

Estado del Arte de Sistemas Multi-agentes Basados en LLM y Predicción Bursátil con Sistema de Lenguaje



Elaborado por:

Salvador Nicolás Sanchez



Índice

Índice	2
Estado del arte de los LLM	3
“A Survey on Evaluation of Large Language Models”	3
“The imperative for regulatory oversight of large language models (or generative AI) in healthcare”	4
“Not What You’ve Signed Up For: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection”	5
Problemática	6
Planteo	6
Estructura	7
Ventajas	8
Desventajas	8



Estado del arte de los LLM

El análisis del estado del arte se centra en explorar las aplicaciones actuales de modelos multiagentes que hacen uso de los 'Large Language Models' (LLM). En esta investigación, se llevó a cabo una exhaustiva revisión de diversos papers y artículos relacionados con la actualidad de los modelos grandes de lenguaje, con el objetivo de contextualizar y comprender a fondo las aplicaciones prácticas que estos modelos ofrecen. A través de esta revisión bibliográfica, se busca proporcionar una visión integral de las características de los LLM y cómo se aplican en la práctica, destacando las innovaciones y tendencias emergentes en este campo.

"Neural Amortized Inference for Nested Multi-Agent Reasoning"

En este Artículo se presenta un enfoque de inteligencia artificial denominado Inferencia Amortizada Neural (NAI), el cual busca acelerar el proceso de razonamiento en entornos con múltiples agentes. La investigación se centra en desarrollar un método que permita realizar inferencias eficientes en escenarios complejos donde varios agentes interactúan entre sí.

El NAI se basa en entrenar redes neuronales para estimar las creencias de los agentes en diferentes niveles de razonamiento. Estos modelos reciben información sobre el entorno y las acciones previas de los agentes para estimar sus creencias actuales.

Una vez entrenados, estos modelos se usan para hacer inferencias más rápidas en cada nivel de razonamiento. En lugar de hacer cálculos exactos en cada paso, se utilizan estos modelos previamente entrenados para hacer estimaciones más rápidas.

Para estimar la distribución posterior de las creencias de los agentes, se utiliza un enfoque de muestreo de partículas. Este enfoque utiliza muestras ponderadas (o partículas) para representar diferentes estados y creencias posibles de los agentes, lo que permite una aproximación más precisa de las creencias actuales.



[“GPT-in-the-Loop: Adaptive Decision-Making for Multiagent Systems”](#)

Nathalia Nascimento, Paulo Alencar, Donald Cowan, Universidad de Waterloo, Canadá. 21/8/2023

El texto aborda el desarrollo e implementación del enfoque "GPT-in-the-loop", que combina los Grandes Modelos de Lenguaje, como "Generative Pre-trained Transformers" (GPT), con sistemas multiagente. Esta integración busca mejorar la capacidad de razonamiento y toma de decisiones de los agentes autónomos, así como aumentar la transparencia en sus procesos de toma de decisiones mediante la explicabilidad de los modelos GPT.

Se realizaron experimentos en dos escenarios: el control inteligente de farolas y la búsqueda de alimentos en la robótica evolutiva. En ambos casos, se observaron mejoras significativas en el rendimiento de los agentes autónomos cuando se aplicaba el enfoque "GPT-in-the-loop". Los resultados demuestran una mejora sustancial en la capacidad de toma de decisiones de los agentes, así como una mayor transparencia en el proceso.

Además, se destaca el potencial de este enfoque para transformar diversos campos, desde el control de sistemas inteligentes hasta la robótica evolutiva. Se plantea la necesidad de investigaciones futuras para explorar aún más estas posibilidades, incluida la diversificación de opciones de modelos de lenguaje grande y la exploración de diferentes interacciones entre GPT y demás modelos.



“MULTI-AGENT COLLABORATION: HARNESSING THE POWER OF INTELLIGENT LLM AGENTS”

Yashar Talebirad, University of Alberta y Amirhossein Nadiri, York University, 05/06/2023

El paper aborda los desafíos y limitaciones de implementar sistemas multiagentes para mejorar los modelos de inteligencia artificial basados en lenguaje. Se destaca la preocupación por la gestión de la adición dinámica de agentes, que aunque ofrece flexibilidad, podría conducir al agotamiento de recursos o a la ineficiencia del sistema. Se propone la implementación de un mecanismo de gestión de recursos y coordinación para abordar este riesgo.

Además, se señala la escalabilidad como un desafío significativo a medida que el sistema crece en tamaño y complejidad, lo que requiere recursos computacionales adicionales para manejar las interacciones y operaciones de un gran número de agentes. También se destaca la necesidad de nuevas métricas y metodologías de evaluación para medir con precisión el rendimiento del sistema y sus agentes individuales.

Por último, se subrayan las consideraciones éticas, ya que el sistema podría tomar decisiones o acciones con un impacto significativo en individuos o en la sociedad en general. Por lo tanto, es crucial diseñar el sistema de manera ética y garantizar que respete los derechos e intereses de todos los usuarios. En conclusión, aunque los sistemas multiagentes ofrecen oportunidades emocionantes, es fundamental abordar estos desafíos y limitaciones para garantizar su éxito y su funcionamiento ético y eficiente.



Problemática

En el estudio de los mercados existe un constante objetivo de lograr predecir cómo van a fluctuar los precios en la bolsa, esto ha llevado a muchos análisis de la economía con un enfoque técnico con modelos que en base a los valores anteriores de las acciones, bonos o cotizaciones buscan predecir su futuro precio, sin embargo este ámbito es muy sensible a los rumores, noticias y hechos de las personas que lo rodean, lo que hace importante un modelo que logre aportar al análisis del mercado un análisis de lo humano. En este ámbito planteamos un modelo que en base a lo que se hable en la sociedad logre predecir su impacto sobre la economía.

