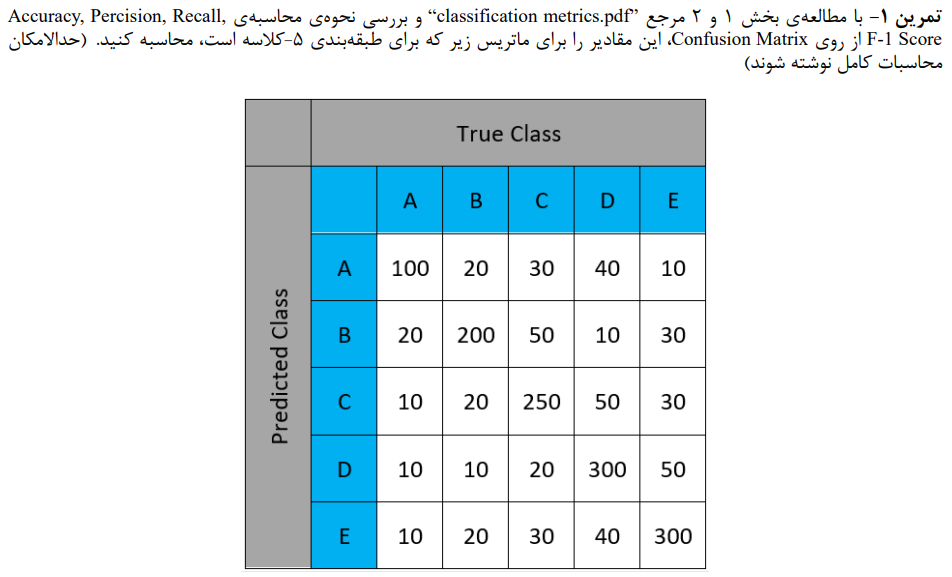
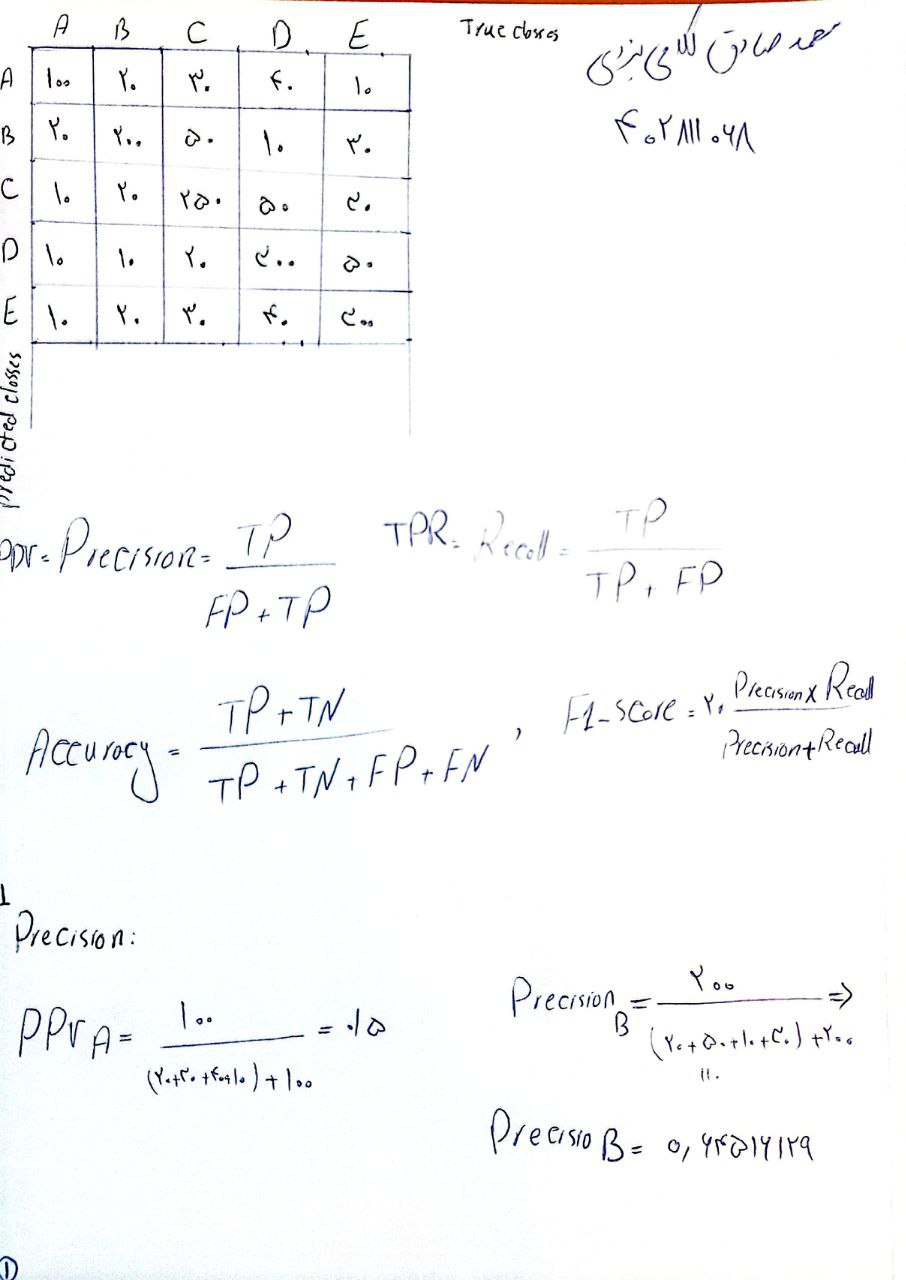
