Maastricht University

Explainable and Reliable AI

Research Project Presentation 2

Group 11

Supervisor

Evgueni Smirnov

Sotirios Ftiakas Dimitris Feizai

Giannis Triantafyllou

Coordinator

Menica Dibenedetto

Problem Description

Problem Description

Scenario:

Person **A** and Person **B** feel slight discomfort in the chest area, and they request **hospitalization**. (The hospital has only **1 slot** available.)

- Which patient should the hospital **accept** and which one to **reject**?
- How can the doctor provide an **explainable** and **reliable** decision?

Decision Tree

ML model that uses data features as decision rules in a tree like structure.

- **Internal Nodes:** Store features as decision rules
- **Branches:** Path based on feature value
- **Leaf Nodes:** Store prediction outcome

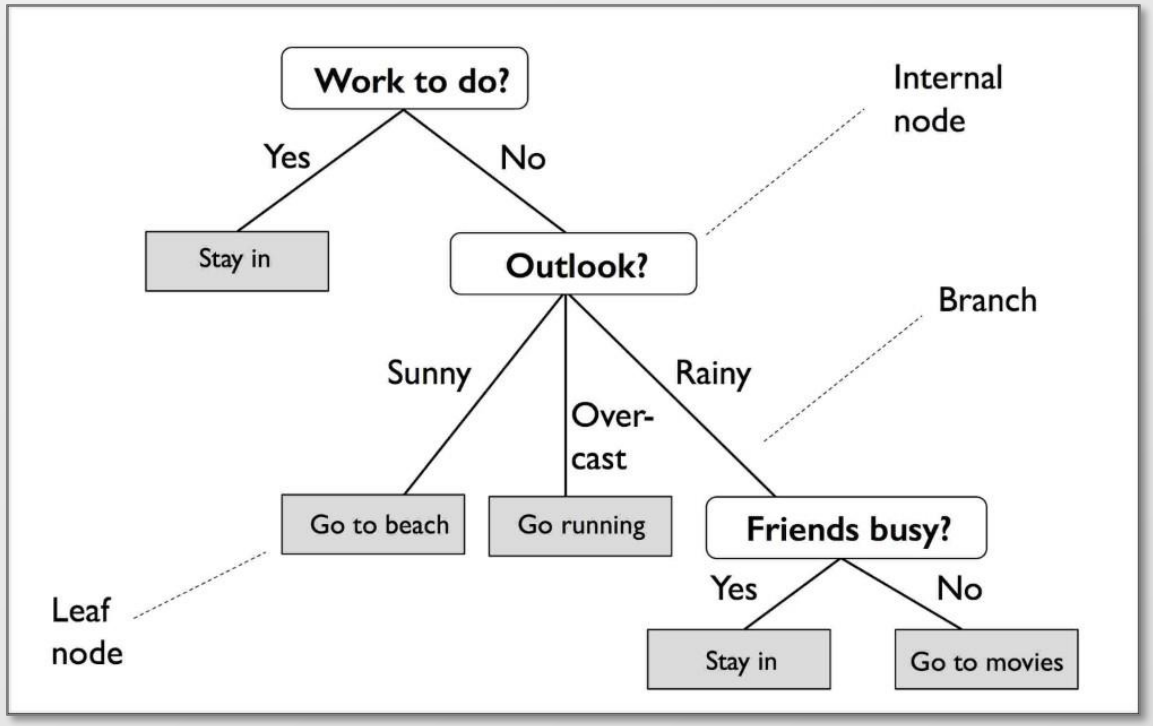


FIGURE: Decision Tree structure example

Inductive Conformal Predictor (I.C.P.)

