

Proposal

Sofia Guttman

2024-04-13

Proposal for Final Project

I plan to expand the scope of the project by exploring emerging trends and innovations within the tech and AI industry. This expansion will involve investigating recent developments, partnerships, and technological advancements among key players in the market. By incorporating real-time data and industry updates, we can provide a more nuanced and up-to-date analysis of the evolving landscape.

Exploratory Topic Chosen: For this extended project, I will incorporate text analysis, network science, and unsupervised learning to delve deeper into the complexities of the tech and AI industry. Here are some key questions and potential avenues of exploration for each topic:

- Text Analysis:** How can we extract actionable insights from vast amounts of unstructured textual data, such as customer reviews and financial reports? I plan to employ natural language processing techniques to analyze sentiment, identify key themes, and extract valuable information from textual sources. Additionally, sentiment analysis will enable us to gauge customer perceptions and market sentiment towards different companies and products.
- Network Science:** What are the underlying connections and relationships between tech and AI companies in the market? By applying network science methodologies, we can construct and analyze complex networks of interdependencies, collaborations, and competitive relationships among companies. This network analysis will uncover hidden patterns, central players, and structural properties within the industry, providing valuable insights into market dynamics and competitive strategies.
- Unsupervised Learning:** Can we identify distinct clusters or patterns within the tech and AI industry based on financial performance, market presence, and other relevant metrics? Through unsupervised learning techniques such as clustering and dimensionality reduction, we can uncover hidden structures and groupings within the data. This segmentation will enable us to better understand the diverse landscape of tech and AI companies, identify emerging trends, and tailor strategic recommendations accordingly.

Final Product Deliverables: The final product will encompass a range of deliverables aimed at providing comprehensive insights and actionable recommendations for stakeholders in the tech and AI industry. These deliverables include:

Comprehensive Report: A detailed report summarizing the methodology, findings, and recommendations of the analysis. The report will provide an in-depth exploration of key trends, patterns, and insights derived from the data analysis, supported by visualizations and descriptive narratives.

Published Shiny Application: An interactive Shiny application for exploring and visualizing company data, sentiment analysis, network visualizations, and predictive modeling. The Shiny application will offer stakeholders an intuitive platform for interactive exploration and analysis of key metrics and trends within the tech and AI industry.

- **Interactive Map:** An interactive map showcasing the geographical distribution of company operations, market reach, and strategic partnerships. The map will provide a visual representation of the global footprint of tech and AI companies, highlighting key regions of activity and expansion.
- **Predictive Model:** A predictive model forecasting future market trends, company performance, and technological innovations within the tech and AI industry. The predictive model will leverage integrated data and advanced analytics techniques to generate forecasts and scenario analyses, enabling stakeholders to anticipate and adapt to changing market conditions.
- **Executive Summary:** A concise executive summary summarizing the key findings, insights, and recommendations of the analysis. The executive summary will serve as a high-level overview for busy executives and decision-makers, providing actionable insights and strategic guidance.

Project Progress Schedule: To ensure timely progress and successful completion of the final project, I have outlined a detailed schedule with specific checkpoints and milestones leading up to the May 6 deadline:

Week 1 (April 17 - April 23): Gather additional data sources (financial reports, customer reviews, social media data) and conduct preliminary data cleaning and preprocessing.

Week 2 (April 24 - April 30): Conduct initial exploratory data analysis, including descriptive statistics, data visualization, and preliminary insights generation.

Week 3 (May 1 - May 6): Implement text analysis techniques for sentiment analysis, topic modeling, and key insights extraction from textual data. Compile findings into a comprehensive report for submission by May 6.

By adhering to this schedule and allocating sufficient time for each phase of the project, I aim to ensure thoroughness, rigor, and quality in the analysis and deliverables. Regular progress

updates and communication will be maintained to address any challenges or adjustments needed along the way.