Fundamentals of Artificial Intelligence
and
Knowledge Representation

Covid-19
Bayesian Network

Marini Luca
University of Bologna
Artificial Intelligence Master Degree course

Contents

1 Problem Description **1**

2 Conditional Probability Tables (CPT) **3**
2.1 CPT - Loss_Taste . . . 3
2.2 CPT - Loss_Smell . . . 3
2.3 CPT - Fever . . . 3
2.4 CPT - Persistent_Cough . . . 3
2.5 CPT - Over_Eighty . . . 3
2.6 CPT - Vaccinated . . . 4
2.7 CPT - Live_With_Other_People . . . 4
2.8 CPT - Potential_Symptoms . . . 4
2.9 CPT - Positive_To_Covid . . . 7

3 Independence **9**

4 Inference **10**
4.1 Exact Inference . . . 10
4.1.1 Probability of being positive to Covid-19 . . . 10
4.1.2 Probability of being positive to Covid-19 having potential symptoms . . . 10
4.1.3 Probability of being positive to Covid-19 being vaccinated . . . 11
4.1.4 Maximum A-posteriori Probability . . . 11
4.2 Approximate Inference . . . 12

References **14**

1 Problem Description

Since 2020 Covid-19 spread all around the World. Many computer scientists, supported by medical experts, have developed online platforms in which it is possible to assess the probability of being positive based on multiple factors.

This project aims to recreate a simple online Covid-19 test that outputs the probability of being positive by leveraging a Bayesian Network. This Network considers users that live in Italy because some assumptions (the number of persons over 80 years old and the number of vaccinated inhabitants) were made considering Italian data.

Please note that the probabilities that this system outputs should **not** be considered as a medical diagnosis. As previously stated, this project aims to simulate an online Covid-19 test. Therefore, it is just an experiment, and its aim is **not** to provide reliable data.

The Bayesian Network is composed of nine boolean variables:

- **Persistent_Cough**: the user has a persistent cough.
- **Loss_Taste**: the user has lost taste.
- **Loss_Smell**: the user has lost smell.
- **Fever_Above_38**: the user has a fever higher than 38°C.
- **Over_Eighty**: the user is over 80 years old.
- **Vaccinated**: the user is vaccinated.
- **Live_With_Other_People**: the user lives with other people in the same house.
- **Potential_Symptoms**: this variable is conditioned on Persistent_Cough, Loss_Taste, Loss_Smell, and Fever. It tells if the user has potential symptoms or not. It is introduced to group all the symptoms together and not overwhelm the other variables that affect Positive_To_Covid (Vaccinated and Live_With_Other_People).
- **Positive_To_Covid**: this variable is used to model the positiveness of the user to the virus. It is directly conditioned on Vaccinated, Live_With_Other _People, and Potential_Symptoms variables.

