
c) Implement Time-series forecasting on suitable dataset.
d) How to evaluate the above algorithm?
Use Dataset: Use any suitable dataset
2 Journal [5]
3 Viva [5]

1 Perform following tasks [40]
c) Create any R markdown document implementing any machine learning algorithm of your choice.
d) Upload it in your RStudio account

2 Journal [5]
3 Viva [5]

1 Perform following task using MongoDB [40]
f) Create suitable database in MongoDB.
g) Create suitable collection in database.
h) Insert 3 documents in above collection.

i) Perform CRUD operation on documents inserted in collection.
j) What is use(s) of MongoDB database?

2 Journal [5]
3 Viva [5]

1 Perform following task using MongoDB [40]
a) Create suitable database in MongoDB.
b) Create suitable collection in database.
c) Insert 3 documents in above collection.
d) Perform CRUD operation on documents inserted in collection.
e) What is use(s) of MongoDB database?

2 Journal [5]
3 Viva [5]

1 Perform following tasks [40]
e) Create any R markdown document implementing any machine learning algorithm of your choice.
f) Upload it in your RStudio account

2 Journal [5]
3 Viva [5]