CLASS PRESENTATION

AP

DATA SCIENCE: TECHNIQUES AND APPLICATIONS (DSTA)

A NEW[ISH] MODULE

- Designed for MSc Data Science students
- Contents carefully dovetail with existing MSc DS modules

WE ARE IN THIS TOGETHER

MAIN TOPICS

- Python modules for Data Science, e.g., `Numpy`
- The geometric view of multidimensional data
- The data-as-network view (graph algo.)
- the information-theoretic view of data, e.g., text-as-data

- a motivating DS problem: sports ranking
- selected topics: matrix *slicing*: finding latent dimensions to datasets.
- [discontinued: handling exponentially-distributed data]

IMPORTANT ASPECTS

The inevitable *overfitting* to MSc DS may make this module less appealing to MSc ACT students than the title suggests

Machine Learning topics are selected so as not to overlap the essential Machine Learning and Applied ML modules.

The table of contents is varied and may appear syncretic wrt. traditional, textbook-based modules..

This is a *final* module in charge of synthetising a large, fast-moving area.

so, expect a more seminarial (than textbook) approach.

ORGANIZATION

IN LAB

- Alessandro Provetti
- Abul Hasan
- Paschalis Lagias

ONLINE

Standard Moodle/Teams provision

Please check your email and MyBirkbeck calendar for (unlikely) time/place amendments.

Extra effort: if you are familiar with the Markdown presentation language and the GitHub software repository you can follow material preparation live in there.

PLEASE SEE YOUR

- PG admins, e.g., Yeti Wan, for support in navigating school: cs-pg@bbk.ac.uk
- TAs, Abul or Paschalis, for support with lab experience


AND FINALLY SEE YOUR

- module coordinator, Alessandro, about the study materials and general *'meaning of life'* questions

TOO MANY E-MAILS

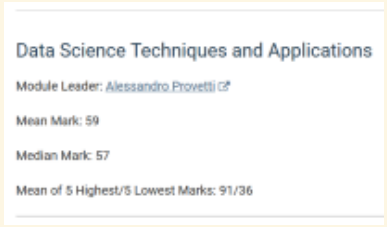
Your feedback and questions are always welcome.

Help avoiding inbox overflow by please contacting us via MS Teams or Moodle class forum

ASSESSMENT

PAST MARKING TRENDS

DSTA vs. ML in 2019:



Data Science Techniques and Applications

Module Leader: Alessandro Provetti

Mean Mark: 59

Median Mark: 57

Mean of 5 Highest/5 Lowest Marks: 91/36

Machine Learning

Module Leader: George Magoulas

Mean Mark: 55

Median Mark: 54

Mean of 5 Highest/5 Lowest Marks: 84/22

The exam paper was harder with more challenging questions
than in previous years but in line with MSc-level modules. In th

COURSEWORK

For the second time, coursework will take the form of an in-class, multiple-choice quiz via Codio/ Moodle.

Please see the details under the ‘Assessment’ tile on Moodle.

In view of the novelty of the approach, marks will be awarded generously, but always at the discretion of the School’s Exam board.

FINAL TEST

For the second time, the final test will take the form of an in-class, multiple-choice quiz via Codio/ Moodle.

Please see the details under the ‘Assessment’ tile on Moodle.

In view of the novelty of the approach …