AI Text Completion Project Summary

Part 1: Building the Application
Environment Setup
- **Provider Chosen:** Hugging Face
- **Model Used:** gpt2 via transformers.pipeline
- **API Key:** Not required for gpt2 (public model)
- **Libraries Installed:** transformers, torch, accelerate

Application Details
- Built in Jupyter Notebook format
- Accepts user input (prompt, temperature, max_length)
- Sends input to gpt2 model via Hugging Face Transformers
- Displays readable responses
- Supports:
  - Running multiple prompts
  - Input validation (e.g., empty prompt)
  - Customizing temperature and max_length

Prompts Tested
- **Creative:** "Continue this story: Once upon a time…"
- **Informational:** "Summarize the water cycle…"
- **Instructional:** "Explain recursion like I'm five."
- **Poetic:** "Write a haiku about the ocean."
- **Opinion-based:** "Pros and cons of social media for teenagers?"
- **Technical:** "Difference between supervised and unsupervised learning."
- **Fun:** "What would happen if cats ruled the world?"

Part 2: Debugging and Improvements
API Errors Handled
- GPU errors on macOS (MPS) fixed by forcing CPU (device=-1)
- Deprecated args (like use_auth_token) removed

Input Handling
- App avoids crashing on empty input
- Length and temperature inputs are validated

Parameter Tuning
- Interactive control of creativity (temperature)
- Controlled output length (max_length)

Part 3: Experimentation and Evaluation
Prompt Evaluation

Prompt Type	Coherence	Creativity	Limitations
Creative	High	High	Repetitive patterns
Informational	Moderate	Low	Lacks detail
Instructional	Good	Medium	Clear but simplified
Poetic	High	High	Strong performance
Opinion-based	Moderate	Medium	Balanced, but shallow
Technical	Moderate	Low	May include factual inaccuracies
Fun	High	High	Entertaining and coherent

Observations

- Performs best with creative, poetic, or fun prompts
- Struggles slightly with technical precision and abstract logic
- Output tone is highly influenced by temperature

Suggested Improvements

- Add a fact-checking filter for critical content
- Use newer models like falcon-7b-instruct for better logic