MATHS 1004 Mathematics for Data Science I

Lewis Mitchell

maths.adelaide.edu.au/lewis.mitchell
@lewis_math

Visit the main page

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

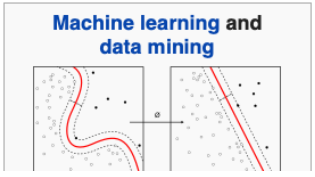
Data science

From Wikipedia, the free encyclopedia

Not to be confused with information science.

Data science is a multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data.[1][2] Data science is the same concept as data mining and big data: "use the most powerful hardware, the most powerful programming systems, and the most efficient algorithms to solve problems".[3]

Data science is a "concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyze actual phenomena" with data.[4] It employs techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, and information science. Turing award winner Jim Gray imagined data science as a "fourth paradigm" of science (empirical, theoretical, computational and now data-driven) and asserted that "everything about science is changing because of the impact of information technology" and the data deluge.[5][6]



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

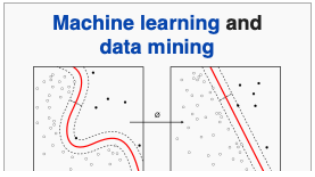
Data science

From Wikipedia, the free encyclopedia

Not to be confused with information science.

Data science is a multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data.[1][2] Data science is the same concept as data mining and big data: "use the most powerful hardware, the most powerful programming systems, and the most efficient algorithms to solve problems".[3]

Data science is a "concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyze actual phenomena" with data.[4] It employs techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, and information science. Turing award winner Jim Gray imagined data science as a "fourth paradigm" of science (empirical, theoretical, computational and now data-driven) and asserted that "everything about science is changing because of the impact of information technology" and the data deluge.[5][6]



