**Quiz 4 – Spring 2017**

**CS583: Data Ming and Text Mining**

Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ UID\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
|  | **Marks** |
| Q1 |  |
| Q2 |  |
| Q3 |  |
| **Total** |  |

1. (10 marks) Answer the following questions. **Multiple answers may be correct** for each question.

a) Which of the following sentences are factual and which are subjective?

1. Everyone will die someday, which is sad.
2. I brought an iPhone yesterday afternoon
3. This guy is disgusting
4. I think he went home after the dinner.

b) Which of the following sentences express comparative opinions?

1. This car is very nice
2. This car beats the rest
3. This car is better than my Honda Civic
4. He love the shape of this car

c) When we say that opinion summary need to be quantitative, what do we mean?

1. We mean that the system needs to count the number of opinion sentences
2. We mean that the system needs to extract all positive and negative opinions
3. We mean that the system needs to count the number of positive opinions and the number of negative opinions
4. We mean that the system needs to produce the percentage of positive opinions and the percentage of negative opinions.

d) Which algorithm is called the memory-based recommendation method?

1. kNN
2. Association rules
3. Clustering recommendation
4. Matrix factorization

e) Comparing users based on their ratings of items.

1. The approach that uses the above idea is called item-based CF
2. The approach that uses the above idea is called user-based CF
3. The approach that uses the above idea is called integrated CF
4. The approach that uses the above idea is called user-item based CF
5. (5 marks) produce all opinion quintuples from the following review.

**Id: Abc123 on 5-1-2008 --** “*I bought an iPhone yesterday. It is such a nice phone. The touch screen is really cool. The voice quality is great too, although the battery life is not long. However, my mom was mad at me as I didn’t tell her before I bought the phone. She thought the phone was too expensive”*

1. (5 marks) Each set of numbers below represents a set of books bought by a customer.

{1, 4}

{1, 4, 5, 7}

{4, 5, 7, 3}

{5, 6}

{1, 4, 5, 6}

{4, 3}

{4, 3, 7}

{4, 3, 7}

If the new customer has purchased both book 1 and book 3 and we want to recommend two books for him, what would you recommend and why? Please use association rule based recommendation to solve this problem.