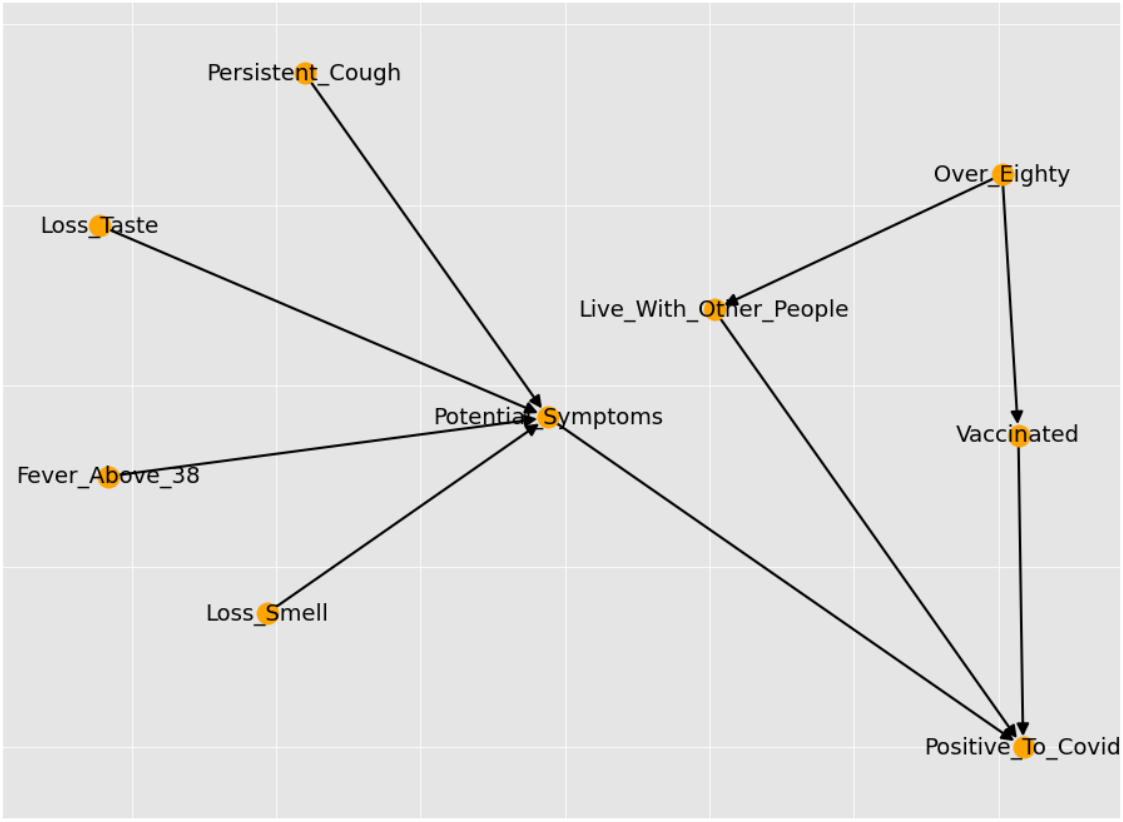

Figure 1: Graphical view of the Bayesian network

1. Persistent_Cough, Loss_Taste, Fever_Above_38, and Loss_Smell will determine the outcome of Potential_Symptoms.

2. Being over eighty years old conditions the probability of living with other people in the same house.

3. The variable Over_Eighty also conditions the probability of being vaccinated. Indeed, this age group was given priority to get the vaccine.

4. The positiveness to Covid-19 is conditioned by the presence of three variables that models respectively:

    - The presence of potential symptoms;
    - The fact that a person lives with other people in the same house;
    - And the fact that a person has received or not the vaccine.

2 Conditional Probability Tables (CPT)

In this section are reported the Conditional Probability Tables for each node in the Bayesian Network.

Please note that most of the following data is not accurate, while some statistics are taken from the *Italian National Institute of Statistics* (ISTAT).

2.1 CPT - Loss_Taste

Loss_Taste(No)	0.9
Loss_Taste(Yes)	0.1

2.2 CPT - Loss_Smell

Loss_Smell(No)	0.85
Loss_Smell(Yes)	0.15

2.3 CPT - Fever

Fever_Above_38(No)	0.82
Fever_Above_38(Yes)	0.18

2.4 CPT - Persistent_Cough

Persistent_Cough(No)	0.75
Persistent_Cough(Yes)	0.25

2.5 CPT - Over_Eighty

Over_Eighty(No)	0.93
Over_Eighty(Yes)	0.07

In March 2021, 4.442.048 people in Italy have more than 80 years old [3]. The total number of Italian inhabitants in 2021 is 59.641.488, as visible on the ISTAT website [2].

2.6 CPT - Vaccinated

Over_Eighty	Over_Eighty(No)	Over_Eighty(Yes)
Vaccinated(No)	0.92	0.57
Vaccinated(Yes)	0.08	0.43

The percentage of over 80 years old Italian people that are vaccinated is 42.6% (1.893.746 out of 4.442.048), while just 8.35% of the whole Italian population has been vaccinated [3].

2.7 CPT - Live_With_Other_People

Over_Eighty	Over_Eighty(No)	Over_Eighty(Yes)
Live_With_Other_People(No)	0.35	0.55
Live_With_Other_People(Yes)	0.65	0.45

In Italy, 44.5% of the elderly do not live alone [1].

2.8 CPT - Potential_Symptoms

Loss_Taste	Loss_Taste(No)	Loss_Taste(No)
Persistent_Cough	Persistent_Cough(No)	Persistent_Cough(No)
Fever_Above_38	Fever_Above_38(No)	Fever_Above_38(No)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.98	0.33
Potential_Symptoms(Yes)	0.02	0.67

Loss_Taste	Loss_Taste(No)	Loss_Taste(No)
Persistent_Cough	Persistent_Cough(No)	Persistent_Cough(No)
Fever_Above_38	Fever_Above_38(Yes)	Fever_Above_38(Yes)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.6	0.1
Potential_Symptoms(Yes)	0.4	0.9

