Data Mining

2022-2023 Winter Semester

Titanic Dataset Project

Emir Mirza YILDIRIM

Questions

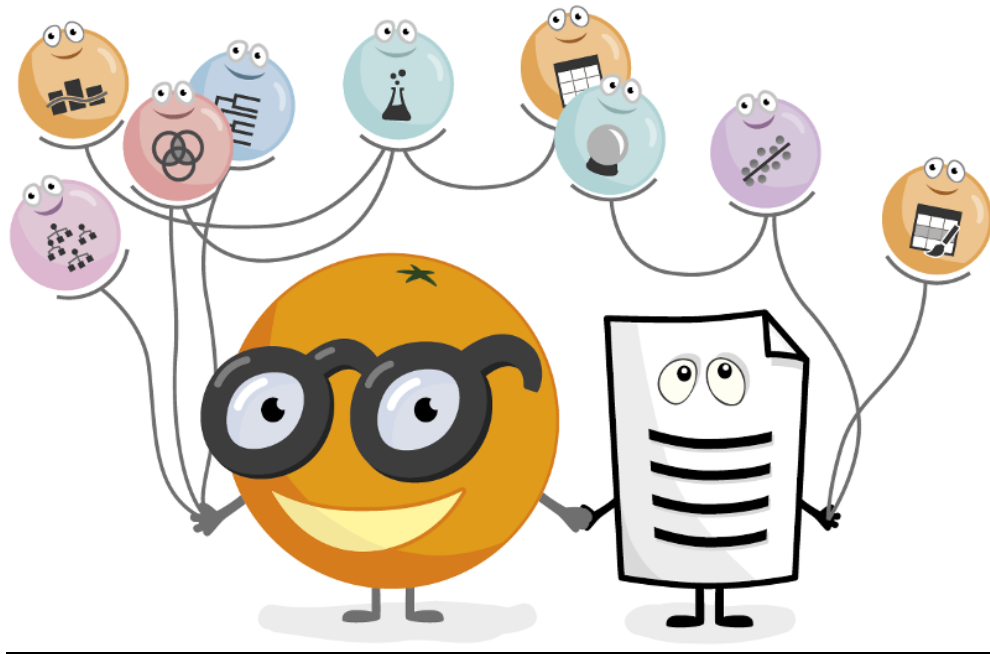1- Which factors are crucial to survive?
- Age
- SibSp
- PClass
- Sex

2- Who has a higher chance to survive? Jack, as a young male, traveling alone as a passenger of 3rd class, or Rose, as a young female, traveling with family and passenger of 1st class ?
- Jack is 90.7% dead.
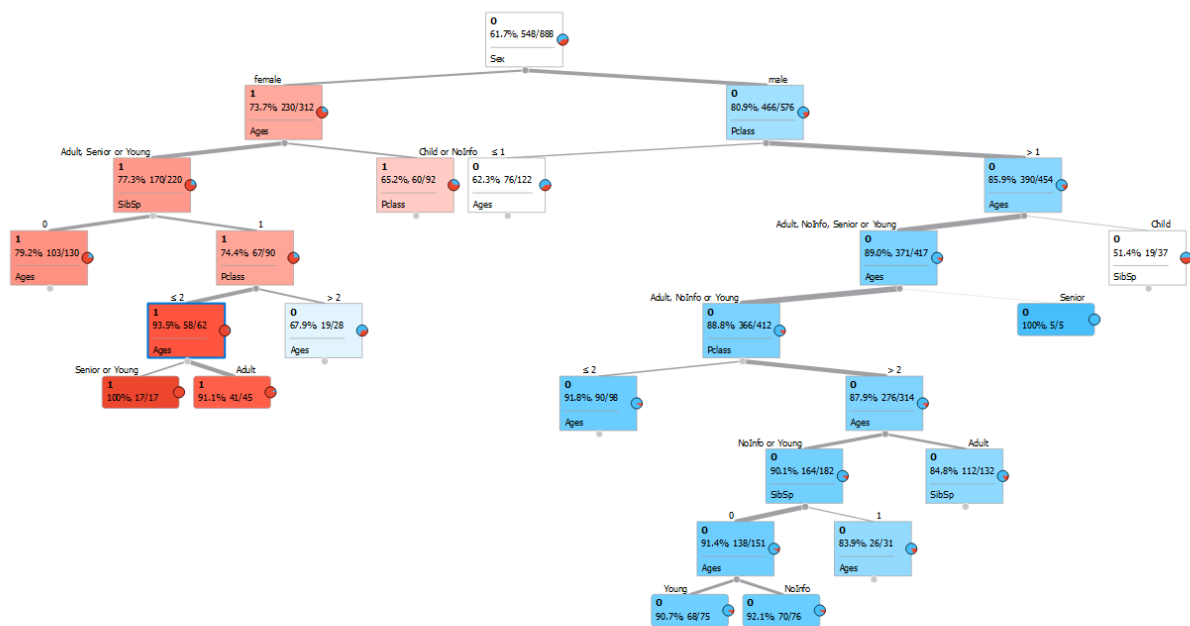- Rose is 100% alive.
- Rose has a higher chance to survive.

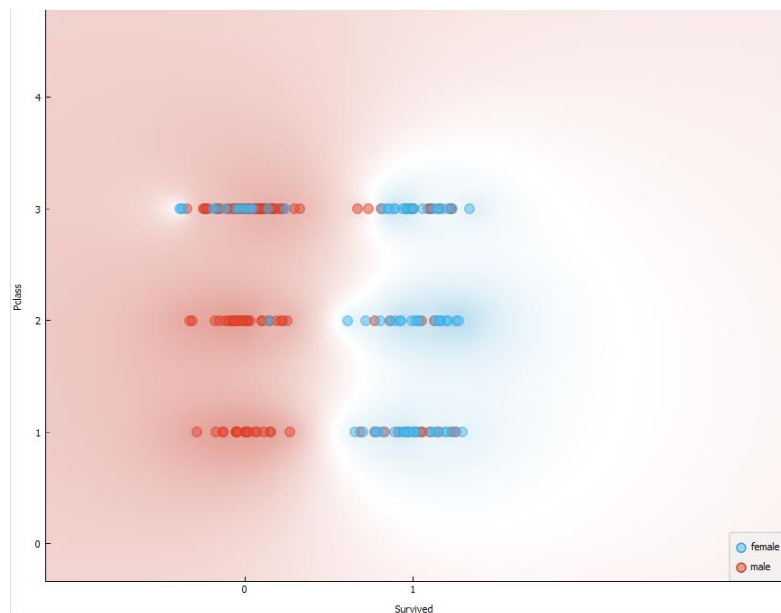# Tools and Some Information about the tools that I have used :

Orange3

Jupyter Notebook

- I have used tree view (Orange3) to find out who has a higher chance to survive because it's easy to find your way.

- Also I have used Scatter Plot (Orange3) to find if I can find some examples similar to Rose or Jack.

- And I have used other graphical tools in Orange3. But I have used jupyter notebook to change something in the dataset on the final step of cleaning. Because in my old version of dataset I had some different datatypes or some different approachoes to the data. I have changed something in the dataset to use Orange more efficient using Jupyter Notebook.

- But I will not talk about the changes that I did, because I wrote comments in the Jupyter File.