

Post 1

minimize for cost function الى بتعمل Params علطول عشان اجيب ال Gradient Descent مش لازم اروح اعمل ديه بدل ما اعمل Params ال Direct اروح اشتغل عادى جدا واجيب Normal Equation لا أنا ممكن عن طريق ال Applicable ديه ولكن ده دايمًا مش Params انا مش عارف عددها ممكن يكون كام عشان اوصل لافضل قيم لل Iterations بحيث انى فى بعض الحالات ده مش هينفع زى مثلاً لو عندى داتا خاصه بالصور وحجم الصور ده كبير الصورة الواحدة ممكن هنا بتكون قريبه من Complexity وده لان ال Gradient Descent تكون اكثر من 10000 بيكسل مثلاً فى الحالة ديه بروح لل الى عندك فى الحاله Data مناسب وفى نفس الوقت الميمورى بتاعتك تقدر تشيل ال Features لكن لو كان عندك عدد $O(n^3)$ Params تجيب ال Direct ديه ممكن تروح

Theta_hat: is the value of theta that minimize the cost function very well.

T: transpose means make convert rows to columns and columns to rows

-1: Is the inverse of the matrix.

https://www.linkedin.com/posts/abdelrahman-rezk_machinelearning-mathematics-activity-6800158010577612800-GunP

Post 2

الى عنده معلومات زيادة ياريت ينورنا فى الكومنتات ^^

الى model paramters لل Tweak بيروح فيها يعمل iterative process هو عبارة عن Gradient Descent ال cost function بتقلل ال

وبعد كده بيحاول tweak الجديدة بعد ما عملتلها params بالنسبة لل cost function فى كل خطوه هو بيحاول يشوف ال كمان بتفرق فى function وطبيعة ال slope يشوف الخطوه الجاية المفروض ينحدر فى انهى اتجاه وده بيختلف على حسب ال Global يروح لل model فده بيساعد ال Convex shape هى اصلاً عبارته عن Mean Square Error حاجة زى ال لكن بيظهر عندى مشكلة فى حجم الخطوه الى بحاول اخدها عشان مخدش وقت Local Minium لان مفيش اصلاً minium وده الى احنا Global وهو انى انتهى فى اماكن بعيدة عن ال diverge او انى يحصل Global كبير جدا عشان او اصل لل graph عشان حجم الخطوه يقل زى ما واضح فى ال Global وفى الغالب احنا بنقله كلما اتجهنا لل learning rate انسمية بيساعد الموديل يوصل اسرع بعيدا عن لو فيه قيم كبيرة اثناء العمليات الى بتحصل scaled تكون features كمان انى اخلى ال Arithmetic overflow ممكن يحصل عندك

إن شاء الله Gradient Descent المختلفة فى ال Approaches الجاي هيكون عن ال

https://www.linkedin.com/posts/abdelrahman-rezk_machinelearning-linearregression-activity-6800638120002932736-pNng

Post 3

هو learning rate الى هي مرتبطة بالموضوع ده ال model paramters بتوفر بعض ال sklearn ال كده وده الفائدة بتاعته هو التحكم فى summation وهو عامل زى علامه ال tolerance وال eta المعروف ب حاجه لا تذكر تغير بسيط cost function بمعنى انى بعد عدد معين الموديل بتاعك بيغير فى ال iterations عدد ال بتاعتك وهنا ال iterations جدا لانه بيكون قرب جدا جدا من امثل حل وفى الحاله ديه انتا محتاج توقف ال والى جايه او الى قبلها ويشوف iteration بتاع two gradient vectors ده بيروح يشوف الفرق بين tolerance الجميلة بيروح problem solving ايام ال epsilon الفرق لو كان اظغر من قيمة معينه الى احنا كتير بنسميها على iteration يوقف الموديل بدل ما يكون بيعمل الفاضى

https://www.linkedin.com/posts/abdelrahman-rezk_machinelearning-linearregression-gradients-activity-6800876316737589248-Kq2L