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

- **Two alternatives for fusion**

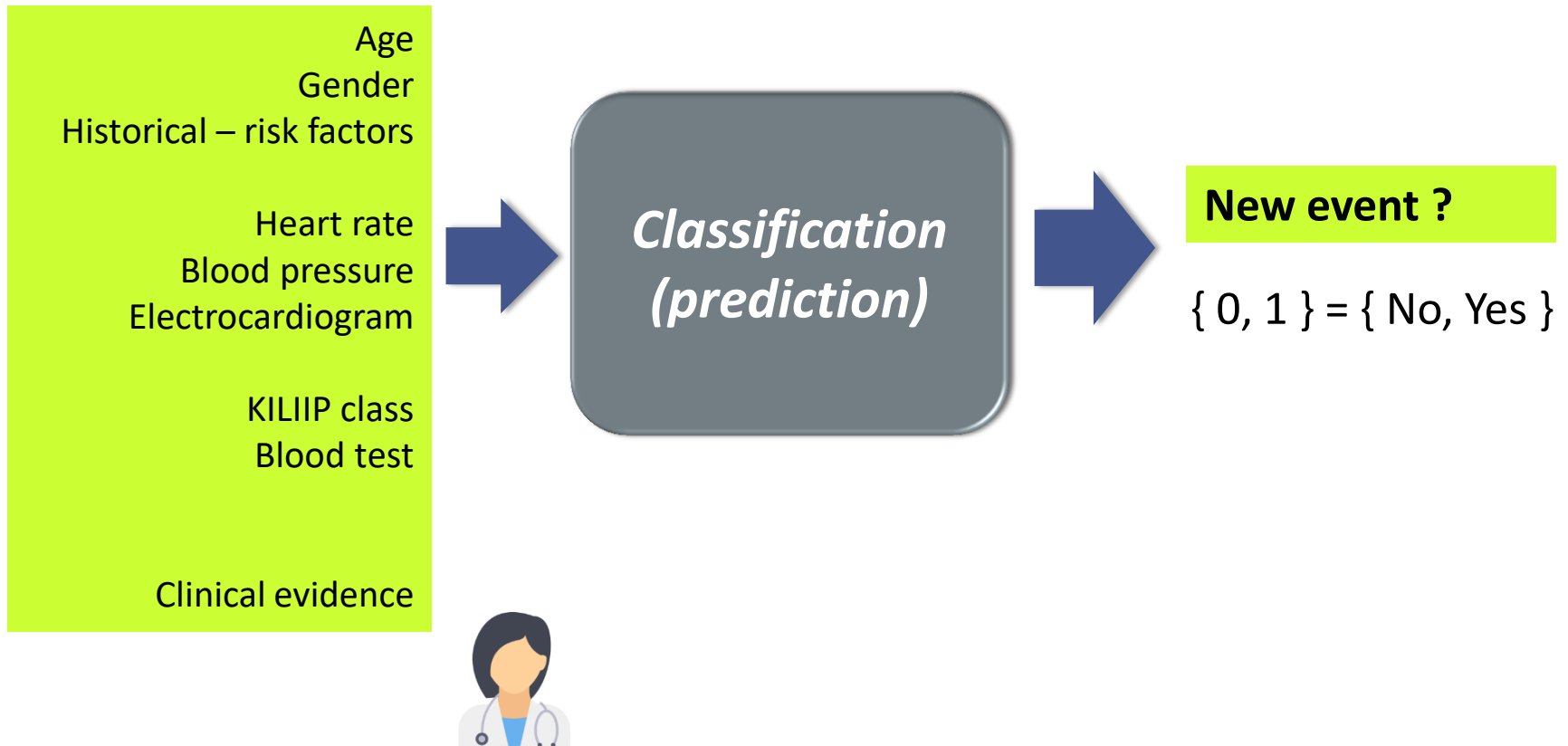 - **1|** Bayesian : cardiovascular risk assessment

  - **2|** Fuzzy: **

▶ **Goal**

- To develop a risk model, applicable to artery coronary syndrome (ACS) patients that have been admitted to the emergency unit with an episode of myocardial infarction (MI)

- The model should be able to predict if a new event will occur in the next 30 days.

▼ **Risk assessment of a new event**

Age
Gender
Historical – risk factors

Heart rate
Blood pressure
Electrocardiogram

KILIIP class
Blood test

Clinical evidence

Classification (prediction)

New event ?

{ 0, 1 } = { No, Yes }

▶ **Risk assessment of a new event : X - INPUTS**

- **1| Historical**
  - **GD** |Gender | { female, male } = { 0, 1 }
  - **AG** |Age | [33 .. 90]
  - **RF** |Risk Factors | { noRisk, risk } = { 0, 1}
    - RF - Related with the clinical history of the patient
    - Family, past events, …
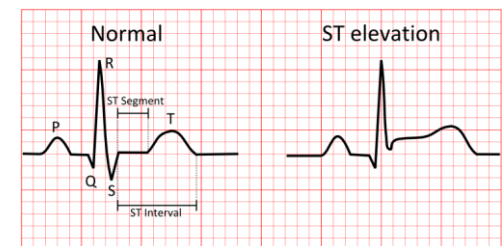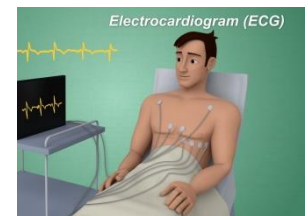
▰ **Risk assessment of a new event : X - INPUTS**

- **2| Measurements**
  - **SBP** |Systolic blood pressure | [60 .. 221]
  - **HR** |Heart rate | [40 .. 153]
  - **ST** |ST elevation (ECG) | {0, 1}

