

# Practical Application II

Cristea-Olaru Ștefan & Neagu Alex-Ștefan

17 January 2025

## 1 Introducere

Acest document are rolul de a prezenta abordările pe care le-am avut noi pentru o problemă de învățare automată, mai exact, realizarea unui ranking a activităților ce pot genera cel mai mare profit pentru un proprietar de lanț hotelier.

## 2 Intelegerea setului de date

Un lucru pe care l-am dedus noi, și este de asemenea un lucru rațional, este că profitul generat de fiecare locație este, în medie, direct proporțional cu celelalte atribute, mai exact:

Cu cât este ratingul mai mare  $\Rightarrow$  Cu atât este profitul mai mare.  
Cu cât este numărul de vizitatori mai mare  $\Rightarrow$  Cu atât este profitul mai mare.  
Cu cât este confortul mai mare, sau cazare,  $\Rightarrow$  Cu atât este profitul mai mare.

Putem să deducem o creștere oarecum **liniară** a revenue-ului în funcție de aceste atribute. Asadar, am încercat doi algoritmi care sunt separatori liniari, mai exact:

1. **Regresia Liniară**
2. **ID3**(*modificat pentru Arbori de Regresie*)

## 3 Algoritmii

### 3.1 Regresia Liniară

Regresia liniară este foarte bună pentru această situație deoarece setul nostru de date are caracteristica menționată anterior.

### 3.2 ID3

#### Arbori de Regresie

În realitate, arborii de decizie sunt și ei niște separatori liniari, asadar nu este de mirare că dau rezultate la fel de bune ca și regresia liniară.

Motivul pentru care am selectat și acest algoritm este că regresia liniară suferă de o sensibilitate atunci când atributele datelor de antrenament se regăsesc în domenii diferite. Acest lucru poate să fie reparat, standardizând/normalizând domeniile atributelor, dar a fost dificil și nu am reușit să o facem în mod eficient.

### 3.3 Adaboost

Adaboost, la fel ca ceilalți doi algoritmi precedenți, este foarte bun în situația aceasta. În realitate, el este o combinație ponderată a mai multor clasificatori slabi, cum ar fi compasii de decizie de la ID3.

Asadar, și acesta dă rezultate foarte bune pentru datele din acest db.

## 4 Rezultate

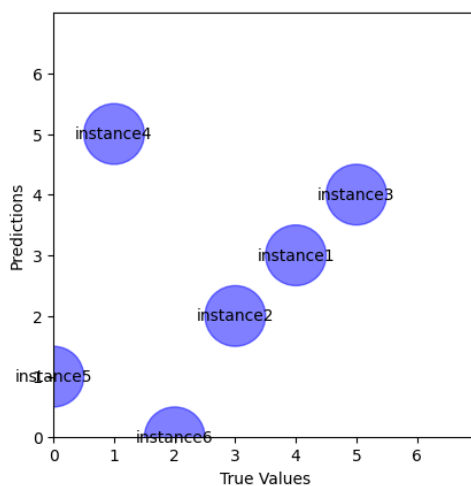


Figure 1: Brazil

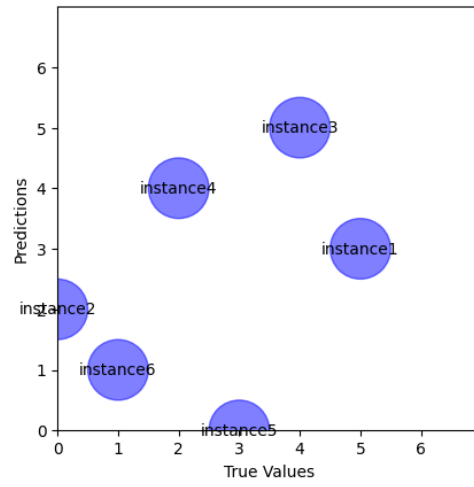


Figure 2: USA

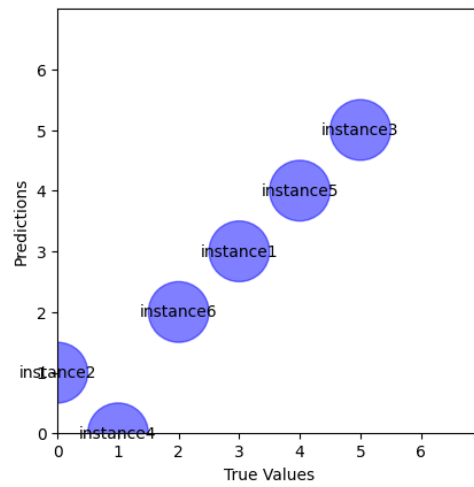


Figure 3: Egypt

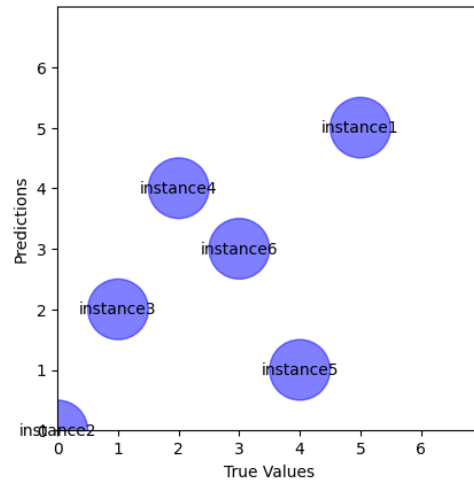


Figure 4: France

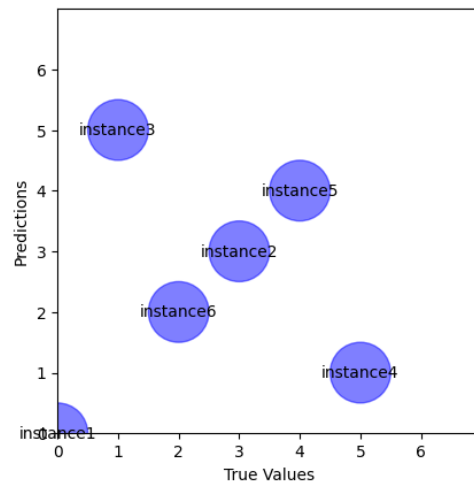


Figure 5: China