- Returns **prediction regions – intervals** for **regression** problems and **sets of labels** for **classification** problems.
- **Training – Calibration – Test set:** The percentage split of each set can differ.
- **Nonconformity function:** Measures how unusual an example looks relative to previous examples

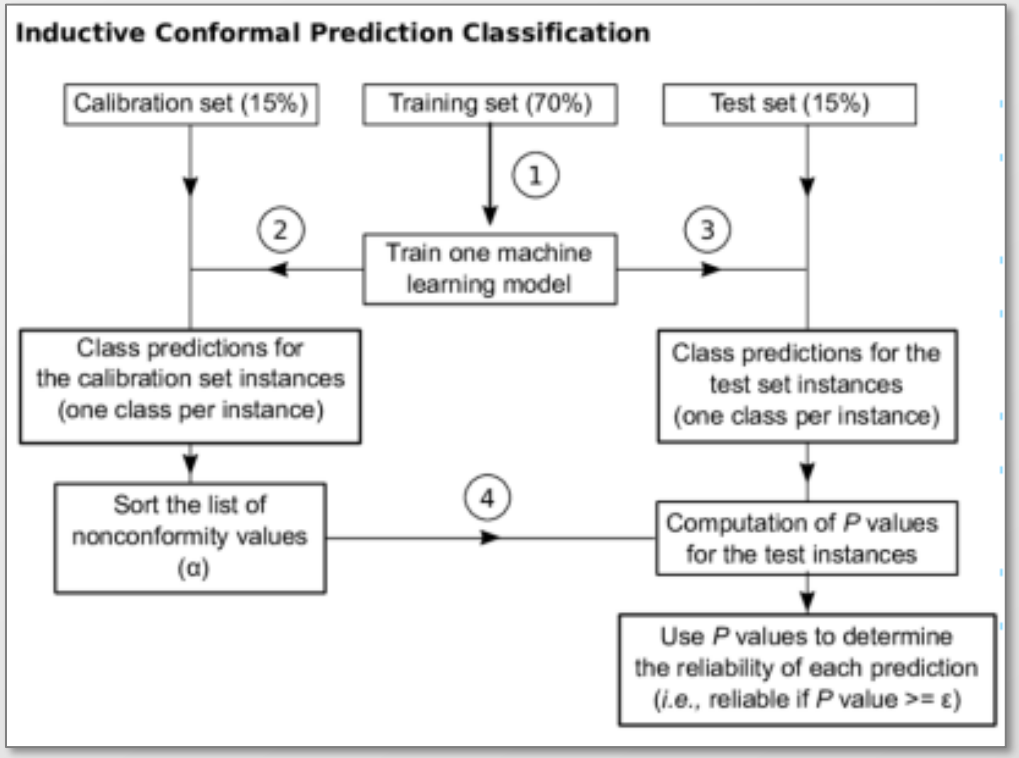


FIGURE: Inductive Conformal Predictor structure example for classification

Validity of I.C.P.

To test the **validity** of a conformal set predictor we use the **error rate e.**

The **error rate e** for a **significance level ε** is defined as the proportion of test instances whose predicted prediction-sets do not contain the correct class.

To test the I.C.P. validity, we need to show that **for any significance level ε ∈ [0, 1]** we have **e ≤ ε.**

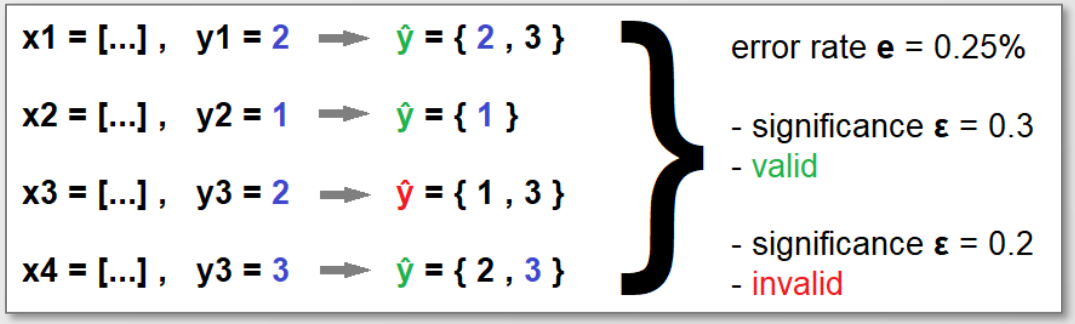


FIGURE: Validity example

Information Efficiency of I.C.P.

