Linear Discriminant Analysis in R

Spoken Tutorial Project
National Mission on Education through ICT

Yate Asseke Ronald

Debatosh Chakraborty

FOSSEE, IIT Bombay

16 February 2024

fossee
better
education

Learning Objectives

We will learn about:

Learning Objectives

We will learn about:

- **Linear Discriminant Analysis (LDA) and its implementation**

Learning Objectives

We will learn about:

- **Linear Discriminant Analysis (LDA) and its implementation**
- **Assumptions of LDA**

- **Limitations of LDA**

Learning Objectives

- **Limitations of LDA**
- **LDA on a subset of Raisin dataset**

Learning Objectives

- **Limitations of LDA**
- **LDA on a subset of Raisin dataset**
- **Visualization of the LDA separator and its corresponding confusion matrix**

System Specifications

- **Windows 11**

System Specifications

- **Windows 11**
- **R v 4.3.0**

System Specifications

- **Windows 11**
- **R v 4.3.0**
- **RStudio v 2023.06.1**

System Specifications

- **Windows 11**
- **R v 4.3.0**
- **RStudio v 2023.06.1**

System Specifications

- **Windows 11**
- **R v 4.3.0**
- **RStudio v 2023.06.1**

It is recommended to install R version 4.2.0 or higher

To follow this tutorial, the learner should know:

Pre-requisites

To follow this tutorial, the learner should know:

- **Basics of R Programming**

Pre-requisites

To follow this tutorial, the learner should know:

- **Basics of R Programming**
- **Basics of Machine Learning using R**

Pre-requisites

To follow this tutorial, the learner should know:

- **Basics of R Programming**
- **Basics of Machine Learning using R**
- **If not, please access the relevant tutorials on R on this website https://spoken-tutorial.org**

Linear Discriminant Analysis

Linear Discriminant Analysis

LDA is a statistical method

Linear Discriminant Analysis

LDA is a statistical method

- **It is used for classification**

Linear Discriminant Analysis

LDA is a statistical method

- **It is used for classification**
- **It constructs a data driven line that best separates different classes**

Linear Discriminant Analysis

LDA is a statistical method

- **It is used for classification**
- **It constructs a data driven line that best separates different classes**
- **It is based on maximization of likelihood function to classify two or more classes**

Applications of LDA

- **LDA technique is used in several applications like:**

- **LDA technique is used in several applications like:**
- **Fraud Detection**

Applications of LDA

- **LDA technique is used in several applications like:**
- **Fraud Detection**
- **Bio-Imaging classification**

Applications of LDA

- **LDA technique is used in several applications like:**
- **Fraud Detection**
- **Bio-Imaging classification**
- **Classify patient disease state**

Assumptions for LDA

- **All data entries are continuous, Gaussian, with equal covariance matrix for all the classes**

Assumptions for LDA

- **All data entries are continuous, Gaussian, with equal covariance matrix for all the classes**

- **Mean vectors for each class are different**

Assumptions for LDA

- **All data entries are continuous, Gaussian, with equal covariance matrix for all the classes**

- **Mean vectors for each class are different**

- **Data records are independent and identically distributed among each class**

Limitations of LDA

Limitations of LDA

- **Departure from Gaussianity can increase misclassification probability in LDA**

Limitations of LDA

- **Departure from Gaussianity can increase misclassification probability in LDA**
- **LDA may perform poorly if data has an unequal class covariance matrix**

Implementation of LDA

- **Now let us implement LDA on the raisin dataset with two chosen variables**

Implementation of LDA

- **Now let us implement LDA on the raisin dataset with two chosen variables**

- **More information on raisin data is available in the Additional Reading material on this tutorial page**

We will use:

We will use:

- **A script file LDA.R**

Download Files

We will use:

- **A script file LDA.R**
- **Please download this file from the Code files link of this tutorial**

We will use:

- **A script file LDA.R**
- **Please download this file from the Code files link of this tutorial**
- **Make a copy and then use it for practising**

Summary

In this tutorial we have learnt:

- **Linear Discriminant Analysis (LDA) and its implementation**
- **Assumptions of LDA**

Summary

- **Limitations of LDA**
- **LDA on a subset of Raisin dataset**
- **Visualization of the LDA separator and its corresponding confusion matrix**

- **Perform LDA on inbuilt** `PlantGrowth` **dataset**
- **Evaluate the model using a confusion matrix and visualise the results**

About the Spoken Tutorial Project

- **Watch the video available at**
  https://spoken-tutorial.org
  /What_is_a_Spoken_Tutorial
- **It summarises the Spoken Tutorial project**

About the Spoken Tutorial Project

- **Watch the video available at https://spoken-tutorial.org /What_is_a_Spoken_Tutorial**

- **It summarises the Spoken Tutorial project**

- **If you do not have good bandwidth, you can download and watch it**

Spoken Tutorial Workshops

The Spoken Tutorial Project Team

- **Conducts workshops using spoken tutorials**
- **Gives certificates to those who pass an online test**
- **For more details, please write to contact@spoken-tutorial.org**

Answers for THIS Spoken Tutorial

- **Questions in THIS Spoken Tutorial?**

- **Visit**
  https://forums.spoken-tutorial.org

- **Choose the minute and second where you have the question**

- **Explain your question briefly**

- **The Spoken Tutorial project will ensure an answer**

- **You will have to register to ask questions**

Forum to answer questions

- **Do you have any general/technical questions?**

- **Please visit the forum given in the link**

  https://forums.fossee.in/

Textbook Companion Project

- **The FOSSEE team coordinates the coding of solved examples of popular books and case study projects**

- **We give certificates to those who do this**

- **For more details, please visit these sites**

  https://r.fossee.in/

  https://fossee.in/

Textbook Companion Project

- **The FOSSEE team coordinates the coding of solved examples of popular books and case study projects**

- **We give certificates to those who do this**

- **For more details, please visit these sites**

  https://r.fossee.in/

  https://fossee.in/

Acknowledgements

- **Spoken Tutorial project was established by the Ministry of Education(MoE), Govt of India**

Thank You

- **This tutorial is contributed by Yate Asseke Ronald and Debatosh Chakraborty from IIT Bombay**
- **Thank you for joining**