



# Hyrje në shkencën e të dhënave

Pjesa 6 – R

Prof. Asoc. Dr. Ermir Rogova

# What is R?

- A data analysis platform
- Can do scientific operations involving math and stats
- Compare to Matlab, Stata, SAS
- Open-source, free, widely used

# Where is R?



DOWNLOAD FOR FREE FOR ANY  
PLATFORM



LOTS OF PACKAGES AND DATASETS  
AVAILABLE AS WELL



R SOFTWARE CAN DO IT ALL, BUT R  
STUDIO (AN IDE) IS BETTER!

# Basics of R

- One-line commands
- Arithmetic operators
- Logical operators
- Control structures
  - Condition checking with 'if' and 'else'
  - While loop
  - For loop
- Functions
- Importing data

# Data frames

- An array or a matrix that contains rows and columns
- Available in R without loading any external packages
- Addressing a row or column by their names

# Visualizations with R

- ggplot2 package
- Plotting histogram, bar chart, pie chart, line chart

# Basic statistics and machine learning



Descriptive  
statistics



Correlation



Regression



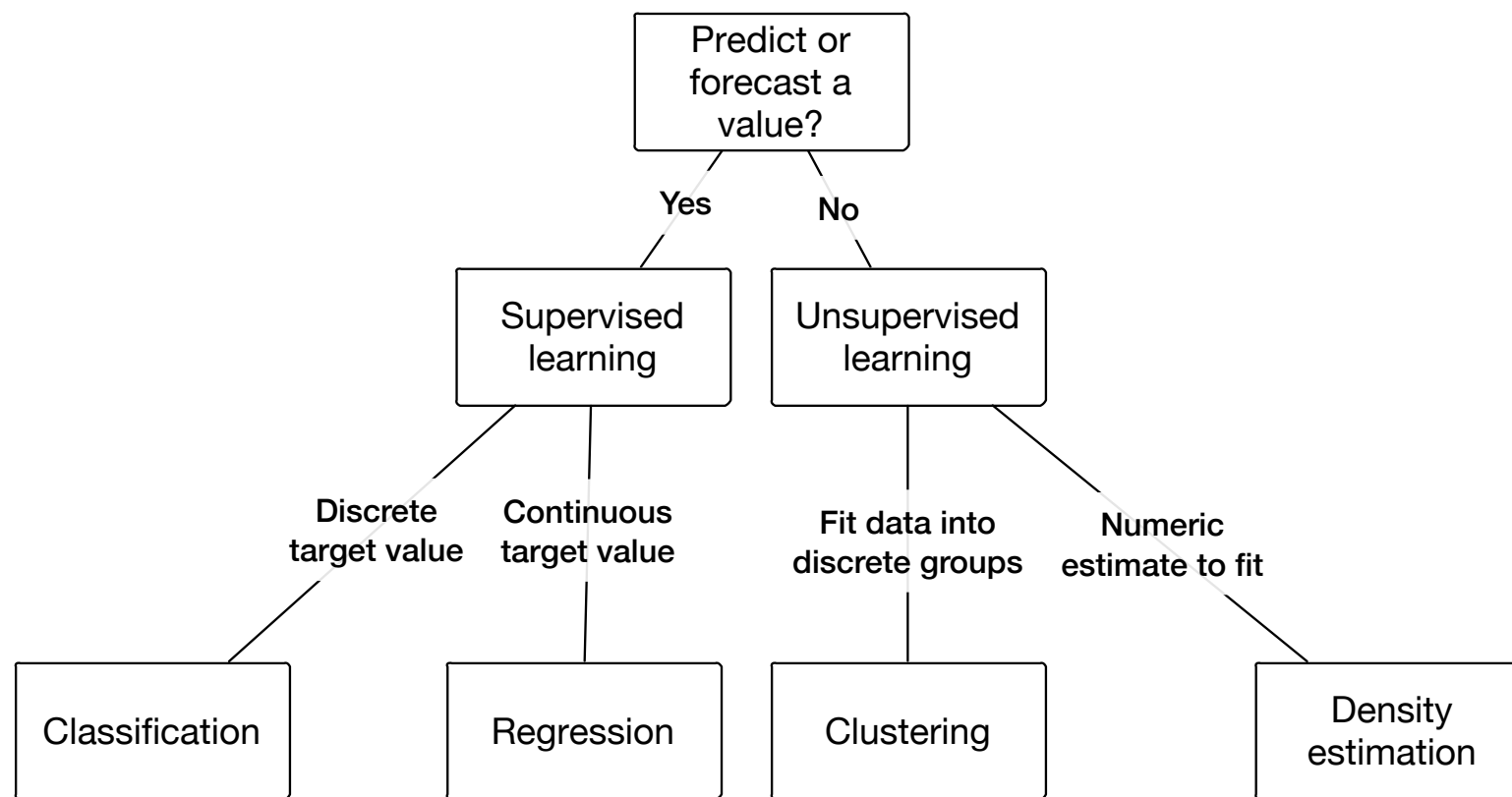
Classification



Clustering

# Outline of core ML problems

## Machine Learning Outline





# Summary

- R: open-source, free, very powerful
- Excellent choice for quick analysis and visualization
- Most popular interfaces to use: R Studio, Jupyter notebook

# Temat e javës së shtatë

- MySQL
- Getting Started with MySQL
- Creating and Inserting Records
- Retrieving Records
- Searching in MySQL
- Accessing MySQL with Python
- Accessing MySQL with R
- Introduction to Other Popular Databases



# Pyetje???