

Logistic Regression using R

Spoken Tutorial Project

<https://spoken-tutorial.org>

National Mission on Education through ICT

<https://sakshat.ac.in>

Tanmay Srinath

Madhuri Ganapathi

IIT Bombay

8 March 2022



Learning Objectives



Learning Objectives

We will learn about:



Learning Objectives

We will learn about:

► **Logistic Regression**



Learning Objectives

We will learn about:

- ▶ **Logistic Regression**
- ▶ **Applications of Logistic Regression**



Learning Objectives

We will learn about:

- ▶ **Logistic Regression**
- ▶ **Applications of Logistic Regression**
- ▶ **Implementation of Logistic Regression in R**



Learning Objectives

We will learn about:

- ▶ **Logistic Regression**
- ▶ **Applications of Logistic Regression**
- ▶ **Implementation of Logistic Regression in R**
- ▶ **Drawbacks of Logistic Regression**



System Specifications



System Specifications

► **Ubuntu Linux OS version 20.04**



System Specifications

- ▶ **Ubuntu Linux OS version 20.04**
- ▶ **R version 4.1.2**



System Specifications

- ▶ **Ubuntu Linux OS version 20.04**
- ▶ **R version 4.1.2**
- ▶ **RStudio version 1.4.1717**



System Specifications

- ▶ **Ubuntu Linux OS version 20.04**
- ▶ **R version 4.1.2**
- ▶ **RStudio version 1.4.1717**



System Specifications

- ▶ **Ubuntu Linux OS version 20.04**
 - ▶ **R version 4.1.2**
 - ▶ **RStudio version 1.4.1717**
- R version 4.1.0 or higher**



Pre-requisites



Pre-requisites

► Basics of R Programming



Pre-requisites

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



Pre-requisites

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



Pre-requisites

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

If not, please access the relevant tutorials on

<https://spoken-tutorial.org/>



Logistic Regression



Logistic Regression

- It is a popular statistical model



Logistic Regression

- ▶ It is a popular statistical model
- ▶ It is used for regression of binary responses and classification tasks



Logistic Regression

- ▶ It is a popular statistical model
- ▶ It is used for regression of binary responses and classification tasks
- ▶ It is used to predict a binary outcome based on a set of independent variables



Applications of Logistic Regression



Applications of Logistic Regression

- **Predict if a patient has diabetes or not**

```
library(faraway)  
data(pima)
```



Applications of Logistic Regression

- Classify various plant species

<https://archive.ics.uci.edu/ml/datasets/One-hundred+plant+species+leaves+data+set>



Applications of Logistic Regression

- Detect fraudulent credit card transactions

<https://www.kaggle.com/mlg-ulb/creditcardfraud>



Applications of Logistic Regression

We will implement **logistic regression** on the **iris** dataset



Download Files



Download Files

We will use:



Download Files

We will use:

- ▶ **A script file LogisticRegression.R**



Download Files

We will use:

- ▶ A script file **LogisticRegression.R**



Download Files

We will use:

- ▶ A script file **LogisticRegression.R**

Download this file from the **Code files** link of this tutorial

Make a copy and then use it for practising



Drawbacks of Logistic Regression



Drawbacks of Logistic Regression

- ▶ **Logistic regression works well when the response variables are independent**



Drawbacks of Logistic Regression

- ▶ Logistic regression works well when the response variables are independent
- ▶ Logistic regression doesn't perform well when the independent variables are non-linearly related

Summary

We have learnt about:

- ▶ **Logistic Regression**
- ▶ **Applications of Logistic Regression**
- ▶ **Implementation of Logistic Regression in R**
- ▶ **Drawbacks of Logistic Regression**



Assignment



Assignment

- ▶ **Apply logistic regression on the Wine dataset**



Assignment

- ▶ **Apply logistic regression on the Wine dataset**
- ▶ **This dataset can be found in the HDclassif package**



Assignment

- Install the package and import the dataset using the `data()` command



Assignment

- ▶ Install the package and import the dataset using the `data()` command
- ▶ Measure its accuracy by tabulating 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
- ▶ 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 FOSSEE project will ensure an answer

You will have to register to ask questions



Forum to answer questions

- ▶ Questions not related to the Spoken Tutorial?
- ▶ Do you have general/technical questions on the Software?
- ▶ Please visit the FOSSEE Forum
<https://forums.fossee.in/>
- ▶ Choose the Software and post your question



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

- ▶ **The Spoken Tutorial and FOSSEE projects are funded by the Ministry of Education, Govt. of India**



About the Contributors

- ▶ **This tutorial is contributed by Tanmay Srinath and Madhuri Ganapathi, IIT Bombay**

