

Spark, Recommender Systems, and Network Graphs

1. What is Spark?
2. What is PySpark?
3. What are some advantages and limitations of PySpark?
4. What is MapReduce?
5. What is YARN?
6. What is Resilient Distributed Dataset (RDD)?
7. When it comes to recommender systems, we use the following methods:
 - a. Collaborative Filtering
 - b. Content-Based Filtering
 - c. Matrix Factorization (ALS)

Briefly explain how these methods work along with their limitations.

8. Why is non-negative matrix factorization called so?
9. Briefly explain the following:
 - a. Degree centrality
 - b. Closeness centrality
 - c. Betweenness centrality
10. Describe a scale-free network.