

## Post 4

بمعنى انى من خلاله ققدر اشتغل out-of-core model هو عبارته عن Stochastic Gradient Descent ال feed فى كل مره وبيعمله one-instance وده لانه سريع جدا بسبب انه بياخد online learning model على stochastic بس فى نفس الوقت كلمة feed الى كانت بتاخذ كل الداتا وتعملها Batch للموديل بتاعك على عكس ال فى نفس الوقت هنلاقي ال overlapping فى كل مرة ويمكن يحصل random instance ديه بتعنى انى بياخد بتعلى وتوطى لكن بعد فترة معينه هنلاقيها بتنزل وده بسبب انى الموديل لسه مشفش داتا كثير وال cost functions فى الحالة ديه بتكون غير منتظمة طبيعى للتعلم غير منتظم ولكن ده بيساعد الموديل انه يهرب من ال cost function learning rate ولكن مش احسن حاجه وده من خلال ال global minimum بس بيقترب لل local minimum قلة الخطوة بتاعته ده بيساعد كثير global بمعنى انى فى كل ما تقترب من ال learning schedule وانى اعمل وامشى فى ال shuffle ده عن طريق انى اعمل randomness ويمكن اتجنب موضوع ال global انى اوصل لل بحيث ميكونش shuffle لآخر واحد ورا بعض ولكن اخلى بالى من عملية ال instance بتاعته من اول process ما مثلا بييجى ورا بعض بشكل متكرر كثير جدا class فيه

```
sgd_reg = SGDRegressor(max_iter=1000, tol=1e-3, penalty=None, eta0=0.1)
```

ده بيوضح انى لما tol ال regularization ولكن من غير اى train for SGD regressor كده انا ققدر اعمل على الفاضى وال iterations بتقل بقيمة اقل من 001. اوقف الموديل بدل ما اعمل loss function اوصل انى ال eta learning rate هى ال

[https://www.linkedin.com/posts/abdelrahman-rezk\\_machinelearning-regression-activity-6805212107353714688-Vpj9](https://www.linkedin.com/posts/abdelrahman-rezk_machinelearning-regression-activity-6805212107353714688-Vpj9)

## Post 5

لا منك هتاخذ كل الداتا مرة واحدة فيكون Stochastic وبين ال Batch هو مرحلة وسط ما بين ال Mini-batch ال hardware بحيث يكون سريع جدا وكمان فى ميزة وديه ليها علاقة بال one-instance بطيء جدا ولا منك هتاخذ او مضاعفات ال 32 زى instances بيكون اسرع مع 32 train فمثلا ال matrix operations مع ال design ال global minimum يقرب اكثر لل Stochastic على عكس ال Mini-batch فى نفس الوقت ال ram بظبط ال اكثر ولكن ده ممكن global فى وقت اقل وده لانى باخذ جزء من الداتا وكل ما كان الجزء اكبر كان قربة سريعا لل بينما minimum عطلول راح لل Batch وهنلاحظ هنا الفرق بين ال 3 وازاى ال local minimum يخلية يقع فى global minimum الاتنين التانيين حوالين ال لانه بيرن على كل الداتا iteration بياخذ وقت كبير جدا فى كل Batch ولكن طبعا ال الى اكلنا عليهم فى الكام پوست الى فاتوا ولسه مكملين بإذن الله linear model والجدول ده ملخص لل SVD standard for Singular value decomposition ال Normal equations مع ال singular matrix فى حالة انى ال inverse الى بتحاول تلاقى

[https://www.linkedin.com/posts/abdelrahman-rezk\\_machinelearning-linearregression-activity-6805430258016301057-ORMo](https://www.linkedin.com/posts/abdelrahman-rezk_machinelearning-linearregression-activity-6805430258016301057-ORMo)