Loss_Taste	Loss_Taste(No)	Loss_Taste(No)
Persistent_Cough	Persistent_Cough(Yes)	Persistent_Cough(Yes)
Fever_Above_38	Fever_Above_38(No)	Fever_Above_38(No)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.8	0.28
Potential_Symptoms(Yes)	0.2	0.72

Loss_Taste	Loss_Taste(No)	Loss_Taste(No)
Persistent_Cough	Persistent_Cough(Yes)	Persistent_Cough(Yes)
Fever_Above_38	Fever_Above_38(Yes)	Fever_Above_38(Yes)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.4	0.06
Potential_Symptoms(Yes)	0.6	0.94

Loss_Taste	Loss_Taste(Yes)	Loss_Taste(Yes)
Persistent_Cough	Persistent_Cough(No)	Persistent_Cough(No)
Fever_Above_38	Fever_Above_38(No)	Fever_Above_38(No)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.31	0.13
Potential_Symptoms(Yes)	0.69	0.87

Loss_Taste	Loss_Taste(Yes)	Loss_Taste(Yes)
Persistent_Cough	Persistent_Cough(No)	Persistent_Cough(No)
Fever_Above_38	Fever_Above_38(Yes)	Fever_Above_38(Yes)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.09	0.04
Potential_Symptoms(Yes)	0.91	0.96

Loss_Taste	Loss_Taste(Yes)	Loss_Taste(Yes)
Persistent_Cough	Persistent_Cough(Yes)	Persistent_Cough(Yes)
Fever_Above_38	Fever_Above_38(No)	Fever_Above_38(No)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.26	0.1
Potential_Symptoms(Yes)	0.74	0.9

Loss_Taste	Loss_Taste(Yes)	Loss_Taste(Yes)
Persistent_Cough	Persistent_Cough(Yes)	Persistent_Cough(Yes)
Fever_Above_38	Fever_Above_38(Yes)	Fever_Above_38(Yes)
Loss_Smell	Loss_Smell(No)	Loss_Smell(Yes)
Potential_Symptoms(No)	0.04	0.02
Potential_Symptoms(Yes)	0.96	0.98

2.9 CPT - Positive_To_Covid

Live_With_Other_People	Live_W_Other_People(No)	Live_W_Other_People(No)
Potential_Symptoms	Potential_Symptoms(No)	Potential_Symptoms(No)
Vaccinated	Vaccinated(No)	Vaccinated(Yes)
Positive_To_Covid(No)	0.85	0.98
Positive_To_Covid(Yes)	0.15	0.02

Live_With_Other_People	Live_W_Other_People(No)	Live_W_Other_People(No)
Potential_Symptoms	Potential_Symptoms(Yes)	Potential_Symptoms(Yes)
Vaccinated	Vaccinated(No)	Vaccinated(Yes)
Positive_To_Covid(No)	0.25	0.75
Positive_To_Covid(Yes)	0.75	0.25

Live_With_Other_People	Live_W_Other_People(Yes)	Live_W_Other_People(Yes)
Potential_Symptoms	Potential_Symptoms(No)	Potential_Symptoms(No)
Vaccinated	Vaccinated(No)	Vaccinated(Yes)
Positive_To_Covid(No)	0.7	0.92
Positive_To_Covid(Yes)	0.3	0.08

Live_With_Other_People	Live_W_Other_People(Yes)	Live_W_Other_People(Yes)
Potential_Symptoms	Potential_Symptoms(Yes)	Potential_Symptoms(Yes)
Vaccinated	Vaccinated(No)	Vaccinated(Yes)
Positive_To_Covid(No)	0.05	0.64
Positive_To_Covid(Yes)	0.95	0.36

3 Independence

In this section are shown the local independence of the three variables Vaccinated, Potential_Symptoms, and Positive_To_Covid.

```
(Vaccinated _|_ Live_With_Other_People, Potential_Symptoms, Fever_Above_38,
Loss_Smell, Loss_Taste, Persistent_Cough | Over_Eighty)
```

```
(Potential_Symptoms _|_ Over_Eighty, Live_With_Other_People, Vaccinated |
Loss_Taste, Loss_Smell, Fever_Above_38, Persistent_Cough)
```

```
(Positive_To_Covid _|_ Over_Eighty, Fever_Above_38, Loss_Smell, Loss_Taste,
Persistent_Cough | Vaccinated, Live_With_Other_People, Potential_Symptoms)
```

4 Inference

It is possible to infer the probabilities of specific nodes of the Bayesian Network given other variables' values. Moreover, there are two main Inference methods:

- Exact Inference;
- Approximate Inference.

