

به نام خدا



دانشگاه تهران  
پردیس دانشکده‌های فنی  
دانشکده برق و کامپیوتر



درس پردازش زبان های طبیعی

تمرین ۴

سجاد پاکدامن ساوجی

۸۱۰۱۹۵۵۱۷

اردیبهشت ۱۴۰۰

## فهرست مطالب

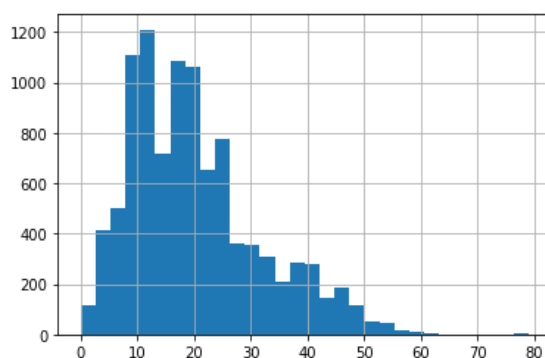
3	مقدمه
3	پیش پردازش
3	۱. BERT
5	۲. XLNet
7	مقایسه و نتیجه گیری
7	اجرای کد ها

## مقدمه

در این تمرین قصد داریم که طبقه بندی توییت های توهین آمیز را با استفاده از مدل های زبانی بر اساس transformers انجام دهیم. بدین منظور از representation های دو مدل BERT و XLNet استفاده خواهیم کرد.

## پیش پردازش

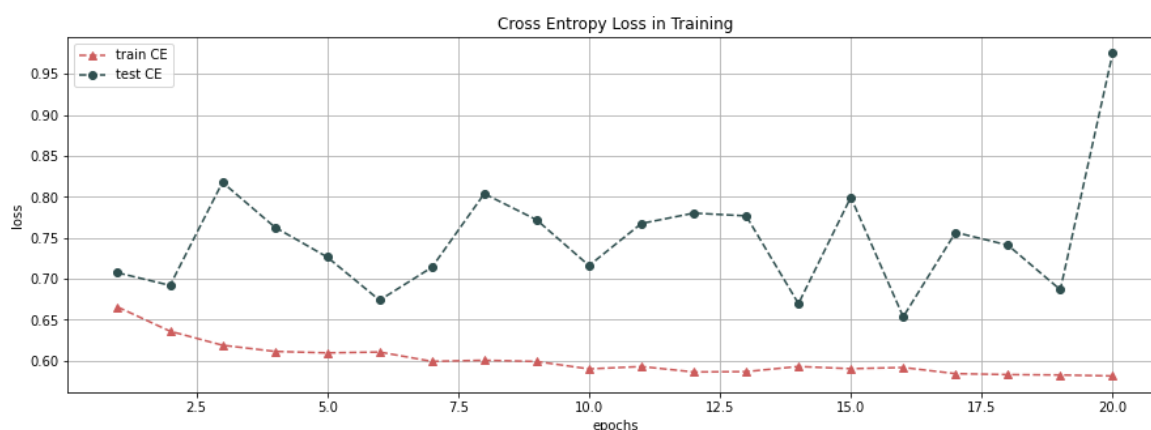
پیش از آن که مدل های زبانی را برای طبقه بندی آموزش دهیم، لازم است که پیش پردازش های لازم را بر روی توییت ها انجام دهیم. در این قسمت مشابه تمارین ۱ و ۲ عمل شده است. همچنین طول هیستوگرام طول جملات (تعداد کلمات در هر جمله) نیز بدست آمده است. با توجه به این نمودار، همان تعداد ۱۲۸ کلمه برای طول استاندارد انتخاب شده است.



