

Master in Data Science

Departments of Informatics & Mathematics

Programme board

Riccardo De Bin (Chair, MI)

Thordis Thorarinsdottir (MI)

Martin Giese (IFI)

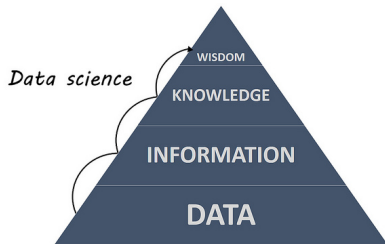
Ole Christian Lingjærde (IFI)



Master in Data Science at UiO

► What is Data Science?

The science of extracting knowledge or insight from various types of data.



- This program: Combines **mathematics**, **statistics** and **informatics** making you ready for meeting the Data revolution.
- Goal of the program:

Data Science is the science of extracting knowledge or insight from various types of data. The study of Data Science combines mathematics, statistics and informatics making you ready for meeting the Data revolution.



UiO • Universitetet i Oslo

Background

- ▶ Why such a program?
 - ▶ Data Science methodology partly included in many other programmes
 - ▶ Lack of studies focusing on **methodology** in Data Science
- ▶ Why at UiO?
 - ▶ Two centers for research-based innovation
 - ▶ BigInsight: **Statistics**
 - ▶ Sirius: **Informatics**
 - ▶ Both based on problems related to **Big Data**
 - ▶ Both financed by
 - ▶ The Norwegian Research Council (excellent researchers)
 - ▶ The industry (useful research)
 - ▶ long story of great collaboration, that led to **InteGreat**:
 - ▶ Norwegian Centre for Knowledge-driven Machine Learning
 - ▶ Centre of Excellence funded by the Research Council of Norway
 - ▶ Also many other excellent researchers/groups within topics related to Data Science



Programme structure

An interdisciplinary background in **statistics** (probability theory, inference, machine learning) and **computer science** (algorithms, visualization, database) in addition to a solid foundation in general **mathematics**.

4. semester	Thesis		
3. semester	Thesis		Course
2. semester	Thesis	Course	Course
1. semester	Course	Course	Course

- ▶ 1. Semester:
 - ▶ Two **compulsory** courses
STK-IN4300 - Statistical learning methods in Data Science
IN-STK5000 - Responsible Data Science
 - ▶ You need to **choose** one more extra course
 - ▶ Many possible courses in both statistics and informatics, depend a bit on specialisation!
 - ▶ Master thesis/supervisor



Specializations/research groups

- ▶ Possibility for both theoretical and applied projects
- ▶ Statistics and Data Science
- ▶ Biomedical informatics
- ▶ Language Technology
- ▶ Analytical Solutions & Reasoning
- ▶ Digital Signal Processing and Image Analysis
- ▶ Collaboration with external partners.



Recommended courses within specialisations

Examples:

- ▶ **Statistics and Machine Learning**
 - ▶ STK4051 - Computational Statistics
 - ▶ STK4290 - Probabilistic Graphical Models
- ▶ **Database Integration and Semantic Web**
 - ▶ IN5040 - Advanced Database Systems for Big Data
 - ▶ IN4070 - Logic
 - ▶ IN4060 - Semantic Technologies
- ▶ **Data Science and Life Science**
 - ▶ IN4030 - Introduction to Bioinformatics
 - ▶ STK4051 - Computational Statistics
- ▶ **Language Technology**
 - ▶ IN4080 - Natural Language Processing
 - ▶ IN5550 - Advanced Topics in Natural Language Processing
- ▶ **Digital Image Processing**
 - ▶ IN5520 - Digital Image Analysis
 - ▶ IN5400 - Machine Learning for Image Analysis



UiO • Universitetet i Oslo

How to choose?

- ▶ Courses
 - ▶ Many choices within informatics and statistics
 - ▶ Interesting topics
 - ▶ Database systems, Parallel computing, High-level programming/scripting, statistics
 - ▶ Be aware: Many new courses within machine learning/Data Science.
 - ▶ All may not suit within this program!
 - ▶ Master thesis:
 - ▶ Specialization
 - ▶ Supervisor and project
 - ▶ Possibilities for master projects connected to external companies/startups
- ▶ If unsure: **Ask!**



Process

- ▶ Early October: Presentation of supervisors/projects
- ▶ October: Talk to possible supervisors
- ▶ October 30th (tentatively): Send to debin@math.uio.no a ranked list of at least 3 possible supervisors/topics
- ▶ Early November: Decision on allocation of students to supervisors
- ▶ November: In collaboration with supervisor, make plan for master study, both courses and master project, see <https://www.mn.uio.no/math/english/studies/masterstudies/index.html>
 - ▶ For short thesis, there will be no project plan.
- ▶ December 1st (tentatively): Deliver master plan.



Statistics and Machine learning

- ▶ Focus on statistical aspects related to data science and machine learning methods
 - ▶ Implementation/testing of a new idea
 - ▶ Evaluation of newly proposed methods
 - ▶ Can lead to new ideas!
 - ▶ Both algorithmic based and model based approaches
 - ▶ Modification of existing methods to big data
 - ▶ Modification of existing methods to new applications
- ▶ Central courses
 - ▶ STK-IN4300 - Statistical learning methods in Data Science
 - ▶ STK4051 - Computational statistics
- ▶ Examples:
 - ▶ <https://www.mn.uio.no/math/english/people/aca/debin/master-data-science/supervisors-and-master-topics.html>

