Web Research Agent - Technical Report

Overview

This notebook implements a **ReAct-style Web Research Agent** that performs autonomous internet research using:

- **LLM reasoning (Gemini 1.5 Flash)** for question generation.

- **Search tool (Tavily API)** for finding answers.

- **Structured summarization** and **report generation** in Markdown and PDF.

In this project, the **reasoning step** is done by Gemini when generating research questions, and **acting** is done using Tavily's web search.

Step-by-Step Explanation

1. Installation and Imports

Install and import the required libraries for:

- LLM reasoning: google-generativeai, langchain-google-genai

- Web search tool: tavily-python

- Display and export: reportlab, fpdf, colab

2. Setting API Keys

API keys are securely retrieved and set in the environment for:

- Google Generative AI

- Tavily Web Search

3. Reasoning:

generate_research_questions()

This method prompts Gemini to generate 5 deep research questions about the given topic. It uses LLM for reasoning to break down the topic into meaningful sub-questions.

```
prompt = f"""
You are a helpful AI researcher.
Generate 5 important research questions that cover different aspects of the topic: '{self.topic}'.
"""
response = llm.invoke(prompt)
```

4. Acting:

search_web_for_answers()

For each question generated, this step:

- Sends it to Tavily's web search

- Gets top 3 relevant results

- Extracts content summaries

- Maps summaries and sources to each question

5. Report Generation:

generate_report()

The report is structured in Markdown format with:

- Introduction

- Sections per research question

- Key Points and Source Links

- Conclusion

6. Orchestration:

run()

Main method to run everything in order:

1. Generate Questions (Reason)

2. Search Web (Act)

3. Compile Report (Summarize)

7. Exporting as PDF

Optional step to export the Markdown report as a styled PDF using reportlab:

● Adds headings, spacing, and paragraphs.

● Downloads directly in Colab.

Sample Interaction Flow

User Input:
- Enter the topic you want to research: "Artificial Intelligence in Education"

Agent Actions:
- Generates 5 questions
- Searches answers using Tavily
- Creates structured report

User Prompt:
- Do you want to download the report as PDF? yes

PDF downloaded successfully

How Reasoning Works

Step	Role of LLM	ReAct Role
Generate Questions	Breaks topic into sub-areas	Reason
Search Web	Not done by LLM	Act
Summarize	Not directly used here, but could	Reason
Report	Combines results into coherent structure	Synthesis