Planteo

Este modelo se basa en que una noticia, rumor o dicho va a generar una reacción en los remitentes que se refleja en la confianza de los bonos y por consiguiente en su valor. También suponemos que disponemos de la información de las noticias de diferentes portales, chat de diferentes grupos públicos en diferentes redes sociales y de diferentes ámbitos u otro tipo de canal donde se comparta información hacia las personas como lo pueden ser la televisión, canales de youtube y demás.

La capacidad para analizar datos provenientes de estas múltiples fuentes nos permite obtener una perspectiva más completa y contextualizada del entorno social, permitiendo evaluar no solo la magnitud de la información, sino también su impacto potencial en la confianza del mercado. En última instancia, esta aproximación integral busca mejorar la precisión de las predicciones financieras al considerar no sólo los eventos en sí, sino también la complejidad de las interacciones y percepciones que tiene la sociedad actual.

Estructura

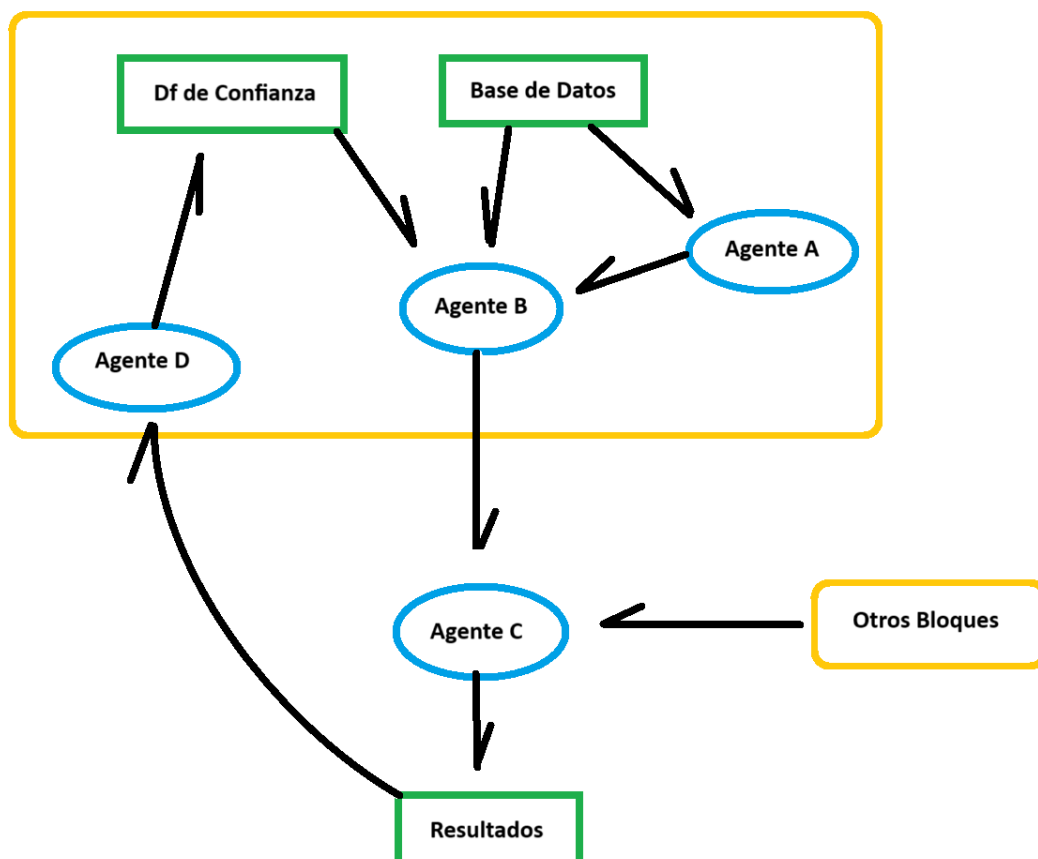
La evaluación de cada fuente, en normas generales, va a constar de tres base de datos, la principal (noticias, videos de noticieros, chats de redes sociales, etc.), una base de datos con los valores de confianza de cada ente e individuo a la hora de realizar las predicciones y una base de datos de los resultados. También dispondremos de cuatro agentes para realizar las predicciones:

Agente A: Dado unos datos de la base de datos predice en qué cantidad va a ser afectada cada tipo de cotización.

Agente B: Dado la predicción del agente A, los datos y los valores de confianza de los entes involucrados predice cuánto va a variar el valor de cada cotización individual.

Agente C: Dada la predicción de los agentes A de cada fuente disponible retorna las variaciones de cada cotización y las guarda en una base de datos de resultados.

Agente D: Dado los resultados obtenidos y la confianza de los entes involucrados actualiza estos últimos valores.





Ventajas

Este modelo posee varias ventajas:

- Escalabilidad al permitir adaptar varios tipos de fuentes al mismo proyecto.
- Permite la múltiple ejecución del código gracias a permitir complementar varios módulos.
- Al tener varios agentes permite reducir la complejidad de estos ya que parte de la información ha sido procesada previamente.
- Al tener un df de las confianzas de las fuentes permite eliminar las redundancias en futuras ejecuciones.
- Permite identificar las fuentes de información más redundantes.

Desventajas

Las desventajas que posee sería:

- Alto consumo de recursos.
- Dificultad en la adaptación de diferentes tipos de fuentes de datos.
- En el caso de usar aprendizaje automático obtenemos una dificultad en el entrenamiento.

Pese a las ventajas que brindaría en la predicción de tendencias inmediatas del mercado su alto requerimiento computacional y posibles cuellos de botella retrasarían la predicción lo suficiente para no ser rentable en este ámbito.