project

January 14, 2025

```python
[142]: # Packages
       import os
       import cv2
       import numpy as np
       import pandas as pd
       import base64

       from pdf2image import convert_from_path
       from openai import OpenAI
       from dotenv import load_dotenv
       from IPython.display import Markdown, display

       # For api keys
       load_dotenv()
```

```
[142]: True
```

```python
[155]: def pdf_to_images(pdf_path):
           images = convert_from_path(pdf_path, dpi=300)
           return images

       # Preprocess image for better OCR
       def preprocess_image(image, pdf_path,i):
           img = np.array(image)
           gray = cv2.cvtColor(img, cv2.COLOR_RGB2GRAY)
           thresholded = cv2.adaptiveThreshold(
               gray, 255, cv2.ADAPTIVE_THRESH_GAUSSIAN_C, cv2.THRESH_BINARY, 11, 2
           )
           # Save the preprocessed image and create folders to store them
           os.makedirs(f"./data/{os.path.splitext(pdf_path)[0]}", exist_ok=True)
           cv2.imwrite(f"./data/{os.path.splitext(pdf_path)[0]}/
        ↪preprocessed_page_{i+1}.png", thresholded)


       folder_path = "./data/"
       pdf_files = [file for file in os.listdir(folder_path) if file.endswith(".pdf")]

       for pdf_path in pdf_files:
           images = pdf_to_images(folder_path + pdf_path)
```

```python
    # Preprocess and save image
    for i, image in enumerate(images):
        preprocess_image(image, pdf_path, i)
```

```python
[157]: client = OpenAI(
    api_key=os.environ['OPENAI_API_KEY'],
)

def ask_gpt_to_parse_financials():
    results = []
    for pdf_path in pdf_files:
        images = pdf_to_images(folder_path + pdf_path)
        png_files = []
        for i in range(len(images)):
            png_files.append(f"./data/{os.path.splitext(pdf_path)[0]}/
    ↪preprocessed_page_{i+1}.png")

        def encode_image(image_path):
            with open(image_path, "rb") as image_file:
                return base64.b64encode(image_file.read()).decode('utf-8')

        # Encode the .png files to base64
        base64_images_list = []  # List to store the base64 encoded images
        for png_file in png_files:
            try:
                encoded_string = encode_image(png_file)
                base64_images_list.append(encoded_string)  # Add to the list
            except Exception as e:
                print(f"Error encoding {png_file}: {e}")  #

        messages = [
            {"role": "system", "content": "You are a financial analyst."}
        ]

        # Iterate through the list of base64-encoded images
        for j, base64_image in enumerate(base64_images_list, start=1):  #␣
    ↪start=1 for proper part numbering
            messages.append(
                {
                    "role": "user",
                    "content": [
                        {
                            "type": "text",
                            "text": f"""
                            This is part {j} of the data for the bank statement.
```

```
                              Go through the financials and write a report on␣
    ↪whether we should give them a business loan.
                              Find insights on the total monthly deposits and␣
    ↪withdrawals.
                              Also insights on their big regular bills like rent,␣
    ↪employee salary, utilities, etc.
                              Also catch if they have any other outstanding loans!
    ↪ But overall just use common sense on if you would give them a loan or not␣
    ↪based on how they spend their money.
                              """
                    },
                    {
                        "type": "image_url",
                        "image_url": {
                            "url": f"data:image/jpeg;base64,{base64_image}"
                        }
                    }
                ]
            }
        )

        # Send the messages to the model
        response = client.chat.completions.create(
            model="gpt-4o",
            messages=messages
        )
        results.append(response.choices[0].message.content)

    return results

results = ask_gpt_to_parse_financials()
```

[158]:
```
# data_1
display(Markdown(results[0]))
```

Financial Analysis Report

Overview:

The financial statement for Bethlehem Home Health Care Agency covers the period from October 2, 2017, to October 31, 2017. It provides detailed insights into the business's cash flow, including deposits, withdrawals, and outstanding checks.

Deposit and Withdrawal Summary:

- **Total Deposits:** $87,684.11
- **Total Withdrawals:** $47,024.45
- **Card Withdrawals:** $586.27
- **Checks Paid:** $30,416.69

This results in a net positive cash flow for the month, with ending balance of $20,144.38 from an opening balance of $10,487.68.