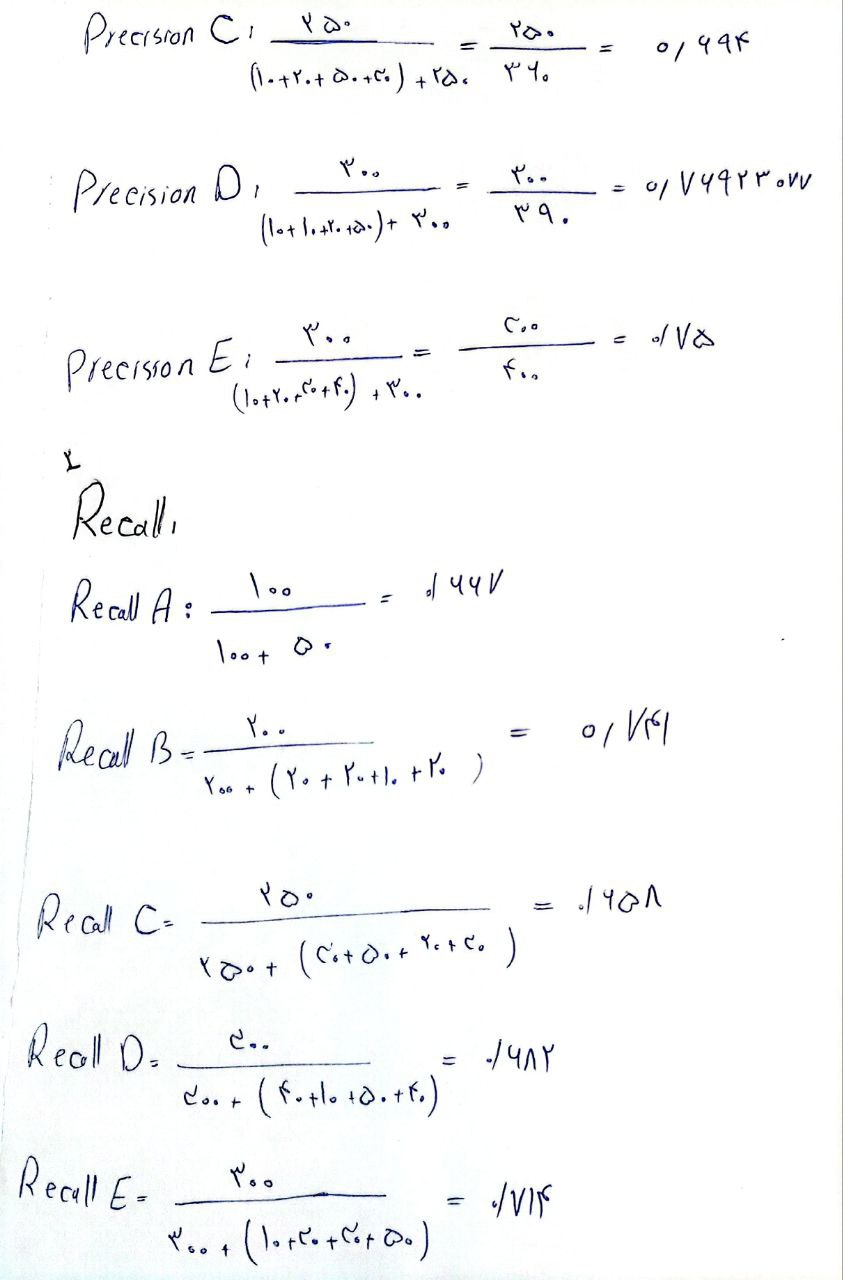
