



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

1. Mark and his family are planning to move to a new city and are in the market for a new home. They have been searching online for homes in their desired area and have found several properties that meet their requirements. However, they are not sure about the prices of these homes and want to get a rough estimate before making an offer. How will you help Mark to buy a new house?

- a) Read the house dataset using the Pandas module; b) Print the first five rows
c) Perform basic statistical computations on the dataset or show distribution of data
d) Print the columns and their data types; e) Detect null values in the dataset. If there are any, replace them with the mode value; f) Explore the dataset using a heatmap; g) Split the data into training and testing sets; h) Predict the price of a house.

2. Implement a Python program for the most specific hypothesis using Find-S algorithm for the following given dataset and show the output:

Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes

3. Develop a Python code for implementing Linear regression and show its performance.

4. Develop a Python code for implementing the EM algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

1. Can the breast cancer classification problem be solved using Naive Bayes classification?

- a) print the 1st five rows; b) Basic statistical computations on the data set or distribution of data
c) the columns and their data types; d) Detects null values in the dataset. If there is any null value, replace it with mode value; e) Split the data into test and train
f) Evaluate the performance of the model by evaluation metrics such as confusion matrix.

2. Implement a Python program for the most specific hypothesis using Find-S algorithm for the following given dataset and show the output:

Size	Color	Shape	Class
Big	Red	Circle	No
Small	Red	Circle	No
Small	Red	Triangle	Yes
Big	Blue	Circle	No
Small	Blue	Circle	Yes

3. Develop a Python code for implementing Polynomial regression and show its performance.

4. Develop a Python code for implementing the KNN algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

1. You are a data scientist at a retail company and your manager has asked you to create a model to predict future sales. The company has been collecting data on sales and advertising expenditures for the past 5 years. Your manager wants to use this information to forecast sales for the next quarter and make informed decisions about advertising and inventory. Your task is to build a predictive model that takes into account past sales data and advertising expenditures to forecast sales for the next quarter. You decide to use linear regression to build your model because it is a simple and interpretable method for predicting a continuous outcome.
 - a) Print the 1st five rows;
 - b) Basic statistical computations on the data set or distribution of data
 - c) the columns and their data types;
 - d) Explore the data using scatterplot
 - e) Detects null values in the dataset. If there is any null value, replace it with mode value
 - f) Split the data into test and train
2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.

Size	Color	Shape	Class
Big	Red	Circle	No
Small	Red	Triangle	No
Big	Red	Circle	Yes
Small	Red	Circle	No
Small	Blue	Circle	Yes

3. Develop a Python code for implementing Logistic regression and show its performance.
4. Develop a Python code for implementing the Naive Bayes algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

1. Julia is a botanist who is studying the Iris genus. She has collected the data of different sepal length, sepal width, petal length, and petal width of various Iris flowers and wants to classify the flowers into their respective species based on their physical characteristics. How is the Perceptron algorithm applied to the Iris flower classification problem?
2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.

Example	Citations	Size	Library	Price	Editions	Buy
1	Some	Small	No	Affordable	Few	No
2	Many	Big	No	Expensive	Many	Yes
3	Many	In	Medium	No	Expensive	Few
4	Many	Small	No	Affordable	Many	Yes

3. Develop a Python code for implementing Polynomial regression and show its performance.
4. Develop a Python code for implementing the KNN algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LABORATORY EXAM ITA06 : Machine Learning

1. Julia decides to compare the performance of different machine learning algorithms for this task. She splits her data into a training set and a test set and trains several models, including Decision Tree Classifier, Logistic Regression, KNN Classifier. Julia wants the performance measures based on accuracy and speed of execution. Help her do the comparison of the classification algorithms. Use Itir.csv flowers dataset.
2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.

Shape	Size	Color	Surface	Thickness	Target Concept
Circular	Large	Light	Smooth	Thick	Malignant (+)
Circular	Large	Light	Irregular	Thick	Malignant (+)
Oval	Small	Dark	Smooth	Thin	Benign (-)
Oval	Large	Light	Irregular	Thick	Malignant (+)

3. Develop a Python code for implementing Logistic regression and show its performance.
4. Develop a Python code for implementing the Naive Bayes algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LABORATORY EXAM ITA06 : Machine Learning