Harvard Business Review

DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

Big Data Borat
@BigDataBorat

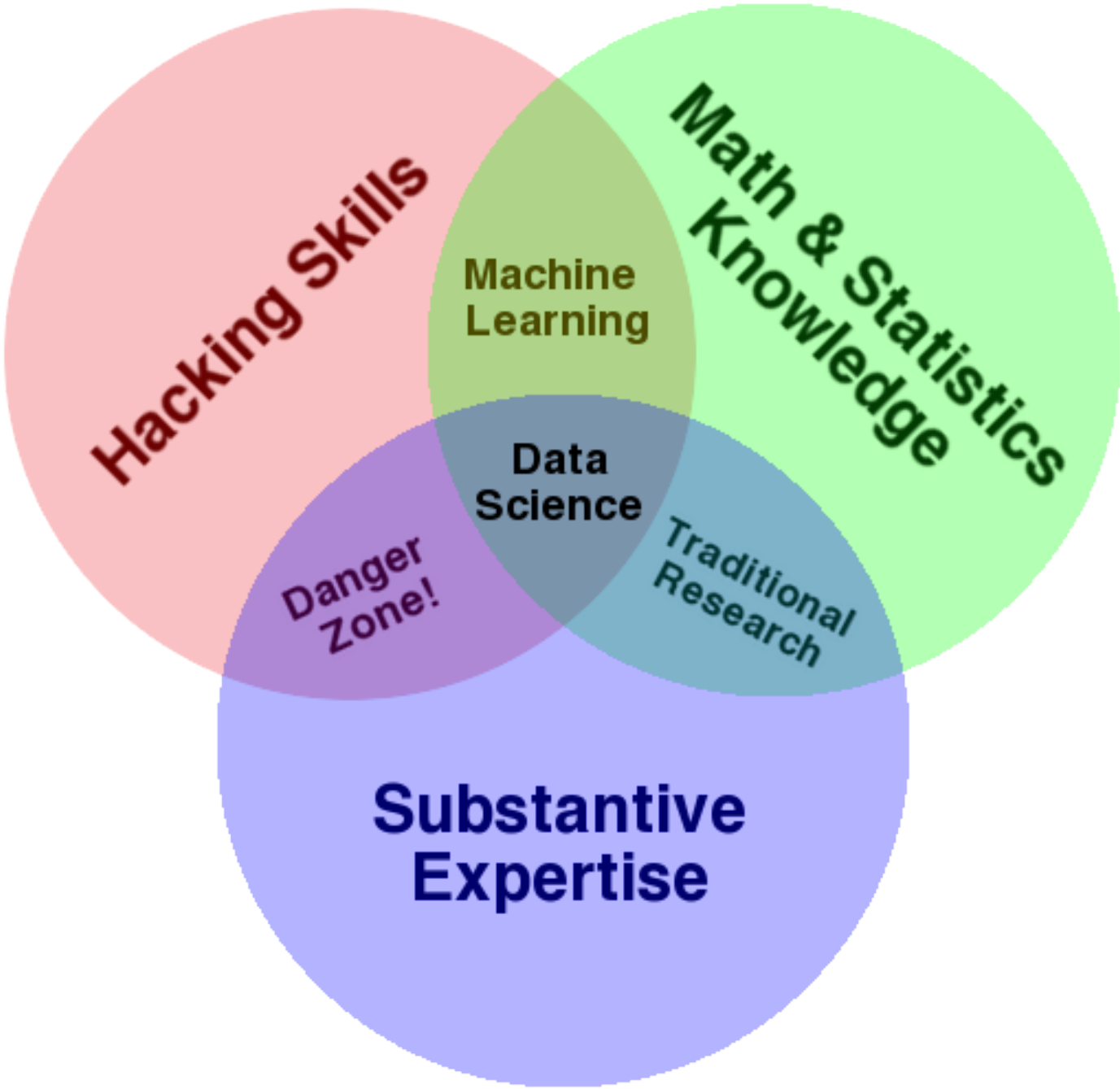
Follow

Data Science is statistics on a Mac.

