Designs of Algorithms and Programming for Massive Data

DS8001

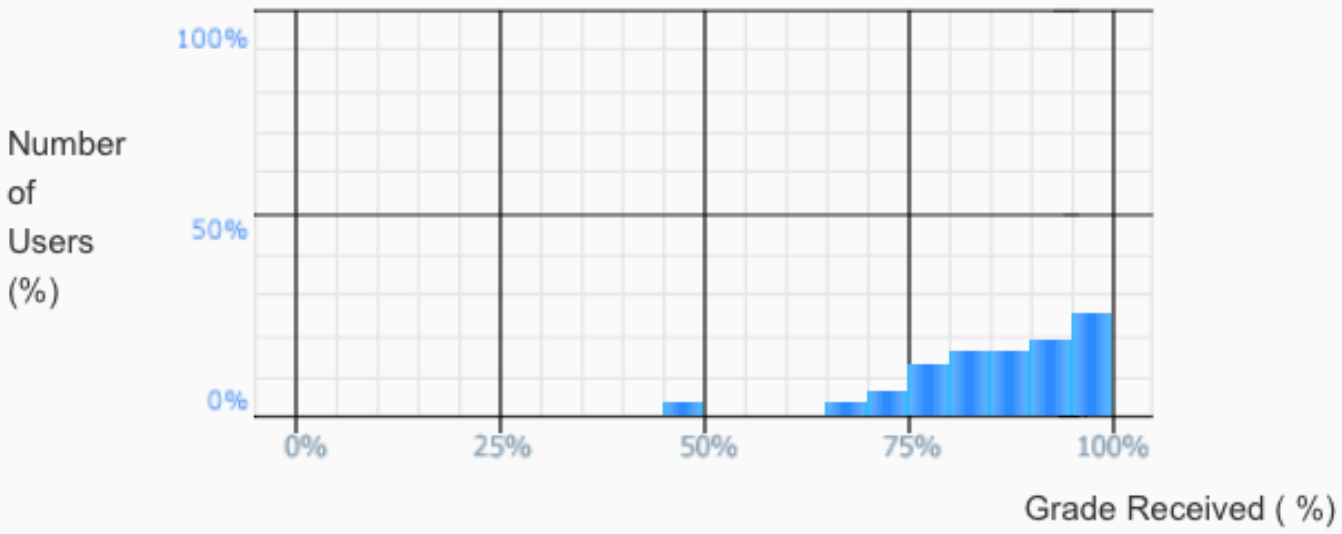
Andriy Miranskyy

Nov. 7, 2016

Midterm stats

Number of submitted grades:	32 / 32
Minimum:	47.65 %
Maximum:	100 %
Average:	86.46 %
Mode:	83.53 %, 88.24 %, 91.18 %
Median:	88.24 %
Standard Deviation:	10.79 %

Grade Distribution



Project

- Project submission date: December 14th
  - Grades have to be submitted by December 21st
  - Gives me a week to mark
- Submit a report (5 pages max, given Times New Roman size 10 font)
  - Background / Introduction
    - Why what I am doing is important
  - Goal
    - What am I trying to achieve?
  - Related literature
    - What others have done?
  - Method
    - How am I achieving it?
  - Results
    - What is the outcome?
  - Summary
- Code listings can be appended separately and do not count toward 5 pages quota.

Outline

- Parallelization
- Case study training of Gaussian Naïve Bayes.

Parallel Algorithms Design

Cont’d

Suggested reading

- A. Grama, A. Gupta, G. Karypis, and V. Kumar, Introduction to Parallel Computing, 2nd ed., Addison Wesley, 2003
  - Chapters 3, 5

Continued Tasks