Supervised Learning-Regression Page > phenomenon that skews result of an algorithm in favor or against an Idea · systematic error in ML models -> incorrect assumptions.
· Diff b/w average prediction of our model & correct value.
· Bias & variance -> components of reducible error. Low Bias -> fewer assumptions -> form of target function -> by model

High Bias -> more assumptions -> unability capture imp features

of dataset -> can't perform well on new data
mamples -> Low Bias -> decision trees, k-nearest nighous & SVMnamples -> High Bias -> linear discriminant analysis & logistic organ. ways to Reduce High Bias-· Increase input features as model is underfitted. · Decrease regulatization time period · Use more complicate model -> that have polynomial features Increase too wing data. Inclase regularization term

Variance -

- changes in model when voing different portions of training dataset.
- how much random variable differs from its expected value-based on single towning set.
- measures (inconsistency of diff predictions using diff training sets.
- Model -> high variance -> low bias & vicevessa
- comes with highly complet model with lot of features.
 - contributes to flexibility the model.

high valiance signs -> Noise in dataset

- · everfiffing
- · complexity
- · Policing data points together.

• Low variance -> low variation in prediction of target furtions.

• High variance -> large variation in prediction of target furtions.

To Reduce High Variance

- Reduce input feature or no of pareineters as a model is overfitted.
- Do not use a much complex model
- Inculase training data.
- Incluse the regularization teem-

& Bias & Variance trade-off_ Inadoff bliv bias & variance is inversely proportional to variances, tackle by either increasing a) complexity of model

b) Inaining dataset. Overfitting statistical model fits against its matching data. · Algo/ model can't perform well on unseen data-. low bias of high variance. model - esistent model - at complex enought: match all datapoint & performs well · Reasons - · noisy data · training data is too small · large no. of featuresto avoid Overfitting Cross Validation · Train with mole data. Removing features Early stopping the baining legilar Latien emeling

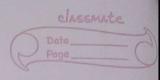
Underfitting -* · Data model is unable to capture relationship blu input & output. occurs whem model is too simple High bias low variance. · less dals - for model.

Jo avoid underfitting. > o growing educational time of model
> by increasing the wide variety of functions Rogsession -. H · functional relationship between two or more variables

that often emplically determined from data & vscd topocalist

values of one variable given value of others. · variable -> influence on output > known as emplomates, indept.
input or predictor variable-· Variable -> entrone repends on what variable -> dependent variable of dataset in which target values all known. (correlation b/w dependent & indepent variables). desired ouput -> one or more continous input -> regression-

