Identifying Jobs Impacted by AI: A Data-Driven Approach

Leveraging NLP Techniques for Insightful Analysis

Agenda

Executive Summary

Data Overview

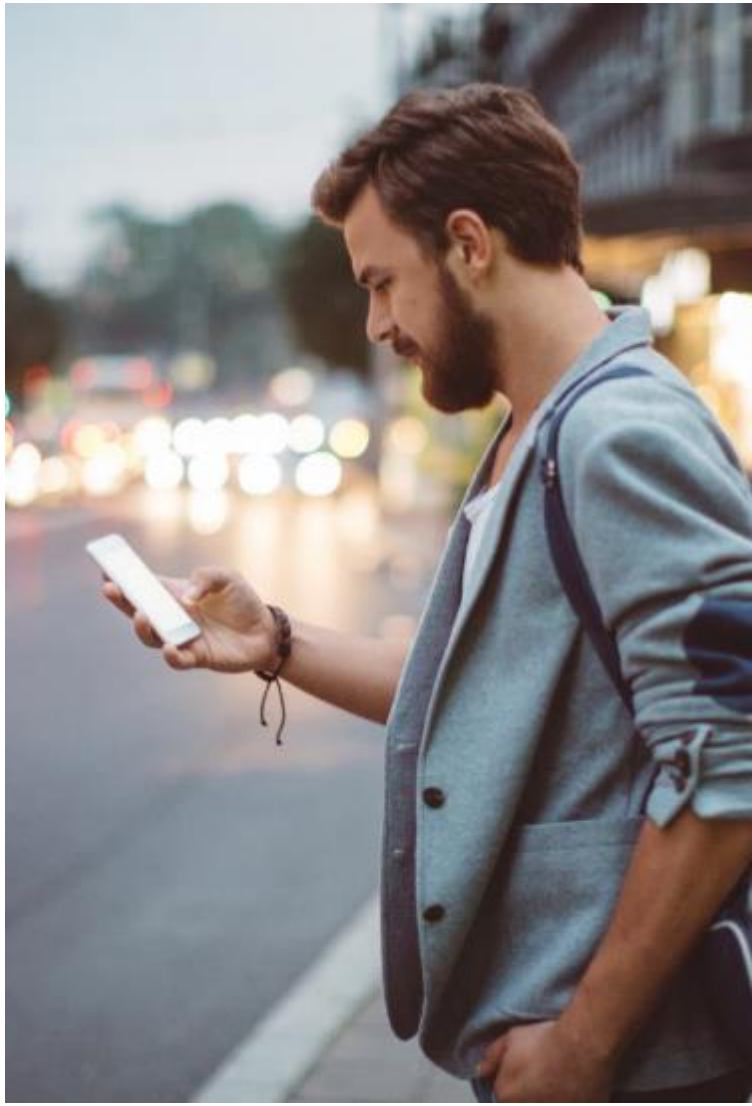
Methodology

Key topics

Sentimental Analysis

AI Job Impact Assessment

Recommendation and Conclusion



Executive Summary

•**Problem:** The need to identify jobs susceptible to disruption by AI.

•**Data:** Analysis conducted on a dataset comprising 200K news articles related to AI, Machine Learning, and Data Science

•**Methods:** Employed techniques such as Topic Modeling, Sentiment Analysis, and Named Entity Recognition (NER) for comprehensive analysis.

•**Key Insights:**

•Most prevalent discussions revolve around the automation of office jobs.

•Positive sentiment towards AI integration has been steadily increasing over time.

•**Recommendations:**

•Companies should prioritize investment in employee retraining programs.

•Policymakers should allocate funds towards transitional programs to mitigate job displacement.

Data Overview

Data Source: The dataset comprises 200K news articles focusing on AI, Machine Learning, and Data Science.

Volume: Approximately 200,000 articles are included in the dataset.

Content: Articles cover various topics within the realms of AI, ML, and data science.