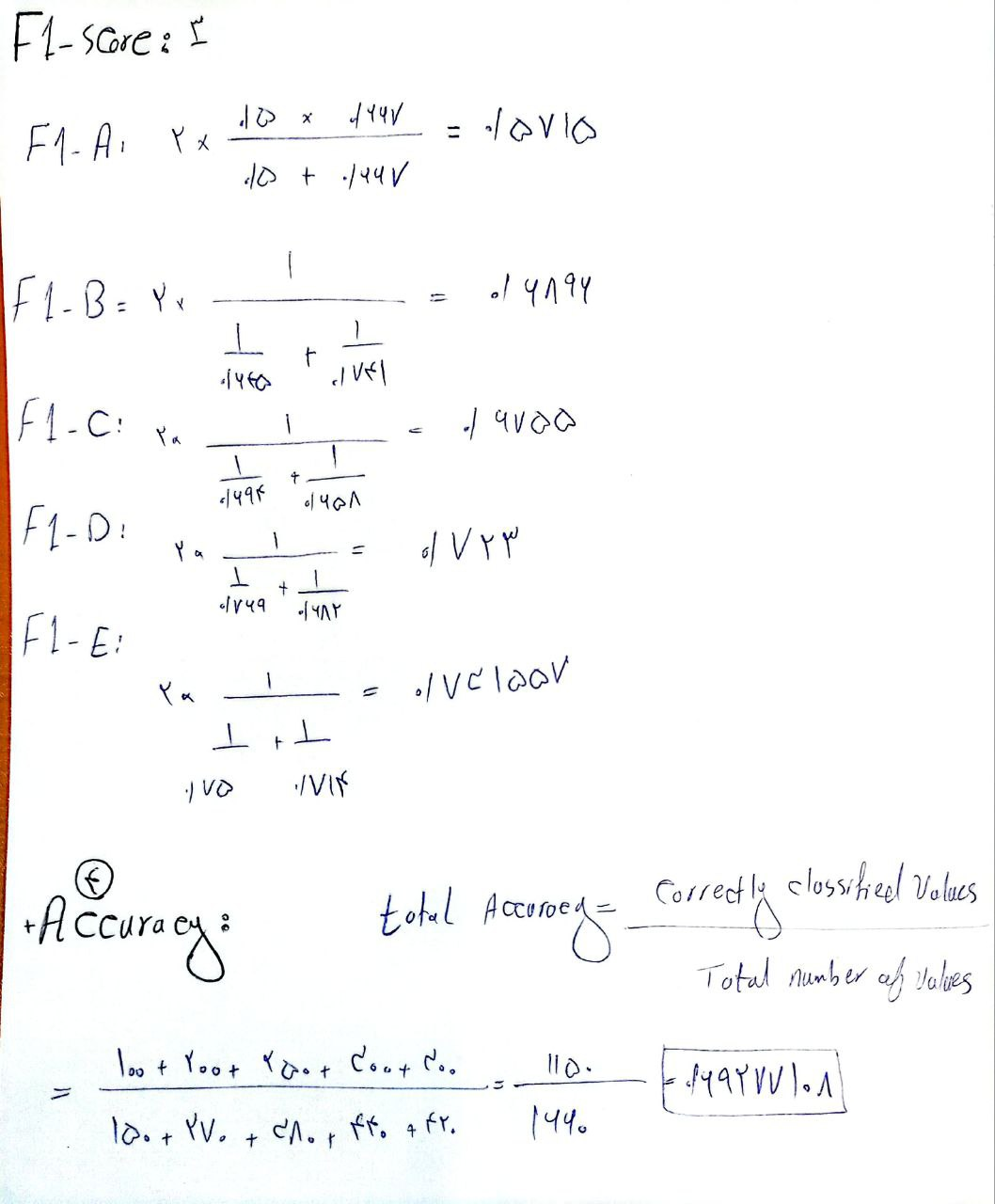
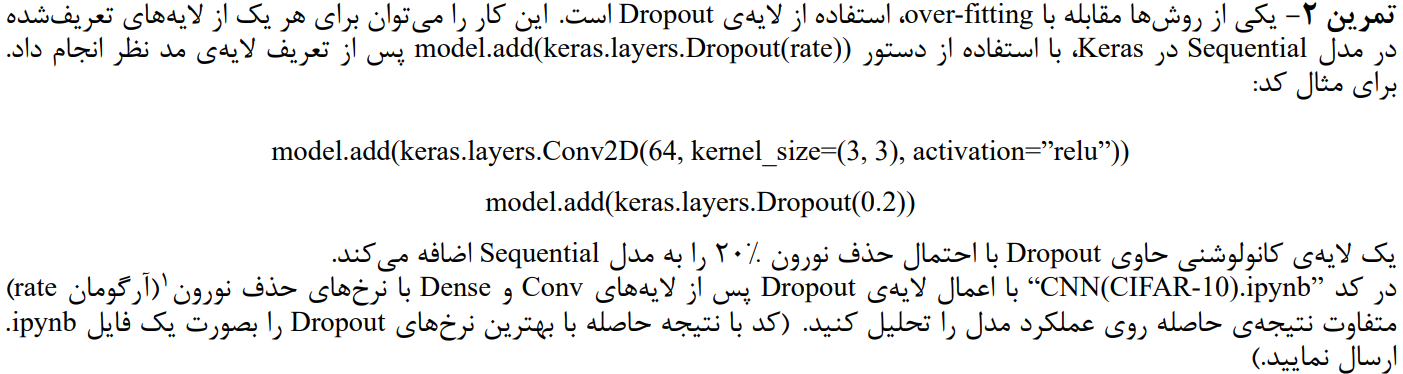
**محمّدصادق کلامی یزدی ۴۰۲۸۱۱۰۶۸**











