Prediction of Lung Cancer

**Algorithms Used:**

* Random Forest
* Decision Tree
* Support Vector Classifier

**Colab Notebook:** <https://colab.research.google.com/drive/1vWFIg3zzFj34Ey2O2bQzh5tImViBUUfp?usp=sharing>

**Source:**

<https://www.kaggle.com/code/ihabsherbiny/lung-cancer-detection/notebook>

**Dataset:**

<https://www.kaggle.com/code/ihabsherbiny/lung-cancer-detection/input>

**Rows, columns:** (284, 16)

**About:**

The effectiveness of the cancer prediction system helps people to know their cancer risk with a low cost and it also allows people to make the appropriate decision based on their cancer risk status.

The data is collected from the website, which is an online lung cancer prediction system.

Total no. of attributes:16 No .of instances:284

**Attribute information:**

1. Gender: M(male), F(female)

2. Age: Age of the patient

3. Smoking: YES=2, NO=1.

4. Yellow fingers: YES=2 , NO=1.

5. Anxiety: YES=2, NO=1.

6. Peer\_pressure: YES=2, NO=1.

7. Chronic Disease: YES=2, NO=1.

8. Fatigue: YES=2 , NO=1.

9. Allergy: YES=2, NO=1.

10. Wheezing: YES=2, NO=1.

11. Alcohol: YES=2, NO=1.

12. Coughing: YES=2, NO=1.

13. Shortness of Breath: YES=2, NO=1.

14. Swallowing Difficulty: YES=2, NO=1.

15. Chest pain: YES=2, NO=1.

16. Lung Cancer: YES, NO.

**Algorithms Model and Accuracy**

|  |  |
| --- | --- |
| **Models** | **Scores** |
| Random\_forest | 0.875 |
| Decision\_Tree | 0.892857 |
| SVC | 0.928571 |

Best: **Support Vector Classifier**