

❖ در حالتی که مدل، الگوهای مرتبط با مسئله را در داده‌های آموزشی به خوبی یاد نگیرد، Under Fit شده‌است. با توجه به این تعریف، و در نظر گرفتن تعاریف Bias و Variance می‌توان گفت که، Bias در این حالت می‌تواند زیاد باشد، چرا که آموزش درست انجام نشده، بنابراین احتمال بدست آوردن جواب نزدیک به پاسخ صحیح کم است، و Bias زیاد. اما Variance احتمالاً زیاد نیست، چرا که پاسخ‌ها به طور کلی به درستی نیستند، و حساس به پارامتری نیز نشده‌اند، بنابراین در یک Range خواهند بود.

در حالتی که مدل، داده‌های آموزش را به خوبی یاد گیرد، و بر روی برخی از Featureها حساس شود، که باعث شود برای تشخیص جواب شبکه گمراه شود، Over Fit شده‌است. در این حالت، Bias نسبتاً کم است، زیرا پاسخ‌ها اغلب در نزدیک پاسخ صحیح است، اما Variance می‌تواند به دلیل حساسیت زیاد به برخی Featureها و عدم یافتن General Pattern زیاد شود.

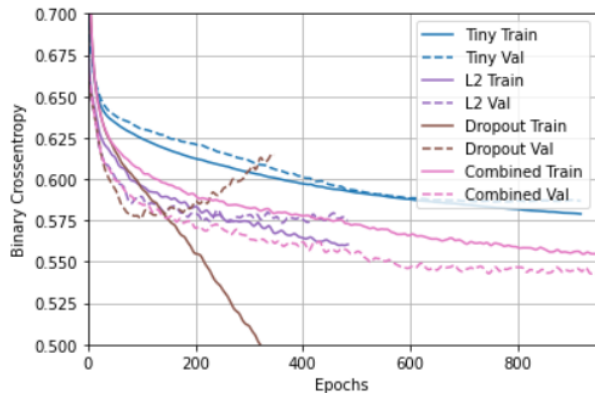
یکی از راه‌های بهبود Under Fitting افزایش تعداد داده‌های ورودی است. تعداد Parameterهای مدل را افزایش دهیم. همچنین می‌توان مدل را پیچیده‌تر کرد.

برای بهبود Over Fitting نیز چند راه موجود است. داده‌های Train را به دو قسمت Train و Validation تقسیم کنیم، که در طول آموزش دقت پاسخ قابل بررسی باشد. Regularization یکی دیگر از روش‌های است که می‌توان برای جلوگیری از Over Fitting استفاده کرد. همچنین می‌توان آموزش را زودتر متوقف کرد.

❖ به طور کلی ابتدا در رابطه با تعداد Parameterهای شبکه، پیچیدگی آن مثال‌های مختلف می‌زند، و نیجه آن است که در صورت پیش از حد بزرگ بودن تعداد Parameter، شبکه به سرعت شروع به Over Fit شدن می‌کند. همچنین می‌توان تعداد داده‌های بیشتری را به ورودی داد، تا آمادگی بیشتری نسبت به ورودی‌ها ایجاد شود (یادگیری بهتر).

در Strategy اول برای جلوگیری از Over Fitting مبحث Weight Regularization مطرح شده‌است، به این ترتیب که پیچیدگی مدل را با محدود کردن مقادیر وزن (به مقادیر کوچک) کنترل کنیم و وزن‌ها را Regular می‌کند. در مثال زده شده، پس از Regularization به روش L1, L2 مدل Large که Over Fit شده بود و مقادیر دقت برای Train و Test بسیار متفاوت بود، مقادیر دقت بسیار نزدیک بهم شد (Over Fit صورت نگرفته) و همچنین Loss نیز به سرعت شروع به افزایش دوباره نکرد. روش دیگر Regularization حالت Dropout است، به این ترتیب که برخی از نوروها را با احتمال مشخص شده‌ای، صفر می‌کند. در این حالت هم عملکرد مدل Large با مقایسه مقادیر دقت Test و Train می‌بینیم بهتر شده‌است، و Over Fitting زیاد اتفاق نیافتاده‌است، اگرچه پس از زمان اندکی مقادیر Loss دوباره کمی زیاد می‌شود در حالی که Train در حال

کاهش است. با این حال هیچ کدام از دو روش به اندازه مدل Tiny که تعداد Parameter کمی داشت مناسب نیست. در Technique بعدی ترکیب دو روش اول را داریم یعنی L2 + Dropout عملکرد این روش با اختلاف از حالت Tiny نیز بیشتر است و دقت Train, Test تقریباً برابر است، Loss نیز به طور مداوم کم می‌شود. بنابراین بهترین تاثیر را برای جلوگیری از Over Fitting دو روش L2 و Dropout به طور همزمان دارند، نمودار به شکل زیر است:



البته این نمودار Binary Crossentropy را نمایش می‌دهد.

❖ در ابتدا تعدادی لایه جدید با تعداد نرون بیشتر اضافه می‌کنیم. پس از پیچیده‌تر شدن مدل و افزایش Parameters شبکه دچار Over Fitting می‌شود. برای حل این مشکل، با توجه از ایده سوال قبلی به مدل Dropout و L2 اضافه می‌کنیم. مرحله اول فقط Dropout اضافه شد، گرچه عملکرد بسیار خوبی داشت، زمان بسیار زیادی برای آموزش صرف شد، اما نتیجه بدون Over Fitting بود.

Epoch: 1300, accuracy:0.6702, binary\_crossentropy:0.5780, loss:0.5780, val\_accuracy:0.7070, val\_binary\_crossentropy:0.5543, val\_loss:0.5543,

یکی دیگر از روش‌های جلوگیری از Over Fitting اضافه کردن داده‌های ورودی است، که این امکان در اینجا فراهم نیست (هرچند می‌شود Data Augmentation انجام داد، اما در اینجا قرار ندادیم). در آخر L2 را نیز به Loss Function اضافه کردیم.

Epoch: 600, accuracy:0.6222, binary\_crossentropy:0.6216, loss:0.6361, val\_accuracy:0.6490, val\_binary\_crossentropy:0.5937, val\_loss:0.6082,

دقت بسیار پایین آمد، انگار Under Fit شد.

با ادامه همین روند، متوجه شدم بازده L2 زیاد می‌ست، برای همین تنها لایه های Dropout قرار داده شد، هرچند زمان زیادی می‌برد.

❖ محاسبات سوال دو در ابتدا به شکل کامل، اما برای Epoch دوم به صورت کد انجام شد برای افزایش سرعت. اما به طور دقیق تک تک مراحل انجام شده.

❖ منابع:

- ✓ <https://www.analyticssteps.com/blogs/l2-and-l1-regularization-machine-learning>
- ✓ <https://medium.datadriveninvestor.com/how-to-handle-overfitting-and-underfitting-470a1f7389fe>

