Tracking the Increasing Complexity of Deep Learning Research

Team members: Vinit Nagap, Liam Pavlovic, Aniket Gupta

Problem Statement:
Deep learning, as a research field, has seen significant growth, both in complexity and volume of publications, during the last decade. This rapid growth makes it difficult for researchers in the field to keep up with current developments, leading to feelings of inadequacy and burnout. We want to visualize how the field has developed over the last decade so researchers can better understand how the field grows and impacts the people working in it. There have been some explorations into the growing volume of publications https://medium.com/voxel51/neurips-2023-and-the-state-of-ai-research-2cd2d678e264. These are good preliminary results but they do not depict the level of detail (i.e. showing growth of individual sub-fields sorted by keywords) and timescale that we will use in this project.

Proposed Solution:
The proposed solution is an interaction visualization tool that tracks the evolving complexity of deep learning research. The visualization will have a web-like structure that grows/changes with a time-based slider and shows the development of, and connections between, different subfields in deep learning research. These subfields will be identified via keywords associated with different publications. Users can select a specific subfield to see more detailed charts for that subfield (i.e. line graph of number of pubs per year, spatial map showing where authors are located, number of citations papers with this subfield receive, etc).

Data:
Will be scraped from publication repositories for various conferences (i.e. NeurIPS, CVPR, ICML, etc.) and google scholar. Attributes scraped will be fields like authors, citations, keywords, title, year of publication, and acceptance decision that are available for all papers and do not require processing their actual text.

Domain Questions:
1. How has the distribution of research topics in deep learning evolved over the years?
2. Are there specific years where certain sub-fields experienced rapid growth/rapid decline?
3. What are the common characteristics of papers accepted at major conferences? Are there subfields with higher than average acceptance/rejection rates?
4. Are there new subfields that have gained prominence in recent years?
5. Are certain subfields more popular in specific geographic regions (coarse i.e. continents)? institutions?
6. What are the relationships between different sub fields (e.g. keywords that commonly co-occur on publications)?

Roles and Responsibilities:
Vinit - Team Leader and Project Manager:

Vinit, our team leader and project manager, is in charge of project coordination and work delegation, ensuring that timelines are followed and goals are met as we construct the deep learning research visualization tool. His responsibilities also include promoting team communication and fostering a collaborative environment, all of which are critical to project success.

Aniket - Technical Lead and Data Manager:

Aniket manages the technical parts of our project, including data collecting, cleaning, and preparation. He uses advanced data analysis techniques such as topic modeling and network analysis, and he manages the technical development of the visualization tool to ensure that insights are seamlessly integrated into the project's framework.

Liam - Design Lead and User Experience Coordinator:

Liam is in charge of designing our interactive visualization tool, focusing on a user-friendly interface and receiving feedback to ensure that it meets the needs of our users. His emphasis on accessibility and visual appeal greatly contributes to the tool's success in deep learning research.

Shared Responsibility - Communication and Collaboration Specialist:

Team members collaborate on internal and external communication, providing a smooth flow of information. They also administer collaborative tools, which enable effective cooperation and information exchange.

Meeting Schedule and Communication:

The team meets regularly for updates, task assignment, and issue resolution, which promotes regular communication. Ad hoc meetings are held as needed for special issues, ensuring rapid problem resolution. Communication is primarily conducted through a dedicated Slack channel, with formal documentation and announcements sent via email.

Accountability and Participation:

Each team member is responsible for performing given duties within the agreed-upon timeframe, providing individual accountability. Regular progress updates during team meetings, as well as proactive communication about issues or delays, all help to ensure effective project management.

Expectations and Commitment:

All team members are required to actively participate in discussions and commit to setting aside enough time for individual tasks outside of meetings. Encouraging immediate assistance or clarification is critical to avoiding bottlenecks and maintaining a smooth operation.

Unavailable Dates:

Vinit - Unavailable on Tuesday, Wednesday, and Thursday.
Aniket - Unavailable on Tuesday and Thursday.
Liam — Unavailable Thursday.