تمامی کد‌ها در Kaggle ران شده‌اند و کد‌های آن‌ها به Colab انتقال داده شده است. در صورت نیاز به خود Kaggle اطلاع بدید تا به شما دسترسی بدهم.

در تمرین 2 با استفاده از Keras-Tuner بهترین مقادیر Dropout برای هر لایه بدست آورده شده ولی کماکان مدل مشکلی که در تمرین 3 گفتید را داشت، پس به صورت گام به گام موارد زیر اضافه شد تا اثر هر کدام دیده و گزارش شود. در نهایت بهترین مدل، مدلی است که از EarlyStopping، LR-Schecular، Dropout، Batchnorm و L2-Regularization استفاده می‌کند و مقدار بهینه‌تمامی این پارامتر‌ها توی هر بخش گزارش شده است.

* L2 Regularization
* Dropout
* EarlyStopping and Learning Rate Scheduler
* Batchnormalization

# 1. مدل بهینه با استفاده از Dropout

|  |  |
| --- | --- |
|  |  |
|  |  |

# 2. مدل بهینه با استفاده از EarlyStopping، LR-Schecular و Dropout

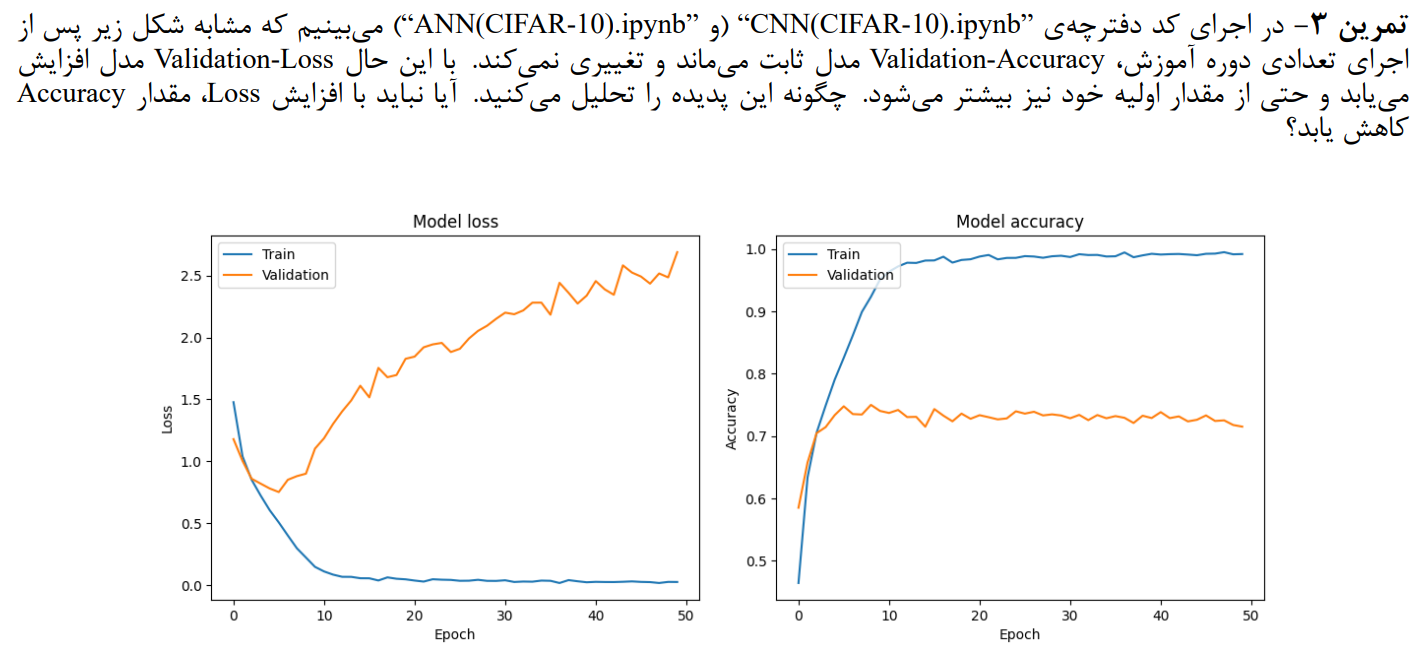
|  |  |
| --- | --- |
|  |  |
|  |  |

# 3. مدل بهینه با استفاده از EarlyStopping، LR-Schecular، Batchnorm و Dropout

|  |  |
| --- | --- |
|  |  |
|  |  |

# 4. مدل بهینه با استفاده از EarlyStopping، LR-Schecular، Batchnorm، L2 regularization و Dropout

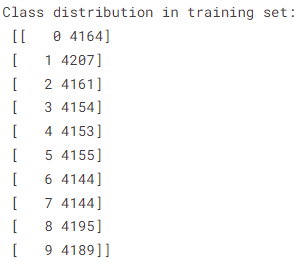
|  |  |
| --- | --- |
|  |  |
|  |  |



نتیجه‌گیری:

با اضافه کردن هر کدام از این پارامتر‌ها (در سوال2) به مدل، از Overfitting جلوگیری کردم که در نهایت مدل هم کاهش loss بسزایی دارد و هم accuracy بالاتری، یعنی در عین افزایش دقت، اطمینان (Confidance) مدل هم افزایش داده شده است.

بررسی بالانس بود:



چون مدل می‌تواند دقت‌ تقریباً یکسانی داشته باشد اما Confidanceهای متفاوتی داشته باشد که هرچقدر Confidance کمتر شود Loss بیشتر می‌شود.

این اتفاق به این دلیل می‌افتد که ما در آخرین لایه شبکه از **softmax** استفاده کرده‌ایم. softmax به تنها چیزی که نگاه می‌کند این است که در آخرین لایه کدام کلاس احتمال بیشتری دارد و آن به عنوان کلاس خروجی انتخاب می‌کند و از این رو اگر Confidance آنقدر کم شود که کلاس را اشتباه تشخیص ندهد (کلاس درست بیشترین احتمال را بین مابقی کلاس‌ها داشته باشد) و فقط اطمینان (Confidance) از کلاس انتخابی کاهش یابد، در نتیجه روی Accuracy اثری ندارد و مقدار Loss افزایش می‌یابد.

