Distributed Data Mining Lab

Winter Term 2014/15
Chair for Bioinformatics
L. Richter

DDM Lab WS14/15
About the Course

- Wednesday, 10-12am
- 15 meetings from Oct. 8\textsuperscript{th} to Jan 28\textsuperscript{th}
- Highly innovative
- Exploration of distributed data mining technologies
- First time of this class
- No pre-arranged script available

DDM Lab WS14/15
Contents

Topics, a.o.:
- Hadoop File System
- Hadoop
- Spark
- Mahout
- MLlib
Grading

What is Graded?:

- The grade is based on: the whole semester performance i.e. the complete wiki entries, all presentations of a team
- Each group reports weekly in the wiki
- If there are no objections the group is awarded the same grade, minor differentiation is possible
- Individual grading/report only on request

DDM Lab WS14/15
Grading

Criteria:

- Presentation criteria: focus to things that matter, communicate your message in an easy and understandable way, it is important, that the audience can follow

- Wiki: volume, conciseness, clearness.
Grading

Criteria in Detail:

- The volume should reflect the amount of time and work you spend on the topic
- The conciseness refers to the level of details and the precision (e.g. do things really work using the given instructions)
- The clearness: Is the red line of the work visible, is it easy to understand.
Realization

Conduction Hints:

- Therefore the wiki entries should have meaningful structure and cover different levels of abstraction. Referenced white papers, manuals, tutorials and further reading.

- Since this course is new, problems and challenges should adequately be reflected.

- We plan to record weekly notes which leads then to the final grade.

DDM Lab WS14/15
Organization

- Register participants
- Form teams
- Discuss the meeting time
- Identify interest in / find slots for coding sessions
- Answer questions about prerequisites
Infrastructure

- Bunch of virtual machines running Debian
- you have root access to these machines
- Provide a wiki for reports and documentation
- Provide a forum for off-line discussion
Conclusion

- New course
- No previous experience with matter yet
  -> you are free to explore the opportunities given by the technology
  -> you are free to suggest additional topics as the course goes on
- you have to do reading and technology research on your own
Conclusion

- We are a kind of project manager where you are the technical specialist.
- The supervisors work in collaboration with the students and refine strategies and directions on the fly as the course is going on. This is based on the feedback from the students and done in collaboration with the students -> clear communication is essential on both sides.

DDM Lab WS14/15
Have fun!