

ML learning requires HUGE DATASETS in order of Terrabytes.

I saw DIDI's cab data around 3TB somewhere on the web, if I am to train a self driving car model using data from Tesla, GM, Uber, BMW. How would I process these many TB of data on the blockchain?

How would the training take place? Could anyone please delineate as to how it will happen? (step by step that is...)

On the Discord I saw that current SGX limitation is 4gb, how would training happen if the raw data is in TBs ?