Inter kignession— med also > successed learning predict dependent variable based on Enderon dant. finding straight line that best approximates set of pointer on graph. simple linear regression > only 1 independent variable x > independent variable. x > independent variable. y > dependent variable. v > entitiers affect over simplifies real world problems. vse case > o Demand Fercasting of Healthcare of other predictions. degistic Regression. und to model & estimate probability of an event estimate problem probability of an event estimate problem probability of an event signal curve or S-curve in short signal curve or S-curve in short ent come data points > 0 or 1 out come data points > 0 or 1 solve clarification issues predict sekelshood of categorial dependent variable.		Classmate	1
multiple linear regression -> only 1 independent variable x -> independent variable based an Endependent simple linear regression -> only 1 independent variable x -> independent variable x -> independent variable y -> dependent variable y -> entitiers affect oversimplifies seal world problems be cases -> o Demand Ferecasting of feelthcare of their predictions be cases -> o Demand Ferecasting of an event befinate problem probablish of an event signation curve or S-curve in short out come data points -> O as 1 solve clarification issues predict lekelihood of categorial dependent variable.		Joseph Repressione	
multiple linear regression -> only 1 independent variable x -> independent variable based on Endependent simple linear regression -> only 1 independent variable x -> independent variable x -> independent variable y -> dependent variable y -> entitiers affect · over simplifies seal world problems be cases -> o Demand Ferecasting -> Healthcare - Other predictions lise cases -> o Demand Ferecasting -> Healthcare -> Other predictions degistic Regression. wed to model & estimate probability of an event whimate problem probability of an event sigmaid curve or S-curve in short out come data points -> O as 1 solve clarification issues predict lekelihood of categorial dependent variable.	*	Amen reguestion -	
x = independent variable y = dependent variable. Output. v = \(\text{N} = \text{P} \) + \(\text{P}_1 + \text{P}_2 \text{x}_2 + \text{P}_3 \text{x}_3 + \text{ Bn} \text{x}_n + \text{E} \text{Disadv} = \cdots \text{outliers affect} \(\text{ovels implifies real world problems} \) \(\text{Vse casts} = \cdots \text{Demand Ferecasting} \text{Healthcase} \cdot \text{Other predictions}. \(\text{Vse casts} \text{ independent} \text{problems} \text{Pealthcase} \text{Other predictions}. \(\text{Vse casts} = \text{Verecasting} \text{Healthcase} \text{Other predictions}. \(\text{Vse casts} \text{Vse of model & estimate probablisty of an event \text{event beauting} \text{valse (independent)}, \text{s a curve}. \(\text{s a curve} \) \(\text{s and come data points} = \text{O on 1} \) \(\text{out come data points} \) \(\text{o on 1} \) \(\text{out come data points} \) \(\text{Solve} \) \(\text{classification} \) \(\text{s rues} \) \(\text{predict behelihood of categorial dependent variable}. \end{area}	1	ml alon	
x = independent variable y = dependent variable. Output. v = \(\text{N} = \text{P} \) + \(\text{P}_1 + \text{P}_2 \text{x}_2 + \text{P}_3 \text{x}_3 + \text{ Bn} \text{x}_n + \text{E} \text{Disadv} = \cdots \text{outliers affect} \(\text{ovels implifies real world problems} \) \(\text{Vse casts} = \cdots \text{Demand Ferecasting} \text{Healthcase} \cdot \text{Other predictions}. \(\text{Vse casts} \text{ independent} \text{problems} \text{Pealthcase} \text{Other predictions}. \(\text{Vse casts} = \text{Verecasting} \text{Healthcase} \text{Other predictions}. \(\text{Vse casts} \text{Vse of model & estimate probablisty of an event \text{event beauting} \text{valse (independent)}, \text{s a curve}. \(\text{s a curve} \) \(\text{s and come data points} = \text{O on 1} \) \(\text{out come data points} \) \(\text{o on 1} \) \(\text{out come data points} \) \(\text{Solve} \) \(\text{classification} \) \(\text{s rues} \) \(\text{predict behelihood of categorial dependent variable}. \end{area}		· m ago -> supervised learning	
x = independent variable y = dependent variable. Output. v = \(\text{N} = \text{P} \) + \(\text{P}_1 + \text{P}_2 \text{x}_2 + \text{P}_3 \text{x}_3 + \text{ Bn} \text{x}_n + \text{E} \text{Disadv} = \cdots \text{outliers affect} \(\text{ovels implifies real world problems} \) \(\text{Vse casts} = \cdots \text{Demand Ferecasting} \text{Healthcase} \cdot \text{Other predictions}. \(\text{Vse casts} \text{ independent} \text{problems} \text{Pealthcase} \text{Other predictions}. \(\text{Vse casts} = \text{Verecasting} \text{Healthcase} \text{Other predictions}. \(\text{Vse casts} \text{Vse of model & estimate probablisty of an event \text{event beauting} \text{valse (independent)}, \text{s a curve}. \(\text{s a curve} \) \(\text{s and come data points} = \text{O on 1} \) \(\text{out come data points} \) \(\text{o on 1} \) \(\text{out come data points} \) \(\text{Solve} \) \(\text{classification} \) \(\text{s rues} \) \(\text{predict behelihood of categorial dependent variable}. \end{area}		o predict dependent variable based	-
