

Semester test, 2 unit 2 noise up 2024.

Question 1 SVM model 'mtcars' Splitting data Prediction Accuracy

- * Short description & printing structure of dataset
- * Splitting dataset into Training & test.
- * Fitting SVM - radial & polynomial kernel
- * Predicting 'mpg' values \rightarrow test data.
- * Prediction Accuracy \rightarrow evaluate and comment.

Question 2 'mtcars' SVM model contour plot

- * Fitting SVM model with dependent & independent variable
- * Creating a contour plot for 'mpg' values against the 'wt' and 'hp' values

Question 3 'mtcars' linear regression regression line plot model

- * Fit a linear regression model with 'vs' as dependent variable. Obtaining summary of the model output with brief summary
- * Creating a regression line plot for 'mpg' values against the 'wt' values.

Question 4 ANOVA model

- * Review data structure
- * Fit ANOVA MODEL
- * Are there significant diff b/w yields across treatments.

Question 5.

Normalize data

ANN model

of interest

usage

* Review structure and brief description.

* Normalize dataset.

* Fit ANN model

* Plot ANN model.

Question 6. LOOPS + FUNCTION

For 1) Creating a function with following

loop that computes the factorial

of a given number, n.

Question 7. LOOPS

Repeat loop that continues to add random samples drawn from a standard normal until sum exceeds 10.

labor drama avoidance

survived club weird

labor drama weird

black with this transformation right side

Computer Application

Semester test 2023-24

2023.

Question 1

LOOPS

Write a while loop to find the first 10 Fibonacci numbers.

Question 2

LOOPS

Write a repeat loop to find the factorial of a given number. If the number is less than 0, it should print an error message.

Question 3

LOOPS

Identifying mistake in code

Question 4

Identify the mistake in the code.

- * Scale the dataset around mean and. use standard deviation
- * Split the dataset into training & test.
- * Fit ANN model with a hidden layers with 4 neurons in each.
- * Measure accuracy of predictions i.e. the MSG
- * Plot the ANN model.

Question 5 'mtcars' (linear model) gim (generalized)
CSOS

- * Fit generalized linear model
 - * Use Gaussian function for the family of the gbm fit
 - * Split data into training and test.
 - ↳ use training model for fitting.
 - * Obtain predicted values on test data.
 - * Calculate prediction Accuracy i.e. MSE

Question 6: **FOR Loops**
Write a program that calculates and prints the remaining balance on a loan each year for 5 years.

Question 7 LOOPS. all about them (part 1)

Loop that calculates and prints the balance in a bank account after each year for 10 years showing interest fact alone?

Tutorials

2023S

Tutorial 1

BASIC STATS

- Basic Stats (mean / variance)
- MATRIX Summation

Tutorial 2

LOOPS

- Calculating AVG temps with \Rightarrow while loop
- Count days above avg temp \Rightarrow for loop
- Find max temp swing \Rightarrow repeat loop

Tutorial 3

BASIC STATS ANALYSIS

Linear regression plots

- Loading and Cleaning data.
- Basic Summary
- Removing rows with missing values
- Summary after cleaning
- Imputing missing values with column means
- Replace actual data with imputed data

- Histograms (each quantitative variable other than 'month')
- Boxplots (by month)
- Scatterplot with smoothing
- Pairwise Scatterplot matrix
- Time Series plot
- Linear regression model.
- Fit linear regression model.
- Face Scatter plot
- Scatter plot (GDP vs unemployment)
- CG Pairs.

Tutorial 4. 'mtcars' Data Analysts.

Linear regression / plots / Shapiro-Wilk

- Scatterplot with linear smoothing

Histogram (horsepower)

Boxplot (mpg) grouped by cylinder

Density plot (mpg)

Correlation Analysis

- pairwise plot

Linear regression modeling via pairwise()

Extracting residuals and fitted values

- residuals vs fitted values plot

Normal Q-Q plot (residuals).

Scale-Location Plot (spread vs fitted)

Histogram & density of residuals

Shapiro-Wilk Normality Test

- log-transform the response

- check model presummarily

- Recheck residuals/normality fit using

qqplot after new residuals

standardized vs observed quantile from empirical

(dotted red) qqplot

alg name 7007 • printnorm diag function

alg name 7002 • Examine diagnostic summary

standardized vs

qq diag quant

6 2009 0 • labda 0.012232 per unit

alg name 7002 7007 quant 1.21

0

Tutorial 5 SVM model implementation

- * Load and preview dataset.
- * Create binary response variable.
- * Train - test split.
- * Train SVM model with linear kernel.
- * Predictions on test data.
- * Calculate Training error.

Assignment 1: Drawing random samples / Vectors / Matrices / mean / SD

Assignments

2025) Drawing random samples / Vectors / Matrices / mean / SD

Assignment 1: Drawing random samples / Vectors / Matrices / mean / SD

- * Drawing random samples [excluding endpoints].
- * Calculating the grand / mean / end / standard deviation.
- * Drawing random samples [including endpoints].
- * Creating vectors.
- * Combining vectors / transpose / trace / determinant.
- * Creating a matrix with suitable R functions.

