**Using K nearest neighbors classifier (KNN) to detect breast cancer**

**Introduction**

Machine Learning algorithms have been used for several applications. One of them in the medical field. This project applies a Machine Learning algorithm (in Python) to perform tumor classifying according to the tumor characteristics.

**Part I: the dataset**

The following libraries were imported:

*# importing libraries*

**import** sklearn

**import** pandas **as** pd

**import** numpy **as** np

**from** sklearn.datasets **import** load\_breast\_cancer

**from** sklearn.model\_selection **import** train\_test\_split

**from** sklearn.neighbors **import** KNeighborsClassifier

**from** sklearn.metrics **import** accuracy\_score

**import** matplotlib.pyplot **as** plt

The dataset was loaded from scikit-learn databases using the following code:

*# load a toy dataframe from https://scikit-learn.org/stable/datasets/toy\_dataset.html*

cancer **=** load\_breast\_cancer()

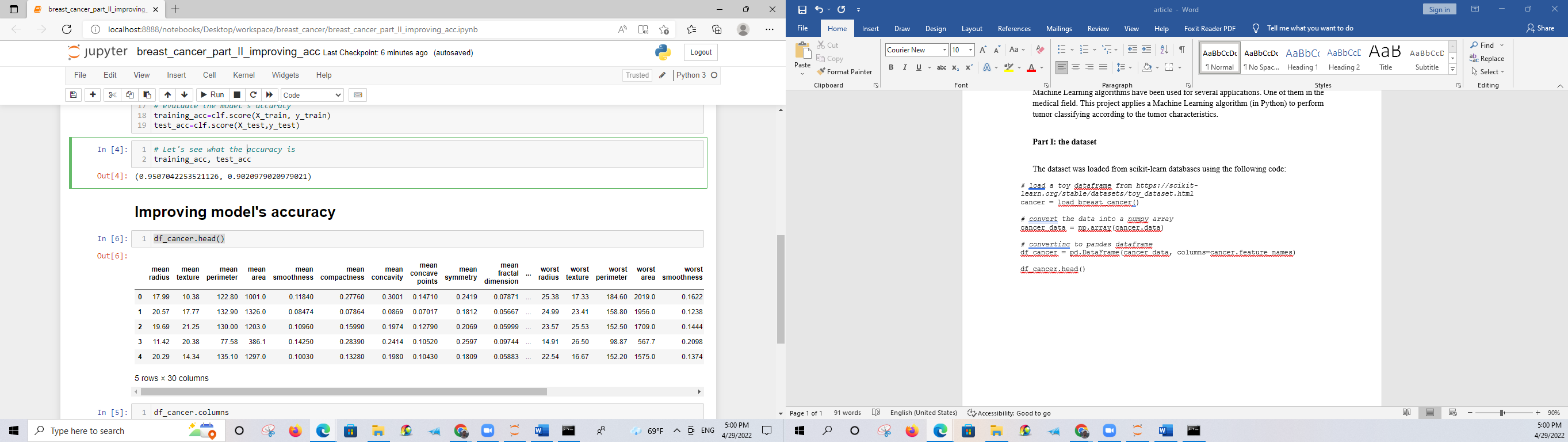
*# convert the data into a numpy array*

cancer\_data **=** np**.**array(cancer**.**data)

*# converting to pandas dataframe*

df\_cancer **=** pd**.**DataFrame(cancer\_data, columns**=**cancer**.**feature\_names)

df\_cancer**.**head()



df\_cancer**.**columns

Index(['mean radius', 'mean texture', 'mean perimeter', 'mean area',

'mean smoothness', 'mean compactness', 'mean concavity',

'mean concave points', 'mean symmetry', 'mean fractal dimension',

'radius error', 'texture error', 'perimeter error', 'area error',

'smoothness error', 'compactness error', 'concavity error',

'concave points error', 'symmetry error', 'fractal dimension error',

'worst radius', 'worst texture', 'worst perimeter', 'worst area',

'worst smoothness', 'worst compactness', 'worst concavity',

'worst concave points', 'worst symmetry', 'worst fractal dimension']

**Part II: the model**

Then the model was built using scikit-learn