PROACTIVE AGENT: SHIFTING LLM AGENTS FROM REACTIVE RESPONSES TO ACTIVE ASSISTANCE

Yaxi Lu1, Shenzhi Yang2, 1 Department of Computer Science and Technology, Tsinghua University 3 Dec 2024  
2 Gaoling School of Artificial Intelligence, Renmin University of China

**Problem:** Large language models (LLMs) have shown remarkable capabilities in solving complex tasks,

But most existing LLM-based agents operate **reactively**, requiring explicit human instructions to initiate actions.

**Solution:** The authors argue that LLM agents should be **proactive**, capable of autonomously initiating tasks by understanding and responding to their environment

The authors propose a comprehensive data-driven approach to develop proactive agents.

**Data Collection and Annotation**: Real-world human activities are collected in three settings: coding, writing, and daily life.

**ProactiveBench Dataset**: diverse dataset named ProactiveBench is created, containing 6,790 events for training and 233 events for testing. The dataset is used to train and evaluate the proactivity of LLM agents.

**Environment Gym:** simulates events within a specified background setting and maintains the state of the environment. It generates events based on historical events and the current environmental state.

**Proactive Agent**: Predicts tasks that the user might assign based on the event history. It interacts with tools to complete specific tasks assigned by the user.

**User Agent**: This component simulates user activities and responses based on predefined user characteristics. It decides whether to accept and execute the tasks proposed by the **proactive agent.**

The most important point” development of a data-driven approach to enhance the proactivity of LLM agents, this approach enables agents to anticipate and initiate tasks autonomously, significantly improving their effectiveness in human-agent interactions.”

The fine-tuned models demonstrate significant improvements in proactiveness. The Qwen2-7B-Proactive model achieves an F1-Score of 66.47%.

Limitations

Firstly, the environment settings we have explored are still limited.

contexts in this paper provide foundational understanding but broader application areas need to be investigated to fully establish the versatility and robustness of the proactive agent.

Cons:

**Overproactiveness and User Annoyance:** If the agent initiates too many actions or interrupts at inappropriate times, it can frustrate users rather than help them.

**Computational and Resource Intensity:** **constant monitoring of user behavior, increasing computational costs and energy consumption.**

Running an always-on proactive system demands **high processing power**

**Privacy and Security Risks:** Proactive behavior often involves **continuous data collection** from users, raising **privacy concerns**.

Over-reliance on proactive agents may reduce **human autonomy**, making users dependent on AI-driven decisions.

Future research should focus on enhancing the precision and timeliness of task predictions to improve the efficacy of the proactive human-agent interaction.

**Personal conclusion**: the overuse of the proactive LLM agents will push back the user’s natural reaction on the platform. And redirect the user to different goal.

We should purposely reduce the proactive agent so the human keeps the need for it.

While testing we should favor:

• Missed-Needed: Nt = 1, Pt = ∅, the user needs help, but the agent does not provide it. 20%  
• Non-Response: Nt = 0, Pt = ∅, the user does not need help, the agent does not prompt. 25%

• Correct-Detection: Nt = 1, Pt ̸= ∅, and the user accepts the task predicted by the agent. High more than 50%

• False-Detection: Nt = 0, Pt ̸= ∅, the user does not need help but agent prompts. Should be low 5%

N = 1 (user need assistance) N = 0 (not ).

R = 1 (user accept assistance) R = 0 (user reject)

A high proactive agent will increase the expectancy of the user and will dissatisfy the user’s impression about the agent’s efficiency

ProactiveBench is a comprehensive dataset designed to enhance the proactive capabilities of large language model (LLM) agents. It comprises 6,790 events across three distinct settings: coding, writing, and daily life.

Question: how can we calculate the proactive agent’s success of the user based on the agree or disagreement of the proposed task by the agent.

For example, if the user agrees with proposed task:

In What kind of activity is favorable to use a proactive agent? And when its better to stay neutral?