Details

In March of 2023, Goldman Sachs released a report indicating that approximately 25% of tasks in the US and Europe could be automated using AI. However, as depicted in the visualization (link), not all industries will be impacted equally. Sectors such as office tasks, legal, architecture, and social sciences show potential for over 30% automation, whereas roles in construction, installation, and building maintenance are likely to be less affected.

Supporting evidence from a Facebook Research paper (link) reinforces this notion with Moravec’s paradox, suggesting that AI's biggest challenges lie in sensorimotor skills rather than abstract thought or reasoning, aligning with Goldman Sachs' predictions.

Despite the impressive findings of these papers, they are heavily influenced by recent advancements in Large Language Models (LLMs). Hence, the focus of this final project is on analyzing a collection of approximately 200K news articles (about 900 MB) on Data Science, Machine Learning, and Artificial Intelligence. The objective is to identify industries and job lines most likely to be impacted by AI over the next several years, extracting meaningful insights from unstructured text.

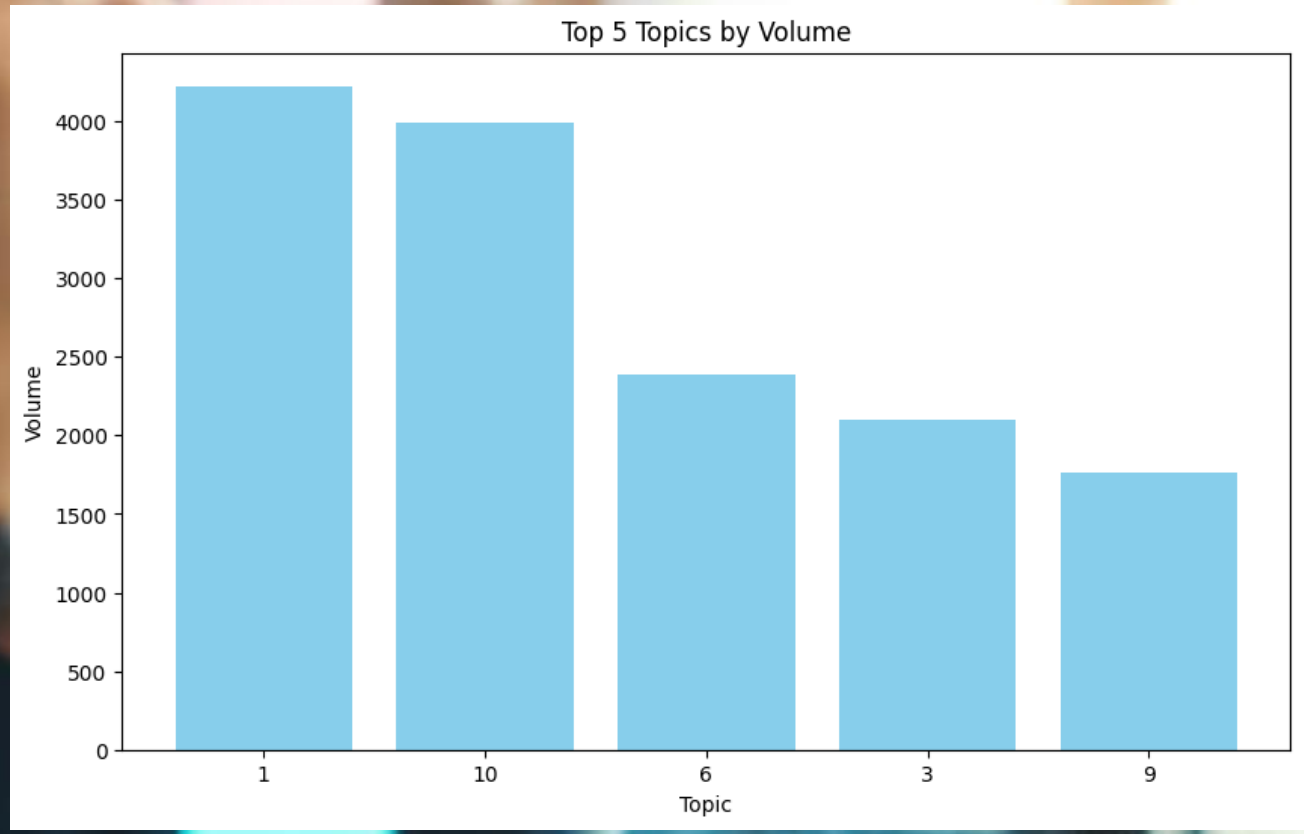
The project aims to provide actionable recommendations for leveraging AI to automate tasks and improve employee productivity, considering the introduction of novel technologies and algorithms such as AI for image generation and Conversational AI, which represent a paradigm shift in the adoption of AI technologies and data science.

