

توضیح پیاده سازی MNIST:

برای پیاده سازی شبکه مورد نظر از 4 لایه استفاده کرده ایم که 3 لایه اول شامل 256 نورون هستند و از تابع فال سازی relu در آن ها استفاده شده. برای لایه آخر به تعداد کلاس ها یعنی 10 نورون استفاده شده با تابع فعال سازی softmax. برای آموزش مدل از 20 تکرار استفاده کرده ایم و 10 درصد داده ها را به عنوان داده ارزیابی استفاده کرده ایم. همان طور که میبینید در تکرار آخر دقت داده آموزش 100 و ارزیابی برابر 98 درصد است. دقت داده تست هم برابر 97.9 درصد است. خروجی confusion matrix نشان میدهد که قطر اصلی که تعداد حدس های درست ما هست زیاد است.

توضیح CIFAR 10:

برای آموزش این مدل دقیقاً از مدل قبلی استفاده کرده ایم. همانطور که میبینید دقت روی داده تست 51 درصد است. دقت روی داده ترین هم در تکرار آخر به 57 درصد رسیده است. دلیل اینکه دقت این مدل برای این دیتاست کمتر به دلیل پیچیدگی بیشتر این داده ها است. داده CIFAR 10 به دلیل تنوع گوناگون خروجی ها مسئله را پیچیده کرده و استفاده از شبکه های fully connected کار ساز نیست. چون مدل پیچیده است باید شبکه پیچیده تری انتخاب کنیم که این باعث پیچیدگی محاسباتی شده و عملاً نمیتواند مسئله الگو های پیچیده را یاد بگیرد.