

گزارش پیاده‌سازی مدل Transformer Financial

۲ اسفند ۱۴۰۴

در این گزارش، به بررسی و پیاده‌سازی یک مدل ترنسفورمر برای تحلیل احساسات در حوزه مالی می‌پردازیم. از مجموعه داده Bank Phrase Financial استفاده شده است که شامل جملات مالی با برچسب‌های احساسی است. هدف نهایی آموزش مدلی است که بتواند با دقت بالا احساس نهفته در اخبار و جملات مالی را تشخیص دهد.

۱ مقدمه

تحلیل احساسات در بازارهای مالی نقشی حیاتی دارد. مدل‌های زبانی مبتنی بر معماری Transformer توانسته‌اند تحولی عظیم در پردازش زبان طبیعی ایجاد کنند. در این پروژه، ما تلاش کردیم تا یک مدل سفارشی را بر روی داده‌های مالی آموزش دهیم.

۲ مجموعه داده (Dataset)

داده‌های مورد استفاده از مجموعه Financial Phrase Bank استخراج شده‌اند. این مجموعه شامل جملات برچسب‌گذاری شده توسط متخصصان مالی است. برای افزایش حجم داده‌ها و بهبود دقت مدل، ما از پیکربندی‌های مختلفی که درصد توافق متخصصان را نشان می‌دهند استفاده کردیم:

- sentences_allagree (توافق ۱۰۰٪)
- sentences_75agree (توافق ۷۵٪)
- sentences_66agree (توافق ۶۶٪)
- sentences_50agree (توافق ۵۰٪)

این داده‌ها ترکیب شده و برای آموزش مدل آماده‌سازی شدند.

۳ معماری مدل

مدل پیشنهادی FinancialTransformer نام دارد که یک مدل مبتنی بر معماری استاندارد ترنسفورمر است. مشخصات کلیدی مدل به شرح زیر است:

- اندازه نهان (Embedding Dimension): ۲۵۶
- تعداد هدها (Attention Heads): ۸
- لایه پیش‌خور (Feed Forward): ۵۱۲
- تعداد لایه‌ها: ۴
- طول دنباله: ۱۲۸

۴ روند آموزش

برای آموزش مدل از تابع زیان CrossEntropyLoss و بهینه‌ساز AdamW با نرخ یادگیری 0.0002 استفاده شد. مدل به مدت ۱۰ دوره (Epoch) آموزش دید و نتایج بر روی داده‌های اعتبارسنجی ارزیابی شد.

۵ نتیجه‌گیری

مدل در نهایت به 0.9939 Accuracy و 0.9934 F1-Score رسید که نشان‌دهنده عملکرد بسیار خوب آن در تشخیص احساسات مالی است.