Group Assignment

Requirements

Before commencing your work, it's essential that the project's theme is finalized and mutually agreed upon by your group and myself. This step ensures that your team's effort aligns with the course objectives and desired learning outcomes.

Deadline

June 21th, 2023

Presentation

Your group will be allotted 20 minutes to present your work in class. You should plan to effectively communicate your project's key points, methodologies, findings, and conclusions. Following your presentation, an additional 5-minute segment for an interactive session will include questions, clarifications, and constructive feedback from the audience and myself.

Options

For the group assignment, you have three possibilities.

Option 1: Application Development

It should address a real-world use case situation, and you must prove a justification for the necessity of your work – why is it needed and where is it applicable? Is there another approach/system that is attempting to tackle this use case? What unique value does your system bring that others lack? In essence, your task is to pitch your project as a must-have solution convincingly.

This project should be a hands-on experience culminating in a functioning product. While I understand that developing a full-fledged system like Alexa in a span of three weeks may not be feasible, your final output should still be operational. This could take the form of a demo or a proof of concept that exhibits the key functionalities of the envisioned system. For any functionality that remains unrealized by the project's end, present them as future development prospects grounded in the work you have already accomplished.

You are encouraged to utilize any tools discussed in class and additional resources available to you. However, your project should display creativity and originality - it should not merely replicate an exercise from a lab session!

An integral part of your project is to conduct a succinct yet comprehensive review of the field or industry where your system would be applied. This should include a broad understanding of the current trends, key challenges, major players, and existing solutions within this space. This review will provide important context for your system and its relevance, contributing to a more compelling argument for its necessity and potential impact.

Option 2: System/technology review

Your second choice for the assignment is to conduct a comprehensive review of a system or technology of your choosing. This review should encompass the following aspects:

- There are no restrictions on the type of system.
- **Usage Evaluation:** Undertake a detailed system examination covering all its operations and features. This exploration should yield a holistic understanding of the system's functionalities and applications.
- **Technical Evaluation:** Delve into the underlying technologies, algorithms, and procedures that drive the system. It's understood that the precise implementation details are often proprietary and thus inaccessible. To navigate this, you are encouraged to research academic literature relevant to the type of system you're reviewing. For instance, if you're studying SIRI, delve into scholarly work on chatbots and conversational systems. This should facilitate understanding common challenges and solutions in the domain, which you can then link to SIRI's functionality. A well-executed usage review could also shed light on the hidden specifics of the system's implementation. Comparing your chosen system to alternatives can further enhance your understanding.
- **Categorization:** Establish a classification system for technologies based on their characteristics and position your chosen system within this framework.
- **Identification of Limitations:** Identify key shortcomings and constraints of the system. Drawing from your user experience and expert perspectives in the field (which can be gleaned from research papers and studies), propose potential solutions to these identified issues.
- **Future Directions:** Anticipate and discuss possible upgrades or enhancements for future system iterations. These projections should be rooted in your user experience and relevant research trends in the field's literature. This future-focused segment can showcase the potential of your chosen system and shed light on the ongoing development trajectory in the given domain

The scope of your chosen system or technology should guide the structure and length of your review. Ensure it logically and thoroughly examines the chosen subject, regardless of page count.

Option 3: State-of-the-art review

Your third option is to conduct an in-depth state-of-the-art review of a specific technology, research field, or NLP system. This comprehensive analysis should include, at the very least:

- **A summary of the task's current state-of-the-art:** This overview should clearly delineate the current position of the chosen field or technology, highlighting the problems that have been solved or are close to resolution, persistent challenges, limitations of existing technology and potential future research directions.
- **Concept/Technology/System Explanation:** Provide a detailed account of the various concepts, technologies, or systems you discuss to understand their functionality, underlying principles, and the diverse approaches and algorithms they employ.
- **System Comparison:** Assess the performance of available systems in relation to the problem at hand, as well as their limitations, requirements, and applicability in real-world, production-grade scenarios.
- **Historical Review:** Compile and explain the various significant works conducted towards solving the task from its inception. This should provide a comprehensive picture of the diverse proposals, their shortcomings, unsuccessful attempts ("dead ends"), and promising research trajectories.

Before embarking on this task, I suggest you peruse some recent reviews or survey articles to get a sense of the structure, organization, and depth of coverage expected in such reviews.

Remember, it is not sufficient to merely compile a list of resources, systems, or citations. Your review must be logically structured and thorough enough to fully grasp the task, various proposals, limitations, diverse approaches to solving the problem, and future potential actions to enhance the current state. Your ultimate goal should be to provide a cohesive narrative that guides the reader through the evolution, current state, and future prospects of the chosen field, technology, or system.

Your review should be as comprehensive as necessary to represent the current state of research, technology, or system accurately and thoroughly you are examining. While brevity is appreciated, it should not compromise the quality and depth of your review. Include all pertinent details, comparisons, and analyses necessary to provide a holistic view of the subject matter. Consider the breadth of your chosen topic, and structure your review in a manner that logically and comprehensively covers all its aspects.

