

# TSEC HACKS 2021

---

## PROBLEM STATEMENT

---



---

# MACHINE LEARNING

---

## **PS 1: Video KYC Application**

Analysis of Covid Vaccine: Earlier KYC verification was done in person, but due to COVID-19, banks are trying to implement a way to perform KYC virtually.

KYC: It means Know Your Customer, in which the organization makes some extra effort to verify the identity, suitability and risks involved with maintaining a professional relationship.

Your task is to create a web application in which the user will upload the government documents.

The ML model will try to perform some type of verification and verify that the person uploading those documents is a legit user.

### **Optional Features:**

- What should happen if the model predicts a false negative?
- Try to add any fraud detection system.

Your solution will be judged on the basis of the uniqueness of your ML approach and different types of verification methods.

## **PS 2: Invoice Processing System**

One major problem in the finance world is that invoice formats are not standard and structured.

Extracting useful information from these documents is easy for humans but a challenging task for machines.

Your task is to create a platform where a user will

- Upload their invoices in PDF/scanned image.
- ML model will extract all of the required information.
- Export the data into a JSON/CSV format.

ML Model: You need to train a model from scratch on the given dataset or you may use your own data.

Optional Features:

- The user should have an option to review and validate the captured information from the model.

Your solution will be judged heavily on the accuracy of your model.

Dataset:

<https://drive.google.com/drive/folders/1ShItNWXYiY1tFDM5W02bceHuJjyeeJI2>

## **PS 2: Invoice Processing System(contd)**

Example of important information from Invoices:

- Invoice Number
- PO Number
- Invoice Date
- Due Date
- Terms
- Subtotal
- Tax
- Amount Due
- Currency
- Vendor Name
- Vendor Address
- Customer Name
- Other

We would appreciate an efficient, scalable and accurate project.