4.1 Exact Inference

The method used to perform Exact Inference is Variable Elimination.

4.1.1 Probability of being positive to Covid-19

The following table shows the probability of being positive to Covid-19 without any given assumption.

```
P(Positive_To_Covid)
```

```
+------------------------+--------------------------+
| Positive_To_Covid      |   phi(Positive_To_Covid) |
+========================+==========================+
| Positive_To_Covid(No)  |                   0.6095 |
+------------------------+--------------------------+
| Positive_To_Covid(Yes) |                   0.3905 |
+------------------------+--------------------------+
```

The probability of being positive without any given assumption (39%) is lower than being negative (61%).

4.1.2 Probability of being positive to Covid-19 having potential symptoms

```
P( Positive_To_Covid | Potential_Symptoms = Yes )
```

```
+------------------------+--------------------------+
| Positive_To_Covid      |   phi(Positive_To_Covid) |
+========================+==========================+
| Positive_To_Covid(No)  |                   0.1806 |
+------------------------+--------------------------+
| Positive_To_Covid(Yes) |                   0.8194 |
+------------------------+--------------------------+
```

The probability of being positive to Covid-19, given the positiveness of variable Potential_Symptoms, is very high (81%). Indeed, the variable Potential_Symptoms is conditioned by four different intense symptoms frequent in Covid-19 patients. For example, the single loss of taste has a strong influence on the positive value of Potential_Symptoms. The only variable that affects Potential_Symptoms in a small quantity is Persistent_Cough.

4.1.3 Probability of being positive to Covid-19 being vaccinated

```
P( Positive_To_Covid | Vaccinated = Yes )
```

```
+------------------------+--------------------------+
| Positive_To_Covid      |   phi(Positive_To_Covid) |
+========================+==========================+
| Positive_To_Covid(No)  |                   0.8726 |
+------------------------+--------------------------+
| Positive_To_Covid(Yes) |                   0.1274 |
+------------------------+--------------------------+
```

If one person is vaccinated, then the probability of being infected with the virus falls to 12%.

4.1.4 Maximum A-posteriori Probability

In this subsection, the Maximum A-posteriori Probability (MAP) of the variable Positive_To_Covid is computed.

It can be noticed that if someone has the virus, then the most probable values of the variables Over_Eighty, Live_With_Other_People, and Vaccinated are respectively No, Yes, and No. Hence, it means that it is more probable that a person is positive to Covid-19:

- If he/she is younger than 80 years old rather than older;
- If he/she lives with other people rather than living alone;
- If he/she is not vaccinated.

```
MAP of Positive_To_Covid = 'Yes'
```

```
{'Over_Eighty': 'No', 'Live_With_Other_People': 'Yes', 'Vaccinated': 'No'}
```

4.2 Approximate Inference

In this section, Approximate Inference is applied to compute the probability of the variable Vaccinated, given that the user is older than 80 years old and that has not the virus.

The same probability is first calculated with Exact Inference (Variable Elimination) to visualize how well the Approximate methods perform compared to the Exact Inference methods.

```
Exact Inference:

P(Vaccinated | Over_Eighty = "Yes", Positive_To_Covid = "No")

+-----------------+-------------------+
| Vaccinated      |   phi(Vaccinated) |
+=================+===================+
| Vaccinated(No)  |            0.4780 |
+-----------------+-------------------+
| Vaccinated(Yes) |            0.5220 |
+-----------------+-------------------+
```

If an older adult does not have Covid-19, it is slightly more probable that he/she has received the vaccine (52%) rather than not (48%). The previous statement makes sense because only 43% of over eighty people in Italy are vaccinated (Subsection 2.6), and the assumption of not having the virus increases the probability of having received the vaccine.

Then, the probability P(Vaccinated | Over_Eighty = "Yes", Positive_To_Covid = "No") is estimated by using three different Approximate Inference methods, which are:

- Weighted Likelihood;
- Gibbs Sampling;
- and Rejection Sampling.

In Figure 2 it can be seen that they are tested in five consecutive steps with an increasing number of samples.

