**امیرحسین احمدی 97522292 – تمرین هفتم**

1. *هر چه مدل ها ساده تر باشند واریانس کمتری دارند، زیرا پیچیدگی خاصی ندارند و زیاد تغییری نمیکنند ولی بایاس بالایی دارند و جواب های خوبی نمیدهند و نیاز به ترین بیشتر دارند که در این حالت مدل underfit شده است. برعکس هر چه مدل پیچیده تر باشد بایاس خیلی کمی دارد و جواب ها خوب است ولی در عین ها واریانس زیاد است و با تغییرات ریز هم نتیجه متفاوت میشود و generalization کاهش میابد و به داده های جدید جواب خوبی نمیدهد و در واقع مدل overfit میشود. به این دلایل بهترس مدلی داشته باشیم که بایاس واریانس در تعادل باشند تا overfit و underfit نشویم.*

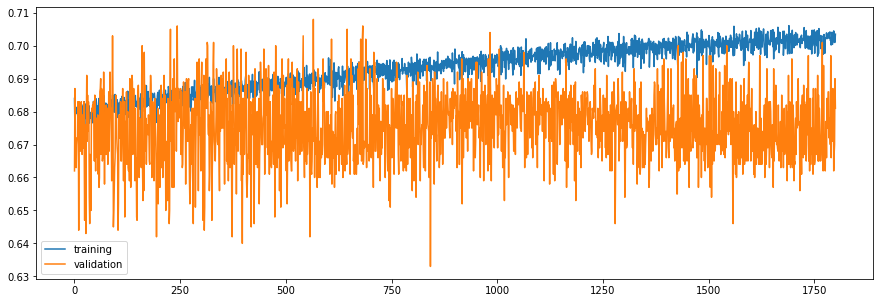
*راه های جلو گیری از ovefitting :*

1. *استفاده از هایپر پارامتر های کمتر و ساده کردن مدل*
2. *استفاده از dropout در مدل*
3. *افزایش داده های آموزشی Augmentation*

*راه های جلو گیری از underfitting :*

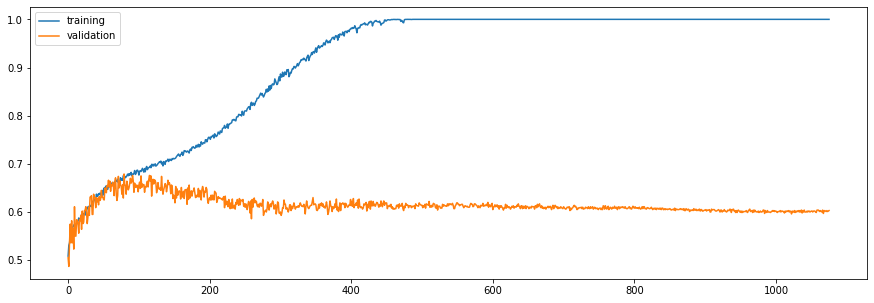
1. *افزایش لایه های مدل*
2. *افزایش تعداد نورون های هر لایه*
3. *افزایش فیچر ها*
4. *در پی دی افی جدا جواب کامل دو آمده است.*
5. *الف) در اینجا مدل اورفیت شده است. در بخش اول با اضافه کردن لایه جدید باعت افزایش اورفیتینگ شده ایم. در بخش بعد با استفاده از روش L2 regularization باعث کاهش اور فیتینگ شده ایم، مدل با 4 لایه پنهان و 512 نورون بدون اورفیت شدن تا 400 ایپوک آموزش را ادامه میدهد. در قسمت بعدی با استفاده از dropout باعث کاهش اورفیتینگ شده ایم ولی با این حال از ایپوک 180 به بعد باز اورفیت میشویم. در حالت بعد dropout و L2 را باهم ترکیب میکنیم و میبینیم که عمل کرد بسیار خوبی دارد و لاس کمتری حتی از مدل های ساده تر دارد.*

