Artificial Intelligence Laboratory 3: Bayesian Network

DT8042, Halmstad University

December 1, 2021

Introduction

This lab is designed to introduce you the use of Bayesian Network and Naïve Bayes algorithm. By the end of this lab session you will:

- Learn how to construct Bayesian Network from data.
- Learn how to implement Naïve Bayes algorithm for classification tasks.
- Learn how to use Bayesian Network and Naïve Bayes for inference.

Environment setup

In this lab, we work with the bnlearn library. Installation instructions of bnlearn and a short description for adding Conda environment to your jupyter notebook is available.

Dataset

In this lab, you will work with an artificial smart grid dataset to construct Bayesian Networks. In the **Lab3.zip**, you can find the records of failures in smart grids with the specific conditions in which the failures have happened: an artificial failure dataset (**smart_grid.xlsx**), containing 700 observations. The attributes in this dataset are: Number_of_Customers, Time, Day, Season, Weather, Demand_Factor, Overload, and Outage_Duration, with values listed in the following. Part of this dataset is shown in Figure 1.

- **Season**: Spring, Summer, Autumn, Winter
- **Outage_Duration**: Less_than_1H, More_than_1H
- **Number_of_Customers**: Low, High

- **Overload**: Yes, No
- **Weather**: Cold, Warm
- **Time**: Morning, Afternoon, Evening, Night
- **Demand_Factor**: Low, Medium, High
- **Day**: Weekdays, Weekend

Season	Outage_Duration	Number_of_Customers	Overload	Weather	Time	Demand_Factor	Day
Autumn	Less_than_1H	Low	Yes	Cold	Morning	Low	Weekdays
Winter	Less_than_1H	Low	No	Cold	Evening	Low	Weekdays
Spring	More_than_1H	Low	No	Cold	Evening	Low	Weekdays
Winter	Less_than_1H	High	No	Warm	Morning	Low	Weekdays
Spring	More_than_1H	Low	No	Cold	Morning	Low	Weekend
Winter	More_than_1H	Low	No	Cold	Morning	Medium	Weekdays
Autumn	More_than_1H	Low	No	Warm	Evening	Low	Weekdays
Spring	More_than_1H	Low	No	Cold	Evening	High	Weekend
Summer	Less_than_1H	Low	Yes	Warm	Evening	High	Weekend
Winter	More_than_1H	High	No	Cold	Night	Low	Weekdays
Autumn	Less_than_1H	Low	Yes	Cold	Night	Low	Weekend
Spring	Less_than_1H	High	Yes	Cold	Afternoon	High	Weekend
Summer	Less_than_1H	High	Yes	Warm	Afternoon	Medium	Weekend
Autumn	Less_than_1H	High	Yes	Warm	Evening	High	Weekend
Autumn	More_than_1H	Low	No	Cold	Afternoon	Medium	Weekdays
Autumn	More_than_1H	Low	No	Cold	Morning	Low	Weekend
Summer	Less_than_1H	Low	No	Cold	Morning	Medium	Weekend

Figure 1: An example of dataset smart_grid.xlsx

Task 1

Read the code in **Lab3_task1.ipynb**. Learn a Bayesian Network of the smart grid data set. Construct the network and compute the following:

1) p(Outage duration=$Otg \leq 1$ | Time = Morning, Demand Factor=Medium)

2) p(Demand Factor=High | Overload=Yes, Weather=Cold)

3) p(Number of Customers=Low | Demand Factor=High)

Task 2

Implement Naïve Bayes algorithm in **Lab3_task2.ipynb**, performing classification by making inference:

a. Compute the Likelihood table of having pet(s), for each categorical feature.

b. Compute the posterior probabilities.

c. Compute the Likelihood of having pet for numerical features.

d. Making inference with Naïve Bayes.

e. Implementing a Naïve Bayes Classifier and perform classification on the Iris dataset.

Grading Criteria

- In order to pass this lab, you need to complete task 1 and 2a-2d.
- Extra credits: complete task 2e.
- Your submission should include **code** and a **report** of what you have done, observed and learned.