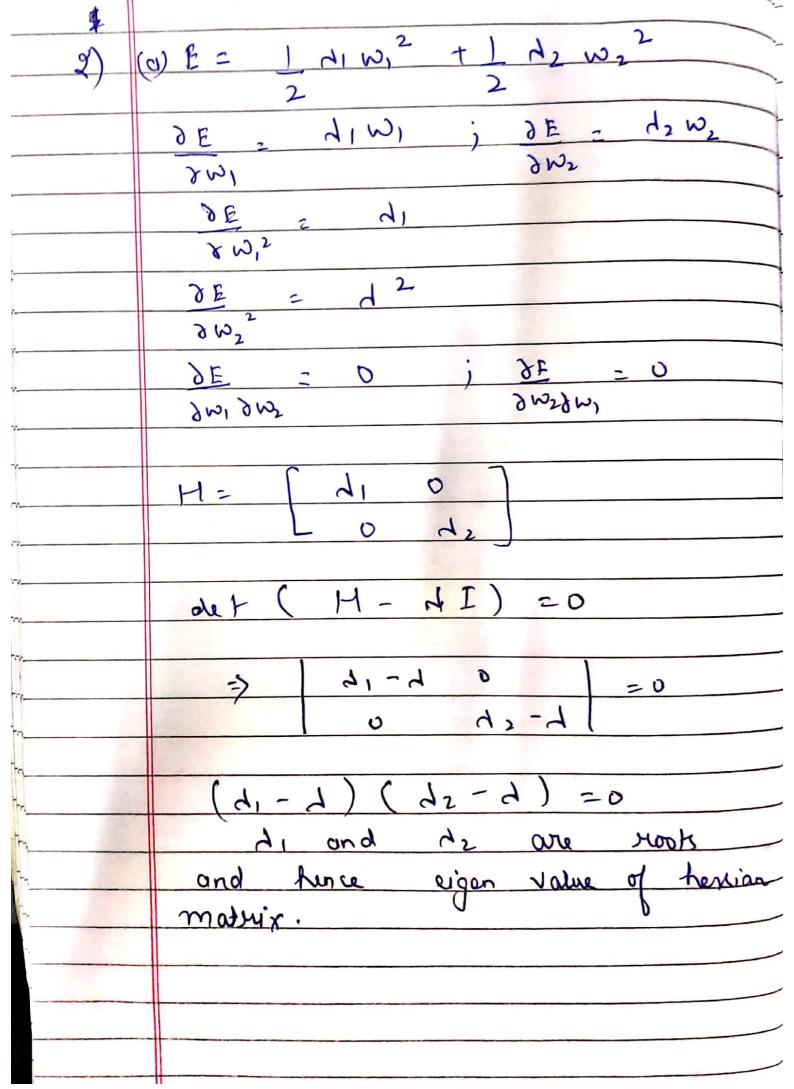
Λ . Υ	the a second subject to the last
nm 1	let a neural network contains
	hidden layer let w, be weights connecting input layer and Ist hidden layer and we be connections.  blu hidden layer and output  layer let X be the input.
	connecting input layer and the
	hidden layer and we se connections.
	blu hideen layer and output
	layour let X be the input.
	Output of this newsal not =
	$=$ $tanh(X W_1) W_2$
	Envior Z (Y - ton (xw1) W2)
	to delegreshed If WI is zono and
	Since funh is zero at 0, the
	This will lead to zero vertou at
	output layor. This will lood to

- atti	this will cause the evaluation of
- 100 1	gradiants equal to zero. Since
Da	gradient l'e revo the backprop
r	-agation will not update our
<u>-</u>	. I weight to and in the model
- Uic	will not learn mant all?
. 10	E set pointed It asker me
- et to	Thorefore; origin in weight space
i	Thorefore; origin in weight space is a stationary point.



(6)	We verify from our implementation
V.	that any bleauning less than
daird.	0.1 leade to convey gence.
ł '	The volue of the weights divolige
	Low learning rate of 0.1.
	We commude the algorith will
	converge if learning yate < 0.1.
NOH	Dil and de and de
	to de house south