Instructions:
(1) No queries will be answered during the quiz.
(2) There will not be any compensatory quiz.
(3) The duration of this quiz is 20 minutes (+5 minutes to upload). Late submissions will not be considered.

Q1. Is there any mistake in the algorithm given on slide-26 of revisiting logistic regression (given below)? Justify your answer with appropriate mathematical explanation.   [10 marks]

Learning as optimization for regularized logistic regression

- Final algorithm: $w^j = w^j + \lambda(y - p)x^j - \lambda 2\mu w^j$
- Initialize hashtables W, A and set $k=0$
- For each iteration t=1,...T
  - For each example $(\mathbf{x}_i, y_i)$
    - $p_i = \ldots ; k{+}{+}$
    - For each feature $W[j]$
      - If $x_i^j > 0$ then
        » $W[j] \mathrel{*}= (1 - \lambda 2\mu)^{k-A[j]}$
        » $W[j] = W[j] + \lambda(y_i - p^i)x_j$
        » $A[j] = k$

k-A[j] is number of examples seen since the last time we did an x>0 update on W[j]

Q.2 In the power iteration method for computing page rank, do we need to perform an additional step of making the sum of entries of 'r' equal to 1 after updating it? Justify your answer with appropriate mathematical explanation.   [10 marks]

Q.3 In TrustRank, suppose we employ some humans to validate the initial set of seed (trusted) pages. It so happens that one of them marks some spam pages as trusted pages. Propose an algorithm to identify such spam pages.   [10 marks]