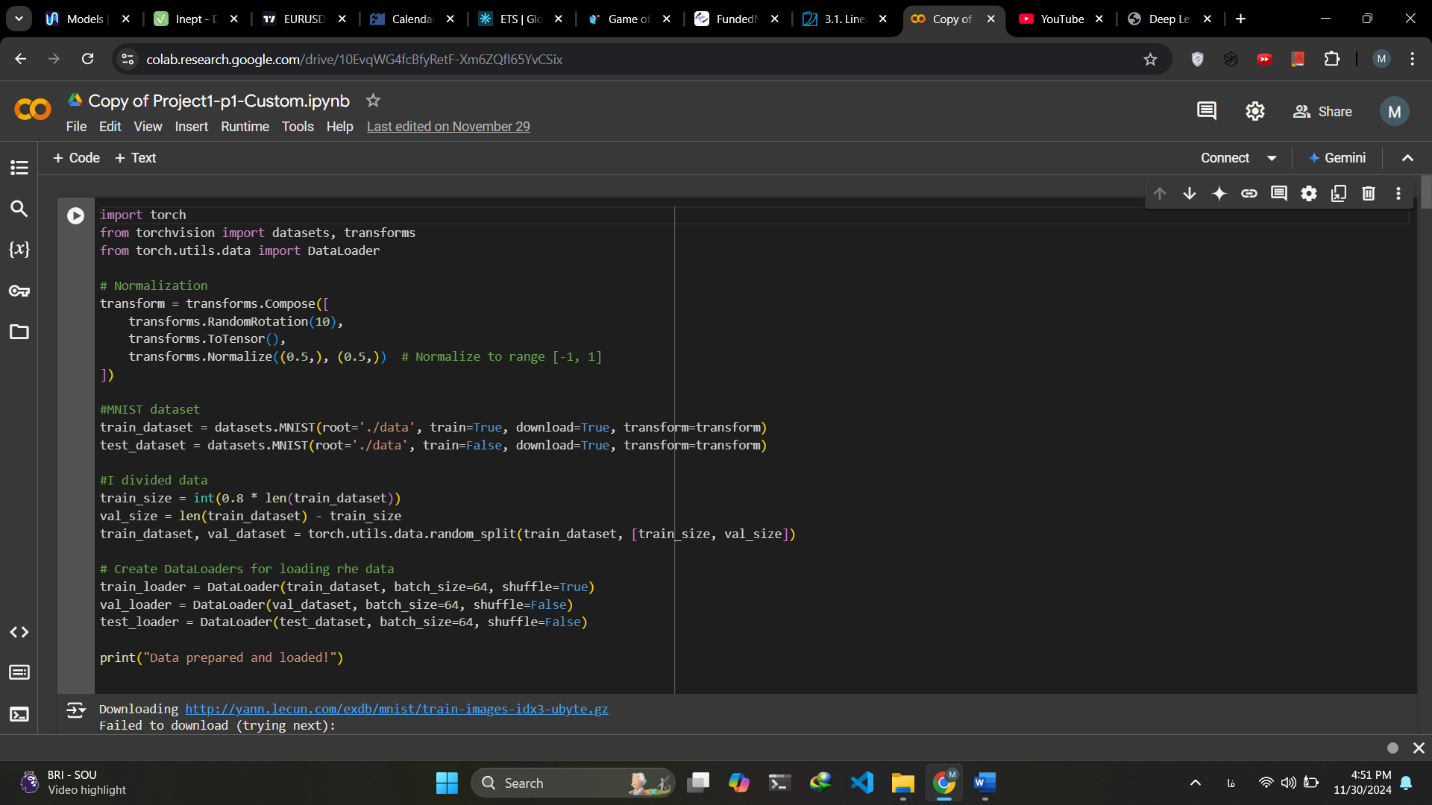
به نام خدا

پروژه اول دیپ لرنینگ

محمد محسن همایونی

**بخش اول:**

در بخش اول من دیتاست MNIST رو برای ترین کردن مدل دلخواه و مدل پری ترین شده انتخاب کردم. ابتدا دیتاست را دانلود نموده و دیتا ها را pre-process کردم.

****

دیتا را به سه بخش train, val, test تقسیم کردم و با کمک ترنسفورم آن را تبدیل به تنسور و سپس نرمال کردم.

با جستجو در وب به یک مدل پیشنهادی رسیدم و آن را 10 ایپاک ترین کردم که به مرور دقت آن بهبود پیدا کرد و در ایپاک آخر به دقت 98.84% رسید.

برای ResNet ابتدا ورودی را از تصویر تک بعدی grayscale تبدیل به یک تصویر 3 بعدی کردم که با ورودی معمول آن مطابقت داشته باشد. سپس Max-pooling را حذف کردم و خروجی آن را که در حالت عادی 256 عدد است به 10 عدد (برای اعداد 0 تا 9) تغییر دادم.