Assignment 2: Analysing correlations / PLOTS

System of linear equations.

- * Solving system of linear equations.
- * Drawing random sample.
- * Calculating mean.
- * Analysing data and preparing a report.
 - Pick up data clean + check histogram 'mpg'.
 - Boxplots for 'mpg' (with car weights).
 - Correlation matrix (apart from 'car' column).

Assignment 3 DATA REPORT

- * Preview data.
- * Histogram (murder rate)
- * Scatterplot (murder vs assault)
- * Line chart (rape over time)
- * Correlation: (urban pop vs rape rate)
↳ bit of explanation.

Assignment 4. Multiple line charts.

Single plot multiple line charts.

Plot using ggplot with facet_wrap
Melting dataset to a long format.

- * Converting variable to numeric
- * Split dataset into training and test
- * Scale dataset
- * Create a neural network model
- * Predict 'medv' using nn model
- * Plot the nnmodel
- * Calculate the (MSE) to measure predict

Q2:

- * Defining different networks for the neural networks
- * Initialise list to store results
- * Loop through each configuration

memory num. run un v... needed

- * Predict mpg using neural network.
- * calculate MSE
- * Store results
- * Print MSE results

Multiple plots exercise, provided before test.

Q1

- Histogram: 'iris' dataset (Sepal length, Sepal width, Petal length, Petal width in single plot.)

Q2

- Histogram: numeric variables in 'mtcars'

Q3

- Using Airquality to plot histograms (Ozone, Solar.R and Wind (ignoring NA))

Q4.

- Visualize histograms of len grouped by supp using facers

Q5 'uscrime'

- Plots crime rate over income and educ

Q6 'mcycle'

- Head acceleration curves

Q7 "Boston" Plotting variables over variables

- Plotting median hove over other variables

- variable
- Q8
use wage dataset and plot wages by exp and edu.
 - Q9
BudgetFood dataset plot food exp by income group

Q10

- Plot wage density from wages dataset

Q11

- Plot density of Avg income from cuschol dataset

Q12

- Plot density of nonlabor income over from SwissLabor dataset (AER package)

Q13

- Density plots for Age of women

Q14

- Show avg wage by edu, marital status from wages data.

Q15

- Compare no of cars by number of cylinders and year types from 'mtcars' dataset

Machine learning tools provided just before the test

Q1

Logistic regression on Binary outcome
(Swiss Labor dataset)

Q2

Poisson regression for count data
(wrablearns dataset)

Q3
Ordinary linear regression 'mtcars'

Q4

Applying SVM on wages data (Binary classification) to model the income based on edu, exp, marital status and sex

Q4

Applying SVM on Swiss Labor data (Binary classification) to model the participation status based on edu, age, inc

Q5

Housing price Categorization using Boston data from MASS package with SVM

Q6

Applying ANN on Swiss Labor data to model participation status based on edu, age, income

Q7: building about parallel grid search fit ANN model on mtcars to predict mpg values based on the hp, wt, disp values.

error metric to decide your model.

Q8: Analyze relationships bw wage-related numeric variables with PCA

Q9: Reduce socio-economic predictors

predictors of labor participation using PCA

(numerical data doesn't have perfect linear relationship or multicolinearity)

PCA is a technique to find the principal components of the data which are orthogonal axes explaining maximum variance.

PCA transformation does not change the data distribution.

What is other disadvantage of PCA project?

Q2

Tutorial

2024

"Tutorial - memo 2024"

Tutorial 4. LOOPS

- * While loop that prints out the Fibonacci sequence up to the 10th term.
- * While loop calculates the sum of their first 50 odd numbers
- * For loop that creates a vector of the first 20 multiples of 3.
- * While loop that checks if a given number is prime
- * For loop creates a matrix of size 5x5

Tutorial 5 Functions / Data analysis

- Write a function to calculate a tip amount based on a bill amount and a tip percentage.
- Write a function to find missing values or consecutive nans and return a vector of the missing values.

Create data partition

Fit a SVM model

Predict values

New training/test data on random subset

RMS, RMSE values

Partition 60:30

(Untar plus including tar partitions)

Tutorial 6

clusters. \Rightarrow Scaled data.

- * Create clusters using scaled data.

cluster plot.

- * Partition data 70:30

- * SVM model \Rightarrow Polynomial kernel.

- * SVM plot

Tutorial 7: MANOVA ANALYSES

Tutorials

2023.

"ALL CAPS NAMING"

Tutorial 5 memo: creating functions

- * fo function to calculate a tip
- * function to calculate salary.
- * function `find_mode()` that takes a numeric vector as input and return the mode. or the vector of norm 0 min
- * calculate npct function
- * find missing values function.

Tutorial 6 memo: Machine learning basic

- a) Mtcars dataset

↳ create neural network

↳ calculate MSE

- * Predicting the species of iris flowers based on four input variables to create neural network

Plot Iris(iris) Plotnet(iris)