

Advanced Methods in Data Science and Big Data Analytics

1. MapReduce and Hadoop

- The MapReduce Framework
- Apache Hadoop
- Hadoop Distributed File System
- YARN

2. Hadoop Ecosystem and NoSQL

- Hadoop Ecosystem
- Pig
- Hive
- NoSQL--Not only SQL
- HBase
- Spark

3. Natural Language Processing

- Introduction to NLP
- Text Preprocessing
- TFIDF
- Beyond Bag of Words
- Language Modeling
- POS Tagging and HMM
- Sentiment Analysis and Topic Modeling

4. Social Network Analysis

- Introduction to SNA and Graph Theory
- Most Important Nodes
- Communities and Small World

- Network Problems and SNA Tools

5. Data Science Theory and Methods

- Simulation
- Random Forests
- Multinomial Logistic Regression

6. Data Visualization

- Perception and Visualization
- Visualization of Multivariate Data