

Python and Gen-AI Assignment

Dear Candidate,

Thank you for your application to the Junior AI/ML Engineer role at QNext AI. After reviewing your qualifications, we are pleased to inform you that you have been shortlisted for the assignment round.

Objective:

Develop a RESTful API using Flask, (You can host it on PythonAnywhere for free - <https://help.pythonanywhere.com/pages/Flask/>. You can use any other hosting provider as well.), that can summarise earnings call transcripts for companies using Gemini 1.5 Flash. The API should accept a company transcript in JSON format and return a structured summary with specific categories.

Task:

Create an endpoint in Flask at "/earnings_transcript_summary" that takes a JSON payload containing a company's name and transcript text. The API should use language processing capabilities to analyse and summarise the transcript text into defined categories.

Requirements:

1. **Endpoint:** '/earnings_transcript_summary'

2. **Input JSON format:**

```
{
  "company_name": "Reliance Industries",
  "transcript_text": "YOUR TRANSCRIPT TEXT HERE"
}
```

- company_name: Name of the company (string).
- transcript_text: Transcript text of the earnings call (string, max 20,000 tokens).

3. **Output JSON format:**

The output should contain a summarised structure as follows:

```
{
  "company_name": "Reliance Industries",
  "financial_performance": "SHORT SUMMARY OF FINANCIAL PERFORMANCE HERE",
  "market_dynamics": "SHORT SUMMARY OF MARKET DYNAMICS HERE",
  "expansion_plans": "SHORT SUMMARY OF EXPANSION PLANS HERE",
  "environmental_risks": "SHORT SUMMARY OF ENVIRONMENTAL RISKS HERE",
  "regulatory_or_policy_changes": "SHORT SUMMARY OF REGULATORY OR POLICY CHANGES HERE"
}
```

4. **Summary Categories:**

- **Financial Performance:** Summarise key financial metrics or statements about the company's recent performance.
- **Market Dynamics:** Summarise any commentary on market trends, demand shifts, competition, etc.
- **Expansion Plans:** Summarise any information on the company's plans for growth or expansion.
- **Environmental Risks:** Summarise references to environmental issues, sustainability, or ESG concerns.
- **Regulatory or Policy Changes:** Summarise any information on recent or upcoming regulatory or policy changes affecting the company.

Steps to Complete:

1. Set up Flask:

- Use Flask to build the RESTful API.
- Host the API on PythonAnywhere.

2. Processing the Transcript:

- Use a Generative AI model (such as Gemini 1.5 Flash model available via Google's AI API - <https://ai.google.dev/api?lang=python>) to parse and summarise each category from the transcript text.

3. Error Handling:

- Ensure the API handles errors gracefully, such as missing fields or empty transcript text.

4. Documentation:

- Document the API endpoint, input/output format, and any dependencies or setup requirements.

Deliverables:

- **URL to API:** Provide the URL to the hosted API (you can use any host other than PythonAnywhere). We would test the working API.
- **Codebase:** Upload the complete codebase.
- **Documentation:** Submit the documentation mentioning URL to the hosted API.

Evaluation Criteria:

- **Code Quality:** Clarity, organisation, and use of best practices.
- **Functionality:** Whether the API accurately summarises transcripts into the required categories.
- **Documentation:** Clarity and completeness of the documentation and code comments.
- **Error Handling:** Robustness in handling unexpected input or edge cases.
- **Efficiency:** Speed and resource management.

Submission Deadline:

- Submit URL to Hosted API, Codebase and Documentation by November 5, 2024

Good luck, and happy coding!