Data
Science

Information fusion in data analysis

Data Fusion – Project (2)

Alberto Cardoso, Jorge Henriques
alberto@dei.uc.pt, jh@dei.uc.pt

Departamento de Engenharia Informática
Faculdade de Ciências e Tecnologia
UNIVERSIDADE DE COIMBRA

Objective

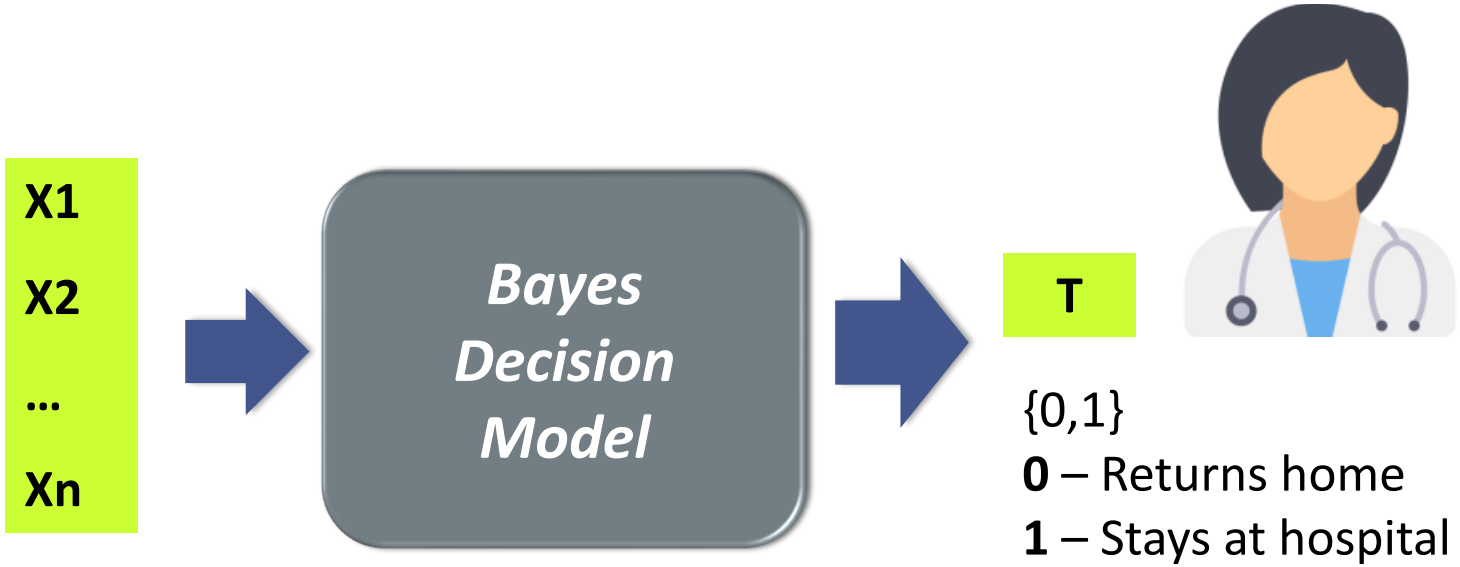
Problem – Stratification of COVID patients

- Individuals with suspected COVID are admitted to the hospital emergency room
- At the time of admission, several variables/parameters are acquired (low cost and simple to acquire)

- Based on these variables, the health professional must decide whether the individual remains hospitalized for additional examinations or should return home.

Objectivo

Stratification of COVID patients



Objectivo

Data set

- **X1** *Gender* {0,1} = { Female, Male}
- **X2** *Age* [34 .. 99]
- **X3** *Marital status* {0,1} = { single, married}
- **X4** *Vaccinated* {0,1} = { No, Yes}
- **X5** *Breathing difficulty* {0,1,2,3} = { none, some, moderate, high}
- **X6** *Heart Rate* [38.. 272]
- **X7** *Blood pressure* [115.. 164]
- **X8** *Temperature* [36.00 .. 38.98]
- **X9** *Clinical Guidelines* A rule based on the breathing difficulty and the temperature

