User Guide

Introduction

This is a simple user guide for the project. It will help you to understand the project and how to use it.

Installation

To run the project, you need to install the following R dependencies as well as the R language itself.

```{r}
install.packages("ggplot2")
install.packages("gridExtra")
install.packages("emojifont")
install.packages("igraph")
install.packages("openxlsx2")
install.packages("yaml")
```

Usage

The project is composed of multiple scripts. The main script is main.R. This script is responsible for running the entire project. It will generate the plots and prepare the environment for the other scripts.

The other scripts are used for specific tasks.

- excel.R: This script is responsible for generating the excel file with the results.
- hierarchy.R: This script is responsible for managing the decision tree.
- results.R: This script is responsible for storing the results in custom data structures.
- weights.R: This script is responsible for managing the different ways to generate the weights.

Configuration

The project has a configuration file called config.yaml inside the config folder. This file is used to configure the project. You can change the parameters in this file to customize the project.

In this file, you can change the following parameters:

- **Distribution**: Select the distribution used to generate the weights, it can be uniform, normal or exponential. The uniform distribution can be sampled using $n - 1$ or n degrees of freedom.

The normal distribution can be customized with the mean μ and standard deviation σ. The exponential distribution can be customized with the rate parameter λ.

- **Filtering**: Select the filtering method used to mark a weight vector as invalid. The filtering is optional and can be disabled by setting this parameter to NULL. Other value that can be used is a number between 0 and 1. This number represents the threshold for the difference between the largest and the second largest weight. Lastly, there are two filtering methods from the literature that can be used: weak and strong, these methods state that all the weights must be satisfy:

$$\text{weak: } 0.05/n \leq w_i \leq 0.7 + 0.3/n$$
$$\text{strong:} 0.1/n \leq w_i \leq 0.6 + 0.25/n$$

- **Simulation parameters**: This section groups parameters to adjust the simulation. You can change the number of instances to run, the number of alternatives, the maximum number of branches (the simulation will run from 2 to this number) per node and the maximum number of levels (also from 2 to this number).

- **Hit ratio instances**: This parameter allows to check for the first k values of the ranking, i.e., if the first k values of the ranking given by the method match the first k values of the ranking given by the true weights the hit ratio takes value 1 and 0 otherwise.

Running the project

To run the project, you need to execute the main.R script. This script will run the entire project.

Results

The results of the project are stored in the output folder. The results are stored in an excel file called results.xlsx. This file contains the results of the simulation in multiple sheets, each sheet contains the results of a distribution.
Also the plots generated by the project are stored in the charts folder.