Specialized Master in data science, Big data

This programme is taught in English.

Job opportunities

The present master has been created to deepen your knowledge and understanding of emerging, state-of-the-art database technologies. Indeed, the intensive use of computers and the internet in the beginning of the present century has a clear impact on the way data have to be collected and treated. In many situations, practitioners have to deal with massive databases (« Big data »).

Data science finds its roots in many applications: genomics and high scale DNA sequencing generate tons of data at many different biological levels; the use of social networks, mobile phones, tablets generate data every single second; robots and industrial equipments are nowadays equipped with sensors that provide a huge amount of information and therefore huge databases. In economics and in finance, practitioners have to deal with real-time forecasts based on high-frequency data (production, trade, market data).

The master is a natural preparation for the following jobs: “data scientist”, “data manager”, “analytics manager” or simply “statistician” or “computer scientist” that are increasingly demanded by companies.

Programme objectives

You have already a master degree and good knowledges in computer sciences or in statistics and you are interested by their applications. Then the present master is a natural choice to improve your skills and become a specialist in massive data analysis. The program we propose here is fully taught in English and therefore opens to the international job market.

In particular, the objective of the master is to improve the following skills:

- Be able to perform a research project or an applied innovation in computer sciences or in statistics.
- Be able to be rigorous, independent, ethic, creative and aware of the impact of the results obtained for a company or for the society in general.
- Be able to be rigorous, independent, ethic, creative and aware of the impact of the results obtained for a company or for the society in general.
- Design and implement applications based on artificial intelligence and learning techniques.
- Clearly communicate to various types of audiences conclusions or results of a project in computer sciences, statistics or econometrics.
- Be able to develop new skills by yourself.
- Be able to be rigorous, independent, ethic, creative and aware of the impact of the results obtained for a company or for the society in general.

Benefits of the training

The specialized master in data science, big data provides an interdisciplinary training in data analysis (model choice, forecast, inference, learning) of big data. The program has been constructed in order to teach both statistical and computer sciences techniques. We furthermore propose lectures in econometrics to let students deal with quantitative practical aspects. The student who wants to complete his/her master by an internship will clearly benefit from the fact that Brussels is full of companies interested by the profile.

Several faculties are involved in the master: the Faculty of Sciences, the Brussels School of Engineering and the Solvay Brussels School of Economics and Management from ULB and also partners for the VUB. This is clearly an asset since it reinforces the interdisciplinary aspect of the master which is supported by various important teams of researchers from the ULB and the VUB:

- ECAERS, Solvay Brussels School of Economics and Management
- IB2 (Interuniversity Institute of Bioinformatics in Brussels), ULB/VUB
- IRIDIA, Brussels School of Engineering
- LISA, Brussels School of Engineering
- Machine Learning Group, Faculty of Sciences
- Mathematical Statistics Group, Faculty of Sciences
- WIT, Brussels School of Engineering
Programme title: Specialized Master in data science, Big data
Programme mnemonic: MS-BGDA
Studies level: Advanced masters
Learning language:
Schedule:
Campus: Solbosch and Plaine
**Unique year**

**Cours obligatoires**

<table>
<thead>
<tr>
<th>COURSE</th>
<th>TITLE</th>
<th>COORDINATOR(S)</th>
<th>CREDITS</th>
<th>CONTACTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFO-H600</td>
<td>Computing foundations of data sciences</td>
<td>Hugues BERSINI</td>
<td>5</td>
<td>24h lecture, 12h tutorial classes, 12h practical work</td>
</tr>
<tr>
<td>MEMO-F541</td>
<td>Mémoire/Stage en entreprise</td>
<td></td>
<td></td>
<td>240h mfe/tfe</td>
</tr>
<tr>
<td>STAT-F600</td>
<td>Multivariate and high-dimensional statistics</td>
<td>Thomas VERDEBOUT</td>
<td>5</td>
<td>24h lecture, 12h practical work</td>
</tr>
<tr>
<td>STAT-S502</td>
<td>Data management and analytics</td>
<td>Pierre DEVILLE</td>
<td>5</td>
<td>36h lecture, 24h tutorial classes</td>
</tr>
</tbody>
</table>

**Module 1 : Statistique**

Two courses chosen from the following

<table>
<thead>
<tr>
<th>COURSE</th>
<th>TITLE</th>
<th>COORDINATOR(S)</th>
<th>CREDITS</th>
<th>CONTACTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFO-F422</td>
<td>Statistical foundations of machine learning</td>
<td>Gianluca BONTEMPI and Bernard MANDERICK</td>
<td>5</td>
<td>24h lecture, 12h tutorial classes, 12h practical work</td>
</tr>
<tr>
<td>STAT-F405</td>
<td>Time series analysis 1</td>
<td>Christophe CROUX</td>
<td>5</td>
<td>24h lecture, 12h tutorial classes</td>
</tr>
<tr>
<td>STAT-F408</td>
<td>Computational statistics</td>
<td>Maarten JANSEN</td>
<td>5</td>
<td>24h lecture, 12h tutorial classes</td>
</tr>
<tr>
<td>STAT-F420</td>
<td>Topics in mathematical statistics</td>
<td>Thomas VERDEBOUT</td>
<td>5</td>
<td>24h lecture, 12h tutorial classes, 12h project</td>
</tr>
<tr>
<td>STAT-F423</td>
<td>Statistical learning</td>
<td>Christine DE MOL</td>
<td>5</td>
<td>24h lecture, 12h tutorial classes</td>
</tr>
</tbody>
</table>

**Module 2 : Informatique**

Two courses chosen from the following

<table>
<thead>
<tr>
<th>COURSE</th>
<th>TITLE</th>
<th>COORDINATOR(S)</th>
<th>CREDITS</th>
<th>CONTACTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFO-F424</td>
<td>Combinatorial optimization</td>
<td>Michaël POSS</td>
<td>5</td>
<td>24h lecture, 12h tutorial classes, 12h practical work</td>
</tr>
<tr>
<td>INFO-F524</td>
<td>Continuous optimization</td>
<td>Michaël POSS and Bernard FORTZ</td>
<td>5</td>
<td>24h lecture, 24h project</td>
</tr>
<tr>
<td>INFO-H423</td>
<td>Data Mining</td>
<td>Mohamed Hatem HADDAD and Mahmoud SAKR</td>
<td>5</td>
<td>24h lecture, 12h practical work</td>
</tr>
<tr>
<td>INFO-H501</td>
<td>Pattern recognition and image analysis</td>
<td>Olivier DEBEIR and Christine DECAESTECKER</td>
<td>5</td>
<td>36h lecture, 24h practical work, 24h project</td>
</tr>
</tbody>
</table>
INFO-H512  Current trends in artificial intelligence  Hugues BERSINI (Coordinator)
5 credits [lecture: 24h, project: 30h]  English

INFO-H515  Big Data : Distributed Data Management and Scalable Analytics  Stijn VANSUMMEREN (Coordinator), Gianluca BONTEMPI and Ann NOWE
5 credits [lecture: 24h, practical work: 18h, project: 45h]  English

Module 3 : Econométrie

One course chosen from the following

ECON-S428

GEST-S503  Financial econometrics  Olivier SCAILLET (Coordinator)
5 credits [lecture: 24h, tutorial classes: 12h]  English