x = independent variable y = dependent variable. Output. v = \(\text{N} = \text{P} \) + \(\text{P}_1 + \text{P}_2 \text{x}_2 + \text{P}_3 \text{x}_3 + \text{ Bn} \text{x}_n + \text{E} \text{Disadv} = \cdots \text{outliers affect} \(\text{ovels implifies real world problems} \) \(\text{Vse casts} = \cdots \text{Demand Ferecasting} \text{Healthcase} \cdot \text{Other predictions}. \(\text{Vse casts} \text{ independent} \text{problems} \text{Pealthcase} \text{Other predictions}. \(\text{Vse casts} = \text{Verecasting} \text{Healthcase} \text{Other predictions}. \(\text{Vse casts} \text{Vse of model & estimate probablisty of an event \text{event beauting} \text{valse (independent)}, \text{s a curve}. \(\text{s a curve} \) \(\text{s and come data points} = \text{O on 1} \) \(\text{out come data points} \) \(\text{o on 1} \) \(\text{out come data points} \) \(\text{Solve} \) \(\text{classification} \) \(\text{s rues} \) \(\text{predict behelihood of categorial dependent variable}. \end{area}		, finding straight line that	
x = independent variable y = dependent variable. Output. v = \(\text{N} = \text{P} \) + \(\text{P}_1 + \text{P}_2 \text{x}_2 + \text{P}_3 \text{x}_3 + \text{ Bn} \text{x}_n + \text{E} \text{Disadv} = \cdots \text{outliers affect} \(\text{ovels implifies real world problems} \) \(\text{Vse casts} = \cdots \text{Demand Ferecasting} \text{Healthcase} \cdot \text{Other predictions}. \(\text{Vse casts} \text{ independent} \text{problems} \text{Pealthcase} \text{Other predictions}. \(\text{Vse casts} = \text{Verecasting} \text{Healthcase} \text{Other predictions}. \(\text{Vse casts} \text{Vse of model & estimate probablisty of an event \text{event beauting} \text{valse (independent)}, \text{s a curve}. \(\text{s a curve} \) \(\text{s and come data points} = \text{O on 1} \) \(\text{out come data points} \) \(\text{o on 1} \) \(\text{out come data points} \) \(\text{Solve} \) \(\text{classification} \) \(\text{s rues} \) \(\text{predict behelihood of categorial dependent variable}. \end{area}		best approximates get al	1
x = independent variable y = dependent variable. Output. v = \(\text{N} = \text{P} \) + \(\text{P}_1 + \text{P}_2 \text{x}_2 + \text{P}_3 \text{x}_3 + \text{ Bn} \text{x}_n + \text{E} \text{Disadv} = \cdots \text{outliers affect} \(\text{ovels implifies real world problems} \) \(\text{Vse casts} = \cdots \text{Demand Ferecasting} \text{Healthcase} \cdot \text{Other predictions}. \(\text{Vse casts} \text{ independent} \text{problems} \text{Pealthcase} \text{Other predictions}. \(\text{Vse casts} = \text{Verecasting} \text{Healthcase} \text{Other predictions}. \(\text{Vse casts} \text{Vse of model & estimate probablisty of an event \text{event beauting} \text{valse (independent)}, \text{s a curve}. \(\text{s a curve} \) \(\text{s and come data points} = \text{O on 1} \) \(\text{out come data points} \) \(\text{o on 1} \) \(\text{out come data points} \) \(\text{Solve} \) \(\text{classification} \) \(\text{s rues} \) \(\text{predict behelihood of categorial dependent variable}. \end{area}	-	simple linear regression - and	1
x = independent variable y = dependent variable. Output. v = \(\text{N} = \text{P} \) + \(\text{P}_1 + \text{P}_2 \text{x}_2 + \text{P}_3 \text{x}_3 + \text{ Bn} \text{x}_n + \text{E} \text{Disadv} = \cdots \text{outliers affect} \(\text{ovels implifies real world problems} \) \(\text{Vse casts} = \cdots \text{Demand Ferecasting} \text{Healthcase} \cdot \text{Other predictions}. \(\text{Vse casts} \text{ independent} \text{problems} \text{Pealthcase} \text{Other predictions}. \(\text{Vse casts} = \text{Verecasting} \text{Healthcase} \text{Other predictions}. \(\text{Vse casts} \text{Vse of model & estimate probablisty of an event \text{event beauting} \text{valse (independent)}, \text{s a curve}. \(\text{s a curve} \) \(\text{s and come data points} = \text{O on 1} \) \(\text{out come data points} \) \(\text{o on 1} \) \(\text{out come data points} \) \(\text{Solve} \) \(\text{classification} \) \(\text{s rues} \) \(\text{predict behelihood of categorial dependent variable}. \end{area}		murale linear seasons independent variable	1
Disadu > • cutliers affect. • oversimplifies real world problems: Use case > • Demand Ferecasting • Healthcare • Other predictions. Lagistic Regression. Lagistic Regression.		> = soding > more than I independ	4
Disadu > • cutliers affect. • oversimplifies real world problems: Use case > • Demand Ferecasting • Healthcare • Other predictions. Lagistic Regression. Lagistic Regression.		malpendant variable variable	13.
Disadu > • cutliers affect. • oversimplifies real world problems: Use case > • Demand Ferecasting • Healthcare • Other predictions. Lagistic Regression. Lagistic Regression.		y - appointment variatete. Output.	1
Disadu > • cutliers affect. • oversimplifies real world problems: Use case > • Demand Ferecasting • Healthcare • Other predictions. Lagistic Regression. Lagistic Regression.		Cooperate to the second	1
vse cases > 0 permand Farecasting of the predictions. Vse cases > 0 permand Farecasting of the predictions. Against Regression. I will to model of estimate probablity of an event occurring > rate (independent) occurring > rate (independent) occurring > rate (independent) occurring > occurrent occurring > occurrent occurring > occurrent occur		· Y = Bo + B1 + B2 X2 + B	
