

Tesi di Dottorato

Knowledge Base Construction Methods for Neuro-Symbolic IA Exploitation

Diego Rincon-Yanez

Abstract

La rappresentazione della conoscenza è stata a lungo considerata un compito chiave nelle scienze dell'informazione. Al giorno d'oggi, con l'ampia portata dell'intelligenza artificiale in tutti gli aspetti della vita quotidiana, capire come rappresentare, visualizzare e analizzare la conoscenza umana è diventato sempre più vitale. Questa tesi si basa sulla capacità intrinseca dei sistemi computazionali di fornire risposte complete e accurate utilizzando una combinazione di apprendimento induttivo, comunemente noto come deep learning, e apprendimento logico; l'argomento è l'IA neuro-simbolica (NeSyAI). Per utilizzare appieno le caratteristiche del sistema NeSyAI, è necessario costruire basi di conoscenza, che sono una raccolta di istruzioni ordinate per formare una struttura di rete di tipo grafo, nota anche come knowledge graph (KG) nella comunità semantica. I knowledge graph possono offrire informazioni di base su uno o più domini, aggiungendo ulteriore contesto all'attività da svolgere. Questa tesi offre numerosi metodi per integrare la conoscenza contestuale nei sistemi di intelligenza artificiale neuro-simbolica producendo basi di conoscenza semanticamente migliorate per un successivo sfruttamento.

Inizialmente è stato esplorato l'Internet of Things (IoT) reale, dove le caratteristiche della rete di sensori sono state mappate in un'ontologia tradizionale di tipo OWL utilizzando e costruendo una base di conoscenza del sensore, degli eventi e delle osservazioni con l'obiettivo di prevedere eventi e luoghi. A questo scopo, è stata creata un'ontologia all'avanguardia per scopi di riconoscimento degli eventi IoT, seguita dall'implementazione di una rete di neuroni convoluzionali, in grado di prendere il KG e scoprire nuove informazioni sulle onde. Questi due compiti, combinati in un unico flusso, producono e convalidano un nuovo modello architettonico per l'IA neuro-simbolica. Per convalidare l'approccio proposto, sono stati implementati due casi d'uso: un vulcano che scopre eventi sismici e un ambiente di monitoraggio di una camera bianca per scoprire eventi di inquinamento.

Nella seconda parte, gli incorporamenti dei grafici della conoscenza (KGE) sono stati sfruttati per ampliare i confini del compito di costruzione della base di conoscenza in diversi percorsi, come la convalida della

conoscenza, la risposta alle domande e la previsione dei collegamenti dalla prospettiva dell'IA neuro-simbolica. Questi compiti sono stati utilizzati insieme a metriche ben note per testare gli approcci di costruzione di testi prosaici sfruttati dalla conoscenza esterna di uno dei più grandi open KG disponibili, come DBpedia e modelli di tipo trasformatore, arrivando a risultati convincenti rispetto all'attuale SOTA, consentendo di testare anche modelli di architettura AI NeSy nuovi e migliorati. D'altro canto, è stata sollevata la necessità di esplorare la generazione di dati sintetici considerando la necessità di sperimentazione continua e di automazione per testare gli scenari di sperimentazione. Infine, è stato concepito un approccio iniziale di spiegabilità per KGE e applicato al caso d'uso della previsione dei flussi commerciali di materie prime.