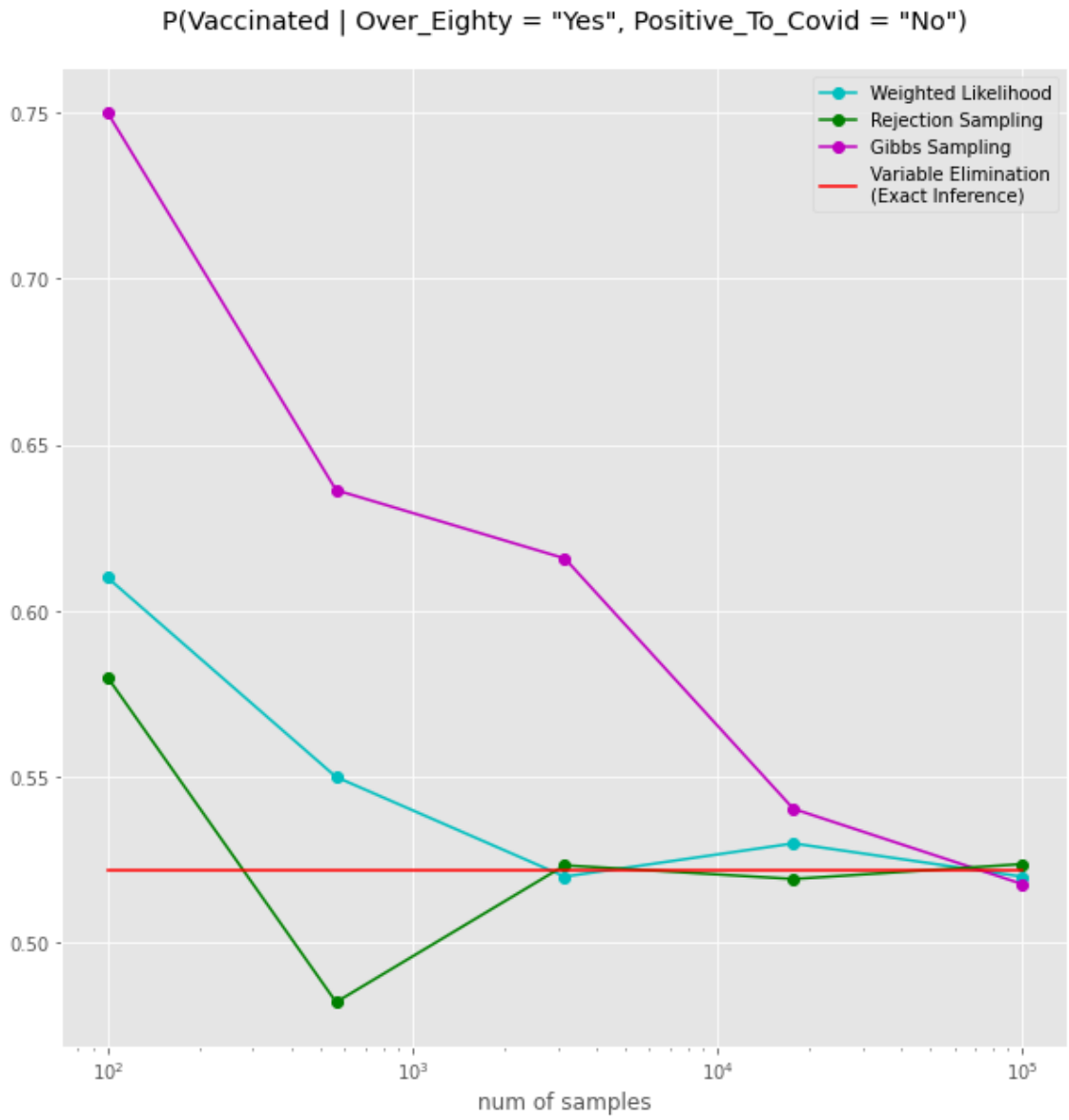

Figure 2: Test of three different Approximate Inference methods

It is visible that, in 5 steps, all the three Approximate Inference methods bring great results close to the Variable Elimination one.

The most accurate (the one with the minor error) is the Rejection Sampling method, which approximates the probability P(Vaccinated | Over_Eighty = "Yes", Positive_To_Covid = "No") to 0.52185.

References

[1] *Aspetti di vita degli over 75*. URL: https://www.istat.it/it/files/2020/04/statisticatoday_ANZIANI.pdf.

[2] *Popolazione Italiana residente al 1° gennaio*. URL: http://dati.istat.it/Index.aspx?DataSetCode=DCIS_POPRES1.

[3] *Vaccini in tempo reale*. URL: https://lab24.ilsole24ore.com/numeri-vaccini-italia-mondo.