

عنوان پروژه: تحلیل احساس کاربران نسبت به خدمات بانک‌ها یا هر خدمت دلخواه دیگر در شبکه‌های اجتماعی

🎯 هدف:

استخراج و تحلیل احساس (مثبت، منفی، خنثی) از نظرات کاربران درباره خدمات بانکی (یا هر صنعت مشابه)، برای کمک به بهبود تجربه مشتری.

📋 مراحل پروژه:

۱. جمع‌آوری داده‌ها:

- استفاده از API توییتر، تلگرام، کامنت‌های سایت‌ها یا داده‌های موجود (مثلاً مجموعه داده‌ی Sentiment140 یا Bank Customer Reviews Dataset)
- یا ساخت دیتاست دستی شامل ۱۰۰۰ نظر واقعی از کاربران.

توجه: می‌توانید از دیتاست‌های آماده و یا دیتای گردآوری شده توسط خودتان استفاده کنید، داده‌ها می‌توانند در مورد هر موضوع دلخواه و از هر شبکه اجتماعی دلخواهی باشند حداقل حجم داده ۱۰۰ نظر است. (گردآوری داده توسط خودتان امتیاز ویژه دارد)

۲. پیش‌پردازش متن:

- حذف نویز (URL)، ایموجی، نشانه‌گذاری‌ها)
- توکن‌سازی (Tokenization)
- ریشه‌یابی (Stemming) یا Lemmatization
- حذف کلمات توقف (Stop Words)
- نرمال‌سازی (کوچک‌سازی، اصلاح غلط‌های نگارشی)

۳. نمایش متن: (Text Representation)

- استفاده از روش‌های مختلف:
 - TF-IDF
 - Word Embedding مثل Word2Vec

۴. مدل سازی (طبقه بندی احساس):

- استفاده از مدل های کلاسیک:
 - Logistic Regression
- پیشرفته:
 - شبکه های عصبی ساده (LSTM / CNN)
 - مدل های Transformers مثل BERT (فاین تیون کردن امتیاز ویژه دارد)

۵. ارزیابی مدل:

- متریک های ارزیابی: Accuracy, Precision, Recall, F1-score, Confusion Matrix
- تحلیل خطاها و مثال هایی از پیش بینی های اشتباه مدل

۶. تحلیل نهایی و گزارش:

- نمایش احساس کلی نسبت به هر بانک یا دسته بندی های خاص خدمات (مثلاً: خدمات آنلاین، کارمندان، شعبه ها)
- نمودارها و اینسایت ها از تحلیل داده

-
- امکان انجام با پایتون (با کتابخانه های `sklearn`, `transformers`, `pandas`, `keras` یا `pytorch`)

انجام پروژه روی داده های فارسی امتیاز ویژه دارد.

انتخاب روش ها و مدل های مرحله ۳ و ۴ نتایج متفاوتی را به همراه دارد که باید در انتهای کار با هم مقایسه شوند.

موفق باشید