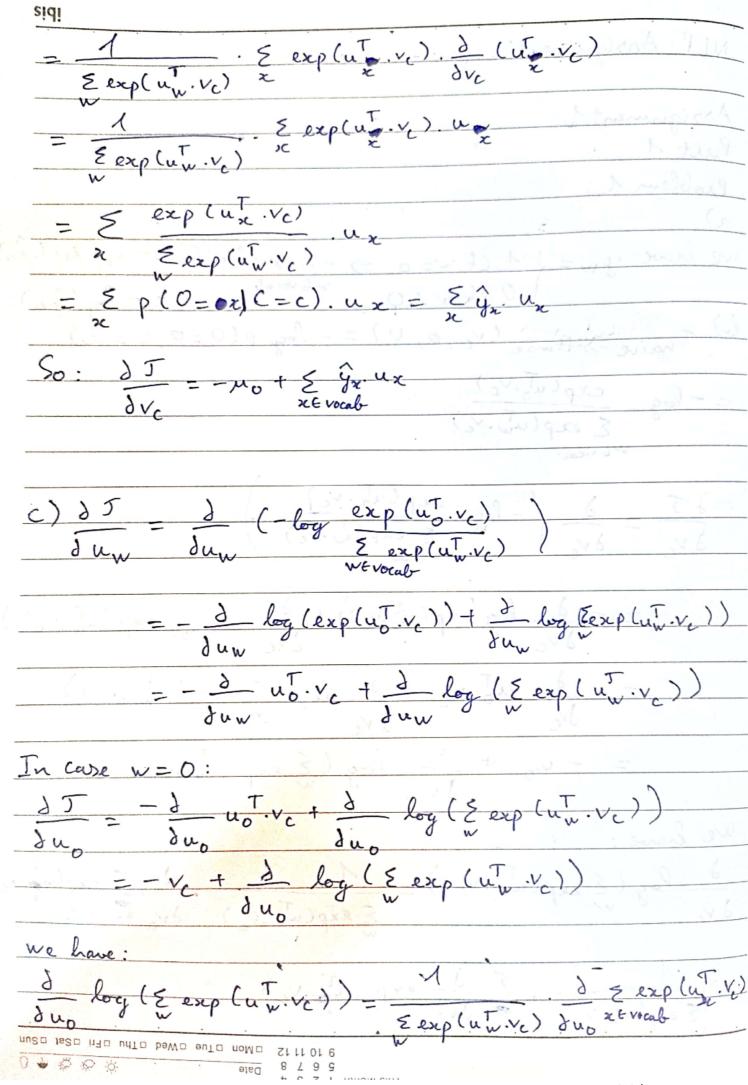
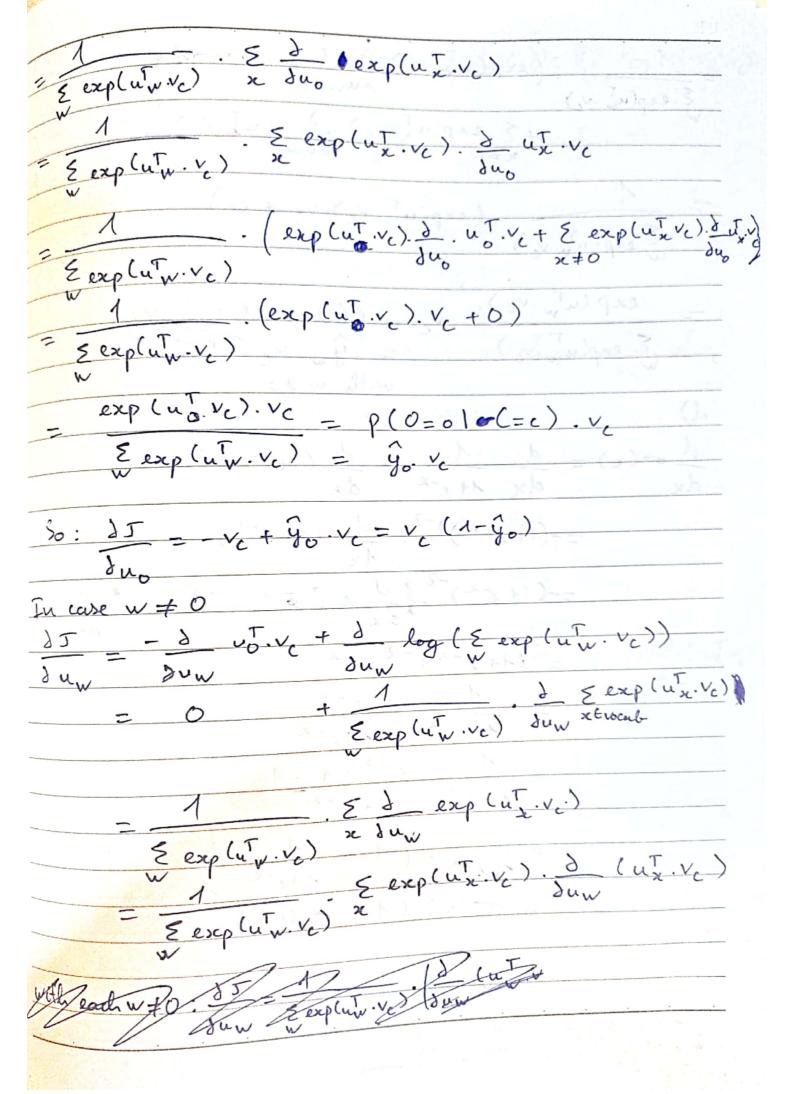
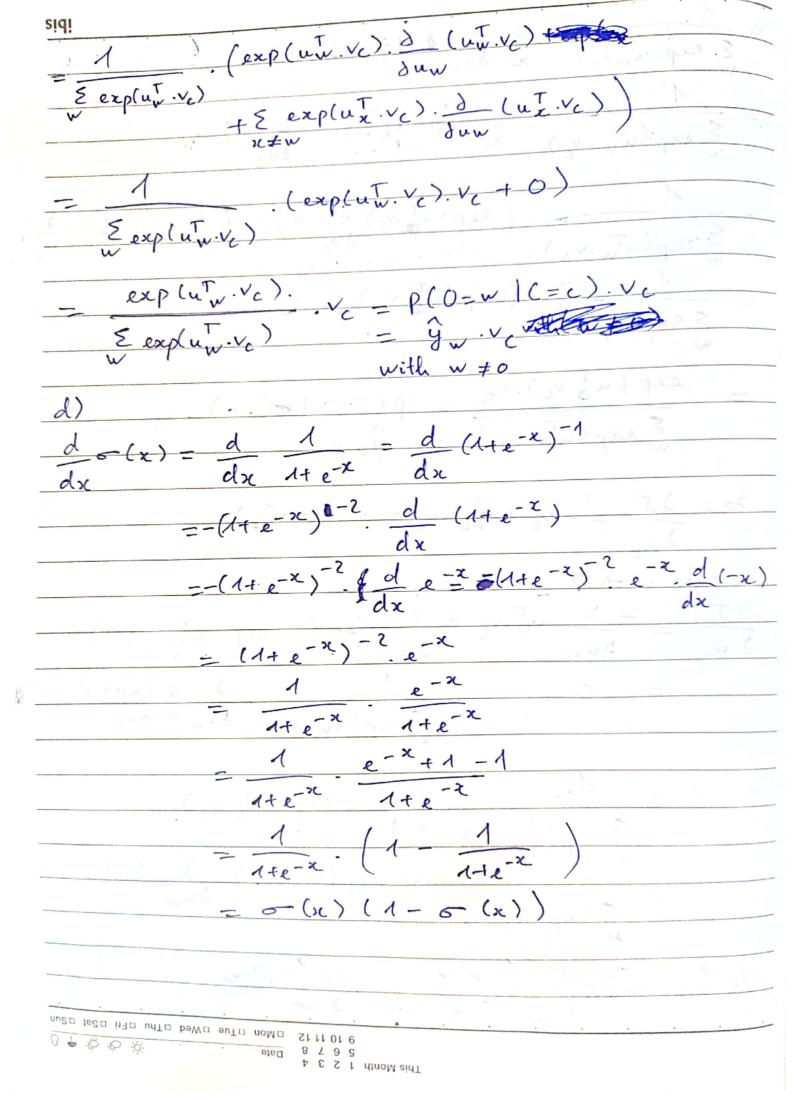
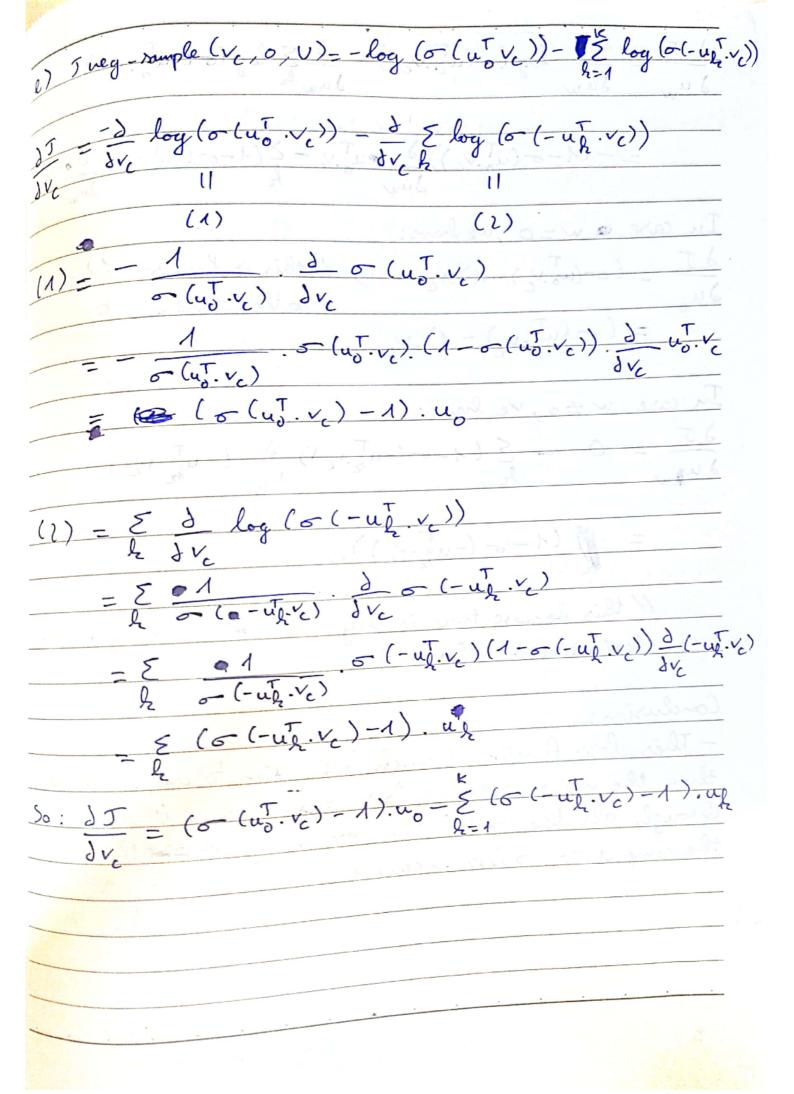
NLP Assign	ment			
0	5v6 1		3 (3V- Ju)	1×3 3
Assignment 1	2 42 1			k
Put	3	7-2-1-1-1-1-1	3 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	3
Problem 1:			The second second second	W
2)			was francisco	STOLL
no have y	- 11/11			
we have you	0 1	wer	ocal	0 10
b) Traine - so-	twax (ve, o	, U/=- le	g p(0=0,C	= ()
n ex	, (ut. vc)	-4 A 3	1 2M 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
= - log exp	xp(uw.ve)	Jan J.K		
WEVOCA	6			
	<i>f</i> • • • • • • • • • • • • • • • • • • •	p(uT.V.)		
= = 9	- log =	0×0(1 V)		660
dre dre	we'u	ocab (aw. vc)	1	<u> </u>
