PROJECT : AI TRENDS IN TECH NEWS

Problem Statement

To analyse trends and patterns in artificial intelligence (AI) news coverage across multiple tech news sources, identifying key topics, publication patterns, and differences between sources to gain insights into the current state and evolution of AI discourse in the media.

Steps Involved:

1. Web scraping
2. Data cleaning
3. EDA
   a. Univariate analysis
   b. Bivariate analysis
4. Text Analytics
   - *An Analysis of Word Frequencies, Trends, Mentions, Sentiments, and Article Rankings.*
5. Topic Modelling
6. Project Insights/Conclusion

Data Attributes:

This data set is collected from 5 different tech news feeds, it has following features

1. Title –   Title of the AI article
2. Link –   HTML link of the AI article
3. Published –   Published date of the AI article
4. Summary –   Brief summary of the AI article
5. Source –   Tech news source of the AI article

Data preview:

Project Approach:

1. Used JupyterNotebook, Installed all the required libraries.
2. Defined customized functions to clean the html data and store it in the desired format.
3. Defined functions to scrape the data and collect all the data into 1 csv file for later use.
4. Exploratory Data Analysis is performed on the output which is now in a data frame.
5. Regular expression, Sentiment Intensity analyser are used to perform various Text analytics.
6. Latent Dirichlet Allocation[LDA model] is used to perform topic modelling.

Author: Rashmi S   9th July 2024

Data Insights and Project Report:
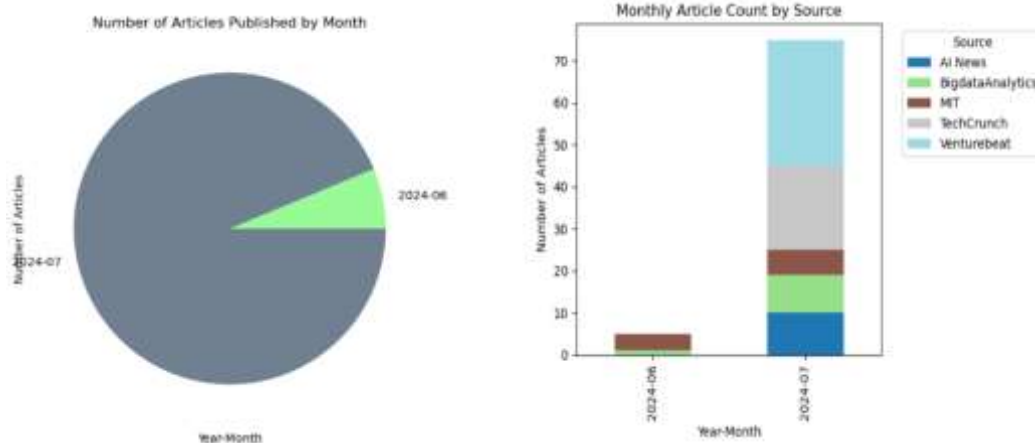
- Number of articles by source:
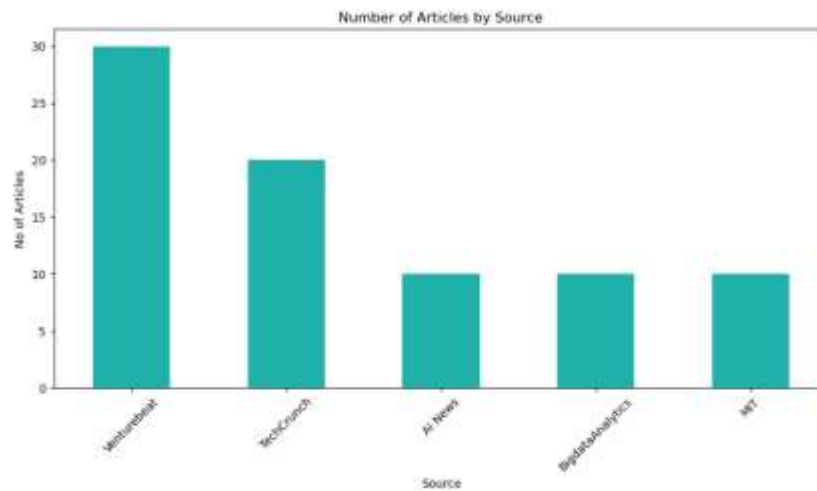
- Number of articles over time[Daily]:
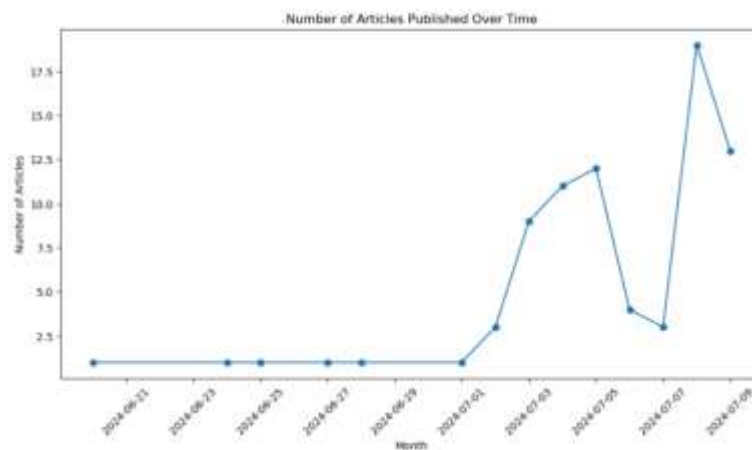
- Monthly articles over time and by source:

Author: Rashmi S   9th July 2024

TEXT ANALYTICS

1. Word Frequency

Top 10 Words in Article Titles
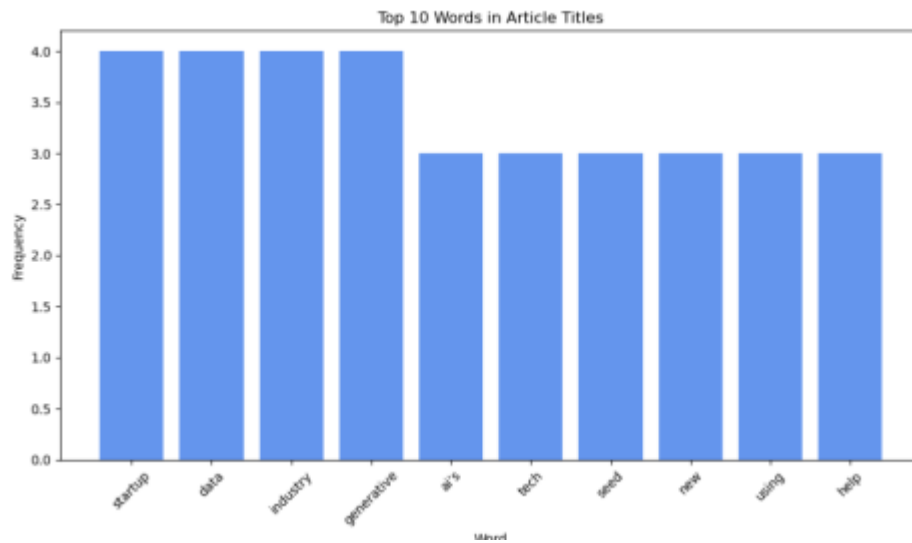
2. Big Company names found in the dataset

Mentions of Big Companies in AI Articles
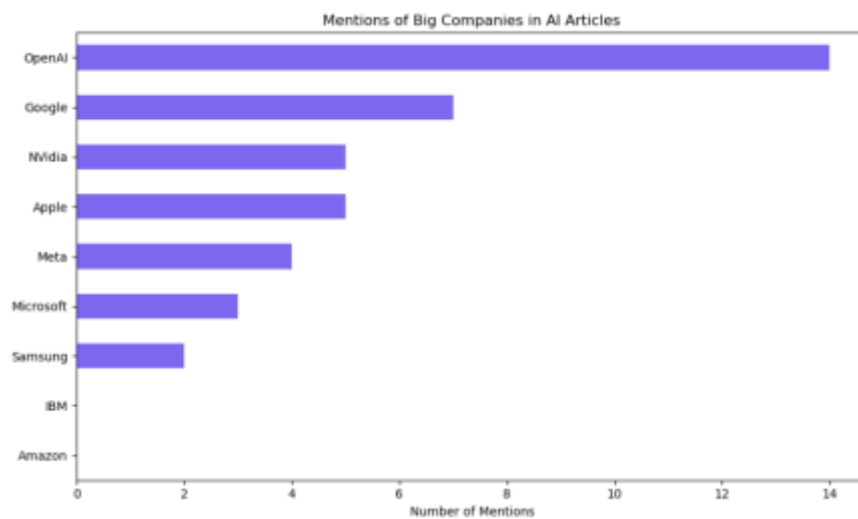
3. Specific Word Trends over time


AI-related Word Trends Over Time
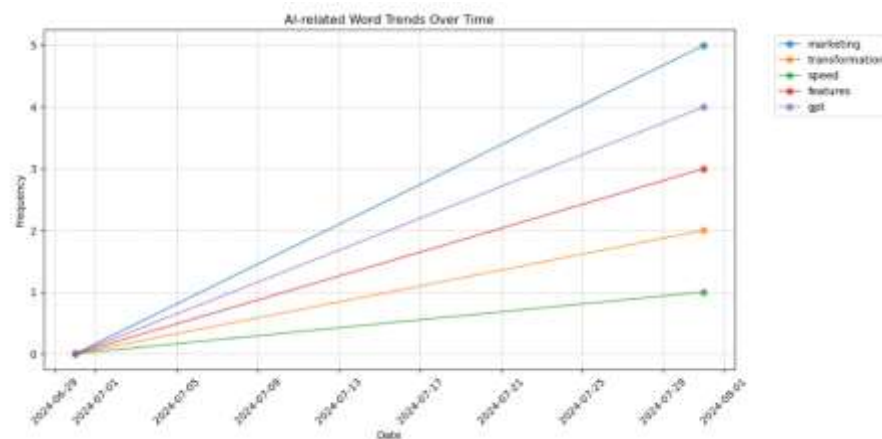
Project Insights:

1. Out of 80 articles collected, we have observed 93% of articles published in the current month (July 2024) compared to the 7% in the previous month (June 2024) indicating an upward growth in the number of articles overall.
2. Top companies[such as OpenAI, Google ,Meta ,Micro soft ,Apple etc.,] have ruled their presence in the articles indicating their involvement in AI.
3. Word Trends :
   a. *Marketing*: indicate a rising interest in marketing strategies or their increasing relevance in AI-related discussions.
   b. *GPT*: shows an increase, indicating a rising interest or mention of this particular technology in the articles.
   c. *Features:* might be reflecting a growing emphasis on discussing specific features of AI technologies or products.
   d. *Transformation*: shows a consistent increase over the given period, indicating a growing focus or interest in this topic in the tech news articles.
   e. *Speed*: This suggests that speed is one of the measure that remains a prominent and stable topic of discussion.
4. Sentiment Analysis:
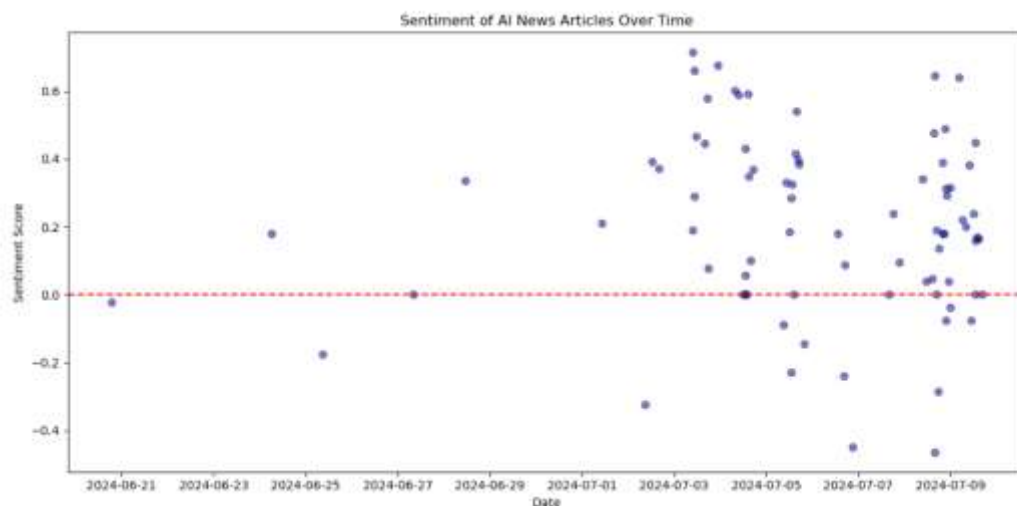
- Overall we have 68% of the articles which has positive sentiment score indicating that the content of articles mostly has appreciation towards AI.
- Neutral scores are at ~18% which is in line with the fact that news articles tend to keep a neutral opinion.

5. <u>Topic Modelling:</u>

Distribution of AI Topics
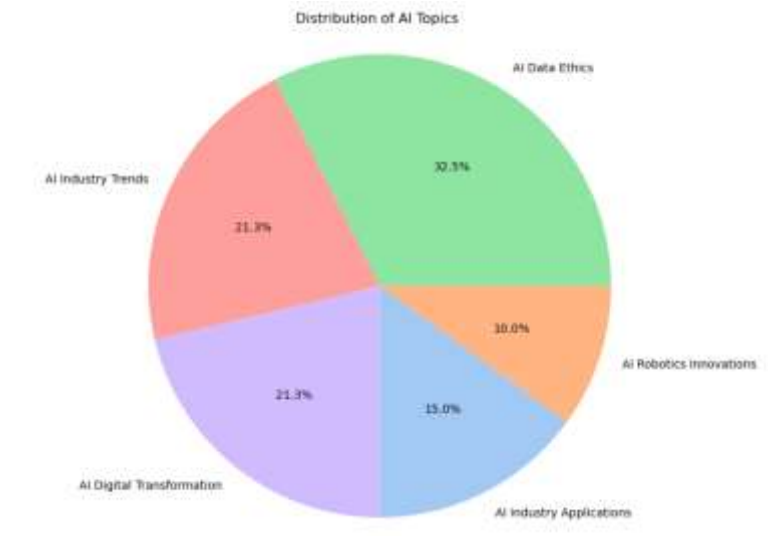
With the help of LDA topic modelling it is evident that the ethical aspect of data with respect to AI is a hot topic.

All the topics showcases a general trend AI's evolution in various areas.

Thank you

Rashmi S - LinkedIn