To test the **informational efficiency** of a conformal set predictor, for a given **significance level ε** we employ **three main metrics**:

- The rate $\mathbf{r^e}$ of **empty** prediction sets
- The rate $\mathbf{r^s}$ of **single** prediction sets
- The rate $\mathbf{r^m}$ of **multiple** prediction sets.

We need **both validity** and **information efficiency**.

Prediction sets **as small as possible** but **NOT empty**.

Our Approach

Combining Decision Tree with I.C.P.

- **Hybrid** prediction model
- Normal decision tree structure
- I.C.P. at **leaf nodes**

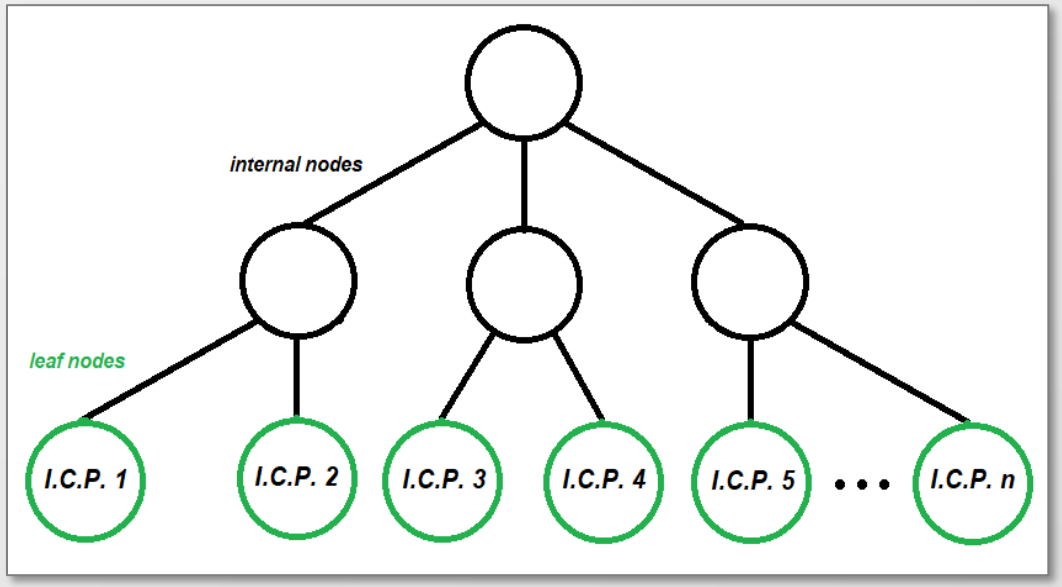


FIGURE: Decision Tree with I.C.P.s at leaf nodes

Training Process

1) **Grow tree** with training set

2) Send **training examples** through the tree

3) Send **calibration examples** through the tree

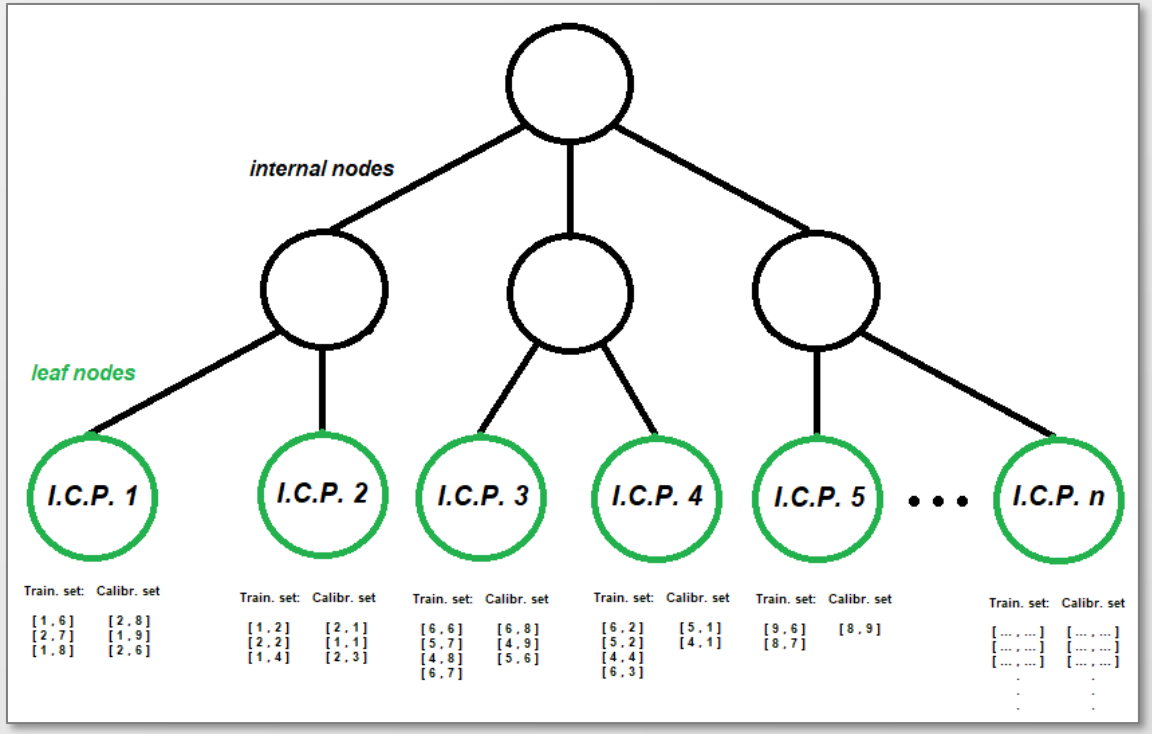


FIGURE: Training and calibration instances at leaf nodes

Classification Process

1) Send **test example** through the tree

2) Once the test instance reaches a **leaf node**, perform conformal prediction among those **training** and **calibration** items **in the same leaf** as the test instance.

