

## عنوان مقاله:

ارائه روشی جهت شباهت‌یابی معنایی متون عربی با استفاده از یادگیری عمیق

## محل انتشار:

نخستین همایش ملی هوش مصنوعی و علوم اسلامی (سال: 1399)

تعداد صفحات اصل مقاله: 1

## نویسندگان:

محمد عبدوس - دانشجوی دکتری هوش مصنوعی دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت ایران

بهروز مینایی بیدگلی - دانشیار، (دانشگاه علم و صنعت ایران) دکتری هوش مصنوعی [b\\_minaei@iust.ac.ir](mailto:b_minaei@iust.ac.ir)

## خلاصه مقاله:

شباهت‌یابی معنایی متون (STS) یکی از زیرشاخه‌های پردازش زبان طبیعی است که در چندسال اخیر توجهات گسترده‌ای را به خود معطوف کرده است. منظور از شباهت‌یابی معنایی، محاسبه میزان شباهت معنایی بین دو سند متنی، پاراگراف یا جمله می‌باشد که به دو صورت تک‌زبان و چندزبان مطرح است. هدف از این مقاله محاسبه میزان شباهت معنایی بین دو جمله در زبان عربی است که با توجه به عربی زبان بودن بسیاری از متون اسلامی، این پژوهش از کاربردهای فراوانی برخوردار است. جهت محاسبه میزان شباهت معنایی بین دو جمله نیازمند نمایش یکجمله به صورت بردار است. در این تحقیق با استفاده از بردارهای از پیش آموزش داده شده بر روی متون عربی موجود در توییتر و ویکی‌پدیا با استفاده از دو روش CBOW و Skip-Gram که از معروف‌ترین روش‌های آموزش تعبیه کلمات می‌باشند بردارهای کلمات استخراج می‌گردد. سپس با استفاده از داده‌های موجود در کنفرانس شباهت‌یابی معنایی سال ۲۰۱۷ که به صورت جفت جمله عربی بودند اقدام به آموزش مدل شبکه عصبی عمیق با نام شبکه سیامی با استفاده از لایه LSTM نمودیم. استفاده از LSTM توانایی یادگیری وابستگی‌های بلندمدت در شبکه را امکان‌پذیر می‌سازد. تاکنون در زمینه شباهت‌یابی معنایی زبان عربی از شبکه سیامی استفاده نشده بود و این اولین پژوهش در این زمینه است. شبکه‌های سیامی در عین سادگی نتایج قابل قبولی را از خود نشان می‌دهد. در لایه آخر شبکه، با استفاده از شباهت کسینوسی بین بردارهای متعلقه دو جمله ورودی، میزان شباهت بین آن‌ها، به دست می‌آید. نتایج بیانگر آن است که با استفاده از روش پیشنهادی میزان همبستگی ۸۳.۴ درصد به دست می‌آید که از سایر روش‌های موجود عملکرد بهتری را از خود نشان می‌دهد.

## کلمات کلیدی:

## لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1347061>