Methodology

- Data Cleaning Techniques Used:
- Noise Removal: Eliminated newlines, tabs, web links, and punctuation marks from the text data to enhance readability and analysis accuracy.
- Filtering: Discarded irrelevant articles based on predefined criteria to focus on relevant content related to AI, ML, and data science.
- **Topic Modeling with LDA (Latent Dirichlet Allocation):**
- Utilized LDA to identify major topics within the dataset.
- LDA helps in uncovering latent topics by assigning each document a distribution over a fixed number of topics, and each topic a distribution over a fixed number of words.
- Resulting topics provide insights into the main themes present in the dataset, aiding in understanding the prevalent discussions.
- Utilized LDA to identify major topics within the dataset.
- LDA helps in uncovering latent topics by assigning each document a distribution over a fixed number of topics, and each topic a distribution over a fixed number of words.
- Resulting topics provide insights into the main themes present in the dataset, aiding in understanding the prevalent discussions.

Key Topics

Topic 1:
1. ai
2. data
3. technology
4. gray
5. learning
6. media
7. group
8. new
9. platform
10. cloud



Sentiment Over Time

Methodology:

Utilized the VADER (Valence Aware Dictionary and Sentiment Reasoner) lexicon for sentiment analysis.

Applied sentiment analysis techniques to compute the sentiment polarity (positive, negative, neutral) of each article using the VADER lexicon.

Calculated the compound sentiment score for each article, representing the overall sentiment.

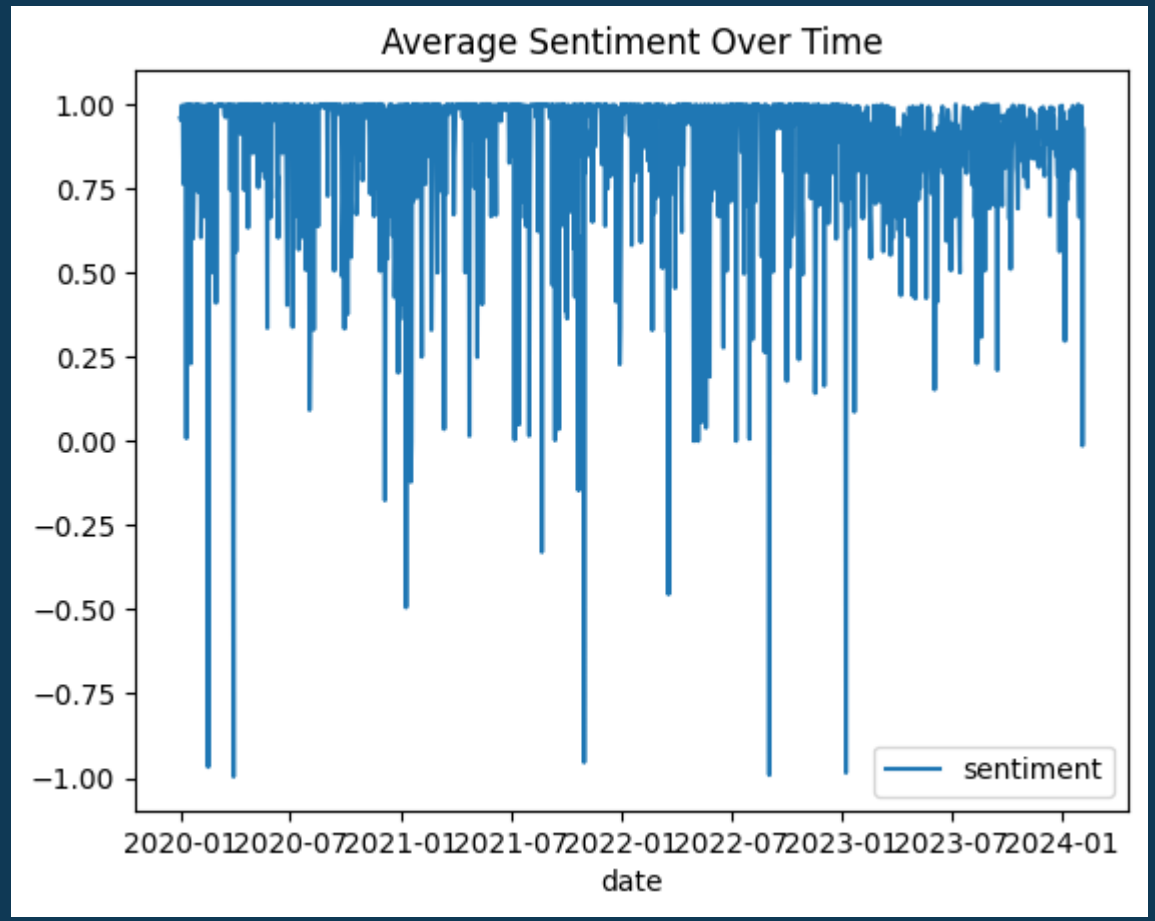
Aggregated the sentiment scores over time to compute the average sentiment per year.

