SDE

You can choose any ONE of the two tasks and share it as per the instructions given below.

Submission:
1. Keep the codebase on Git and share the link with us.
2. Write a short Readme file on Git for the project to explain your approach
3. Share all the necessary documents and links on the same mail chain of this task.

Expectations:
1. Error handling & validation should be robust to handle any scenario.
2. Code Clarity
3. The standards you follow to write your code would play a good role in increasing your chances.
4. Estimated time for task submission is 4-5 days

Tasks:

Option 1. ML/NLP - Correlation of names and number

1. Collect around 200-300 files (images/scanned PDFs) which include ID number (some with multiple names and IDs)
   a. You can use faker package for python, document (ID) template and https://www.thispersondoesnotexist.com/ to generate your own
   b. You can use Indian Aadhar number and PAN number formats (https://www.onedpo.com/download-personal-data-for-testing/)
2. Using OCR extract textual information
3. Extract the Name and ID number linked to that name. (For eg. there is an Aadhar like ID, you need to identify the name and link the Aadhar number to the ID holder)

One way this problem can be solved is using NLP and NER. Feel free to use existing models as well.

Option 2. Crypto

1. Read about Homomorphic Encryption (HE)
   Starting Point: https://blog.nucypher.com/an-engineers-guide-to-fully-homomorphic-encryption/

2. Implement Encryption and Decryption functions in any HE/Partially HE algorithm (Lattice based preferred)
   Feel free to use existing libraries in Python/Golang

3. Demonstrate the additive and multiplicative property of the chosen algo