

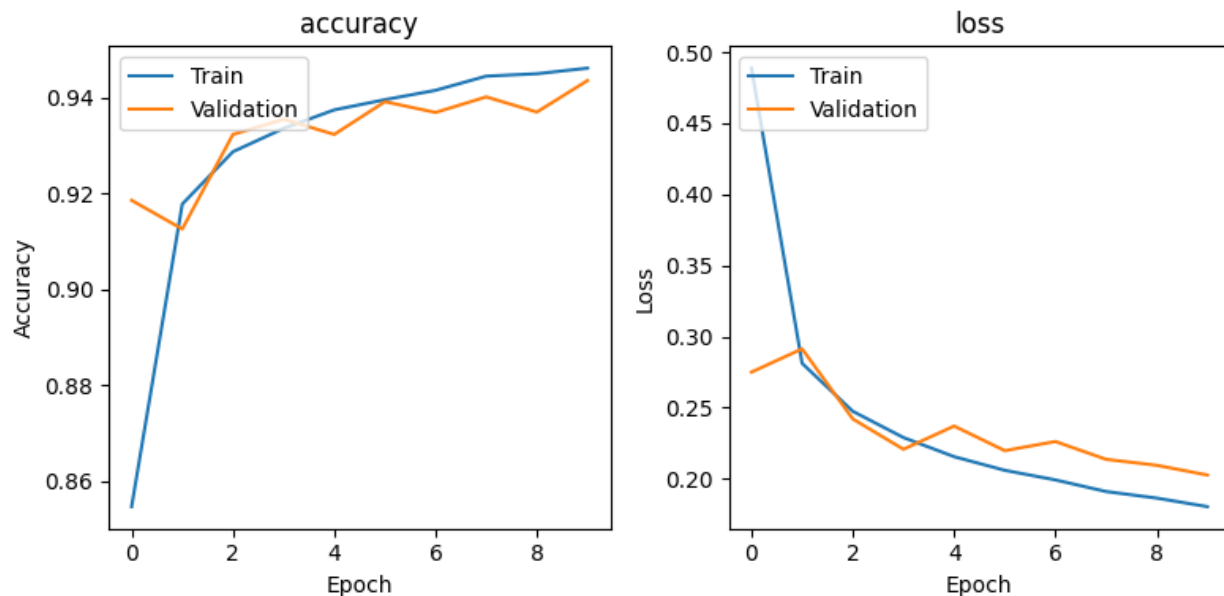
## بینایی کامپیوتر – تمرین سری پنجم

### سوال پنجم – مهرشاد فلاح اسطخزیر

۴۰۱۵۲۱۴۶۲

منطق کد:

ابتدا دیتاست MNIST که متشکل از دست‌نویس اعداد ۰ تا ۹ است را بارگذاری می‌کنم. ۲۵ عدد دیتا رندوم را نشان می‌دهم. در مرحله بعد با استفاده از تابع `to_categorical` برچسب‌ها را به `one_hot` که یک نوع آرایه `sparse` است تبدیل می‌کنم. در مرحله بعد سعی می‌کنم عکس‌ها را به ۲۸ در ۲۸ `resize` کنم. برای این کار ابتدا از تابع `expand_dims` استفاده می‌کنم که آرایه را یک بعدی می‌کند. در مرحله بعد آرایه را به ۲۸ در ۲۸ تبدیل می‌کنم و با تقسیم بر ۲۵۵ نرمالایز می‌کنم و در آخر همه را به `[n_imgs 784 1]` دوباره شکل دهی می‌کنم و عملاً کار تابع `flatten` را که برای گرفتن ورودی کاربرد دارد را انجام می‌دهم. در مرحله بعد شبکه عصبی را تشکیل می‌دهم و برای این کار سه لایه `Dense` با تعداد نوروں ۱۲ تشکیل می‌دهم و بین هر لایه `Dense` هم از تابع فعال‌سازی `ReLU` که معمول‌تر است و فرمول آن  $\max(0, f(x))$  است استفاده می‌کنم و در نهایت برای لایه آخر از تابع فعال‌سازی `softmax` که برای دسته‌بندی چند کلاسه است استفاده می‌کنم. در نهایت برای مدل از بهینه‌ساز آدام استفاده می‌کنم و برای تابع ضرر هم از `categorical_crossentropy` استفاده می‌کنم. در مرحله بعد این مدل را با داده حال حاضر با استفاده از تابع `fit` آموزش می‌دهم. نتایج برای MNIST در عکس پایین قابل مشاهده است. چون که این مجموعه داده بسیار ساده است درصد دقت بالا است و MLP هم قادر به دسته‌بندی این مجموعه داده هست.



برای مجموعه داده CIFAR10 هم از همین روش‌ها استفاده می‌کنم صرفاً چون سه کاناله است یک تبدیل به عکس خاکستری اضافه‌تر می‌زنم. چون این مجموعه داده پیچیده‌تر است و الگوهای سخت‌تری برای یادگیری با MLP دارد نتیجه خیلی خوبی نمی‌گیرد و درصد دقت آن بسیار پایین و حدود ۱۸ درصد است. در پایین نمودار دقت آن قابل مشاهده است.

