

امروزه با توجه به گسترش روزافزون اطلاعات در فضای وب، پاسخ‌گویی به نیاز اطلاعاتی کاربران مانند نمایش اسناد مرتبط به کاربر به یک چالش اساسی تبدیل شده است. سیستم‌های پیشنهاددهنده برای حل این مشکل مورد استفاده قرار می‌گیرند. این روش‌ها را به طور کلی به سه دسته می‌توان تقسیم‌بندی کرد: سیستم‌های پیشنهاددهنده مبتنی بر محتوا، سیستم‌های پیشنهاددهنده هم‌کارانه و سیستم‌های پیشنهاددهنده مبتنی بر اطلاعات شخصی. البته می‌توان سیستم‌های پیشنهاددهنده ترکیبی را هم به عنوان دسته چهارم در نظر گرفت که حاصل ترکیب بیش از یک نوع سیستم پیشنهاددهنده می‌باشد.

تمرکز این پایان‌نامه بر روی سیستم‌های پیشنهاددهنده مبتنی بر محتوا می‌باشد و به دنبال معرفی روشی برای بهبود اینگونه سیستم‌ها در محیط‌های چندزبانه متنی هستیم. اینگونه سیستم‌ها براساس محتوایی که از کاربر بدست می‌آید پیشنهادهای خود را می‌دهند و عملکرد سایر کاربران را در این پیشنهاد در نظر نمی‌گیرند. در محیط‌های چندزبانه باید از میان جریانی از سندهایی به زبان‌های مختلف، سندهایی را که مورد علاقه کاربر می‌باشد را انتخاب کنیم و به کاربر نمایش دهیم. از چالش‌های مهم محیط‌های چندزبانه چگونگی استفاده از منابع ترجمه خواهد بود که این منابع ترجمه در اغلب موارد از کیفیت لازم برخوردار نیستند.

سیستم‌های پیشنهاددهنده در محیط‌های چند زبانه چالشی است که در این پژوهش به آن پرداخته‌ایم. در این پژوهش ابتدا سعی کرده‌ایم با استفاده از شباهت معنایی بین کلمات، تخمین بهتری از پروفایل کاربر به دست بیاوریم. از طرف دیگر، نشان داده‌ایم که استفاده از سندهای بازخورد مثبت و منفی تاثیر بسزایی در عملکرد سیستم‌های پیشنهاددهنده دارد. استفاده از چندپروفایل برای هر یک از کاربران و مقایسه با حالتی که تنها یک پروفایل داشته باشیم و از مدل‌های ترجمه استفاده کنیم از جمله کارهایی است که انجام شده‌است. نتایج آزمایش‌های این پژوهش نشان می‌دهد که استفاده از شباهت معنایی کلمات و سندهای بازخورد مثبت و منفی در بهبود عملکرد سیستم پیشنهاددهنده تاثیر بسازی داشته است.

برای ارزیابی روش‌های پیشنهادی خود از مجموعه داده INFLE که توسط همایش CLEF تهیه شده، استفاده شده است. این مجموعه داده از سه زبان فرانسوی، انگلیسی و عربی فراهم آمده که در این پژوهش تنها از سندهای انگلیسی و فرانسوی آن استفاده کرده‌ایم. همچنین برای ارزیابی کارایی روش‌های پیشنهادی و مقایسه آنها با روش‌های پایه از معیار F1-measure به عنوان معیار اصلی استفاده شده است.