Classmate Date Page

	-	Date Page		
	*	Regularization		
	Ì	x -> features		
		w -> weights		
		$\hat{q} \rightarrow \qquad $		
		Poynomial regression: High order polynomials		
		luss will be less		
		but leads to overfitting the start somerer 2018		
10 80		add penalty termines most arrived pur prostrate		
		(qu prispur)		
		y = ω, + ω, x + ω, x + ω, x + ω, x,		
		Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z		
	-	> High or do - Deliver and it was a line of		
		High order polynomials give better fit and lower data uses		
		data luss. radizzes relamiz a seconi.		
		However complicated hypotheses leads to overfitting		
		existing dota.		
-	100	ide a:		
	· Change the loss firstion to penalize hypothesis			
		renalize hypothesis		
		26107 FX no recent		
	Mary Season	bioh it no test		
TI	_			

*	Ridge Regression mes au amount moissant 321		
	MS E at 2 (norm of we vectors)		
	regularization		
	odursyo moisosels bone seximinosefficient of reaso		
	SUPEROUS AS TO MOUNT OF LOSSENS		
	3		
	10 20		
	if x = 2; y = 1		
	x= 2·1;		
	the state of the s		
	forcing my weight parameters to be small.		
	we have recorded to the		
	-> Also known as 12 or squared value regularization		
C.	1 11 of 11 a Paragraph of		
	-> Tries to reduce the length 11 w11 of the parameter		
	vector, promoting lesser dependency		
u.	predictors (lower model complexity).		
	-> > controls the regularization penalty. > = 0 results		
	in regular MSE		

 Page Page
LASSO Regression (feature selection)
$L_{1ASSO} = \frac{1}{2n} \frac{n}{j=1} (\hat{y_j} - \hat{y_j})^2 + n \frac{1}{2} \omega_i $
$\omega_1 + \omega_2 \leq C$
mide simple as
-> least Absolute Shrinkage and Selection Operator
(LASSO), also known as L, or absolute value
regularization.

change x > 10 to 20 Page Logistic Regression linear classifier: Z = w + w | x + - - - + w d x d · The output above is unbounded, whereas it should be between o and I, ie; $z = \left[0 \leq p(x) \leq 1\right]$ - Changing p by the same amount requires a bigger Changes in x when p is aloready large (or small) than when p is close to 1/2. log P(2) = wo + w, x, + - ... wexe 1-7(2) a sigmoid corve. p(x) = e² = 1 Itez et milte-2 Understanding resignated Corre y = 1 1+ e (ωx+5)

*	Loss function	
	Transaction of the second of t	
	-> In binary classification, t	he cross-entropy loss
	defined as:	
	L(y,p)= - (ylog p + 6	1-4) log (1-p))
	3/1	
	· It is a measure of the	difference between the
	predicted	
		m - 4 1 - w :) seas