



College of Computing and Information Technology – South Valley (Aswan)

CS366-Introduction to Artificial Intelligence  
(Spring-20/21)

# Lung Cancer Prediction

Bishoy Abd El-Massieh Aiad(18108747)

# 1-PROBLEM STATEMENT

- It is the diagnosis of symptoms that effect in lung cancer diagnosis like alkhol , smokes. So , with using a Machine Learning algorithm, we can predict the probability of a person having lung cancer disease or not.

## 2-DATASET

- Features : 1- Name , 2- Surname , 3- Age , 4- Smokes , 5- AreaQ , 6- Alkhol , 7- Result

- Samples :

2	John	Wick	35	3	5	4	1
3	John	Constantir	27	20	2	5	1
4	Camela	Anderson	30	0	5	2	0
5	Alex	Telles	28	0	8	1	0

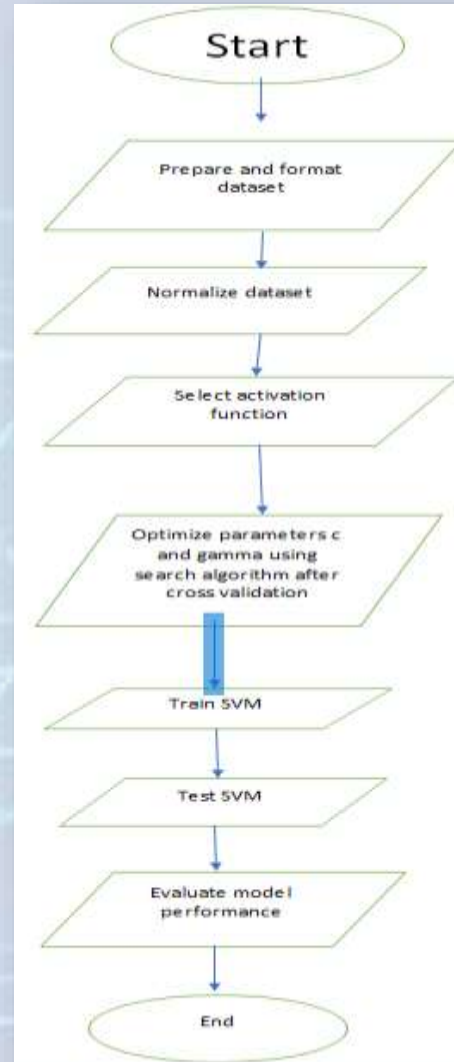
- Labelling : 0(for not having Lung Cancer),1(for having Lung Cancer)
- URL : <https://www.kaggle.com/yusufdede/lung-cancer-dataset>

# MACHINE LEARNING ALGORITHM

- **NAME:** Support Vector Machine using Linear as kernel method
- **REASON OF SELECTION:**
  - It's effective in n dimensional cases.
  - It's memory is efficient as it uses a subset of data points in the decision making functions called support vectors.
  - There are different ways to specify the decision function by using different kernel functions.

# MACHINE LEARNING ALGORITHM

- **FLOWCHART:**



# MACHINE LEARNING ALGORITHM

- **PSEUDOCODE:**

**1- Normalize the data set**

**2- For each  $c$ ,  $\gamma$ .**

- a. Cross validate using leave one out.**

- i. Train and test the SVM.**

- ii. Store the success rate.**

- b. Compute the average success rate.**

- c. Update the best  $c$  and,  $\gamma$  if needed.**

- d. Return to 2.1 with next  $c$ ,  $\gamma$ .**

**3- Choose  $c$ ,  $\gamma$  with best average success rate, and perform step (2) using fine scale around around the selected parameters.**

# PERFORMANCE MEASURES

- Accuracy =  $\frac{TP + TN}{TP + TN + FP + FN}$
- Precision =  $\frac{TP}{TP + FP}$
- Recall =  $\frac{TP}{TP + FN}$
- Specificity =  $\frac{TN}{TP + FN}$

# EXPEPRIMENTAL RESULTS

- <insert here the results (as the values of the performance measures) obtained from implementing the ML algorithm on the dataset>

- Accuracy = 95%
- Precision = 100%
- Recall = 88%
- Specificity = 100%



# URL OF THE PROJECT DESCRIPTION CLIP (60sec.)

- [https://drive.google.com/file/d/1xk6nDzvn\\_AbW5AdbEjnZ4NA97V4MOg0n/view](https://drive.google.com/file/d/1xk6nDzvn_AbW5AdbEjnZ4NA97V4MOg0n/view)



Thank you