در ResNet و در اولین تلاش accuracy خیلی پایینی داشتم در حدود 50% برای حل این موضوع ابتدا تصویر را در بخش resize به اندازه 224\*224 که اندازه خود ResNet است در آوردم ولی در این حالت مشکلی که ایجاد میشد پر شدن سریع رم و از کار افتادن گوگل کولب بود . برای بهبود این شرایط ابتدا اندازه را 64\*64 کردم سپسبچ سایز را از 64 به 16 کاهش دادم و تعداد ایپاک ها را از 10 به 5 کاهش دادم. پس از این تغییرات سرعت آموزش افزایش پیدا کرد و دقت هم به مقدار قابل توجهی بالا رفت. ولی به چیزی که من میخواستم نمی رسید برای همین مقدار learning rate را به طور قابل توجهی کاهش دادم. زمان آموزش بالاتر رفت ولی دقت به چیزی که مد نظرم بود رسید. در آخرین ایپاک دقت مدل به 99.17% رسید.

\*مقایسه ی وزن های دولایه از هر مدل در نوت بوک مربوطه آورده شده است.

در کل مدل کاستومی که ساخته شد زمان بیشتری برای یادگیری داشت و در نهایت حتی با تعداد ایپاک بالاتر هم نتوانست به دقت مدل pre-train برسد. مدل Resnet با اینکه تعداد ایپاک کمتری داشت ولی به علت دارا بود وزن های قبلی توانست دقت بسیار بهتری ارائه دهد.

اوپتیمایز های هر دو مدل از نوع آدام استفاده شده است که به سریع تر شدن یادگیری با تغییر پارامتر کمک میکند.

**بخش دوم:**

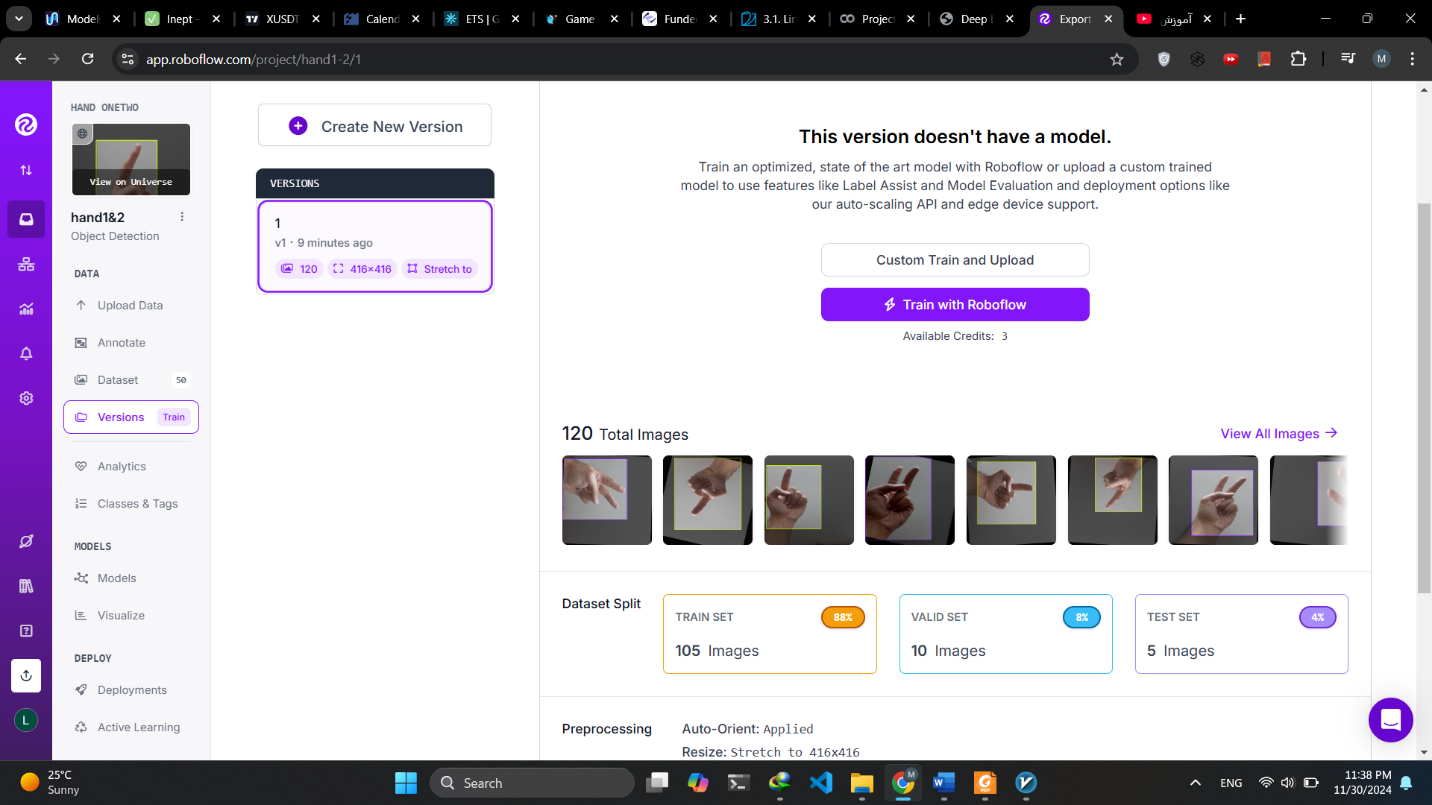
برای این بخش برای راحت تر شدن کار و استفاده از قابلیت های گوگل دیتاست را به صورت مستقیم بر روی درایو دانلود کردم و از آن استفاده کردم.