شکل ۱. هیستوگرام طول جملات

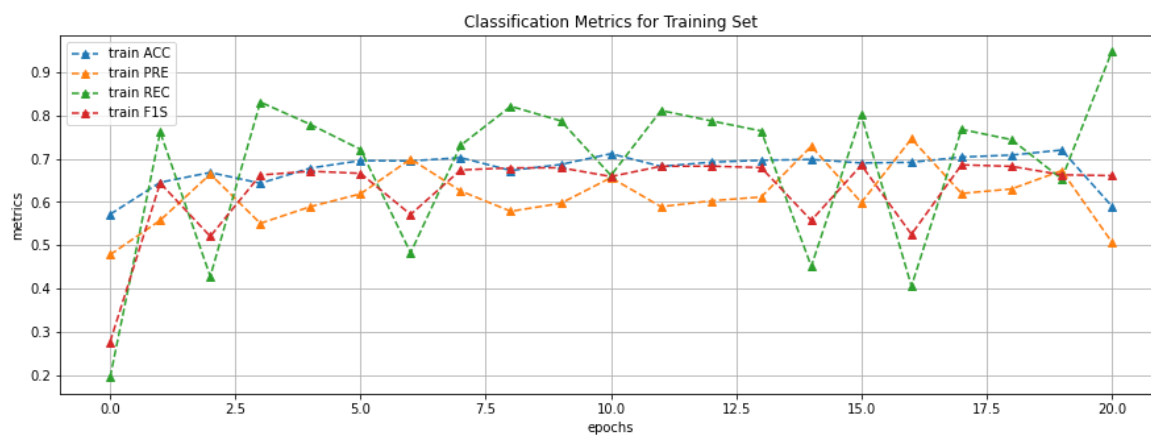
## ۱. BERT

در این قسمت با اضافه کردن دو لایه غیر خطی در انتهای شبکه BERT جملات طبقه بندی شده اند. در شکل ۲ نمودار کاهش تابع هزینه استفاده شده برای داده های آموزش و آزمایش آورده شده است.

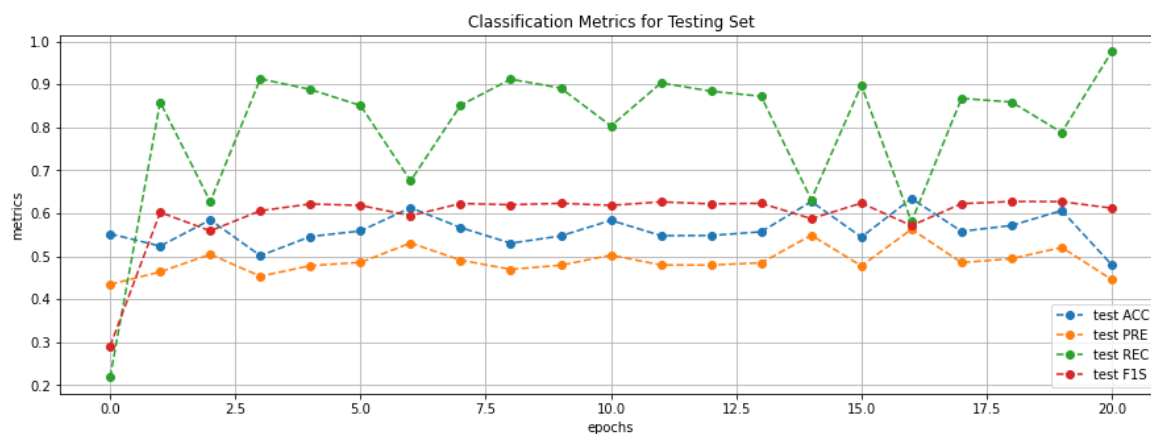


شکل ۲. نمودار تغییرات تابع هزینه در زمان آموزش

در ادامه، در شکل ۳ و ۴ متریک های ارزیابی خواسته شده از جمله accuracy, recall, precision و F1-score آورده شده است.

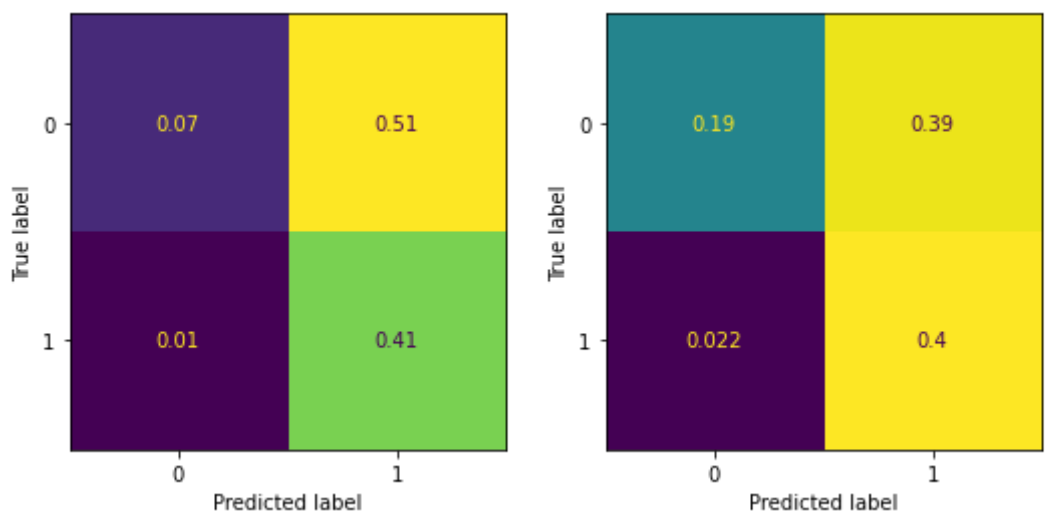


شکل ۳. متریک های ارزیابی در طول آموزش شبکه بر داده های آموزش



شکل ۴. متریک های ارزیابی در طول آموزش شبکه بر داده های آزمایش

در شکل ۵ نیز confusion matrix نرمال شده برای داده های آموزش و آزمایش آورده شده.