Regular Bills and Payments:

1. **Rent:** No specific entries were noted for rent; it might be included in one of the larger check amounts.
2. **Employee Salaries:**
   - Regular checks to individuals (possible employees such as Judy Drinfeld and others) were noted.
3. **Utilities and Other Expenses:**
   - Payment to entities like The Plumbing Pro suggests regular maintenance or utilities.
   - Withdrawals such as $3,000.00 to vendors suggest operational expenses.
4. **Outstanding Loans:**
   - No explicit transactions indicate outstanding loans; however, multiple transfers between accounts could imply balancing various financial obligations.

Additional Observations:

- There are significant single transactions, e.g., $6,000.00 at month-end, potentially indicating capital expenses or inter-company transfers.
- Regular, large deposits suggest a steady inflow possibly from services or reimbursements.

Recommendation:

Based on the analysis:

- The business maintains a positive cash flow with a satisfactory end-month balance.
- The nature of transactions indicates regular business operations with no alarming spending patterns.
- The agency appears capable of handling its financial obligations, suggesting it is a reasonable candidate for a business loan.

Conclusion:

Given the available information, the company demonstrates responsible financial management, making it suitable for consideration for a business loan, assuming continued positive cash flow trends. Further details on the purpose of the loan and any future financial projections would strengthen the approval decision.

```
[159]:  # data_2
        display(Markdown(results[1]))
```

0.0.1 Financial Analysis Report

Overview The bank statement provided spans from September 1, 2019, to November 30, 2019. Here's an analysis based on the data:

Total Monthly Deposits and Withdrawals

1. **September:**
   - **Deposits:**
     - Total: > £540 from visible deposits

- **Withdrawals:**
    - Regular small debits with significant expenses like £960 on 29 Sep.
2. **October:**
    - **Deposits:**
        - Significant deposits include £790 and £860, totaling at least £1650 from major entries.
    - **Withdrawals:**
        - Regular expenses include a £730 cash withdrawal.
3. **November:**
    - **Deposits:**
        - Major deposits include £870 and £960, totaling at least £1830.
    - **Withdrawals:**
        - Regular debits with transactions like £198 on 22 Nov and £262 on 29 Nov.

Big Regular Bills

- **Rent:** There is no clear transaction labeled as rent.
- **Employee Salary:** The payment to "D ROBERTSON" could indicate salary (£960).
- **Utilities/Other Services:** Smaller frequent debits to various vendors could be utilities or operating expenses.
- **Cash Withdrawals:** Several cash transactions suggest ongoing operational cash needs.

Outstanding Loans There is no explicit transaction indicating loan repayments within this statement.

Conclusion on Loan Eligibility Pros: - **Regular Deposits:** The account has regular and significant deposits, indicating consistent income. - **Active Account:** Regular transactions suggest ongoing business activity.

Cons: - **Irregular Large Withdrawals:** Occasional large debits could affect liquidity. - **Low Current Balance:** The balance is relatively low (£2062.05), suggesting limited buffer for expenses.

Decision: Given the inflow consistency but potential cash flow issues due to significant withdrawals and a low end-of-period balance, it would be prudent to evaluate the nature of these large expenditures further. If they are strategic investments or predictable seasonal costs, approving a loan with reasonable terms may be feasible. Additional financial documents would be helpful for a more conclusive decision.

```
[160]: # data_3
       display(Markdown(results[2]))
```

0.0.2 Financial Report on Loan Eligibility

Overview This report evaluates the financial activities of the account holder from October 14, 2017, to January 13, 2018, based on their bank statements. The aim is to assess their eligibility for a business loan by analyzing deposits, withdrawals, and regular expenses.

Monthly Financial Activity

1. **Total Deposits:**

- **Credit**: $3,640.00
- Major credits include regular transfers of $1,000 credited multiple times from an external source.
2. **Total Withdrawals:**
   - **Debit**: $5,461.86
   - Significant debits include payments to various service providers and transfers to other accounts.

