Initial Plan:

A visualisation tool for personal finance management

Student: Pui Chi Angie Wong (C22087025)
Supervisor: Dr Jing Wu

Project Description

Nowadays, spending records come from various sources, such as bank statements, receipts, cash transactions, and bank transfers. While banking apps track spending, they only cover their own bank account-based transactions and ignore other sources, such as receipts, cash payments and other bank account transactions. Some banks provide spending data in Excel format, but most only offer PDF statements, making it difficult to get a full picture. This project focuses on building an application that pulls financial data from multiple sources to give a complete view of personal spending.

Since time is limited, the main focus source will be on extracting and analysing bank statements using image-to-text recognition techniques. This project will evaluate how well these techniques work in financial PDFs to ensure accurate data extraction. The study will assess the effectiveness of various text recognition methods in extracting key financial details like transaction dates, descriptions, and amounts. Additionally, the system will allow manual input to correct any errors that may arise during the extraction process.

If time allows, receipt scanning will be added to increase data coverage. Many expenses are recorded on physical receipts, which are often lost or difficult to track manually. Integrating receipt scanning to convert physical printouts into digital text will help consolidate spending from multiple sources into a single platform.

Once all financial data is extracted, it will be stored in a structured database. Then, an interactive web-based visualisation tool prototype will be created to track spending trends, highlight recurring expenses, monitor savings habits, and set budget targets. The tool will provide graphs and charts that display financial patterns over time, helping users make financial decisions and identify potential savings.

The project will go through key stages: extracting data, cleaning and organising it, analysing image-to-text recognition accuracy, structuring data in a database, and developing an interactive visualisation prototype.

Aims and Objective

Aims: To develop a comprehensive financial management tool that extracts, processes, and visualises personal spending data from multiple sources(focusing on bank statements), evaluates existing image-to-text techniques' performance in financial documents, and enables users to track financial habits, improve budgeting, and enhance financial awareness.

Core Objectives:

1. Develop a financial management tool that integrates spending from multiple sources, focusing on extracting and structuring transaction data from PDF bank statements and allowing manual input for error detection.

2. Evaluate image-to-text data extraction techniques in financial documents by assessing accuracy, strengths, weaknesses, limitations and overall effectiveness both qualitatively and quantitatively.

3. Build a structured database to organise and store financial transactions efficiently while integrating with an interactive visualisation tool.

4. Design and implement a front-end prototype that presents users' spending habits through dynamic graphs and charts, enabling budget tracking. Facilitate user interaction with financial data, including expense tracking, budget planning, and expense forecasting, to improve users' financial awareness and decision-making.

Desirable Features:

1. Offer Data Export and Report Generation – Implement functionality to export financial insights, reports, and visualisations in PDF, CSV, or Excel formats for record-keeping and external use.

2. Receipt Scanning: Enhance data coverage by integrating receipt scanning and converting physical receipts into structured digital records.

3. Multi-Level Database: Support different financial data sources, such as receipts and bank statements, to provide a comprehensive and structured overview of spending.

Risks and Mitigation:

1. Data Extraction Errors: Image-to-text recognition may misinterpret financial data.
   Mitigation: Implement manual input correction and refine algorithms for better accuracy.

2. Time Constraints: Limited time for implementing all planned features.
   Mitigation: Prioritise key functionalities such as bank statement extraction and data visualisation; add receipt scanning later if time allows.

3. Integration Difficulties: Different banks provide data in varied formats, making standardisation difficult.
   Mitigation: Focus on one type of bank statement initially and expand to other structures for greater flexibility if time allows.

Feasibility

Ethical Approval: Using personal bank statement data for extraction, so no ethical approval is required. A small sample of user feedback will be collected, which may only need a simplified ethical approval or not require one. This will be discussed with the supervisor later in the project.

Legal Issues: The project will use open-source libraries for data extraction, processing, and visualisation, ensuring compliance with legal requirements and not for businesses used in this project.

Resources & Software Requirements: No specialised hardware is needed. The project will be developed on standard computing devices. While some image-to-text techniques may require additional libraries, most can be implemented efficiently on a standard computer.

Supervisor Meetings

I plan weekly meetings for my project with my supervisor, Dr Jing Wu. During these meetings, I will discuss the work I have completed and what needs to be done in the following week. Additionally, I will consistently document my findings and progress in my report, capturing key insights from research, the challenges faced, solutions applied, and improvements implemented throughout the project timeline.

Gantt Chart

A visualisation tool for personal finance management

Task	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
Initial Plan												
Researching Image-to-text techniques												
Bank Statement Scanning												
Database Implementation												
Financial Data Analysis												
Visualisation Prototype												
Testing and Evaluation												
Refinement and Final Check												
Report Writing												
Supervisor Meeting												

Colour	Represent
	Main Task
	Weekly Tasks

Work plan

Week 1 (27/1 - 2/2) - Initial Plan

- Making an initial plan.

Deliverable: Initial Plan Document.
Milestone: Completion of project plan.

Week 2 (3/2 - 9/2) - Research on Image-to-text techniques and requirement analysis

- Start drafting the final report.
- Researching various Image-to-text techniques.
- Test and compare their performance in terms of accuracy in financial documents.

Deliverable: Initial drafting of report and research documents.
Milestone: Choose the best techniques after comparing.

Week 3 (10/2- 16/2) -Implementation of Bank Statement Scanning

- Apply the chosen image-to-text technique to extract financial data from PDF bank statements.

Deliverable: Prototype for bank statement scanning.
Milestone: Successful extraction of financial data.

Week 4 (17/2 - 23/2): Improve Accuracy and Implement Manual Input

- Enhance extraction accuracy by refining the algorithm.
- Analysis of the limitations of the current image-to-text techniques in financial documents.
- Develop a manual input feature for error correction.

Deliverable: Enhanced accuracy in extracted data.
Milestone: Data extraction and manual inputting correction.

Weeks 5-6 (24/2 - 9/3): Database Implementation

- Design and implement a structured database.
- Develop methods for querying and retrieving financial data.

Deliverable: Functional database.
Milestone: Database integrated with the backend.

Week 7 (10/3 - 16/3): Financial Data Analysis

- Implement analysis features for extracted data.
- Develop categorisation and trend identification features.

Deliverable: Analytical models for financial insights.
Milestone: Completion of initial analysis module.

Week 8 (17/3 - 23/3): Design the Interactive Visualisation Prototype

- Develop the user interface for financial visualisation.
- Implement dynamic graphs and charts for data representation.

Deliverable: Initial prototype of interactive visualisation.
Milestone: Basic user interface and visualisation features implemented.

Week 9 (24/3 - 30/3): Test and Collect User Feedback

- Conduct user testing on the prototype.
- Gather feedback on usability and functionality.

Deliverable: User feedback report.
Milestone: Completion of the initial testing phase.

Week 10 (31/3 - 6/4): Evaluate and Implement Improvements

- Analyse user feedback and identify key improvement areas.
- Implement necessary refinements to the system.

Deliverable: Improved and refined prototype.
Milestone: Enhanced functionality based on user input.

Week 11 & Easter (7/4 - 13/4): Contingency Week

- Final adjustments and adding desired features for enhancements.
- Write and finalise the project report.

Deliverable: Completed project and documentation.
Milestone: Submission-ready project and report.

Week 12 (14/4 - 20/4): Final Checking and Submission Preparation

- Conduct a final review of all project components.
- Perform debugging and last-minute refinements.
- Prepare submission documents and ensure all requirements are met.

Deliverable: Finalised project submission.
Milestone: Submission of completed project.