# MAJOR COURSES

## 40. Big Data Analytics for Remote Sensing (PH053IU)

| | |
|---|---|
| Module designation | *The aim of the course is to get students familiar with high-performance computing aspects of remote sensing. Students will learn how to discover knowledge from remote sensing data with high-performance computing approaches and data visual analytics tools (Apache Hadoop, parallel Python, R, Google Earth Engine).* |
| Semester(s) in which the module is taught | 1, 2 |
| Person responsible for the module | Dr. Lê Thanh Vân |
| Language | English |
| Relation to curriculum | Compulsory |
| Teaching methods | Lecture, assignment, project. |
| Workload (incl. contact hours, self-study hours) | (Estimated) Total workload: 127.5<br>Contact hours (please specify whether lecture, exercise, laboratory session, etc.): lecture: 37.5<br>Private study including examination preparation, specified in hours: 90 |
| Credit points/ECTS | 3 credits/4.62 ECTS |
| Required and recommended prerequisites for joining the module | Previous Course:<br>Programming for engineers (EE057IU),<br>Earth Observation and Environment (PH027IU),<br>Remote Sensing (PH036IU) |

<table>
<tr><td>Module objectives/intended learning outcomes</td><td>Upon the successful completion of this course students will be able to:<br><table>
<tr><th>Competency level</th><th>Course learning outcome (CLO)</th></tr>
<tr><td>Knowledge</td><td>CLO1. Develop algorithms of analyzing big data in remote sensing using high-performance computing approaches and data visual analytics tools.</td></tr>
<tr><td>Skill</td><td>CLO2. Analyze data to make conclusions to engineering problems in big data and remote sensing.</td></tr>
<tr><td>Attitude</td><td>CLO3. Show abilities of further self-learning and lifelong learning.</td></tr>
</table></td></tr>
</table>

<table>
<tr><td>Content</td><td>The description of the contents should clearly indicate the weighting of the content and the level.<br>Weight: lecture session (3 hours)<br>Teaching levels: I (Introduce); T (Teach); U (Utilize)<br><table>
<tr><th>Topic</th><th>Weight</th><th>Level</th></tr>
<tr><td>Introduction to remote sensing big data</td><td>1</td><td>I, T</td></tr>
<tr><td>Infrastructure and high-performance computing for remote sensing data</td><td>2</td><td>I, T</td></tr>
<tr><td>Hadoop and Map Reduce</td><td>2</td><td>I, T</td></tr>
<tr><td>Distributed database</td><td>1</td><td>T, U</td></tr>
<tr><td>The computing platforms: parallel computing (CPUs and GPUs), Cloud computing</td><td>1</td><td>T, U</td></tr>
<tr><td>Python for Big Data</td><td>2</td><td>T, U</td></tr>
<tr><td>R - advanced spatial statistics for remote sensing big data</td><td>1</td><td>T, U</td></tr>
<tr><td>Remote sensing image handling: Image Enhancement, Data Mining</td><td>2</td><td>T, U</td></tr>
<tr><td>The open platform: Google Earth Engine</td><td>2</td><td>T, U</td></tr>
<tr><td>Final project: Thematic mapping from remote sensing big data</td><td>1</td><td>U</td></tr>
</table></td></tr>
<tr><td>Examination forms</td><td>Short answer question, project.</td></tr>
<tr><td>Study and examination requirements</td><td>Attendance: A minimum attendance of 80 percent is compulsory for the class sessions. Students will be assessed on the basis of their class participation. Questions and comments are strongly encouraged.<br>Assignments/Examination: Students must have more than 50/100 points overall to pass this course.</td></tr>
</table>

| | |
|---|---|
| Reading list | **Textbooks:**<br>[1] *Big Data: Techniques and Technologies in Geoinformatics*, Hassan A. Karimi (editor), 2014, CRC Press.<br>**References:**<br>[2] *High Performance Computing in Remote Sensing, Antonio* J. Plaza and Chein-I Chang (editors), 2008, Chapman & Hall/CRC Computer and Information Science Series.<br>[3] *Hadoop: The Definitive Guide*, 2nd edition, Tom White, 2011, O'Reilly.<br>[4] *An Introduction to R for Spatial Analysis and Mapping (Spatial Analytics and GIS),* Chris Brunsdon, Lex Comber, second edition<br>[5] *Big Data Analysis with Python: Combine Spark and Python to unlock the powers of parallel computing and machine learning*, Ivan Marin, Ankit Shukla, Sarang VK, 2019<br>**Software:** |