	by lexp (	uo.vc)) + -0	- log ( E exp	(uw.ve))
۵	Vc	4	C	
1-1-1-1	ut ve	+ d log (	E explut.	(c))
9	6	dVc	T (	
= -	uo + d	- log (E)	exp (un.ve))	Mel Al
	3 vc	0 W	<u> </u>	TA
wo D	200		- 6	26
ve have:	T	1	7 2	explut.ve
- log ( E e	cp (uw · Vc)	= = ==================================	·ve) dve xE	Vocab
o Ve				ano. A 133
	< \ \	VALUE V.	)	1
2.15	Z	explut.Vc		
E explut. vc)	20			









Problem 2:
a) you maybe - house to purchase you had to
- Neural window - based wodels can be parallelized, but RNN
modeles carmot
E) For example in case in 5, the you will be (1)
- Use a network with fewer layers
- Increase L 2 regularization weight
c) - log (0,2) who had be described and the log of the
wid) to me decentrall milling to the same of the
- True
Love) we when toil combined delice this was to such
- True
and the disculpant is that they and is dead and
Problem 4:
i) 1. Number of outputs: n
Il because each output will correspond to an input at
each time step.
2. g (t) is the probability distribution over 4 categories
including: person, organization, location, none
3. Each input will be a word in the sentence
and will produce an output attacked correspend to
predicted category at each time step
(1) 1 Number of outputs: arbitrary
Il because we don't know how many words the model
will appropriate
Z g (+) is the probability distribution over all words in vocabular
7 Fach in put will be the output of the previous output
and all perfuee a new output which is the next predicted
word for the sentence at each time step