Findings:

Visualization: The bar chart depicts the average sentiment score per year, providing insights into the overall sentiment trend over time.

Positive Sentiment Trend: Observation of the trend reveals a gradual increase in positive sentiment over the years.

Implications: The rising positive sentiment towards AI, ML, and data science suggests growing optimism and acceptance of these technologies, potentially indicating increased adoption and investment in related fields.



AI Job Impact Assessment

Objective: To assess the potential impact of AI on various job sectors based on the analysis of articles from the dataset.

Hypothesized Jobs Likely to be Impacted:

Office Tasks: Administrative roles, clerical work, and routine office operations may face automation due to advancements in AI-powered software and tools.

Legal Services: Certain legal tasks such as document review, contract drafting, and basic legal research could be automated with the implementation of AI algorithms and machine learning models.

Architecture and Design: Jobs related to drafting, planning, and design processes in architecture and engineering fields may witness automation through AI-powered design software and algorithms.

Social Sciences: Roles involving data analysis, research, and predictive modeling within social science disciplines could be affected by AI-driven analytics tools and methodologies.

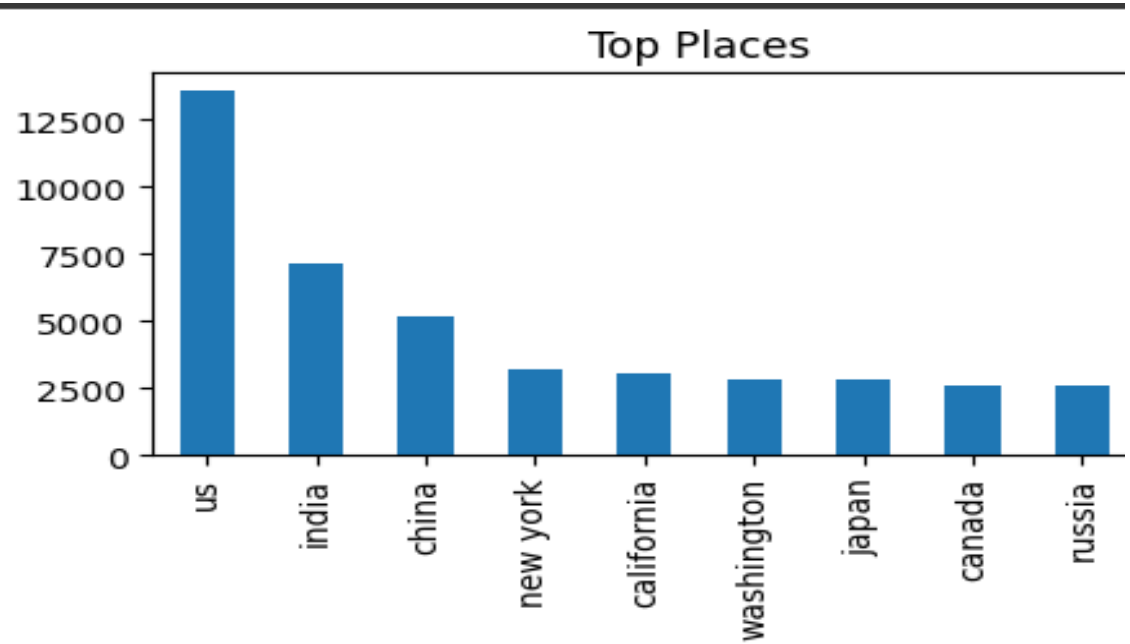
Article Analysis: Examination of articles discussing AI job impact reveals recurring themes and discussions related to the potential automation of certain job roles.