شکل ۵. در راست Confusion Matrix برای آموزش و در چپ برای آزمایش

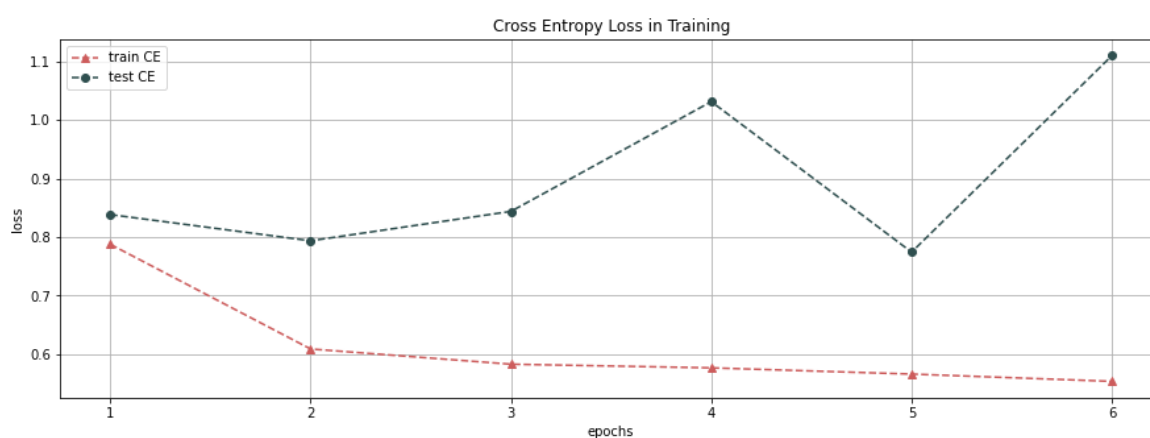
در جدول ۱، متریک های ارزیابی خواسته شده برای داده های آزمایش آورده شده است.

Data Part	Accuracy	Precision	Recall	F1-Score
Train	0.7	0.65	0.95	0.65
Test	0.48	0.44	0.97	0.61

جدول ۱. متریک های ارزیابی برای شبکه BERT

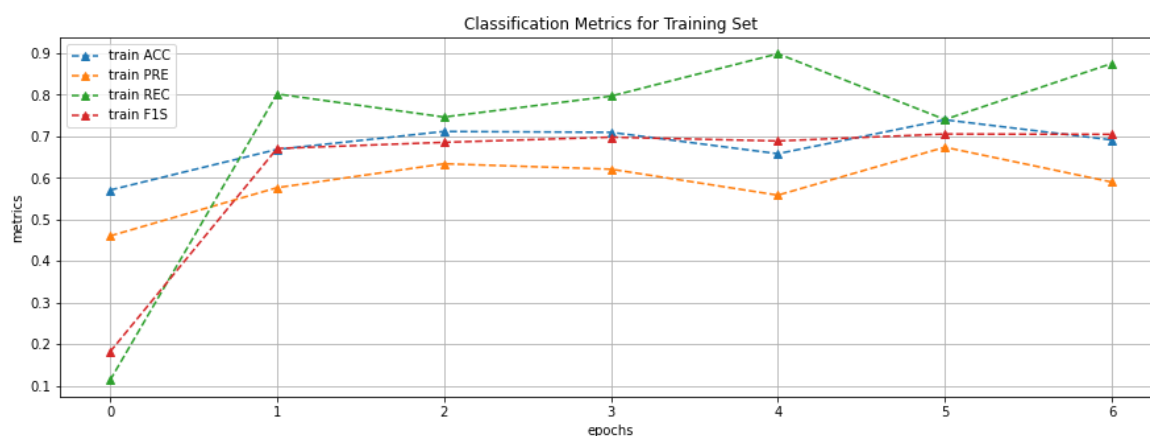
## ۲. XLNet

در این قسمت با اضافه کردن دو لایه غیر خطی در انتهای شبکه XLNet جملات طبقه بندی شده اند. در شکل ۶ نمودار کاهش تابع هزینه استفاده شده برای داده های آموزش و آزمایش آورده شده است.