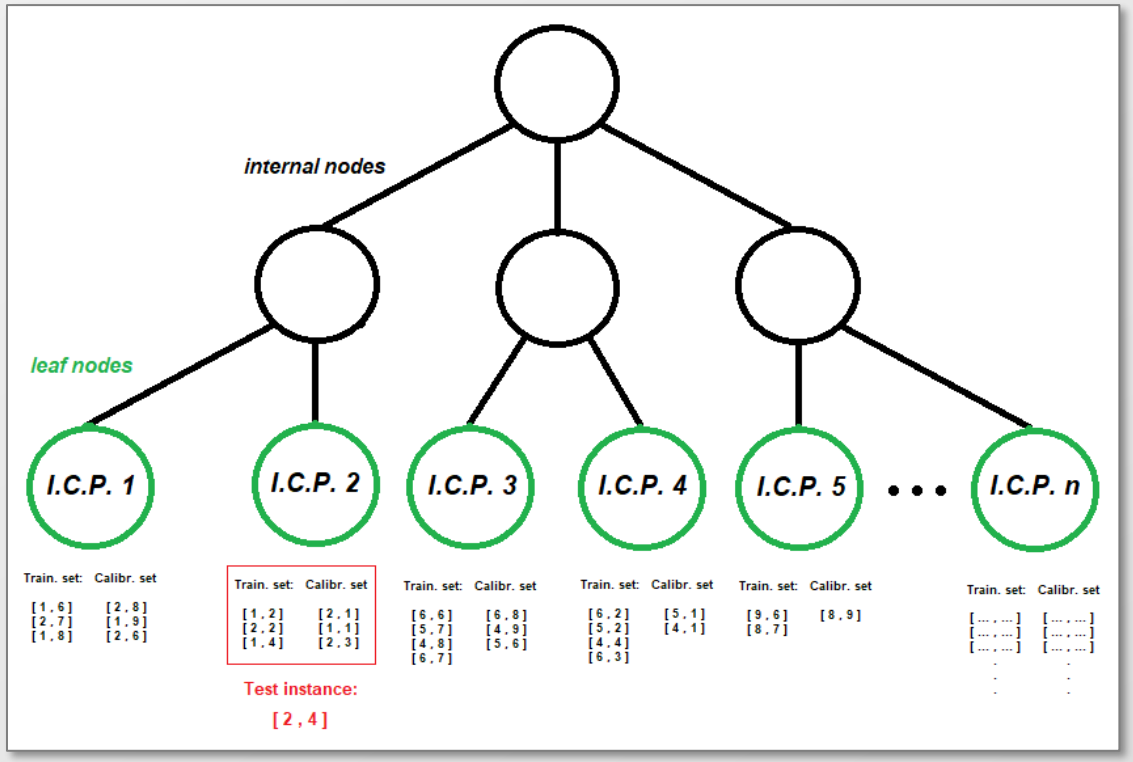


FIGURE: Leaf-specific inductive conformal predictor

First Results

Iris Dataset

First plot:

Error rate **e** / Significance level **ε**

Second plot:

Mean prediction sets **size** /

Significance level **ε**

Third plot:

