**10. Numerical Problem Solving:**

E0123021-Koshree VTop of Form

!pip install boto3

import boto3

import re

textract\_client = boto3.client("textract", region\_name="us-east-1")

document\_path = "bank\_statement.jpg" # Change this to your file name

with open(document\_path, "rb") as file:

image\_bytes = file.read()

# Extract text from document

response = textract\_client.analyze\_document(

Document={'Bytes': image\_bytes},

FeatureTypes=['TABLES', 'FORMS'] # Good for structured data like bank statements

)

# Collect all detected text lines

all\_text = []

for block in response['Blocks']:

if block['BlockType'] == 'LINE':

all\_text.append(block['Text'])

# Print sample extracted text

print("Extracted Lines:")

for line in all\_text[:10]:

print(line)

# Regex to capture amounts (e.g., 123.45 or 1,234.56)

amount\_pattern = re.compile(r'\d{1,3}(?:,\d{3})\*(?:\.\d{2})|\d+\.\d{2}')

transaction\_amounts = []

for line in all\_text:

matches = amount\_pattern.findall(line)

for amt in matches:

# Remove commas and convert to float

amt\_clean = float(amt.replace(",", ""))

transaction\_amounts.append(amt\_clean)

print("Transaction Amounts:", transaction\_amounts)

if transaction\_amounts:

total\_sum = sum(transaction\_amounts)

average = total\_sum / len(transaction\_amounts)

print("Total Transactions Sum:", total\_sum)

print("Average Transaction Amount:", average)

else:

print("No transaction amounts detected.")

Bottom of Form