دانشگاه شهید بهشتی

دانشکده مهندسی و علوم کامپیوتر

اطلاعیه دفاع

نام دانشجو: **امیر حلاجی بیدگلی**		**نام استاد راهنما:** **دکتر حامد ملک**
مقطع: کارشناسی ارشد	**رشته: مهندسی کامپیوتر**	**گرایش: هوش مصنوعی و رباتیکز**
نوع دفاع: • **دفاع پروپوزال** □ • **دفاع پایان نامه** ■ • **دفاع رساله دکترا** □		**تاریخ: ۱۴۰۴/۰۲/۰۹** **ساعت: ۱۵:۳۰** **مکان: اتاق ۲۰۰**

عنوان: استفاده از مدل‌هاي زباني پروتئینی جهت بهبود عملکرد مسائل پیش‌بینی تداخل دارو-هدف

داوران خارجی: دکتر چنگیز اصلاحچی	**داوران داخلی: دکتر آرمین سلیمی بدر**

چکیده:

کشف دارو یکی از فرآیندهای پیچیده، زمان‌بر و پرهزینه در علوم زیستی است. روش‌های سنتی به آزمایش‌های گسترده و سرمایه‌گذاری کلان نیاز دارند که مانعی برای توسعه سریع داروهای جدید محسوب می‌شود. در این میان، هوش مصنوعی به‌ویژه مدل‌های مبتنی بر مبدل، با کاهش زمان و هزینه‌ها، تحول چشمگیری در این حوزه ایجاد کرده‌اند. یکی از چالش‌های اساسی در کشف دارو، پیش‌بینی تعاملات دارو و هدف‌های پروتئینی است، زیرا این تعاملات تأثیر مستقیمی بر اثربخشی و ایمنی داروها دارند.

مدل‌های مبدل در ابتدا برای پردازش زبان طبیعی توسعه یافتند و باعث تحول چشمگیری در درک و پردازش داده‌های متنی شدند. توانایی این مدل‌ها در استخراج ویژگی‌های عمیق و ایجاد نمایش‌های معنایی غنی، باعث شد که در سایر حوزه‌های داده‌محور نیز مورد توجه قرار گیرند. در مسائل مربوط به تداخل دارو و هدف، اطلاعات زیستی داروها معمولاً در قالب رشته‌های مولکولی **(SMILES)** نمایش و پروتئین‌ها نیز با استفاده از ساختار **FASTA** نمایش داده می‌شوند. تحلیل این رشته‌ها به‌صورت متنی امکان استخراج ویژگی‌های مهمی را فراهم می‌کند که می‌توانند در پیش‌بینی تداخل‌های دارو-هدف مورد استفاده قرار گیرند. استفاده از مدل‌های مبدل برای پردازش داده‌های زیستی، امکان تحلیل دقیق‌تر و جامع‌تر تعاملات مولکولی را فراهم کرده است.

برای حل این چالش، یک معماری نوین مبتنی بر مبدل برای مدل‌سازی زبانی مولکولی و پروتئینی توسعه یافته است. در این روش، رشته‌های مولکولی و پروتئینی به‌عنوان ورودی‌های متنی در نظر گرفته شده و با استفاده از مدل‌های زبانی پردازش می‌شوند. این مدل‌ها قادر هستند نمایش‌های جاسازی کارآمدی از ساختارهای مولکولی و پروتئینی استخراج کنند. این بردارهای جاسازی امکان مدل‌سازی وابستگی‌های پیچیده میان داروها و پروتئین‌ها را فراهم کرده و به بهبود دقت پیش‌بینی ضریب وابستگی در مسائل تداخل دارو-هدف کمک می‌کنند. بهره‌گیری از مدل‌های مبتنی بر مبدل امکان تحلیل دقیق‌تر ویژگی‌های زیستی را فراهم کرده و موجب بهینه‌سازی فرآیند کشف دارو می‌شود.

علاوه بر این، یک مدل پیش‌بینی ضریب وابستگی نیز توسعه یافته است که با بهره‌گیری از ویژگی‌های استخراج‌شده توسط مدل‌های زبانی مولکولی و پروتئینی، دقت بالاتری را در پیش‌بینی این ضریب ارائه می‌دهد. عملکرد مدل پیشنهادی با معیارهای میانگین مجموع مربعات و شاخص تطابق ارزیابی شده و نتایج نشان می‌دهد که مدل معرفی‌شده در مقایسه با روش‌های خط پایه بهبود داشته است. این بهبود، نشان‌دهنده کارآمدی استفاده از مدل‌های هوش مصنوعی در تجزیه‌وتحلیل تعاملات دارو و هدف‌های زیستی است.

نتایج این پژوهش نشان می‌دهد که استفاده از مدل‌های مبتنی بر مبدل می‌تواند دقت پیش‌بینی تعاملات دارو-هدف را افزایش داده و به بهینه‌سازی فرآیند کشف دارو کمک کند. این روش نه‌تنها روند توسعه داروهای جدید را تسریع می‌کند، بلکه مسیر تحقیقات آتی در زمینه هوش مصنوعی و بیوانفورماتیک را نیز هموار می‌سازد.