11:02 PM - 27 Aug 2013

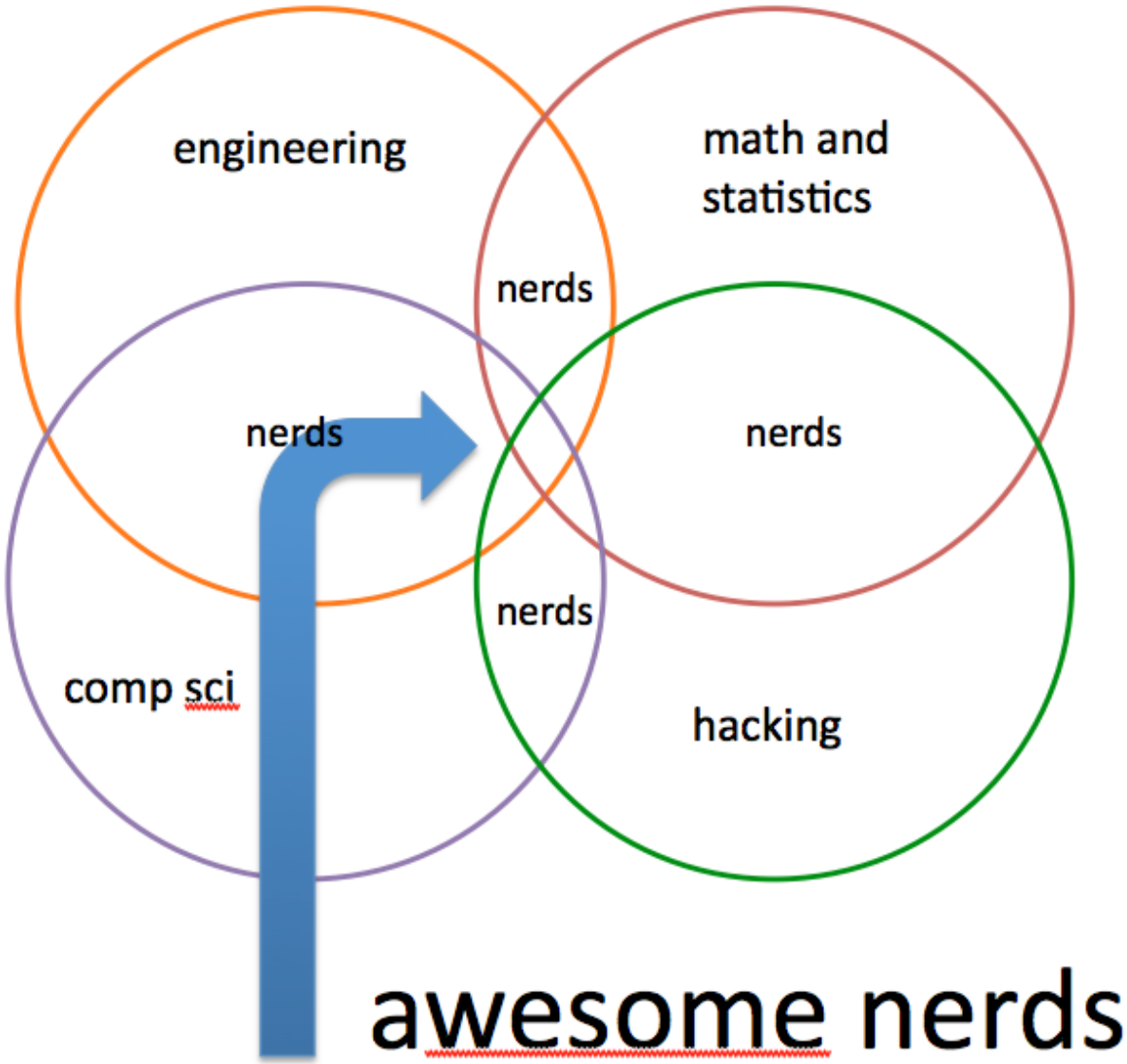
611 Retweets **335** Likes

20 611 335


Hacking Skills
Math & Statistics Knowledge
Substantive Expertise
Machine Learning
Data Science
Danger Zone!
Traditional Research

Drew Conway

Data scientists?