سپس با کمک کد گفته شده در ویدئو دیتا ست را با کمک پسوند های فایل تقسیم بندی کردم و آنها را به سه بخش test, train, val تقسیم کردم .

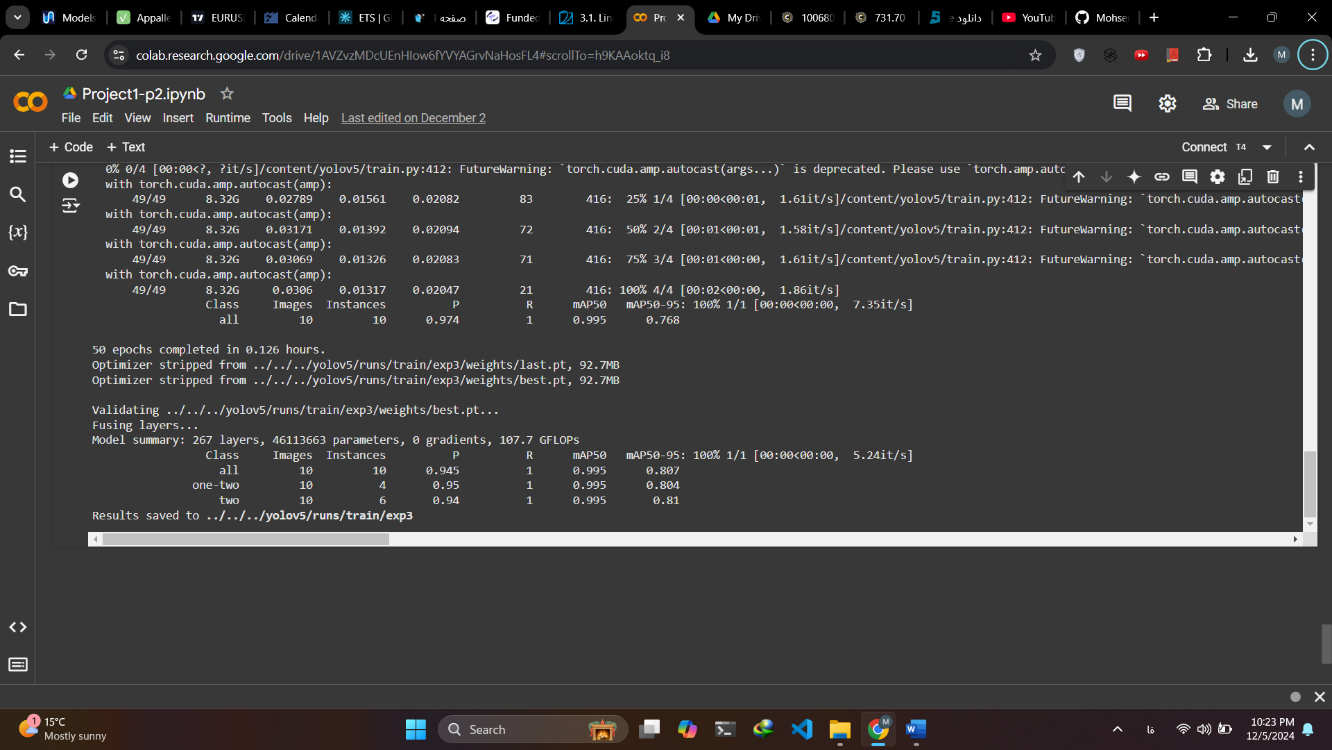
در انتها و قبل از فاین تیون کردن مدل فایل data.yaml را نسبت به آدرس پوشه ها و کلاس های گفته شده در توضیحات دیتاست تغییر دادم وفایل جدید را در پوشه مربوطه قرار دادم.

برای مرحله fine-tune از مدل s در yolo5 استفاده کردم چون دقت خیلی بالایی نمیخواستم. در ابتدا اندازه بچ من 64 و تعداد ایپاک 50 بود که باعث میشد گوگل کولب سشن را قطع کند برای همین اندازه بچ را 32 و ایپاک را 30 قرار دادمو مدل fine-tune شد و دقت قابل قبولی کسب کرد.

برای بخش بعدی که ساختن دیتاست بود حدود 54 عکس از اعداد یک و دو که با دست نمایش داده شده بود گرفتم . ابتدا فرمت عکس ها را به jpeg تغییر دادم و آنها را به کمک roboflow لیبل زدم و 4 نوع data augmentation انجام دادم تا هم تعداد داده ها بیشتر شود و هم دقت بالا تر برود.



**در انتها دیتاست را به مدل یولو دادم و نتیجه مورد قبولی گرفتم. برای این نتیجه با مدل l یولو و با 50 ایپاک ترین کردم.**

****