



HODNOTENIE KVALITY LLM A AGENTICKÝCH SYSTÉMOV: METÓDY A NÁSTROJE

Ciele:

- Preskúmať metódy na hodnotenie kvality Large Language Models (LLM) a agentov.
- Porovnať nástroje ako DeepEval, RAGAS a iné v experimentoch.
- Analyzovať efektivitu metrík pre rôzne aplikácie (napr. RAG).
- Navrhnuť automatizovaný framework na testovanie odpovedí.



AUTONÓMNI AI AGENTI PRE AUTOMATIZÁCIU VÝSKUMU

Ciele:

- Preskúmať AI agentov na analýzu akademických článkov.
- Implementovať agenta na vyhľadávanie a tvorbu reportov.
- Vyhodnotiť presnosť a spoľahlivosť systému.



VYHODNOCOVANIE SPOĽAHLIVOSTI ODPOVEDÍ V AGENTICKÝCH SYSTÉMOCH

Ciele:

- Skúmať metódy hodnotenia kvality odpovedí AI agentov.
- Porovnať metriky: BLEU, ROUGE, METEOR, RAGAS.
- Navrhnuť systém na automatické hodnotenie v multi-agentnom prostredí.