Expert Opinions: Insights from experts in the field of AI and labor economics provide qualitative evidence regarding the likelihood of job automation and its implications for various industries.

Case Studies: Analysis of real-world case studies and examples demonstrating the adoption of AI technologies in job roles supports the hypothesis of job impact across different sectors..

The assessment highlights potential areas where AI is expected to impact job roles and industries, suggesting the need for proactive measures such as upskilling, reskilling, and workforce adaptation to mitigate adverse effects and capitalize on emerging opportunities in the AI-driven economy.

Top candidates for AI integration



Word cloud for success & failures stories

Key Takeaways:

- Word clouds offer an intuitive way to visualize text data, highlighting frequently occurring words.
- Down-sampling allows us to manage computational resources while preserving the essence of the dataset.
- Exploring successful topics sheds light on areas of interest and positivity within the data.
- From the chat there are job that appeared more frequent such as: Market Research, television inc, social media, financial service and more.

Recommendations & Conclusion

Conclusion:

The adoption of AI technologies holds immense potential to transform industries, drive innovation, and improve societal well-being.

However, realizing the full benefits of AI requires proactive measures to address challenges and risks associated with its deployment.

Future Work:

Further research is needed to explore the long-term impacts of AI on employment, economic growth, and social welfare.

Continued collaboration and dialogue among stakeholders are essential to foster a shared understanding of AI's implications and to develop strategies for maximizing its benefits while mitigating risks.

Emphasis should be placed on advancing AI education, promoting diversity and inclusion in AI development, and ensuring ethical AI governance to create a more equitable and sustainable future.

Strategies for Accelerating AI Adoption:

Companies:

Invest in AI research and development to foster innovation and technological advancements.

Implement AI-driven solutions to streamline processes, enhance efficiency, and drive business growth.

Develop training programs to upskill employees and prepare them for the AI-driven future job market.

Academics:

Collaborate with industry partners to bridge the gap between academia and industry needs in AI education and research.

Offer specialized courses and programs focused on AI, machine learning, and data science to meet the growing demand for skilled professionals.

Conduct interdisciplinary research to explore the societal implications of AI and address ethical concerns.

Policymakers:

Establish regulatory frameworks to ensure responsible AI deployment and mitigate potential risks.

Allocate funding for AI research and development initiatives to support innovation and competitiveness in the global market.

Promote diversity and inclusion in AI development to address bias and promote fairness in AI algorithms.

Addressing Risks and Ethical Concerns:

Develop guidelines and standards for ethical AI design and deployment to uphold principles of fairness, transparency, and accountability.

Establish mechanisms for monitoring and auditing AI systems to detect and address biases, errors, and unintended consequences.

Foster collaboration between stakeholders, including industry, academia, policymakers, and civil society, to address ethical challenges and ensure the responsible use of AI technologies.

Thank You