In a perfect scenario

**Paper Title:** Learn-by-interact: A Data-Centric Framework for Self-Adaptive Agents in Realistic Environments

**Title: Differentially Private Data Sharing and Publishing in Machine Learning: Techniques, Applications, and Challenges**

**Authors:** HU Aoting, HU Aiqun, HU Yun, LI Guyue, HAN Jinguang. **Journal:** Journal of Cyber Security, Volume 7, Issue 4, July 2022

SUBJECT: the integration of differential privacy (DP) techniques into machine learning (ML) to protect user privacy while allowing data sharing and publishing.

The authors provide a comprehensive review of differential privacy mechanisms, their applications in ML, and the challenges that arise in implementing these techniques.

**Problem:** Machine learning models often inadvertently leak sensitive information from their training datasets, leading to privacy concerns.

The paper titled "Differentially Private Data Sharing and Publishing in Machine Learning: Techniques, Applications, and Challenges" by Hu Aoting et al., published in the Journal of Cyber Security in July 2022, explores the integration of differential privacy (DP) techniques into machine learning (ML) to protect user privacy while allowing data sharing and publishing.

### Research Background

Machine learning (ML) has become a powerful tool for analyzing large datasets and extracting valuable insights. However, the use of ML raises significant privacy concerns, as models can inadvertently leak sensitive information about the training data. Differential privacy (DP) is a robust framework designed to protect individual privacy by ensuring that the presence or absence of a single data point does not significantly affect the output of an algorithm. This paper aims to provide a comprehensive overview of how DP can be applied to ML to mitigate privacy risks while maintaining model utility.

### Differential Privacy Techniques

The paper begins by introducing the concept of differential privacy and its various definitions and mechanisms. Differential privacy ensures that the output of a query or algorithm does not reveal whether a particular individual's data was included in the dataset. The authors discuss several mechanisms to achieve differential privacy, including:

* **Laplace Mechanism**: Adds noise calibrated to the sensitivity of the query.
* **Gaussian Mechanism**: Uses Gaussian noise, which can be more efficient in certain scenarios.
* **Exponential Mechanism**: Suitable for non-numeric queries by selecting outputs based on a utility function.
* **Randomized Response**: Adds noise directly to the data before it is used in any computation.

### Privacy Threats in Machine Learning

The paper identifies several privacy threats in ML, including:

* **Membership Inference Attacks**: Determine whether a specific data point was part of the training set.
* **Model Inversion Attacks**: Attempt to reconstruct the training data from the model.
* **Attribute Inference Attacks**: Infer sensitive attributes of individuals in the training set.
* **Model Stealing Attacks**: Replicate the model's functionality without access to the original training data.
* **Unintended Memorization**: Models may memorize specific training examples, leading to privacy leaks.

### Differential Privacy in Machine Learning Models

The authors explore how differential privacy can be applied to both discriminative and generative models in ML. They discuss two primary approaches:

1. **Gradient Perturbation (DP-SGD)**: Modifies the stochastic gradient descent algorithm by adding noise to the gradients during training. This method ensures that each update to the model parameters is differentially private.
2. **Knowledge Transfer (PATE)**: Aggregates the knowledge from multiple teacher models trained on disjoint subsets of the data. The aggregated knowledge is then transferred to a student model, ensuring that the student model satisfies differential privacy.

### Challenges and Future Directions

The paper concludes by highlighting several challenges and future research directions:

* **Trade-off between Privacy and Utility**: Achieving strong privacy guarantees often comes at the cost of reduced model performance.
* **Scalability**: Applying differential privacy to large-scale ML models and datasets can be computationally intensive.
* **Adversarial Robustness**: Ensuring that models are robust against both privacy attacks and adversarial attacks simultaneously.
* **Privacy Budget Management**: Efficiently managing the privacy budget to balance privacy and utility over multiple queries or model updates.
* **Application to Specific Models**: Tailoring differential privacy techniques to specific types of ML models, such as Generative Adversarial Networks (GANs), which have unique properties and challenges.

### Summary

The paper provides a thorough analysis of differential privacy techniques and their application to machine learning models. It highlights the importance of protecting user privacy in the context of ML and discusses various mechanisms to achieve differential privacy. The authors also address the challenges and future directions in this field, emphasizing the need for further research to balance privacy and utility, improve scalability, and enhance robustness against various attacks.