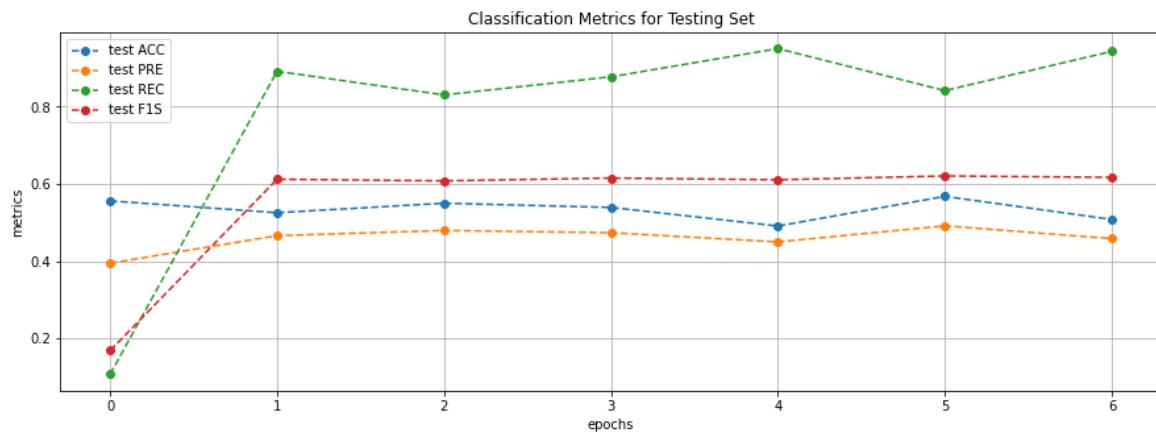


شکل ۶. نمودار تغییرات مقدار تابع هزینه برای داده های آموزش و آزمایش

در ادامه، در شکل ۷ و ۸ متریک های ارزیابی خواسته شده از جمله accuracy, recall, precision و F1-score آورده شده است.

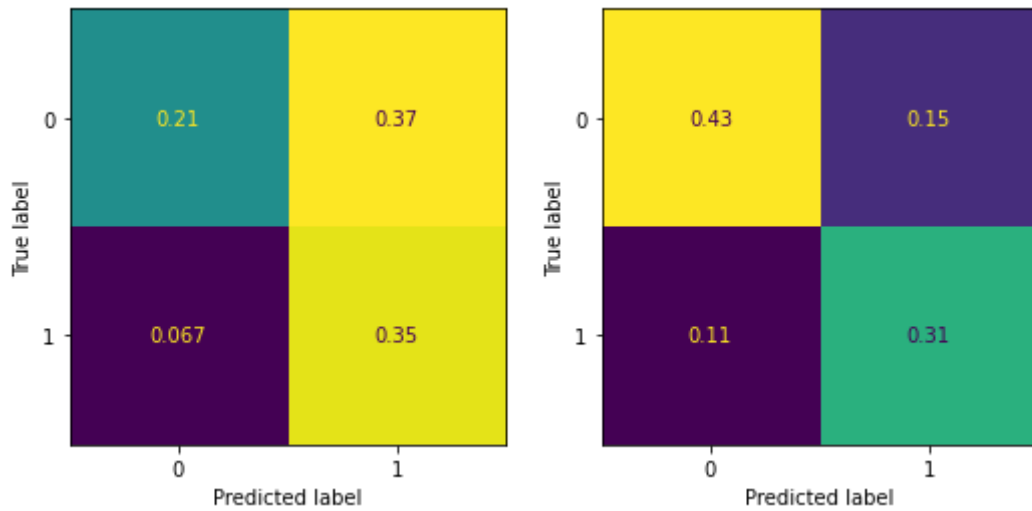


شکل ۷. متریک های ارزیابی در طول آموزش شبکه بر داده های آموزش



شکل ۸. متریک های ارزیابی در طول آموزش شبکه بر داده های آزمایش

در شکل ۹ نیز confusion matrix نرمال شده برای داده های آموزش و آزمایش آورده شده.



شکل ۹. در راست Confusion Matrix برای آموزش و در چپ برای آزمایش

در جدول ۲، متریک های ارزیابی خواسته شده برای داده های آزمایش آورده شده است.

Data Part	Accuracy	Precision	Recall	F1-Score
Train	0.70	0.60	0.88	0.70
Test	0.56	0.49	0.84	0.62

جدول ۲. متریک های ارزیابی برای شبکه XLNet

## مقایسه و نتیجه گیری

با توجه به دقت های بدست آمده در جدول ۱ و جدول ۲، مشاهده میشود که transfer learning با استفاده از XLNet با زمان آموزش کمتر، نتیجه بهتری داده است. این نتیجه گیری با عملکرد کلی شبکه XLNet نیز مطابق است، زیرا که شبکه XLNet که از جمله شبکه های Post Bert است، در تسک های مختلف NLP عملکرد بهتری داشته است و احتمالاً به همین دلیل representation بهتری نیز دارد.

## اجرای کد ها

دو قسمت تمرین در نوت بوک های جداگانه با اسم های NLP\_CA4\_BERT و NLP\_CA4\_XLNet آپلود شده است.