large.

Sub:____

1832 1046

Ars. to the ques. No. 4

Nanishing pradient problem: In RNN to train the network one back-propagate though time and calculate the gradient at each step. Also, the neights beign added at each step. As, here one dealt with a time servier, the network one dealt with a time servier, the network tends to forget the old information when the tends to forget the old information when the gradient value is small. As smaller gradient do not affect the weight updation, the system do not not affect the weight updation, the system do not not affect the weight updation, the system do not problem.

How does LSTM, solve the vanishing gradient

How does LSTM, solve the vanishing gradient

problem: From the cell structure we can see,

problem: From the cell structure we can see,

LSTM has, input vector at line step. I. And

the network cell is denoted by cyl. the output

the network cell is denoted by cyl. the output

here parted through line step 1, It is he . ISTM

here parted through line step 1, It is he . ISTM

tan 3 gates and forget gate controls what information

in cell state to forget and give new information.

f = 6 (m [x+, r+-1]) input gate control impormation so into with a stile of the state of th Icell Astate. Is to halfman out taluation Violet it= 6 (Mi[xi, Lia]) babba verad etterion output gete, haber Ot = 6 (W. [xt-1, xt]) court busteys by = Ot * tanh (Ct) si gradicit value. cell state, ofeel th Et = tanh (We [x, ht-i]) Lene, the system back progragate with time.

The system back progragate with time. atgradient. Problem: the vielson togni and miss have people through time step a till is he . I s His rature start to start and what is Wide medical war wife has deput it state that it