Research papers repositories:

- **Google Scholar:** Google Scholar is a comprehensive online search engine designed to provide access to a wide array of academic literature. It enables users to find papers through an intuitive, Google-style search interface. An additional benefit of Scholar is that it allows you to access all publications by a specific author. This feature can be particularly useful if you discover a work of interest, as it offers a gateway to explore the author's entire body of

research. Moreover, Google Scholar provides information on the publications that have cited a particular paper, which can be instrumental in discovering newer, related works.

  - **Note:** Be aware that while Google Scholar provides free access to many papers, it also lists articles from books and journals that may have restricted access..
  - **Link:** https://scholar.google.es/

- **arxiv:** arXiv is an online repository that hosts pre-prints of scientific papers that have been accepted for publication and numerous early-version papers that are yet to receive approval. These pre-prints are freely available for consultation and download, making arXiv one of the largest open-access research repositories. It has gained significant attention from academics and companies alike, with many notable researchers choosing to share their work on this platform. However, given that there is no review process and anyone can upload a paper, the repository does contain some non-relevant papers. Therefore, it's advised to investigate the authors' credentials and previous research before considering a review paper.
  - **Link**: https://arxiv.org/
- **ACL Anthology:** Managed by the Association for Computational Linguistics, the ACL Anthology is a repository that houses a large collection of noteworthy academic publications in the field of Natural Language Processing. It includes the proceedings from all ACL conferences dating back to the early 1980s, all of which are freely accessible.
  - **Note:** For the most current and relevant insights, it is recommended to focus on conference papers from 2015 onwards.
  - **Link:** http://aclweb.org/anthology/
- **IEEE Xplore**: IEEE Xplore is a research database for discovery and access to journal articles, conference proceedings, technical standards, and related materials on computer science, electrical engineering and electronics, and allied fields. It provides a powerful resource for research in technical fields where AI, including NLP, is applied.
  - **Note:** Some of the IEEE Xplore materials are free, but many require a subscription or purchase.
  - **Link:** https://ieeexplore.ieee.org/Xplore/home.jsp
- **Semantic Scholar**: Developed by the Allen Institute for AI, Semantic Scholar is a free, AI-powered research tool for scientific literature. It's designed to help researchers quickly discover relevant research.
  - **Link**: https://www.semanticscholar.org/

Bits of Advice

Option 1: Application Development:

- Adopt a 'think big, start small' approach: aspire to tackle a substantial problem (an AI doctor is more ambitious than a POS tagger), but avoid trying to establish a startup in three weeks. Begin with a modest proof-of-concept (POC) that demonstrates some features of the intended final system and serves as a testament to your vision and the system's potential.
- Position your system in the market and explain its necessity. Similarly, situate your technology within the state-of-the-art research relevant to your target sector.
- Keep in mind that this exercise focuses on the technical aspects of your system, not its marketability. While it's crucial to identify a need for your system, the primary goal is to demonstrate your competency in NLP.

Option 2: Systems Analysis:

- Investigate the underlying methodology of the system, such as the machine learning algorithm used, the datasets trained on, and the foundational research or papers.
- If the system is well-documented, attempt to implement parts of it.
- Employ the system to examine its behavior, especially in edge cases where it might fail or behave unexpectedly.
- Identify and compare similar systems.
- Conduct a brief state-of-the-art study for the task and attempt to relate it to the system.

Option 3: State of the art review:

- Start by searching for the task's name, and gradually refine your search based on the terminology and references that appear in the results.
- When you find a compelling work, seek additional papers by the same authors. Follow the trail of citations—both those referenced by the paper and those that cite the paper in their own work.
- There exist numerous state-of-the-art review articles (often called surveys) for many tasks and issues. These can provide a comprehensive overview and serve as a valuable starting point for your own review.
- Try to restrict your search to papers from reputable conferences or indexed journals. Avoid work notes and less formal communication methods.
- Remember that conference papers often contain more novel ideas than journal papers.

Submission Guidelines

State of the art review/ System Analysis:

Submit your complete report or analysis via the course platform. Ensure that all files containing your work are uploaded correctly and can be easily accessed.

Application:

Alongside the application code, you are required to submit a concise description of your application. This summary should ideally be contained within one page.

In addition, prepare a user manual that provides straightforward instructions on how to navigate and utilize your application. Think of this manual as a guide you would provide to end users to facilitate their interaction with the app.

If your software requires any specific setup, prerequisites, or configurations to function, please also include a detailed installation and execution guide. This document should cover all the necessary steps that I, or any user, would need to follow to download, install, and configure the application successfully.

If possible, you can also host your application on a server or provide a web link where it can be accessed and run directly. This could provide a more immediate and user-friendly experience for reviewing your work.