	HWI-ML
	TASE
	HP 1 State of pol Louten Salat
	tikelihood tunckon:
	Fol bogottic regrettion, the probability
	P(Inlan) can be given by the Lymond
	function:
3.1	P(yn = 1/nn) = -1 1+e-win
	P(yn = 0/nn) = 1 - 1
	10 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	e-win
	To consoledate 2 equations
	P(m/nn) = ynwTnn
	1+0
1	Given the data, L(w) = This, P(4n/nn)
1	L(w) = #N 1

cog-likelihood function Take natural log on both Ledes (n (L(w)) = & in (-- now Tym) = E-cn (1+e-4nwin) To maninize the libelihood, words manimine log-libelihood & - (log likeliho 1.) Cross _ Entropy Errol Ein (w) fol logistic regression ?s: Ein(w) = - + & (n (-1+e-ynwTny) En (w) = 1 [In (1+ e-yn wTnn)

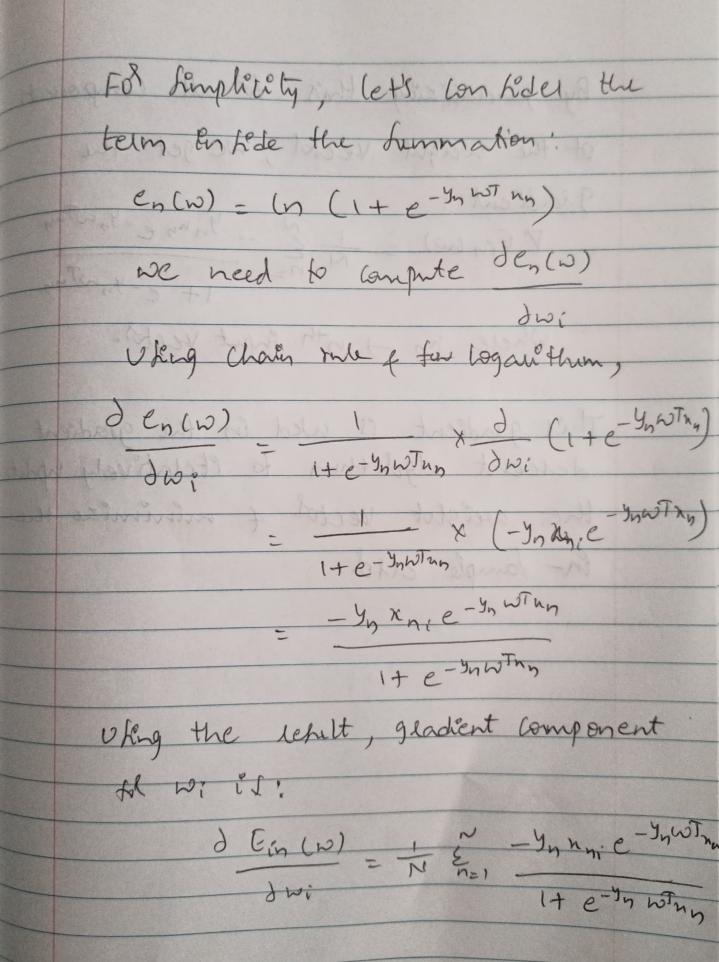
Now temporing the enpuestions, we See that minimizing the cross-Entropy end Ein (w) is equivalent to manimizing the log-likelihood In(L(W)) as both Envolve minimiting manimiting the same summation telm. Thus, we've shown that selecting the hypotheris h that maninizes the likelihood is equivalent to fol logistic legsestion. Companient cottespiending to the the pel house sol ma) jou have

Elin My F 192 (then can generally

MASSI ANDESS NOTES SHE AS 49

HP2 Marine and proportion work To devine the gradient of the in-Sange and want weight weath w, we Start with the defination of the aoss-entropy end for logistic regression Ein (w) = 1 & (n (1+ e-ynwīnn) Fol the Gordent decent algolithm, we need to determine gradient of this class with w.r.t weight becked w. The gradient well be a vector, and the Component collesponding to the its Vector wi can be found by differently Ein wrt wi 4 then can generalize

it for the entire weight vector.



By generalizing this fol all components at the wargut vected, we get the 9 ladient:

V Ein (w) = 1 2 - Innne-ynwTun

The ynwTun

The eynwTun where no - 1 nth Enput vector. This gradient is used in the gradient descent algolithm to Eteratively update the weight vector 4 minimize the en-tample elde trapegores tracked that sout part 127 100 20