**Lab2 Assignment**

The goal of this assignment is to:

Create two attack models using two different shadow models:

One that uses a shadow model with the same architecture of the target model, but different dataset.

Second is that uses a shadow model of a different architecture (a fully connected NN) with the same dataset.

Calculate the precision and recall of both attach models.-

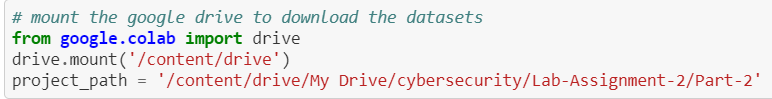
How do these attack models differ amongst each other and with the attack model created in the ICP. Discuss with numbers and statistics.

What are some possible defines mechanisms? Apply at least one that will reduce the precision and recall of the attack model by 15% while keeping the accuracy of the target model as high as possible.

Importing libraries

A screenshot of a cell phone

Description automatically generated



A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

A close up of text on a white background

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

A screenshot of a social media post

Description automatically generated

A screenshot of a social media post

Description automatically generated

A screenshot of a social media post

Description automatically generated

A screenshot of text

Description automatically generated