

یادگیری ماشین برای رابطه‌های بین داده‌هایی حجیمی که روزانه توسط بازیابی اطلاعات، محاسبات زیست‌شناسی، پردازش زبان طبیعی و... تولید می‌شوند، به مساله‌ی سختی تبدیل شده است. روش‌های زیادی برای کار با این داده‌ها معرفی شده است که پایگاه‌های دانش بزرگ را تحلیل و روابط موجود در آن‌ها را استخراج می‌کند. یک دسته از این روش‌ها روش‌های مبتنی بر ویژگی‌های پنهان است که مساله را وارد یک فضای برداری چند بعدی کرده و پس از آن سعی می‌کنند با کاهش تعداد پارامتر، مساله را حل کنند. قدرت این روش‌ها در سادگی آموزش مدل، تعداد کم پارامترهایی که نیاز است آموزش دیده شوند و کار روی پایگاه‌های دانش با مقیاس زیاد است. این مدل‌ها روابط موجود بین موجودیت‌ها را به خوبی تشخیص داده و پس از آموزش به دسته‌بندی خوبی از این موجودیت‌ها دست پیدا می‌کنند اما در مورد پیوندهای بین موجودیت‌ها اینطور نیست و روابط بین پیوندها به خوبی تشخیص داده نمی‌شود. در این پژوهش قصد داریم به کمک قوانین انجمنی روابط بین پیوندها را نیز تشخیص دهیم و در آموزش مدل دخیل کنیم. به کمک راه‌حل ارائه شده در این پژوهش توانستیم روش‌های رسکال، NTN، Structured Embedding، TransE و TransH بهبود داده و این بهبود را توسط دو معیار کیفیت اثبات نموده‌ایم. به‌علاوه چارچوبی ارائه داده‌ایم که به کمک آن می‌توان نقاط ضعف و قوت روش‌های موجود را بررسی کرد. علاوه بر بهبود روش‌های موجود، این راه‌حل منجر به استخراج حقیقت‌های پنهان در پایگاه‌های دانش شده و می‌توانیم یادگیری مدل‌ها را روی پایگاه‌های دانش کامل‌تری انجام دهیم.