

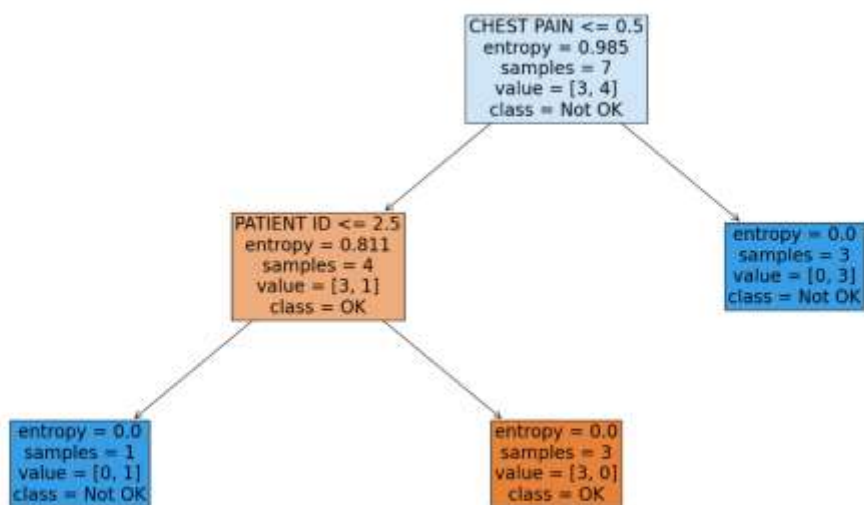
Nhóm 6

Nguyễn Hoàng Xuân Huy – 21520924

Trần Ngọc Thiện – 21521465

1. Dùng entropy để vẽ cây quyết định với dataset cho trong file *hw\_dataset\_1.csv*, biến mục tiêu là **HEART ATTACK**. Lưu ý **PATIENT ID** là thuộc tính dạng rời rạc.

PATIENT ID	CHEST PAIN	GENDER	SMOKES	EXERCISES	HEART ATTACK
1	yes	male	no	yes	yes
2	yes	male	yes	no	yes
3	no	female	yes	no	yes
4	no	male	no	yes	no
5	yes	female	yes	yes	yes
6	no	male	yes	yes	no
7	no	female	no	yes	no



2. Dùng gini để vẽ cây quyết định với dataset cho trong file *hw\_dataset\_2.csv*, biến mục tiêu là **OK?**. Lưu ý **Price** là thuộc tính dạng liên tục.

Type of restaurant	Price	Neighborhood	Restriction	OK?
Fast food	5	Oakland	Vegetarian	Not OK
Ethnic	40	Squirrel Hill	Gluten Free	Not OK
Casual dining	35	Squirrel Hill	None	Not OK
Casual dining	80	Shadyside	Vegetarian	Not OK
Casual dining	11	Oakland	Vegetarian	OK
Fast food	31	Squirrel Hill	None	OK
Ethnic	10	Squirrel Hill	None	OK
Casual dining	22	Shadyside	Gluten Free	Not OK
Fast food	140	Oakland	None	Not OK
Ethnic	68	Shadyside	Vegetarian	OK
Casual dining	57	Shadyside	Gluten Free	OK

