

تارا قشلاقی

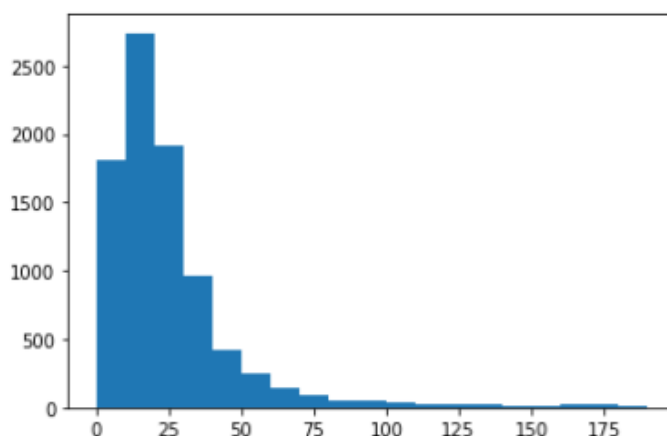
فایل google colab: در این [لینک](#) قرار داده شده است.

در این تمرین ۳ مدل بررسی شد که برای هر کدام یک بار با مدل Glove و یک بار با لایه Embedding که training صورت می‌گیرد انجام شد:

- LSTM_Glove
- GRU_Glove
- simpleRNN_Glove
- LSTM
- GRU_Glove
- simpleRNN_Glove

برای تعیین ماکسیمم مقدار feature ها برای جلوگیری از gradient vanishing در طی back warding با توجه به نمودار به دست آمده ، در بین مقدار ۱۰۰ تا ۱۲۵ توکن، تعداد جمله ها کمتر شده بنابراین مقدار max_feature را برابر با ۱۳۰ در نظر می‌گیریم. بعد از آن فراوانی بسیار کم است و آن‌ها را نادیده می‌گیریم.

با توجه به مقداری که برای بیشترین تعداد توکن به دست آمد بررسی شد که فراوانی بسیار کم است در آن مقادیر بنابراین تنها تا بازه‌ی ۲۰۰ نمودار رسم کردیم.



برای هر مدل با توجه تعداد ایپایکی که در نظر گرفته شده، در جدول قرار داده شده در لینک مقادیر ایپاک آخر گزارش داده شده است.

سپس برای هر مدل نمودار برای مقادیر loss رسم شده است.

تفاوت زیاد نمودار loss و loss-val به دلیل overfit شدن مدل است از طرفی دقت روی داده های validation کمتر از دقت روی داده های آموزش بود. بنابراین مقدار drop out را افزایش دادیم.

در شرایط یکسان این ۶ مدل مقایسه شدند که در لینک آورده شده‌اند.

همانطور که آورده شده است تفاوت بین `loss` و `loss_val` در مدل‌هایی که از `pretraind` استفاده نشده است زیاد است که این به دلیل `overfit` شدن است از طرفی دقت بالاتری دارند.

در طراحی مدل شبکه‌ها سعی شده است `loss` و `loss_val` مقادیر نزدیک‌تری داشته باشند.