Rate $\mathbf{r^s}$ of **single** prediction sets /

Significance level **ε**

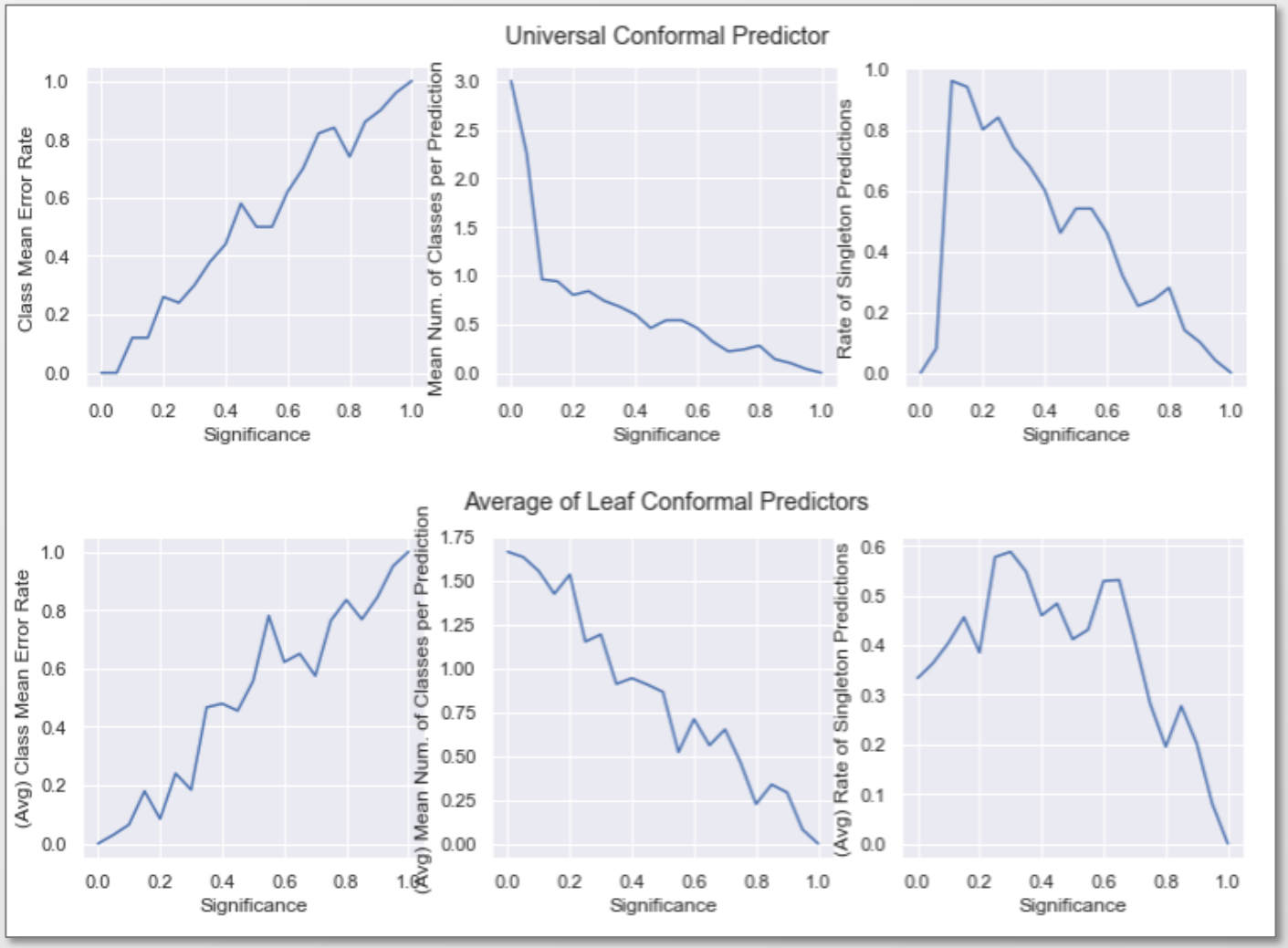


FIGURE: Plots for validity and information efficiency

Challenge

Challenge

What if there is **no calibration** or **training** set on a **leaf node**?

We **can not perform I.C.P.** on a test instance that arrives at such node.

Two possible solutions:

1) Do not perform conformal prediction at all.

