

## عنوان مقاله:

بررسی کارایی مدل های زبانی عمیق مبتنی بر ترنسفورمر

## محل انتشار:

بیستمین سمپوزیوم بین المللی هوش مصنوعی و پردازش سیگنال (سال: 1402)

تعداد صفحات اصل مقاله: 6

## نویسندگان:

مریم مجیدی - دانشجوی دکتری هوش مصنوعی، دانشگاه یزد

محمد قاسم زاده - استاد دانشکده مهندسی کامپیوتر، دانشگاه یزد

## خلاصه مقاله:

مدل های زبانی عمیق مبتنی بر مبدل ها (ترنسفورمرها) در سال های اخیر به عنوان یکی از پیشرفت های بزرگ در حوزه یادگیری عمیق و پردازش زبان طبیعی مطرح شده اند. این مدل ها بر پایه ی معماری مبدل که اولین بار در زمینه ترجمه ماشینی معرفی شد، ساخته شده اند. این پژوهش مدل های زبانی عمیق مبتنی بر مبدل ها را از دو دیدگاه مدل های درک زبان طبیعی (NLU) و تولید زبان طبیعی (NLG) بررسی کرده است. همچنین مهم ترین مدل های زبانی این دو دیدگاه را که در شش سال اخیر مطرح شده اند، معرفی کرده و به تشریح معماری مبدل و اجزای آن پرداخته است. از جمله مدل های مورد بررسی می توان به BERT، XLNet، و خانواده GPT اشاره کرد. این پژوهش به مزایا و محدودیت های هر یک از مدل های ذکر شده و روش های آموزش و انتقال یادگیری در این مدل ها می-پردازد. نتایج آزمایش ها و تحلیل های رسمی نشان می دهند که به کارگیری مبدل ها در معماری مدل های زبانی بسیار بهتر از مدل های مبتنی بر شبکه های عصبی بازگشتی (RNN) و حافظه ی طولانی کوتاه مدت (LSTM) عمل می کنند و مدل های زبانی مبتنی بر مبدل توانایی بالقوه ای برای درک و تولید زبان طبیعی دارند.

## کلمات کلیدی:

پردازش زبان طبیعی، ترنسفورمر ها، مدل های زبانی، یادگیری عمیق

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1948229>

