پروپوزال

|  |  |
| --- | --- |
| علی علی‌محمّدی | 9613027 |

استفاده از شبکه‌های عصبی برای دسته‌بندی متون مربوط به COVID-19

مقدمه

تحلیل احساسات[[1]](#footnote-1) همواره در طی ادوار مختلف از اهمیت بالایی در جامعه‌ی انسانی برخوردار بوده است؛ زیرا انسان موجودی اجتماعی‌ست و نظرات دیگر اعضای اجتماع بر زندگی او تأثیرگذار است. در بسیاری از مسائل دنیای مدرن کنونی نیز دانستن احساس دیگران نسبت به یک پدیده اهمیت خاص خود را دارد؛ مثلاً مهم است که مردم یک کشور رئیس‌جمهور خود را قبول دارند یا خیر؛ یا مثلاً خریداران این محصول از آن رضایت دارند یا خیر. تمام این اطلاعات می‌تواند برای بهبود وضعیت (بالا بردن محبوبیت رئیس‌جمهور یا بهبود کیفیت محصول) به کار گرفته شود.

در این پروژه سعی شده است که با استفاده از تکنیک‌های دسته‌بندی متن[[2]](#footnote-2)، توییت‌های مربوط به بیماری COVID-19 را با به کار گیری رویکرد تحلیل احساسات، با بهره‌گیری از ابزار شبکه‌ی عصبی دسته‌بندی و برچسب‌گذاری کنیم. (Koyel Chakraborty, 2020)

اهمیت این موضوع در آن‌جا مشخص می‌شود که می‌توان با استفاده از نتایج به دست آمده از این دسته‌بندی و Meta-Dataی مربوط به توییت‌ها تشخیص داد که مثلاً در کدام یک از مناطق جغرافیایی، اخباری مثبت درباره‌ی ویروس پخش شده است و یا جو روانی حاکم بر یک ناحیه به چه صورت است. آیا همه‌ی توییت‌ها بسیار منفی هستند و در آن ناحیه تعداد زیاد ابتلا یا مرگ‌ومیر باعث این نتیجه شده است؟ آیا عده‌ای جواب روانی بسیار منفی را ایجاد می‌کنند یا گسترش می‌دهند که روحیه‌ی جامعه را تضعیف می‌کند؟ اخبار امیدبخش کدامند؟ با توجه به اطلاعات به دست آمده، در کدام جامعه نیاز است تا تبلیغات بیشتری برای ترغیب افراد آن جامعه به رعایت پروتکل‌های بهداشتی انجام شود؟

1

مجموعه‌ی داده‌ها

از مجموعه‌ی داده‌های Coronavirus tweets NLP - Text Classification موجود در [Kaggle](https://www.kaggle.com/datatattle/covid-19-nlp-text-classification) برای این منظور استفاده شده است که از قبل برچسب‌گذاری شده‌اند. ممکن است داده‌های دیگری نیز مورد استفاده قرار بگیرند که در حال حاضر هنوز مشخص نشده‌اند.

نحوه‌ی ارزیابی

با توجه به این که برای این منظور از شبکه‌های عصبی استفاده می‌شود، ملاک ارزیابی، دقت[[3]](#footnote-3) و مقدار تابع Loss خواهد بود.

# منابع

Koyel Chakraborty, S. B. (2020, December). Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how popularity is affecting accuracy in social media. *Applied Soft Computing*. doi:https://doi.org/10.1016/j.asoc.2020.106754

2

1. Sentiment Analysis [↑](#footnote-ref-1)
2. Text Classification [↑](#footnote-ref-2)
3. Accuracy [↑](#footnote-ref-3)