vse cases > 0 permand Farecasting of the predictions. Vse cases > 0 permand Farecasting of the predictions. Against Regression. I will to model of estimate probablity of an event occurring > rate (independent) occurring > rate (independent) occurring > rate (independent) occurring > occurrent occurring > occurrent occurring > occurrent occur		P3X3 +	1
vse cases > 0 permand Ferecasting of the predictions. Legistic Regression. Legistic Regression. Lestimate problem probablity of an event. So a curve. Signate curve or S-curve in short. Legistic come data points > 0 or 1. Lettimate problem is short. Legistic come data points > 0 or 1. Lettimate problem is short. Lettimate problem is short. Lettimate problem is short. Lettimate problem is short. Lettimate problem. Lettimate problem. Lettimate problems.		Dready -> 0 pullered 11	The state of the s
degistic Regression. Need to model & estimate probablity of an event estimate problem probablity of an event some or S-curve in short signal curve or S-curve in short out come data points > 0 or 1 nulti > cat, clys, wan solve clarification issues predict excellibood of categorial depardat variable.		anguers affect.	1
degistic Regression. Need to model & estimate probablity of an event estimate problem probablity of an event some or S-curve in short signal curve or S-curve in short out come data points > 0 or 1 nulti > cat, clys, wan solve clarification issues predict excellibood of categorial depardat variable.		les colles vers mplifies real world make	-
degistic Regression. Need to model & estimate probablity of an event estimate problem probablity of an event some or S-curve in short signal curve or S-curve in short out come data points > 0 or 1 nulti > cat, clys, wan solve clarification issues predict excellibood of categorial depardat variable.	N. Section	Use cans -> Demand Feregating at 11	-
degistic legelssion. wed to model & estimate probablity of an event. estimate problem probablity of an event. signate problem probablity of an event. signated curve or S-curve in short. ent come data points > 0 or 1. nulti > cots, class, bron. solve clarification issues. predict lekelihood of categorial departant variable.		Jealthcale other predictions.	-
degistic legelssion. wed to model & estimate probablity of an event. estimate problem probablity of an event. signate problem probablity of an event. signated curve or S-curve in short. ent come data points > 0 or 1. nulti > cots, class, bron. solve clarification issues. predict lekelihood of categorial departant variable.	1	sample make south of the contract of the property of the contract of the contr	-
· used to model & estimate probablity of an event. • estimate problem probablity of an event. • sta curve. • stemate problem probablity of an event. • stemate probablity of an event.	1	Logistic Regression.	1
· used to model & estimate probablity of an event. • estimate problem probablity of an event. • sta curve. • stemate problem probablity of an event. • stemate probablity of an event.	7	Of the state of th	Jan .
ent come data points \Rightarrow 0 or 1. O \Rightarrow totally uncertain $1 \Rightarrow$ Gertain adjos, bon solve clarification issues. Predict ve avaluation algo. predict vekelihood of categorial dependent variable.		Real of the way to me a way the way	1
ent come data points \Rightarrow 0 or 1. O \Rightarrow totally uncertain $1 \Rightarrow$ Gertain adjos, bon solve clarification issues. Predict ve avaluation algo. predict vekelihood of categorial dependent variable.		wed to 110 and of the major of the world	-
ent come data points \Rightarrow 0 or 1. O \Rightarrow totally uncertain $1 \Rightarrow$ Gertain adjos, bon solve clarification issues. Predict ve avaluation algo. predict vekelihood of categorial dependent variable.		estimate robablity of	Long
ent come data points \Rightarrow 0 or 1. O \Rightarrow totally uncertain $1 \Rightarrow$ Gertain adjos, bon solve clarification issues. Predict ve avaluation algo. predict vekelihood of categorial dependent variable.		pelen peoplem of an our	-
ent come data points \Rightarrow 0 or 1. O \Rightarrow totally uncertain $1 \Rightarrow$ Gertain adjos, bon solve clarification issues. Predict ve avaluation algo. predict vekelihood of categorial dependent variable.	1	Is a curve. The lindereder	-
ent come data points \Rightarrow 0 or 1. O \Rightarrow totally uncertain $1 \Rightarrow$ Gertain adjos, won solve clarification issues. Predict ve avaluation algo. predict vekelihood of categorial dependent variable.	1	· sigmaid curve or S- will so a line (Tues)	1
of totally uncertain 1 -> certain adjoshon and the control of categorial dependent variable.	1	out come de la come in short	1
solve clarification issues. Predictive avaluation algo. predict lekelihood of categorial dependent variable.		The state of the s	1
predict lekelihood of categorial dependent variable.			-
predict vekelihood of categorial dependent variable.		clarification issues.	-
the whood of Categorial dependent variable.	1	Predictive evaluation along	to produce to
fly -	1	predict per old and all and and and all and and and all and and and all and	1
f(x) - 1	1	aregor al approant valiable.	1
+ V -	1	01.	1
	VI	f(x) = 1	1
$f(x) = \frac{1}{1 + e^{-x}}$		1+0-2	1



