Economic Development Strategies

Data Science in Practice Fall 2023

Team Members: Hadi Abbas, Abubaker Ahmed, Shricharan Baskaran, Pranay Chopra, Andrew Claros, Yash Sanjay Dalvi, Subhranil Das, Tyler Smith

The Problem

The sponsor presented the problem faced by a lot of counties, especially in the rural areas, tribal lands, or inner cities, which lack the resources required to create Comprehensive Economic Development Strategies (CEDS) document, which is an essential requirement for the counties to obtain and retain Economic Development Districts (EDD) status. The CEDS document helps in acquiring grants from the Economic Development Administration (EDA). To solve this issue, the sponsor wanted to explore the potential of AI in its ability to make insights from several public sources and come up with better CEDS documents aiding counties with resource constraints. The Economic Development Administration (EDA) of the United States Department of Commerce is responsible for allocating funding grants to different districts depending on the information provided in their CEDS report. Our work incorporates retrieval-augmented generation (RAG) and experiments with industry-leading LLMs such as OpenAI's GPT and Google's Bard.

The Approach

The initial approach considered using AI models like GPT and Bard to answer the same question set that the sponsor had provided. Prompt engineering could be the way to go in that regard. After the initial meeting with the sponsor, we devised a better approach to the problem: Retrieval Augmented Generation (RAG). RAG worked because it is a technique for enhancing the accuracy and reliability of generative AI models with facts fetched from external sources. Here, the external sources are CEDS that we obtained from our sponsor.

The Data

The data was obtained from the official website of EDA. The sponsor also provided another set of data directly.

https://www.eda.gov/resources/comprehensive-economic-development-strategy?q=/grant-resources/tools/comprehensive-economic-development-strategy

https://www.sirpc.org/images/pdfs/CEDS-SIRPC_2020_Final_for_EDA.pdf

Detailed Results

Key Learning Experiences

Team members gained valuable insights into economic development strategies, which was crucial for understanding and addressing the sponsor's problem effectively. The team acquired hands-on experience in working with large language models (LLMs) and understanding their capabilities and limitations. Learning how to implement RAG techniques was a key aspect, allowing the team to enhance the accuracy and reliability of generative AI models by incorporating facts from external sources.

Skills Required
- AI and Machine Learning
- Data Retrieval and Processing
- Project Management

Specific Learnings
- Prompt Engineering: Initially considering prompt engineering and then transitioning to RAG demonstrated the team's adaptability and problem-solving skills.
- Cross-disciplinary Collaboration: Collaboration between team members with diverse backgrounds, combining economic development knowledge with AI expertise, was crucial for success.
- Communication Skills: Effectively communicating complex concepts and findings to both technical and non-technical stakeholders, as demonstrated in the project's success with the sponsor.

Impact

Our project was successful in leveraging LLMs to generate CEDS, which was the primary objective our sponsor had in mind at the start of the project.

Findings

One key fact uncovered during the project is the effectiveness of the retrieval-augmented generation (RAG) technique. The project was neither straightforward nor trivial, indicating that the team faced challenges and had a learning curve. This fact underscores the importance of an iterative approach in tackling complex problems, especially when dealing with both economic development strategies and cutting-edge technologies like large language models.

Conclusion and Discussion

Key Points Summary:

1. Project Objective: The project focused on using generative artificial intelligence, to generate comprehensive economic development strategies (CEDS) for Economic Development Districts (EDDs).

2. Approach Evolution: Initially, the team considered using AI models to answer a predefined question set by the sponsor. However, after an initial meeting, the approach

shifted towards retrieval-augmented generation (RAG). RAG involves enhancing generative AI models with facts from external sources, in this case, CEDS obtained from the EDA's official website and additional data provided by the sponsor.

3. Data Sources: Data for the project was sourced from the official EDA website and additional data provided by the sponsor, emphasizing the importance of external information in improving the accuracy and reliability of the generative AI models.

Continuation and Future Areas:

1. Further Model Exploration: Future teams can explore more advanced AI models, building on the success of GPT and Bard. Continuous improvement and experimentation with newer models may enhance the quality and accuracy of CEDS generation.

2. Expanded Data Sources: Consider utilizing additional CEDS for different regions and demographics for more comprehensive and diverse information, potentially improving the overall quality of generated economic development strategies.

3. Enhanced Prompt Engineering: Continue refining prompt engineering techniques to optimize the AI models for specific questions and requirements, thereby improving the relevance and usefulness of generated content.

Issues/Blockers:

1. Initial Approach Limitations: The challenges faced with the initial approach highlight the importance of evaluating different methodologies before settling on a strategy.

2. Data Format: The format or layout for CEDS varies for different regions and demographics; this directly impacts the effectiveness of generative AI models.

3. Model Limitations: While large language models (LLMs) like GPT and Bard are powerful, they have limitations. The models may not fully understand economic development's nuances or generate content lacking practical feasibility. Overcoming these limitations and ensuring that the developed strategies align with real-world economic considerations is a challenge.

Recommendations for Future Work

1. Next Anticipated Questions or Problems:

- How can we enhance the efficiency of CEDS generation by incorporating real-time economic indicators and trends into the models?
- Are there ways to tailor generated strategies for specific types of districts or industries, considering the diverse economic landscapes within different EDDs?

- What metrics and key performance indicators (KPIs) can be established to measure the effectiveness of the generated CEDS documents over time?
- How can the impact on job creation, infrastructure development, and overall economic growth be quantitatively assessed?
- How can the AI models be designed for continuous learning, adapting to new economic trends, and incorporating lessons learned from implementing previous strategies?
- Are there mechanisms to update the models based on feedback and changing socio-economic conditions?
-

2. Skills for Future Teams:

- Proficiency in natural language processing (NLP) and understanding of economic development principles.
- Familiarity with data retrieval techniques, including web scraping and API integration, to continuously update and validate information.
- Strong collaboration skills to engage with domain experts and stakeholders for refining model outputs and ensuring relevance to real-world economic scenarios.
- Understanding of policy analysis and formulation to ensure that the AI-generated strategies align with broader policy objectives at different levels of government. This skill is essential for compliance with regulatory frameworks.

3. Guidance for Successor Teams:

- Emphasize the importance of ongoing collaboration with the sponsor to understand evolving requirements and adapt the models accordingly.
- Explore possibilities for interactive interfaces that allow users to provide feedback and fine-tune generated strategies, enhancing the usability of the AI-generated CEDS.
- Instill a mindset of resilience and flexibility among team members. Economic development is a dynamic field, and the ability to adapt to changing circumstances is crucial for the project's success.
- Transforming the format of data obtained from external sources to be more consistent and easily usable by the RAG would be useful. Establish protocols for data validation and cleansing.
- Celebrate team milestones and successes to foster a positive team culture. Recognition of achievements contributes to motivation and a sense of shared accomplishment.
- Clearly define and communicate metrics for success. Regularly evaluate the project's performance against these metrics to ensure that it continues to meet its objectives and provides tangible benefits to the target counties.