Hilary Mason

This course

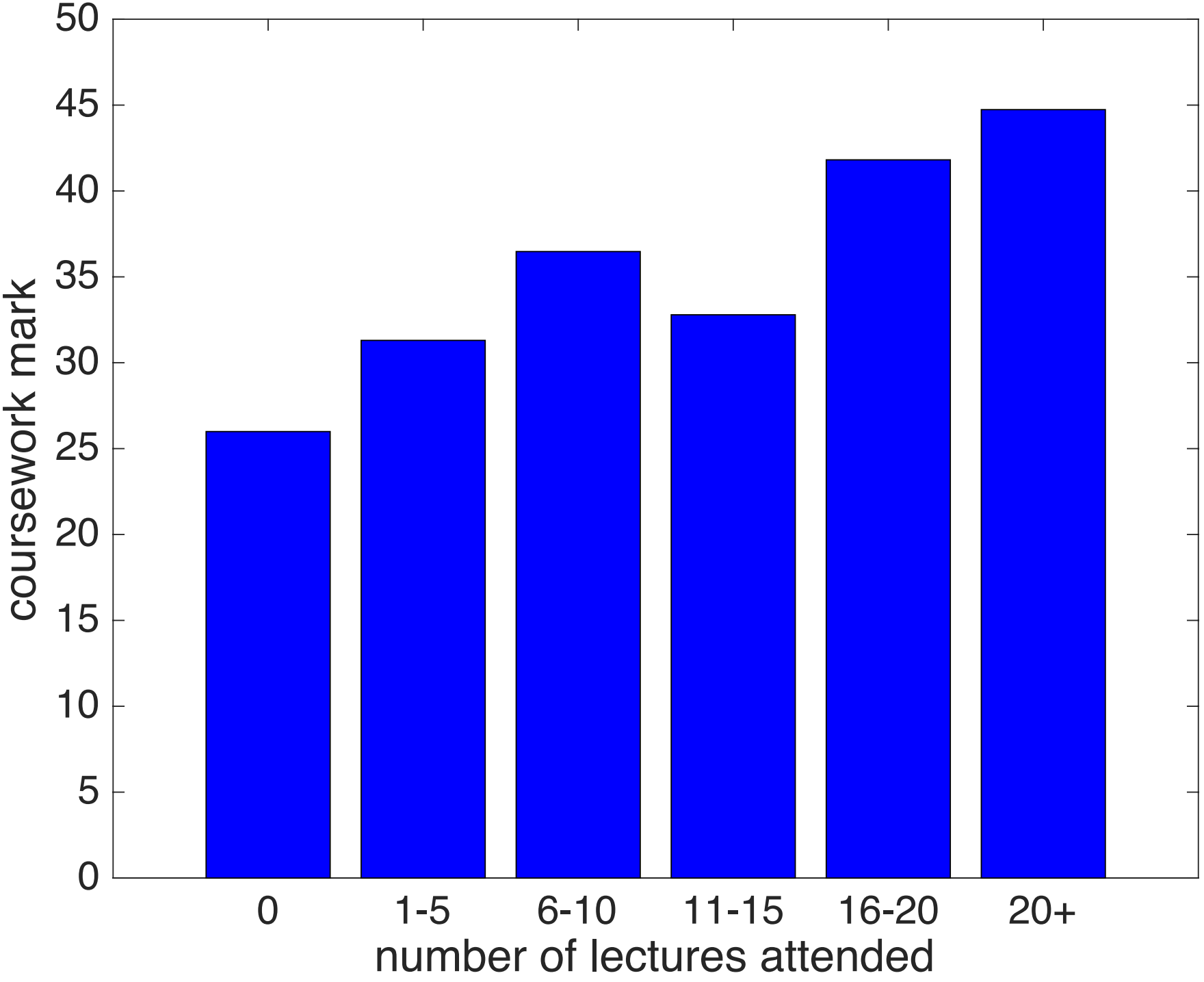
- Focuses on "Math & Statistics Knowledge"
- Introduces "Hacking" (in Python)
- Is motivated by "Substantive Expertise"

Course outline

1. Fundamentals
   - Notation
   - Functions
   - Approximation
2. Series
   - Summation
   - Taylor series
3. Linear algebra
   - Representing big, complex, data
   - Systems of equations
   - Dimension reduction
4. Probability
   - Discrete random variables
   - Continuous random variables & integration
5. Optimisation

Course structure

- 4 classes per week (3 lectures + 1 problem-solving)
  - Please be my audience!
- 6 computer labs (odd weeks)
  - Introduction to Python
- 6 tutorials (even weeks)
  - Best way to master the material for the exam
  - Whiteboard tutorials
  - Marks for (active) participation
  - Feedback on assignments


50
45
40
35
30
25
20
15
10
5
0
coursework mark
0
1-5
6-10
11-15
16-20
20+
number of lectures attended

Computer labs

Echo360 Question: Which programming languages are you familiar with?

- C/C++
- Java
- Julia
- Matlab
- Python
- R

Course assessment

- 5%: Lab & Tutorial participation
- 25%: Assignments (5 assignments, 5% each)
- 70%: Final exam

Assignments

- Due **Tuesdays 11am, odd weeks** (starting week 3)
- Submit via Canvas, **PDF** only
- Lateness/extension policy:
  - Up to 24 hrs late: 40% lateness tax
  - More than 24 hrs late: assignment can't be submitted. Apply for an exemption instead (with documentation)
- Assignments to be returned by next tutorial