Did you find or come across solutions to similar problems by using Generative AI or other sources?

Yes

1. **Generative AI**:
    - **Model Used**: ChatGPT (OpenAI).
    - **Prompts Used**:
        - "Explain the implementation of MapReduce for real-time cryptocurrency analysis."
        - "How to use Flajolet-Martin for unique item counting in a streaming dataset?"
        - "Write a Python implementation for Reservoir Sampling."
    - **Help Provided**: Assisted in understanding the theoretical background of the algorithms, writing code snippets for MapReduce, Flajolet-Martin, Reservoir Sampling, and Differential Privacy, and integrating these algorithms with a real-time pipeline.
2. **Web Sources**:
    - **Websites Referenced**:
        - Apache Kafka documentation (kafka.apache.org) for setting up producers and consumers.
        - AWS EMR documentation (aws.amazon.com/emr) for cluster configuration and running Spark jobs.
        - CoinGecko API documentation (coingecko.com/en/api) for accessing real-time cryptocurrency data.
    - **Help Provided**: Detailed information about configuration, usage, and integration of the respective tools.
3. **Team Collaboration and Mentors**:
    - Discussion with team members and faculty for feedback and insights into privacy-preserving techniques like Differential Privacy and Explainable AI.

These resources provided essential guidance and practical solutions for implementing a real-time big data pipeline and advanced algorithms effectively.