

Lab and Theory Assignment NLP

1. Represent the 2 equation of loss functions MSE and Cross Entropy in Computational Graph?
2. Gradient with respect to all the parameters of LSTM and GRU.
3. When will the LSTM and GRU will be similar to Vanilla RNN?
4. Compute back propagations step in self attention model?
- 5.

	1	2		
	t_1	t_2	y	
x_1	1	2	+ve	1
x_2	1	1	+ve	1
x_3	3	4	=ve	-1
x_4	4	3	-ve	-1

$(x_i = y_i)$
 $x_t \xrightarrow{B, \mu} y_t \in \{+/-\}$
 $N(\mu, \sigma)$
 $x_i^1 \rightarrow$ first feature value of the sample
 Superscript represent the feature id.

Assignment

$$x_t = \langle 3, 3 \rangle \quad y_t = ?$$

Bayes Classifier

Naive Bayes Classifier

6.

Assignment: Randomly initialize A, B & π & using the λ find the following

Data: $x_1: A B C D$ Let there be two possible states
 $x_2: B B C A$
 $x_3: C C C A$ $s_1, s_2 \quad O_t \in \{A, B, C, D\}$
 $x_4: D D A B$

① $x_1: \xi_{ij} = ? \quad \delta(s_i) = ?$ ② $x_2: \xi_{ij} = ? \quad \delta(s_i)$ similarly ③ & ④

7.

Assignment: probability.

vocals = $\{A, B, C, D, E\}$ $q_i \in \{t_1, t_2, t_3\}$

Tags = $\{t_1, t_2, t_3\}$

S1: A A B B C ~~C~~ S2: B C D E A ~~A~~ S3: A A B C D

S4: E E A B C S5: A E B C D

S1: $t_1 t_1 t_2 t_2 t_1$ S2: $t_2 t_3 t_1 t_2 t_1$ S3: $t_1 t_1 t_2 t_3 t_1$

S4: $t_1 t_1 t_1 t_2 t_3$ S5: $t_1 t_1 t_2 t_3 t_2$

(i) Compute $\lambda : A, B, \pi$ (initial probs)

(ii) Assume 1st order HMM where t_i are hidden

(iii) Compute the probability of A A A B C | λ

(iv) Generate the most probable tag set for the following sequence of B B A A E.

Lab Assignment NLP

1. Generate Random data $f(x)$ (0,1) for $Y = 3*f(x) + 0.1$ and train Linear regression model and plot Loss vs iteration plot, and learning rate and loss(each iteration)
2. Plot the function $(2-f(x))^2$, where $f(x) = (1-\text{sigmoid function})^2$.
3. Crawl any corpus wiki around 10 million sentences and find average length of sentence, find the no of unique word in the crawled corpus.
4. Download any text corpus and represent your data using the following method:
 - BOW(Bag of word)
 - TF/TF-IDF
 - One hot vector representation
 - CBOW
 - SkipGram
5. Information extraction in the form of triplet of head, relation and tail of sheared Hindi dataset
6. Perform automatic extraction on the same dataset and use extracted triplets as training and testing data