*ب) دقت مدل اولیه برابر شکل زیر است:*

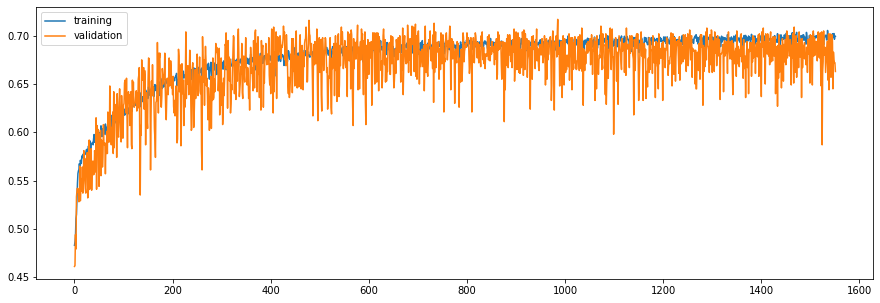
**

*همانطور که مشاهده میشود از ابتدای آموزش دقت ولیدیشن تغییره خاصی نمیکند و به طور میانگین ثابت مانده و به نظر مدل خوب ترین نشده و underfit است.*

*برای مدل دومم سه لایه میانی اضافه کردم تا کمی پیچیدگی افضایش پیداکرده و underfit نشود:*

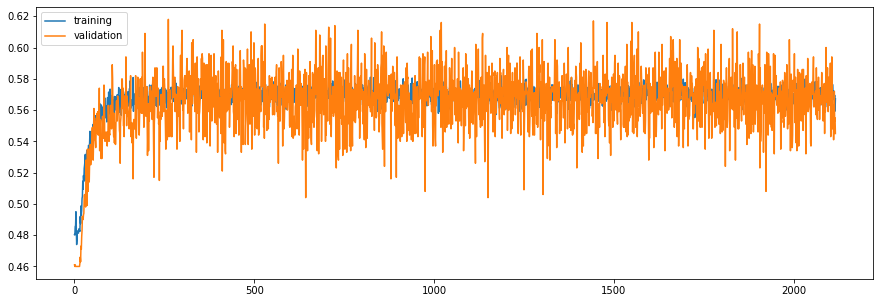
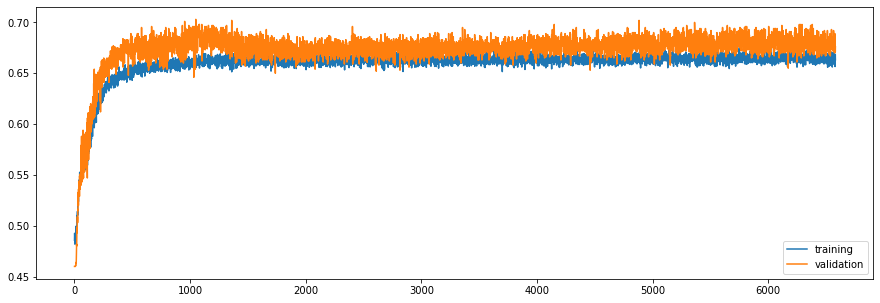
**

*همانطور که میبینیم تا حدود epoch 100 مدل به خوبی ترین میشود ولی از آنجا به بعد دقت ولیدیشن پایین آمده و دقت ترین به سمت 100 مییرود و به وضوح مدل overfit شده است. به همین دلیل در مدل بعدی تصمیم گرفتم از روش regularization L2 استفاده کنم تا overfitting کم شود:*

**

*در این مدل میبینیم که مدل از overfitting درآمده و ترین و ولیدیشن دقتی نزدیکی را ارائه میدهند. ولی دقت مدل پایین و چیزی حدود 70 درصد است.*

*در ادامه به ترتیب از روش dropout و همچنین از ترکیب L2 و dropout برای مدل استفاده کردم:*

**

*همانطور که میبینید از لحاظ اورفیت نشدن بهتر از L2 خالی عملکرده اند و dropout خالی در ولیدیشن حتی بهتر از ترین بوده است، ولی دقت با L2 خالی بهتر است و نشان از این است که این مدل ها به خوبی ترین نشده اند. بنابراین بهترین مدل مدلی است که با استفاده از L2 ساخته ایم.*