

AIM-3 Final Project

General Instructions

1. Students must compose a 1-page proposal (following the structure on the right) and submit it no later than **Wednesday, 27.05.15 at 12 noon**.
2. All proposals are subject to approval.
3. Projects should be of sufficient complexity that it **will take six weeks** to complete.
4. Each team must be comprised of at most two students.
5. Eligible topics may be drawn from Groups A or B.

Proposal Structure (up to one page summary)

1. Project Title
2. Team Members
3. Problem Statement (Well Defined Targets & Scope)
4. Project Plan (Objectives, Planned Methodology, Experiments)

Term Report Structure (10-15 page paper)

1. Cover Page: *Title, Authors, Author Email Addresses, Course Title, Date*
2. Introduction
3. Problem Statement
4. Methodology
5. Experiments
6. Results
7. Conclusion
8. References

Final Presentation Structure

1. Project Title
2. Authors
3. Problem Statement (Well Defined Targets & Scope)
4. Methodology (Describe Approach)
5. Experiments (Rigorous)
6. Results (Interpret and Discuss, Draw Conclusions)
7. Conclusions (Relevance, Importance, Impact)

Group A. Prospective Projects

- [A1] Analyze a scalable analytic, conduct some numerical experiments, and report your findings.
- [A2] Analyze a very large dataset using a scalable analytic and report your findings.
- [A3] Analyze complexity of an analytic, propose a scalable solution, perform experiments, report findings.
- [A4] Compare analytics, discuss the strengths and merits of each and illustrate them on a large dataset.
- [A5] Evaluate an open-source or commercial big data technology using real data, report results.
- [A6] Analyze a very large dataset using visual analytics tools (e.g., Tableau) and present your results.

Group B. Any topic that falls under one of the items listed below

[B1] Distributed Systems / MapReduce

[B2] Apache Flink

[B3] Apache Spark

[B4] Data Mining

[B5] Classification

[B6] Clustering

[B7] Dimensionality Reduction

[B8] Collaborative Filtering

[B9] Streaming

[B10] Network Analysis

[B11] Statistical NLP

[B12] Visual Analytics