- **T** *Decision* Final decision
  {0,1} = { return home, stay at hospital }

1| Data set

- Screening

1| Gender
2| Age
3| Marital status
4| Vaccinated
5| Breathing difficulty

- Measurements

6| Heart rate
7| Blood pressure

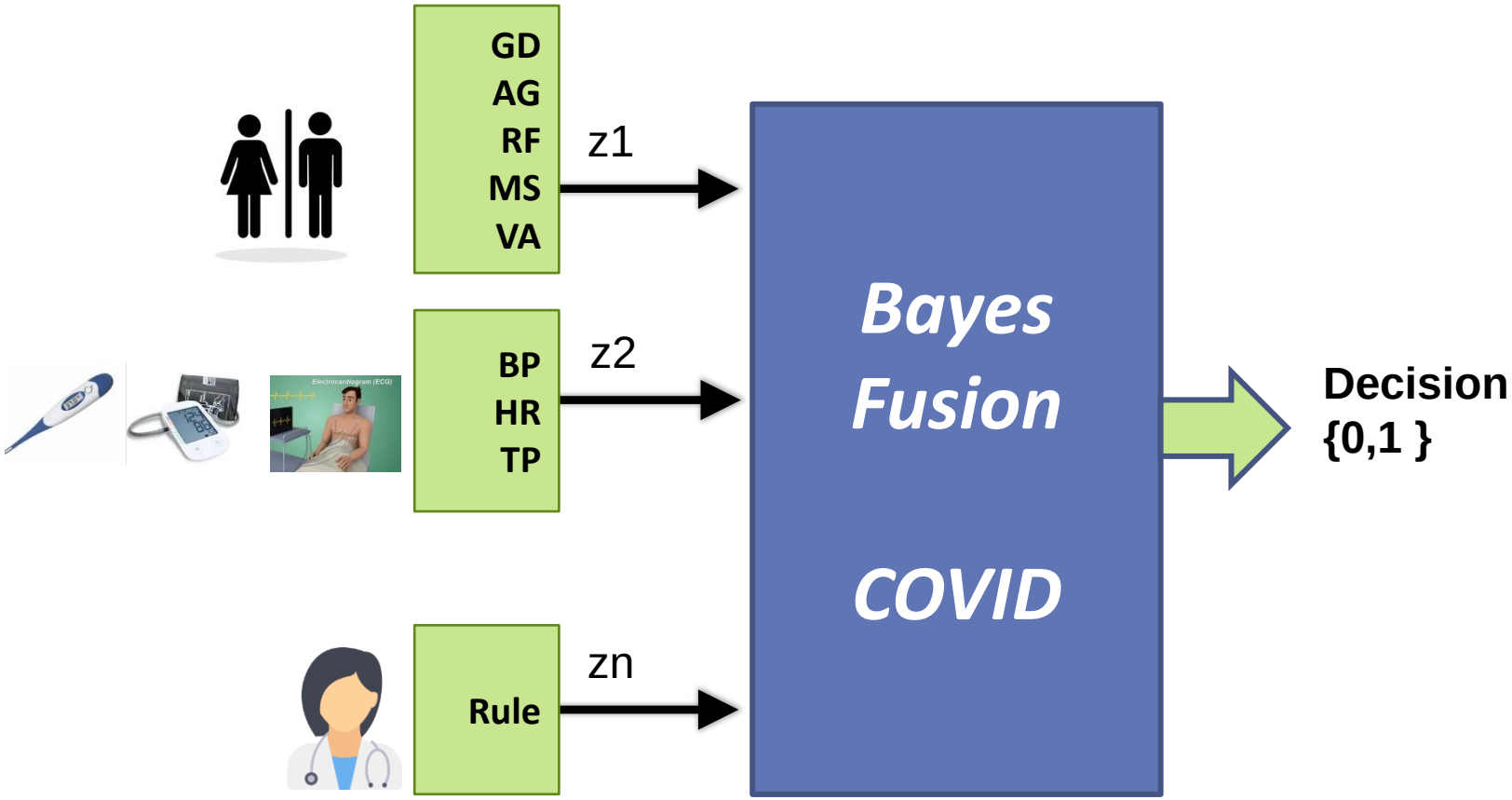
8| Temperature

- Knowledge

9|
IF
    breathing difficulty >= moderate
AND
    Temperature > 37.8
THEN
    Stay at hospital

Information fusion

- Historical, measurements, clinical knowledge (guidelines)



Questions

- Is the performance of the classifier acceptable ?
- Should all information (inputs/variables) be used ?
- Discrete versus continuous variables ?
- Conditional probabilities : normal distribution ?
- ....

Elements for evaluation

- **Code**
  - All code should be provided
- **Report**
  - Maximum 5/6 pages
  - Explain the important decisions
- **Defense**
  - Mandatory
- **Deadline for submission**
  - ??