



Universidade Paulista – UNIP
MBA em Data Science & Machine Learning

Meta Learning - Bayesian Networks

Disciplina: Linguagens de Programação Estatística (R)

Trabalho 02

**UNIP
2018
Prof. Robson Fernandes**

Conjunto de Dados

- Considerar o arquivo **food-sp.csv** disponível na pasta “Trabalho/Trabalho 02”

Objetivo

- Considerar o arquivo **bn_food.r** disponível na pasta “Trabalho/Trabalho 02” e criar um componente baseado no paradigma orientado a objetos utilizando o **R6**.

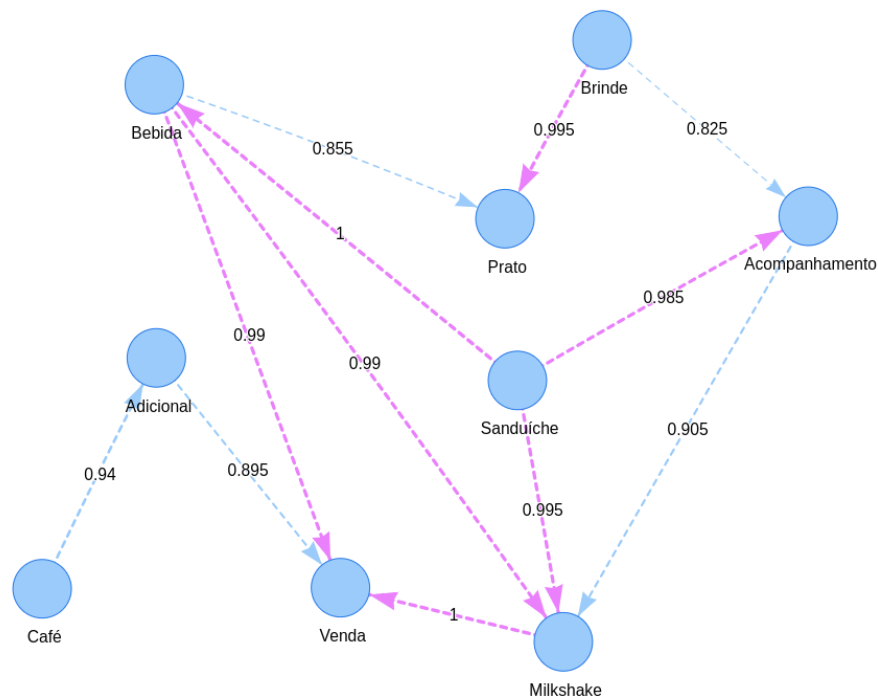
| BayesianNetwork |
|---|
| - dataSet: List |
| - structureBayesianNetwork: Object |
| - fitBayesianNetwork: Object |
| + learnStructure(dataSet) |
| + fit(structureBayesianNetwork, dataSet) |
| + plotBayesianNetwork(structureBayesianNetwork) |
| + plotDistribution(variable) |
| + inference(fitBayesianNetwork, event, evidence) : Double |

Diagrama de Classe : Componente Bayesian Network

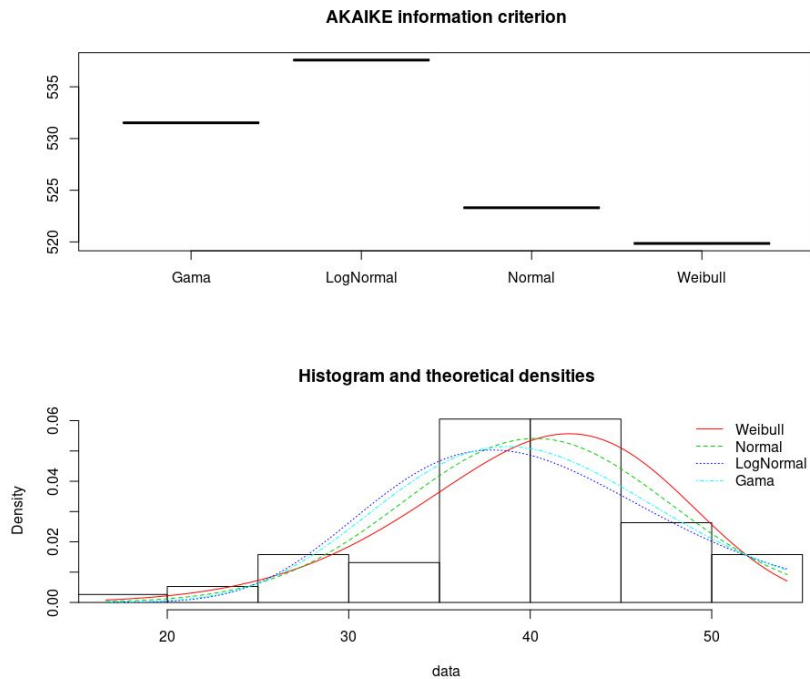
- **Regras**

- o atributo **dataSet** deverá armazenar o *data.frame* do arquivo **food-sp.csv**.
- o atributo **structureBayesianNetwork** deverá armazenar uma rede Bayesiana gerada por meio da função *averaged.network* com base na função *boot.strength baseada* em meta aprendizado com a adoção do algoritmo de aprendizado de de estrutura *Hill Climbing* para encontrar a melhor rede para o problema proposto. Dica: Verificar o arquivo “**Interactive Bayesian Network Strength Viewer.pdf**” na pasta “**Redes Bayesianas**”
- o atributo **fitBayesianNetwork** deverá armazenar a rede Bayesiana ajustada.

- o método **learnStructure** receberá como argumento o **dataSet** e deverá realizar o processo de aprendizado de estrutura da rede para posteriormente armazená-la no atributo **structureBayesianNetwork**.
- o método **fit** receberá como argumento o **structureBayesianNetwork** e **dataSet** e irá ajustar a rede Bayesiana conforme a distribuição de probabilidade do conjunto de dados para posteriormente armazená-la no atributo **fitBayesianNetwork**.
- o método **plotBayesianNetwork** deverá apresentar a rede Bayesiana conforme a relação de forças entre os nós. Ex. rede Bayesiana, representada pela figura abaixo:



o método **plotDistribution** receberá como argumento a variável **variable** e deverá apresentar a distribuição de probabilidade estimada e o critério de AKAIKE para a variável analisada. As variáveis que deverão ser analisadas são: 'GELADO.E.MILK.SHAKE','SANDUICHE','BEBIDA','ACOMPANHAMENTO','ADICIONAL','PROJETO.CAFE', 'BRINDE', 'PRATO','VENDA'. Ex. Análise de Distribuição de Probabilidade estimada, representada pela figura abaixo:



- o o método **inference** receberá como argumento o **fitBayesianNetwork** e **event**, e **evidence**. Sendo, **fitBayesianNetwork** a rede ajustada, **event** (posteriori) a hipótese que se deseja inferir, e **evidence** (priori), o conjunto de evidências que se deseja compor para inferir a hipótese.
- o Exemplo de probabilidade condicional (inferência):

$$P(H \mid E_1, E_2, E_3, E_4, E_5, E_n)$$

Data de Entrega

- Entrega até **24/11/2018 às 23:59** (Código Fonte)
- E-mail robson.fernandes@usp.br / robs.fernandes@outlook.com
- Obs: O Trabalho 01 poderá ser entregue juntamente com Trabalho 02.