1. What is the actual outcome of the product that we are building? Do the users actually upload a file and our model detects whether a file contains malware? What type of malware? Probability of the type/classification of malware?

The app must give multiple solutions as in, it should detect an executable file (if it’s a malware or not ) , it should be able to check an URL , should be able to clean up an affected file just like an antivirus , classify the type and impact of malware.

Do the users actually upload a file and our model detects whether a file contains malware? What type of malware? Probability of the type/classification of malware?

1. What does the dataset look like, a piece of script as an executable file or something else, text based maybe?

Concrete attribute based on the solution we are trying to provide

1. Are there pre trained models to be used / consumed as an API (fitted into the Web App), if so, they would be more accurate and better to use one such model

Client decides on this on that day – any other decisions as well

1. What is meant by hybrid exactly? Is it using multiple models sequentially one after the other or a single model that uses multiple algorithms within itself?

Different models for different solutions

1. What does the attributes of the dataset look like? A concrete example would be better, on what attributes the model is going to be trained on and expected to predict?
2. Who is the main user of the application?  IoT Training Academy, Client under IoT academy?

Admin mainly ,

1. Behaviour Analysis of cyber-Physical system, would need a bit more info on this.
2. What is the Tech Stack to be used – Anaconda Python for AI model, web app – any framework and programming language.

Assignment 1 – Research paper review

Part A

Review and list 5 different research papers

Compare them with each other and synthesize info from them to layout a learning to fill the identified learning gap

Part B

Identify Area of Ethical breach (e.g. Data breach) – case study where there was an ethical breach and then identify questions in canvas.

Use reference templates and scenarios given in canvas.

Kaggle Dataset research more towards more meaningful dataset

Looks for datasets with many columns

Concrete columns data

Lots of records / rows

Assignment 2 Project Brief

Project Scope or Exclusions – deployment, no hands-on training on how to use the system,

Learning issue – actually refers to the different aspects / parts of the project – like web app, model, web app development framework, model training, app and model integration.

Decide on your project specs,

One type of malware or multiple type of malware – decide this and then decide your dataset

Have to justify choosing a specific hybrid model. Comparing the accuracy of different combos is a suggestion for justification