

Tato práce se zaměřuje na vývoj a zdokonalování návrhu tzv. task-oriented dialogových systémů v rychle se rozvíjejícím prostředí výzkumu umělé inteligence a zpracování přirozeného jazyka.

Navrhujeme techniky, které mohou podstatně snížit náklady na vývoj a nasazení těchto systémů, což je motivováno snahou o jejich větší přizpůsobivost a škálovatelnost.

V práci představujeme několik nových přístupů k dosažení těchto cílů.

Nejdříve představujeme automatickou metodu anotace dat, která dokáže extrahovat sadu sémanticky koherentních konceptů (dialogových slotů) z prostého přepisu zaznamenaných konverzací.

Tímto přístupem snižujeme množství manuální anotace potřebné pro porozumění přirozenému jazyku v dané doméně a výrazně tak zefektivňujeme proces vývoje.

Zkoumáme také modely využívající latentní proměnné v modelování task-oriented dialogových systémů.

Tato oblast je do značné míry neprobádána.

Modely využívající latentní proměnné nabízejí možnost využití neanotovaných dat s potenciálem odhalit strukturu vzorců chování pozorovaných v dialogu.

Toho lze dosáhnout prostřednictvím analýzy latentního prostoru a porovnání s akcemi provedenými modelem.

Dále zkoumáme potenciál těchto modelů pro vytváření hierarchických reprezentací pomocí námi navržené architektury.

V návaznosti na nedávný pokrok v této oblasti také využíváme schopnosti předtrénovaných velkých jazykových modelů (LLM) pomocí metody tzv. in-context learning, tedy učení se z kontextu.

Zkoumáme, jak snadné je pro jazykové modely aplikovat znalosti získané tréninkem v jedné doméně na dříve neviděných datech.

Námi navržená metoda založená na učení z kontextu obohaceném o příklady dosahuje pozitivních výsledků s použitím pouze několika tréninkových příkladů.

Ukázala se také jako velmi slibná při hodnocení v interakcích s lidmi.

Tento způsob použití modelů představuje podstatný skok v efektivním využívání výpočetních zdrojů k trénování konverzační umělé inteligence.

To nás přibližuje k flexibilnějším a univerzálnějším systémům.