1. Can the breast cancer classification problem be solved using Naive Bayes classification?
 - a) print the 1st five rows;
 - b) Basic statistical computations on the data set or distribution of data
 - c) the columns and their data types;
 - d) Detects null values in the dataset. If there is any null value, replace it with mode value;
 - e) Split the data into test and train
 - f) Evaluate the performance of the model by evaluation metrics such as confusion matrix.
2. Implement a Python program for the most specific hypothesis using Find-S algorithm for the following given dataset and show the output:

Size	Color	Shape	Class
Big	Red	Circle	No
Small	Red	Circle	No
Small	Red	Triangle	Yes
Big	Blue	Circle	No
Small	Blue	Circle	Yes

3. Develop a Python code for implementing Polynomial regression and show its performance.
4. Develop a Python code for implementing the KNN algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

1. You are a data scientist at a retail company and your manager has asked you to create a model to predict future sales. The company has been collecting data on sales and advertising expenditures for the past 5 years. Your manager wants to use this information to forecast sales for the next quarter and make informed decisions about advertising and inventory. Your task is to build a predictive model that takes into account past sales data and advertising expenditures to forecast sales for the next quarter. You decide to use linear regression to build your model because it is a simple and interpretable method for predicting a continuous outcome.

- Print the 1st five rows;
- Basic statistical computations on the data set or distribution of data
- the columns and their data types;
- Explore the data using scatterplot
- Detects null values in the dataset. If there is any null value, replace it with mode value
- Split the data into test and train

2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.

Size	Color	Shape	Class
Big	Red	Circle	No
Small	Red	Triangle	No
Big	Red	Circle	Yes
Small	Red	Circle	No
Small	Blue	Circle	Yes

- Develop a Python code for implementing Logistic regression and show its performance.
- Develop a Python code for implementing the Naive Bayes algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

- Rani is a botanist who is studying the Iris genus. She has collected data on the sepal length, sepal width, petal length, and petal width of various Iris flowers and wants to classify the flowers into their respective species based on their physical characteristics. Anna decides to use the Perceptron algorithm for this task.
- Implement a Python program for the most specific hypothesis using Find-S algorithm for the following given dataset and show the output;

Example	Citations	Size	Library	Price	Editions	Buy
1	Some	Small	No	Affordable	Few	No
2	Many	Big	No	Expensive	Many	Yes
3	Many	In	Medium	No	Expensive	Few
4	Many	Small	No	Affordable	Many	Yes

- Develop a Python code for implementing and comparing Linear and Logistic Regression.
- Develop a Python code for implementing the Expectation Maximization algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

1. Jack is a car enthusiast and wants to buy a new car. He wants to find the best deal and decides to use machine learning to predict the prices of different car models. Jack collects data on various features such as the make, model, year, engine size, and number of doors, as well as the sale price of each car. He splits the data into a training set and a test set and trains a linear regression model on the training data.

Car Price Prediction with Machine Learning:

- Read the dataset using the Pandas module;
- Print the 1st five rows.
- Basic statistical computations on the data set or distribution of data;
- the columns and their data types
- Detects null values in the dataset. If there is any null value, replace it with mode value
- Explore the data set using heatmap;
- Split the data into test and train
- Fit the model using Naive Bayes Classifier
- Predict the model;
- Find the accuracy of the model

2. Implement a Python program for the most specific hypothesis using Find-S algorithm for the following given dataset and show the output:

Origin	Manufacturer	Color	Decade	Type	Example Type
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chrysler	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive

- Develop a Python code for implementing Polynomial regression and show its performance.
- Develop a Python code for implementing the KNN algorithm with an example.



SIMATS SCHOOL OF ENGINEERING
SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES
CHENNAI-602105



MODEL LBORATORY EXAM ITA06 : Machine Learning

1. John is a young professional who wants to buy his first home. He knows that his credit score is an important factor in determining whether he will be approved for a loan, so he decides to check it. He goes to a financial website that offers a free credit score prediction service based on machine learning algorithms.

- Print the 1st five rows;
- Basic statistical computations on the data set or distribution of data
- the columns and their data types;
- Detects null values in the dataset. If there is any null value, replace it with mode value;
- Explore the data set using boxplot (Credit Scores Based on Occupation)
- Split the data into test and train;
- Fit the model using Naive Bayes Classifier;
- Predict the model

2. Implement a Python program for the most specific hypothesis using Find-S algorithm for the following given dataset and show the output:

Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes

- Develop a Python code for implementing Linear regression and show its performance.
- Develop a Python code for implementing the EM algorithm with an example.