**Towards Proactive Interactions for In-Vehicle Conversational Assistants Utilizing Large Language Models**

Huifang Du 14 mar 2024

explores how large language models (LLMs) can enhance the proactive interactions of in-vehicle conversational assistants (IVCAs) to improve driving safety and user experience.

**Research Background**

In-vehicle conversational assistants (IVCAs) are integral to modern smart cockpits, offering features like navigation, entertainment control, and hands-free phone operation. However, existing IVCAs often lack proactivity, struggling with user intent recognition and context awareness. This limitation results in suboptimal interactions, where IVCAs mostly passively respond to commands rather than proactively assisting drivers. Proactive interactions can significantly reduce driver cognitive load and enhance driving safety. Previous research has identified the need for better proactive behaviors in IVCAs but has not provided comprehensive interaction strategies or explored the potential of LLMs in this context.

### Problem Statement

The primary problems addressed in this paper are:

1. **Lack of Proactivity in IVCAs**: Existing IVCAs mostly react to user commands rather than proactively offering assistance.
2. **Technical Limitations**: Current IVCAs have limitations in intent recognition and context awareness, leading to poor proactive interactions.
3. **Undefined Proactivity Levels**: There is a lack of a clear framework defining different levels of proactivity for IVCAs, making it difficult to systematically study and improve proactive interactions.

### Proposed Solution

To address these problems, the authors propose the following solutions:

1. **Proactivity Framework**: They establish a framework with five levels of proactivity across two dimensions—assumption and autonomy. This framework provides a systematic way to study and evaluate proactive interactions in IVCAs.
2. **Rewrite + ReAct + Reflect Strategy**: To leverage LLMs for proactive interactions, the authors propose a strategy that involves rewriting user questions to be more formal, prompting the LLM to reason and act based on proactive interaction instructions, and reflecting on whether the generated responses meet the desired proactivity level.
3. **Capability and Subjective Experiments**: The authors conduct experiments to validate the feasibility of their approach and assess the impact of different proactivity levels on user perception.

### Proactivity Framework

The framework defines five levels of proactivity:

* **Level 1**: No assumptions; IVCAs passively receive and execute user instructions.
* **Level 2**: Some assumptions; IVCAs suggest possible solutions but require user confirmation.
* **Level 3**: Moderate assumptions; IVCAs take actions with minimal user input.
* **Level 4**: High assumptions; IVCAs offer personalized suggestions and adjust responses based on user preferences.
* **Level 5**: Strong assumptions; IVCAs execute actions automatically with explanations, allowing user intervention.

### Rewrite + ReAct + Reflect Strategy

1. **Question Rewriting**: Convert casual user questions into formal, task-oriented questions using in-context learning (ICL) with LLMs.
2. **ReAct Prompting**: Prompt the LLM to trace reasoning and execute task-specific actions, integrating external knowledge.
3. **Reflect Stage**: Assess whether the response aligns with the desired proactivity level and regenerate if necessary.

### Capability Experiments

The authors construct a dataset covering various in-car scenarios and use gpt-3.5-turbo to simulate user interactions. The success rate and proactivity attainment rate are measured. Results show that gpt-3.5-turbo achieves a success rate of 93.72%, outperforming other state-of-the-art models. Over 78% of conversations meet the desired proactivity levels.

### Subjective Experiments

An IVCA simulator is developed to test user perceptions of different proactivity levels. 40 participants interact with the simulator, and their perceptions are evaluated using a 7-point Likert scale. The results indicate that:

* **Autonomy**: Increases with higher proactivity levels (L5 perceived as most autonomous).
* **Helpfulness, Naturalness, Acceptance, Appropriateness, and Usability**: Highest for Level 4, which balances strong assumptions with user control.

### Conclusions and Future Work

The study demonstrates the potential of LLMs in enhancing proactive interactions for IVCAs. The proposed framework and prompting strategy effectively improve interaction experiences. Future work includes improving response reliability, enhancing decision-making transparency, and extending the testing scenarios and duration to provide more comprehensive insights.

This research offers valuable insights and strategies for developing more effective and user-centered proactive interactions in IVCAs using LLMs.

**Completed tasks**

* Connect Deepseek locally with elevenlab the user can hear the answers
* Create social media platform with mess­anger, create posts, and LLM assistant

Problems

* If there is a lot of users