
گزارش بخش امتیازی

تحلیل و پیش‌بینی ژانر فیلم با استفاده از مدل DistilBERT

مقدمه این پروژه به هدف تحلیل و پیش‌بینی ژانر فیلم بر اساس توضیحات آن‌ها انجام شده است. مدل DistilBERT به عنوان یکی از مدل‌های پیشرفته در پردازش زبان طبیعی برای این کار انتخاب شده است.

داده‌ها و پیش‌پردازش داده‌های استفاده شده شامل دو فایل `dataTrain.csv` و `testTrain.csv` است. این داده‌ها به دو مجموعه آموزشی و اعتبارسنجی تقسیم شده‌اند.

مدل و آموزش مدل DistilBERT با استفاده از کتابخانه Transformers بارگذاری شده و داده‌ها با استفاده از این مدل توکنیز شده‌اند.

سپس داده‌ها به فرمت Dataset تبدیل شده و مدل DistilBERT آموزش داده شده است.

ارزیابی و نتایج مدل آموزش داده شده و نتایج بر روی مجموعه داده‌های آزمایشی ارزیابی شده است. نتایج به صورت دقت مدل گزارش شده است.

نتیجه‌گیری این پروژه نشان داد که مدل DistilBERT قابلیت بالایی در پیش‌بینی ژانر فیلم بر اساس توضیحات آن دارد. در آینده می‌توان از داده‌های بیشتر و مدل‌های پیشرفته‌تر برای بهبود نتایج استفاده کرد.