2) Prune paths that lead to nodes with **no calibration / training set.**

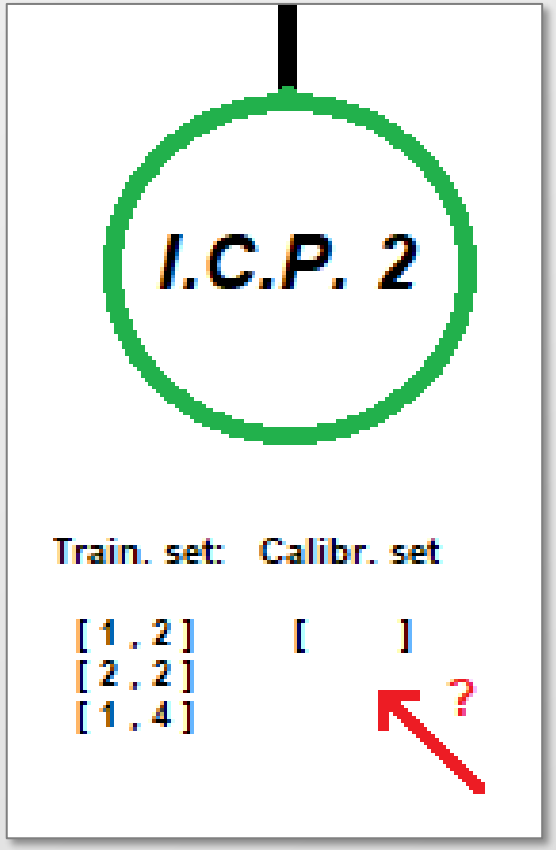
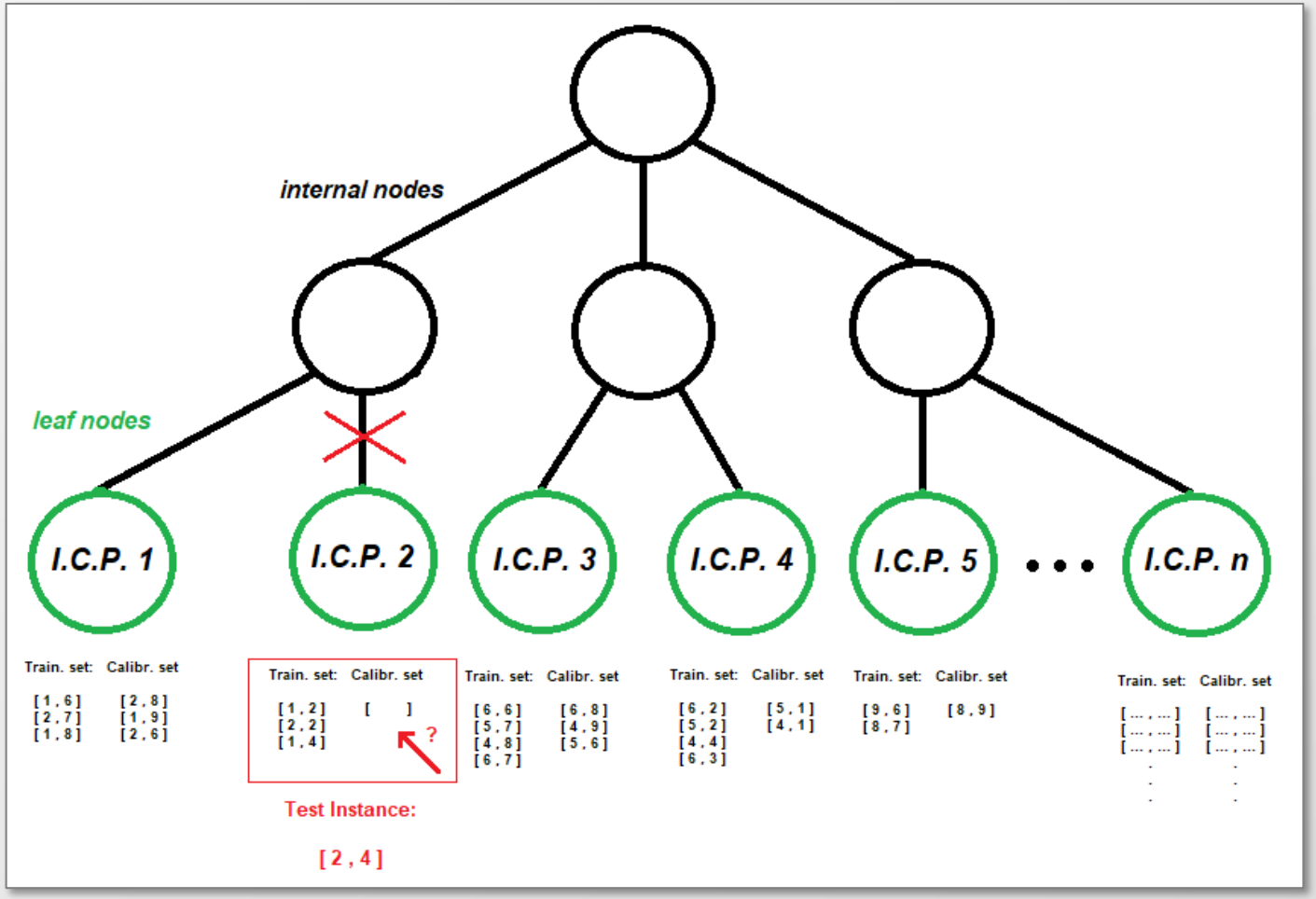
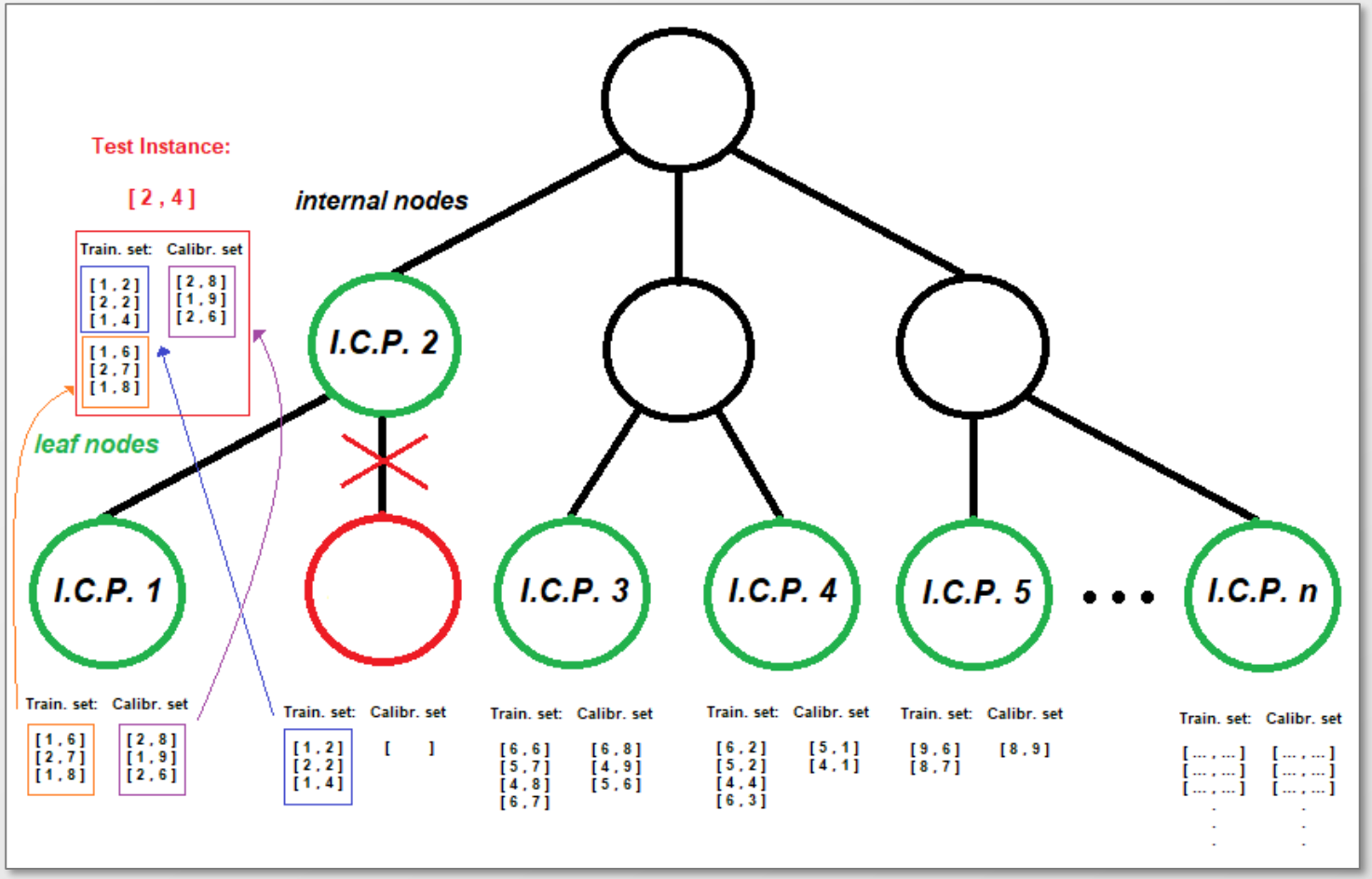


FIGURE: Leaf node with no calibration set

Pruning Method



Pruning Method



Research Questions

Research Questions

- **Question 1**

  Is **Validity** preserved?

- **Question 2**

  Comparison of **global I.C.P.** performance vs **average of leaf I.C.Ps** performance.

- **Question 3**

  Comparison of **global I.C.P.** performance vs **individual leaf I.C.Ps** performance.

Thank you for your attention!

Any questions?

Maastricht University