· dasso Regression -

· lessen complexity of version

· similar to ridge regor except that penally time includes only absolute weight (insted of rectangular of weights.

· . L1 Regularisation

· avoids everlitting of toaining data & useful for feature schehon shrinkage appeared.)

· overcome disadu. of Ridge by not punishing high values of coefficient B > actually setting them to see if not relient.

$$L(x,y) = Min\left(\sum_{i=1}^{n} (y_i - w_i^* x_i^*)^2 + \lambda \sum_{i=1}^{n} |w_i^*|\right)^{\frac{1}{2}}$$

Ridge Regression -

8mall quantity bias delivered > higher long time predictionsquantity added to bias > ridge regression penalty
L2 Regularisationpenalty term with aid of multiply inith landa to the
squared weight of each individual features-

$$L(x,y) = Min\left(z_{i=1}^{n}(y_{i}-w_{i}x_{i}^{n})^{2} + \lambda z_{i=1}^{n}(w_{i}^{n})^{2}\right)$$

Reduce complexity of verist.

. clears up troublif we have greater parameter than samples

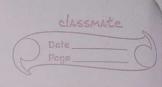
 $\Lambda \propto ie \lambda \rightarrow more coefficient become $20 volonist to collinearily.$

Gradient Descent Algofinds best fit line for giving training datset optimization algo -> minimize cost function.

locate least possible value -> fulfill agiron fee fuction. Batch > process all training ex. for each iteration. Stochastic -> Process 1 toping example per iteration. Faster than Bakeh

Nini Batch -> Mure examples out of total intraining examples are

processed per iteration. & June requirems - film) exist for every x in domain · Sifferential for an univariate funct line segment connecting two fun" paints lays on or above its curve. Non Convex HOLDES LE relatively · slow close to min. for paorly conditioned convex peals gradient points orthogonally to shortest direction tomin speed



Evaluation Metrics/Cost furction-

- establish events or set of value to represent loss of penalty in curred in gaining something.

 Aka loss of old function.
- . The objective f's that need to be maximized as called reward point whility factions.
- Mean Earl (ME)—

 mean of diff b/w achal & predicted value $ME = \underbrace{E(y_i x_i^2)}_{n}$
- . Mean Squaled est of diff b/ns anthetic & expected values extracted by squared the average distinction.

$$MSE = \underline{\leq (y_i - \chi_i)^2}$$

- · Mean Absolute error \(\frac{(y:-zi)}{n} \) → ·
- · RMSE Root Mean Squared River P

RMSF =
$$\int \underline{E}(y_i - x_i)^2$$

$$R^2 = sum of squares due to regression total sum of squares.$$

shows how well data fits regassion model.