

## Post 8

عشان نقدر نعالج مشكلة ال **over fitting** فيه اكثر من طريقة لكنها بتختلف باختلاف ال **model** الى انا مستخدمه ، مثلا لو كنت مستخدم **polynomial degree** عالية تقليلها ممكن يعالج مشكلة ال **over fitting** لانك هتكون قلقك ال **constrain** بتاعت الموديل ، والى بيخلى ال **model** يعمل **over fitt** هو انه بيكون فيه **High Variance** فى الداتا بتاعتك بمعنى تنوع كثير ، والموديل هنا وهو بيعمل **train** بيحاول انه يعمل **tweak** للمتغيرات على اساس كل ال **patterns** او الاختلافات الى بيشفها فى الداتا ، رغم انى بعض الامثلة ممكن تكون لا تمثل اى شىء مجرد **noise** او وجودها فى الداتا بيمثل فى الاخر مجرد **instance** واحد فقط مجرد **row** فى الداتا ، وهنا انا بحتاج انى قتل ال **constrain** بتاعت الموديل درجة الحرية بتعبير تانى فى تعامله مع الداتا عشان تقدر اعمل **generalization** ولما الموديل يشوف داتا جديدة يقدر انه يطلع نتائج كويسة ، لكن ساعات ال **plain regression model** نفسه بيعمل **over fitting** الى هو يعتبر ايسر موديل عندى ، وفى حالة زى كده بحتاج انى اما ازود عدد ال **training data** او انى اعمل ما يسمى ب **regularization** وال **regularization** ده نفسه بيختلف فى عندى **Ridge Regression** وفى **lasso regression** وغيرهم وهو بيختلف فى الجزء الى بضيفه لل **cost function** عشان اعمل **regulaization**.

## Post 9

ال **regularization** ببساطة هو انى بعد ما ال **model** عمل **tweak** للمتغيرات هروح انا اضيف عليها حاجة بسيطة تخليها متكنش حساسة لكل **instance** على حدا عشان ميعصلش **over fitting** ، والجزء الى بضيفه ده بيختلف ، فى البوست ده هنشوف واحد من ال **Regulaizatio term** وهو ال **Ridge Regression** او الى نسمع عنه ال **L2 Norm** وهو انى اضيف الجزء الى فى الصورة بعد علامة + لل **gradient vector of cost** function بتاعى ، وده بيحصل فقط فى ال **training** وساعت اصلا ال **cost function** الى بستخدمها فى ال **training** بستخدم **evaluation method** تانية غيرها فى ال **testing** ، وهنا بقا ال **l2** ده وعن طريق **hyper parameter** تانى هو **lambda** او زى ما هو فى **sklearn** اسمه **alpha** وهنا زاد عندى متغير تانى محتاج اظبطة بجانب ال **learning rate** الى بيتحكم فى ال **step** بتاعت الموديل فى ال **learning** ، و **alpha** هنا لو كان بصفر هنلاقى انى الترم كله بعد ال + بقا سفر وكانى معملتش اى شىء وعبارة عن **linear model** الى عادى الى بستخدمه ، ولو كان قيمة كبيرة جدا هلاقى انى ال **weights** الاخيرة بقت قريبة جدا من 0 وبدل ما كنت فى مشكلة **over fitting** بقيت فى مشكلة تانية وهى ال **under fitting** زى ما هو واضح من ال **graph** هنلاقى ال **weights** بترجع لورا تانى وبتقرب من الصفر ، ال **ridge regression** كمان تقدر اطبقه مع ال **closed form** ال **Normal equation** الى بتجلى ال **weights** النهائية من خطوة واحدة وتقدر تستخدم الاتنين عن طريق اما **Ridge model** او عن طريق **SGDRegressor** من **sklearn**. غير ال **Ridge** فى **lasso regression** وده بيحاول يخلى ال **weights** المرتبطة مع ال **features** الاقل اهمية تكون باصفار وبكدة يعمل **features selection** لل **features** المهمة وفى **regularization** وسط بينهم وهو ال **Elastic Net** وممكن تقدر من خلاله اعمل **Ridge** فقط او **Lasso** فقط او انى يكون **mix** بينهم عن طريق **l1\_ratio parameter** ، كمان فيه طريقة تانية بدل ما اضيف **new hyper parameter** وهى ال **Early stopping** انى اوقف ال **training** لما ال **cost function** توصل ال **minimum** مع ال **validation set** وده لو عملت **plot** هلاقيها بتوصل لنقطة وبعد كده بتبتدى تعلى تانى والفرق بين **cost on training and validation** ابتدى يكبر فى الحالة ديه ابتدى اوقف ال **trainig** .