- Devices
  - **BP device** - Measurement of **BP** and **HR**
  - **ECG** – Measurement of electrical activity
    - **ST** elevation is the most important
    - **HR** – heart rate

▶ **Risk assessment of a new event : X - INPUTS**

- **3| Exams/diagnosis**
  - **CT** |Creatinine | [0.6 .. 11.5]
  - **KL** |KILLIP class | {1, 2, 3, 4}

- Devices
  - **CT** - Blood test - measure how well your kidneys are working.
  - **KL – Killip class -** physical examination – functional capacity
    - 1 - No signs
    - 2 - Mild to moderate signs
    - 3 - Pulmonary edema (Severe)
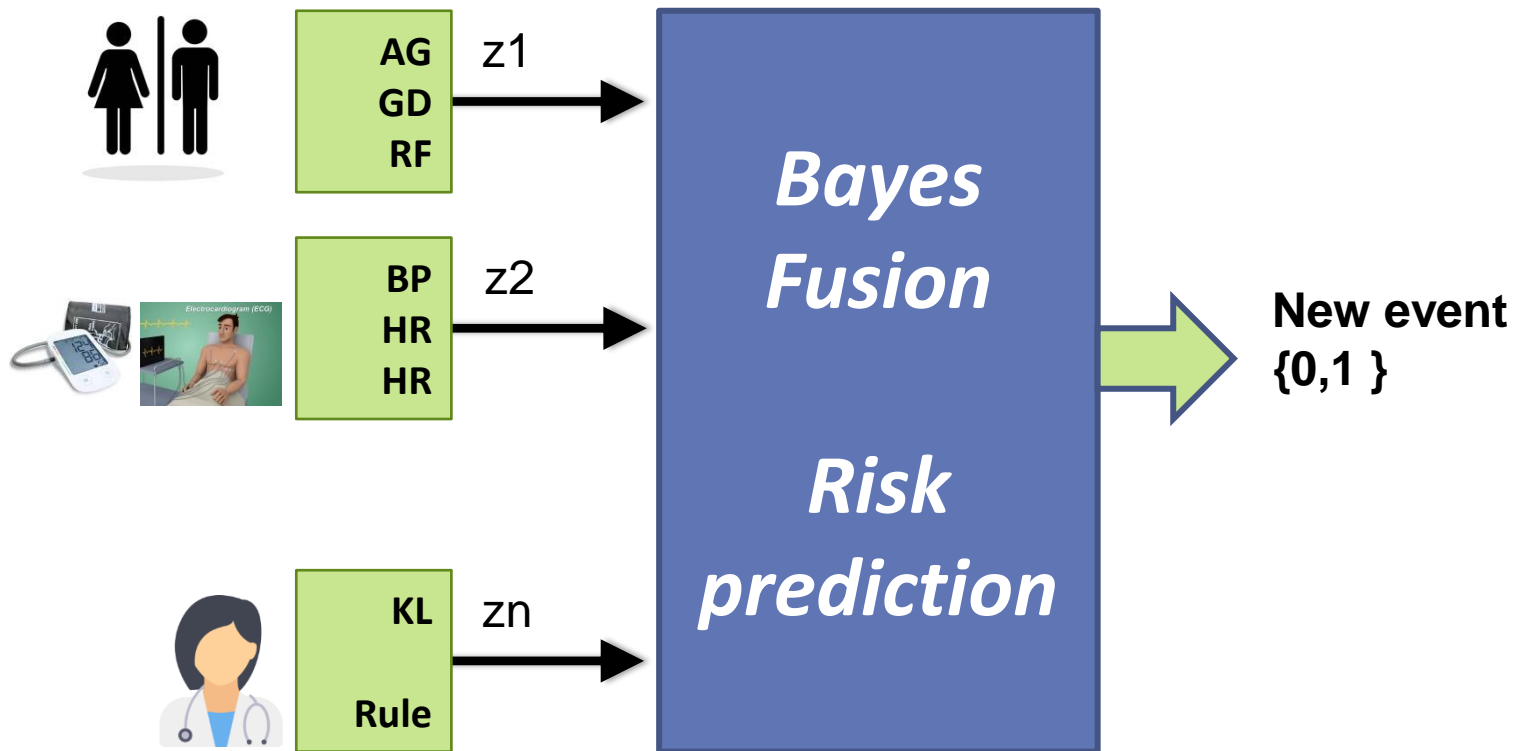   - 4 – Cardiogenic shock (Severe)

Clinical guidelines

If **CT>=1.3 AND ST=1** **risk=1**
If **KL>=2** **risk =1**

◤ Information fusion

- Historical, measurements, clinical knowledge (guidelines)

- **DATASET : DATA.txt**
  {'SEX', 'AGE', 'RF', 'SBP', 'HR1' 'ST', 'HR2', 'CRT', 'KIL', 'EVENT' }

  - **x1** *Gender* {0,1} = { Female, Male}

  - **X2** *Age*

  - **X3** *Risk factors* {0,1} = { No, YES}

  - **X4** *Systolic Blood pressure*

  - **X5** *Heart rate (1)*

  - **X6** *ST deviation* {0,1} = { No, YES}

  - **X7** *Heart rate (2)*

  - **X8** *Creatinine*

  - **X9** *Kilip class* {1,2,3,4}

  - **T** *Target=event* {0,1} = { No event, Event }

- **Questions**

  - Is the performance of the classifier acceptable ?

  - Should all information (inputs/variables) be used ?

  - 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**
    - ????

- **Two alternatives for fusion**

  - **1|** Bayesian : cardiovascular risk assessment

  - **2|** Fuzzy: **