****

**دانشگاه شهید بهشتی**

**دانشکده مهندسی و علوم کامپیوتر**

**اطلاعیه دفاع**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **نام استاد راهنما:**  **دکتر مهرنوش شمس‌فرد** | | | **نام دانشجو:**  **سارا بارونی** | |
| **مقطع:** | **گرایش:**  **هوش مصنوعی** | | | **مقطع:**  **کارشناسی ارشد** |
| **تاریخ: ۲۳ شهریور ۱۴۰۱** | | | **نوع دفاع:**   * **دفاع پروپوزال □** * **دفاع پایان نامه □** * **دفاع رساله دکترا □** | |
| **ساعت: ۱۴ تا ۱۵** | | |
| **مکان:** | | |
| **عنوان رویکردی ترکیبی در ساده سازی متن فارسی** | | | | |
| **داوران داخلی:**  **دکتر علیرضا طالب پور** | | **داوران خارجی:**  **دکتر بهروز مینایی** | | |
| **چکیده:**  زبان طبیعی اغلب از یک سری سازه‌های پیچیده تشکیل می‌شود . این ساختارهای پیچیده نه تنها برای  انسان‌ دشوار است، بلکه برای سیستم‌هایی که ورودی آن‌ها وابسته به زبان طبیعی است هم می‌تواند  منجر به خطا در خروجی شود. ساده‌سازی راه‌حلی برای برطرف کردن این پیچیدگی‌هاست. درواقع ساده‌سازی یک فرایند است که در آن جمله‌های پیچیده از نظر دستوری و یا واژگانی ساده می‌شوند. این  به گونه‌ای است که معنای جمله‌ی اصلی به درستی منتقل  شود. ساده کردن متن به عنوان یکی از فعالیت‌های پردازش زبان طبیعی می‌تواند عملکرد  بسیاری از سیستم‌هایی که ورودی آنها زبان طبیعی است مانند تجزیه‌گرها، ماشین‌های ترجمه، خلاصه‌سازها و  همچنین سیستم‌های بازیابی دانش را بهبود بخشد. در شرایطی که به دلیل پیچیدگی، درک متن برای دسته‌ای از  انسان‌ها مانند زبان آموزان دشوار است، می‌تواند مفید باشد. در این تحقیق ابتدا یک تعریف اولیه برای پیچیدگی، هم از دیدگاه ماشین و هم انسان ارائه شد. سپس الگوریتمی مبتنی بر قاعده برای تشخیص عامل‌های پیچیدگی برای انسان معرفی شد. همچنین یک مدل تشخیص پیچیدگی برای تجزیه‌گر وابستگی استانزا پیاده سازی شده است. این مدل ترکیبی از شبکه‌های با حافظه طولانی کوتاه مدت و شبکه‌های پیچشی است. در ادامه به ساده‌سازی ساختارهای پیچیده در دو بخش واژگانی و نحوی پرداخته شد. در بخش واژگانی مخاطب دانش‌آموز است، که ضمن تشخیص واژه‌های پیچیده، الگوریتمی برای جایگزینی کلمات ساده بجای کلمات پیچیده معرفی شد. در بخش ساده‌سازی نحوی هدف بهبود عملکرد تجزیه‌گرهای وابستگی است، که یک شبکه‌ عصبی عمیق مبتنی بر ویرایش است. این مدل، درواقع بهبودیافته‌ی یکی از بهترین مدل‌های ساده‌سازی در زبان انگلیسی بر روی زبان فارسی است. | | | | |