

بسمه تعالی



دانشگاه صنعتی همدان

گروه کامپیوتر هوش مصنوعی

اطلاعیه دفاع پایان نامه کارشناسی ارشد

دسته بندی اسناد فارسی با استفاده از یادگیری عمیق

ارائه دهنده: رامین عقیقی

مکان: سالن آمفی تاتر ساختمان فرشچیان

زمان: دوشنبه ۱۴۰۲/۱۱/۲۳ - ساعت ۱۱:۳۰

دانشگاه: صنعتی همدان

مرتبه علمی: استادیار

استاد راهنما: دکتر حسن بشیری

دانشگاه: صنعتی همدان

مرتبه علمی: استادیار

استاد داور داخلی: دکتر فاطمه امیری

دانشگاه: بوعلی سینا

مرتبه علمی: استادیار

استاد داور خارجی: دکتر میرحسین دزفولیان

چکیده:

دسته بندی متون به عنوان یک وظیفه اساسی در فرآیند فهم و سازماندهی اطلاعات در زبان های مختلف حائز اهمیت است. مدل های یادگیری عمیق در دهه اخیر، به عنوان یک روش پیشرفته در دسته بندی متن، به دلیل کارایی بالای خود و کاهش نیاز به استخراج ویژگی از متن، در زبان هایی مانند انگلیسی بسیار مورد توجه قرار گرفته است. این روش در زبان فارسی کمتر مورد استفاده بوده و این امر نیازمند بررسی و بهبود در این زمینه می باشد. در این پایان نامه، به بررسی دسته بندی اسناد فارسی با استفاده از روش هایی چون CNN، RNN، LSTM و GRU پرداخته شده و مفروضات مختلفی از جمله امکان استفاده بهینه از مدل های عمیق در زبان فارسی و تأثیر مثبت پیش پردازش دقیق بر دقت مدل ها بررسی شده است. تحلیل نهایی از مدل های دسته بندی متن برای اسناد فارسی و نتایج پیاده سازی بر روی مجموعه داده همشهری، نشان دهنده نقش اساسی در اثربخشی مدل های یادگیری عمیق و اهمیت مراحل پیش پردازش بر عملکرد آنها داشته است. نتایج نشان می دهد که مدل CNN با دقت بالا و مقاومت در برابر تغییرات عملکرد برتری دارد.