Key Financial Insights

1. **Regular Bills:**
   - **Jetts Kingsway Fees**: $27.90, multiple times.
   - **Health Insurance (HBF)**: $122.25, monthly.
   - **Utilities (Telstra)**: $115.00, approximately monthly.
   - Consistent internet fees to JESSICA LAING: $40.00, monthly.
   - **Insurance (SGIO)**: $56.42, regularly.
   - **Tax Payments**: $850.75 in December.
   - **Rent Transfers**: $700 transfers, labeled "Oronsay."
2. **Account Fees:**
   - Monthly account maintenance fee of $4.00.
3. **Transfers:**
   - Regular transfers of $1,000 credited from CommBank app. This indicates consistent external funding.
   - Multiple $700 debit transfers, suggesting a significant ongoing payment.

Current Financial Standing

- **Closing Balance**: $395.43 CR
- The account shows monthly deficits, with withdrawals exceeding deposits by $1,821.86 over the period.

Financial Behavior & Observations

- The account holder has consistent influxes of funds, likely external support or business income.
- There are regular expenses related to utilities, health insurance, entertainment, and potentially rent.
- No evidence of outstanding loans was explicit, though the frequent $700 transfer might indicate ongoing liabilities.
- The account is maintained frequently without overdraft, implying financial discipline despite the monthly shortfall.

Recommendations   Given the analysis, while the account holder shows fiscal responsibility by not exceeding their credit limit and making regular payments, their expenditures surpass their income. This ongoing deficit might pose a risk for loan repayment unless there's a plan to increase income or decrease expenses.

Without a clear indication of other outstanding loans, they appear to be managing their regular obligations. However, for a positive loan recommendation, it would be prudent to address the

income-expenditure discrepancy to ensure the ability to manage additional debt. A detailed look into their business or additional sources of income might provide further insights.

Therefore, it is advisable to gather more information about their business operations and liquidity before approving a business loan.

```
[161]:  # data_4
        display(Markdown(results[3]))
```

0.0.3 Financial Analysis Report

Overview: The bank statement data provided offers a financial snapshot of the account activities from April 2018 to March 2019. Based on the provided transaction details, this report aims to provide insights into the account holder's financial behavior, regular financial obligations, and the feasibility of approving a business loan.

Monthly Deposits and Withdrawals:

- **Deposits:**
    - Regular deposits often occur through IMPS transfers, indicating business income or financial support. Larger deposits are usually between 2,000 to 6,000.
    - Occasional higher deposits (e.g., 9,000 on 18-Sep-2018 and 8,320 on 18-Sep-2018).
- **Withdrawals:**
    - Frequent ATM and cash withdrawals between 100 and 2,000.
    - Regular payments to recurring expenses and personal withdrawals.
    - Transactions indicate the use of the account for daily expenses, business needs, and personal management.

Regular Financial Obligations:

- **Recurring Bills:**
    - Payments to Bajaj Finance, which could indicate an ongoing loan (e.g., 1,912 on several transactions).
    - Regular payments for medical expenses (several 75 to 80 transactions) which might indicate medical insurance or medical bills.
    - Utility payments, including mobile recharges and DTH services (200- 500 monthly).
    - Employee salary payments (e.g., 3,410 on 30-Jan-2019), indicating either small business operations or support staff.

Outstanding Loans:

- **Loan Indicators:**
    - Regular payments to Bajaj Finance suggest existing financial commitments.
    - Several transactions labeled as interest payments could point to other outstanding loans.

Financial Health and Loan Recommendation:

- **Pros:**
    - Consistent income through IMPS deposits suggests stable revenue or personal financial support.
    - Ability to manage regular financial obligations alongside occasional high-value transactions suggests the account holder has a reasonable level of financial literacy and priori-

7

tization.

- **Cons:**
  - The closing balance tends to be low, often under 100, indicating limited liquidity.
  - The presence of existing loan payments (e.g., Bajaj Finance) could restrict further borrowing capacity unless income significantly exceeds the debt obligations.

Conclusion:

Given the steady flow of deposits and the account holder's ability to meet existing obligations, they demonstrate a reasonable capacity to handle finances. However, the low average balance and current loan commitments pose a risk for additional lending. Therefore, any loan approval should factor in these aspects, potential for income growth, and the purpose of the requested loan to ensure alignment with the account holder's financial capability.

Recommendation:

- Approve the loan cautiously, setting a cap based on thorough verification of additional income potential and need for capital.
- Consider offering financial counseling to improve liquidity and savings.
- Regularly monitor loan repayments if approved to mitigate potential default risk.