Start Planning my Capstone Project

As I progress in my Machine Learning journey, I've been exploring different directions for my capstone project. After discussing potential ideas with my mentor, we identified three promising options that align with my interests and expertise: computer vision for environmental monitoring, NLP-based sentiment analysis, and predictive analytics for sustainable urban development. Each of these projects leverages advanced AI techniques and publicly available datasets to address real-world challenges.

While these initial ideas provide a strong foundation, I remain open to refining them as I deepen my understanding and gather more insights throughout the course. The goal is to develop a capstone project that is both impactful and feasible, balancing technical depth with practical application.

1. Computer Vision: Environmental Monitoring

Focus Areas:
- Agriculture: Monitoring crop health, predicting yield, and identifying areas needing intervention.
- Disaster Prevention: Detecting early signs of natural disasters like wildfires, floods, and landslides to enable timely preventive measures.

Data Sources:
-Public datasets such as NASA's Earth Observing System Data and Information System (EOSDIS) and Sentinel-2 from the European Space Agency (ESA).
-Historical data on natural disasters for training predictive models.
Satellite Image Datasets
Satellite Image Deep Learning

2. NLP: Sentiment Analysis

- Sentiment Analysis: Develop an NLP model to analyze sentiment from social media posts, reviews, or comments. This could involve scraping data, preprocessing text, training models like BERT and/or GPT, and visualizing the results.
- Integration with connectors, APIs, and/or ETL processes.

Datasets:
- Twitter Sentiment Analysis Dataset (Kaggle)
- IMDb Reviews Dataset (Kaggle)

3. Computer Vision: Predictive Analytics for Sustainable Urban Development

Focus Areas:
-Urban Growth Prediction: Using historical data and satellite imagery to predict future urban expansion.
-Environmental Impact Assessment: Analyzing the impact of urban growth on local ecosystems and resources, such as air quality, water quality, and green space coverage.

Datasets:
Urban: Global Human Settlement Layer (GHSL), Landsat Satellite Imagery
Environmental: Air Quality Data (OpenAQ), Global Surface Water Explorer
Google Earth Engine
Awesome Public Datasets (GitHub)

Note: as I continue to explore and discover the field, these project ideas might be adjusted and refined throughout the course. This ongoing learning process will help ensure that the final projects could be interesting and also feasible.