# چکیده

موضوع غنی‌کردن متن یکی از مباحث جذاب و چالش‌برانگیز در بین موضوعات مطرح در پردازش زبان طبیعی است که در کاربردهایی مانند زبان‌شناسی مورد استفاده قرار می‌گیرد. اما این کار همواره با چالش‌هایی مانند استخراج معنا از عبارات به صورت غیروابسته‌ به زبان و جمع‌آوری داده‌های مناسب روبرو بوده است. برای استخراج معنا از عبارات، مدل‌های مختلفی وجود دارند که ارزیابی آن‌ها در زبان‌های مختلف امری دشوار است. بدین منظور در پژوهش حاضر، در ابتدا داده‌هایی با ساختارهای مختلف (مانند شعر، نقل‌قول و حکمت) به زبان‌های پارسی، انگلیسی و عربی جمع‌آوری شده‌اند و سپس با استفاده از مدل‌های پیشرو در ساخت امبدینگ، کار غنی‌کردن متن ورودی به کمک این مجموعه داده‌ها و امبدینگ مربوط به آن‌ها انجام شده است. در آخر برای محقق‌کردن هدف این پژوهش، خروجی مدل‌های مختلف برای ورودی‌های مشخص بررسی و به صورت تحت نظارت برچسب‌گذاری شده‌اند تا در مرحله‌ی بعدی بتوان مدل‌های مورد استفاده را ارزیابی کرد. در نتیجه‌ی این ارزیابی، مدل LASER با دقتی بالاتر از ۸۰ درصد در معیار NDCG، عملکرد قابل‌قبولی داشته است و می‌توان گفت برای داده‌های با ساختارهای مختلف و به صورت چندزبانه قابل‌استفاده است.

# مقدمه

توضیح موضوع، لیست کردن کارهای مشابه موجود و ارجاع به آن، نهایتا خلاصه وجه تمایزها و کارهای اصلی انجام شده در این پروژه.

# روش‌ها

در این بخش داده‌ای که جمع‌آوری شده است به همراه روش جمع‌آوری آن شرح داده شده است. سپس مدل‌هایی که برای تولید امبدینگ‌های مورد نیاز استفاده شده اند آورده شده اند. نهایتا آزمایش‌ها و ارزیابی‌هایی که انجام شده‌اند به همراه تقسیم کار آورده می‌شود.

## داده‌ها و نحوه جمع‌آوری آن

...

## مدل‌های مورد استفاده

برای مدل‌های LaBSE، LASER و Fasttext (Tf-idf weighted) مورد استفاده قرار گرفتند.

LaBSE: می‌تواند جملات ۱۰۹ زبان مختلف (از جمله زبان شیرین پارسی) را به یک فضای برداری مشترک ببرد.

## آزمایش‌ها و ارزیابی‌ها

...

## تقسیم کار

...

# نتایج

نتایج آزمایش‌ها و ارزیابی‌ها به همراه نمودارهای مهم

# نتیجه‌گیری

نتیجه گیری و کارهای آینده

# مراجع

1. <https://huggingface.co/sentence-transformers/LaBSE>
2. <https://github.com/facebookresearch/LASER>
3. <https://github.com/bojone/labse>
4. <https://github.com/facebookresearch/fastText>