

Bagging in R

Spoken Tutorial Project

<https://spoken-tutorial.org>

National Mission on Education through ICT

Debatosh Chakraborty

Yate A Ronald

FOSSEE, IIT Bombay

12 July 2024



Learning Objectives



Learning Objectives

We will learn about:



Learning Objectives

We will learn about:

- **Bagging**



Learning Objectives

We will learn about:

- Bagging
- Assumptions for Bagging



Learning Objectives

We will learn about:

- **Bagging**
- **Assumptions for Bagging**
- **Advantages of Bagging**



Learning Objectives

- **Implementation of Bagging using Decision Tree in R**



Learning Objectives

- Implementation of Bagging using Decision Tree in R
- Model Evaluation



Learning Objectives

- Implementation of Bagging using Decision Tree in R
- Model Evaluation
- Limitations of Bagging



System Specifications



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



Pre-requisites



Pre-requisites

To follow this tutorial, learner should know:



Pre-requisites

To follow this tutorial, learner should know:

- **Basic programming in R**



Pre-requisites

To follow this tutorial, learner should know:

- Basic programming in R
- Basics of Machine Learning



Pre-requisites

To follow this tutorial, learner should know:

- Basic programming in R
- Basics of Machine Learning
- If not, please access the relevant tutorials on this website
<https://spoken-tutorial.org>



Bootstrap aggregation (Bagging)



Bootstrap aggregation (Bagging)

- Any classification model fitted across several training data subsets is desired to have consistent decision boundaries



Bootstrap aggregation (Bagging)

- Any classification model fitted across several training data subsets is desired to have consistent decision boundaries
- Large variation in the decision boundaries indicate higher variability of the classification model



Bootstrap aggregation (Bagging)

- Bagging is a commonly used ensemble technique to reduce this variation



Bootstrap aggregation (Bagging)

- Bagging is a commonly used ensemble technique to reduce this variation
- In Bagging, random subsets of the training data are repeatedly chosen to construct multiple classifiers



Bootstrap aggregation (Bagging)

- The Bootstrap classifiers constructed from the chosen subsets are then aggregated



Bootstrap aggregation (Bagging)

- The Bootstrap classifiers constructed from the chosen subsets are then aggregated
- For bagging of the decision tree classifier, the aggregation is done by a majority vote of the class predicted by Bootstrap trees



Assumptions for Bagging

- Each observation is independent
- The assumption of the chosen classifier is satisfied



Advantages of Bagging



Advantages of Bagging

- Bagging reduces the variation of the chosen model



Advantages of Bagging

- Bagging reduces the variation of the chosen model
- Bagging improves the performance (accuracy) of the decision tree classifier in general



Implementation of Bagging



Implementation of Bagging

- We will perform the Bagging of Decision Tree Classifier on the Raisin dataset with two chosen variables



Download Files

We will use:



Download Files

We will use:

- A script file **Bagging-Decision-Tree.R**



Download Files

We will use:

- A script file **Bagging-Decision-Tree.R**
- **Raisin Dataset 'raisin.xlsx'**



Download Files

We will use:

- A script file **Bagging-Decision-Tree.R**
- Raisin Dataset 'raisin.xlsx'



Download Files

We will use:

- A script file **Bagging-Decision-Tree.R**
- Raisin Dataset 'raisin.xlsx'

Please download these files from the **Code files** link of this tutorial

Make a copy and then use them while practicing



Limitations of Bagging

- Bagging is hard to interpret
- Requires more computational time
- Bagging doesn't improve model bias



Summary

In this tutorial we have learnt about:

- **Bagging**
- **Assumptions for Bagging**
- **Advantages of Bagging**



Summary

- **Implementation of Bagging using Decision Tree in R**
- **Model Evaluation**
- **Limitations of Bagging**



Assignment



Assignment

- Apply Bagging using Decision Tree on PimaIndiansDiabetes dataset
- Install the `pdp` package and import the dataset using the `data(pima)` command
- Visualize the decision boundary and measure the accuracy of the model



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



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/>



Acknowledgements

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



Thank You

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

