UNIVERSITAS NEGERI YOGYAKARTA
POSTGRADUATE PROGRAM DEPARTMENT OF ELECTRONICS AND INFORMATICS ENGINEERING EDUCATION
Jalan Colombo Nomor 1 Yogyakarta 55281
Telepon: (0274) 586168 Pesawat 216, 289, 292; Fax. (0274) 586734
Laman: ft.uny.ac.id, E-mail: humas_ft@uny.ac.id

Module name:	Big Data
Module level, if applicable:	Postgraduate
Code:	PTI 8221
Sub-heading, if applicable:	-
Classes, if applicable:	-
Semester:	2nd
Module coordinator:	Nurkhamid, S.Si., M.Kom., Ph.D.
Lecturer(s):	Nurkhamid, S.Si., M.Kom., Ph.D.
Language:	Bahasa Indonesia
Classification within the curriculum:	Elective Course
Teaching format / class Hours per week during the semester:	100 minutes lectures and 120 minutes structured activities per week.
Workload:	Total workload is 90,67 hours per semester which consists of 100 minutes lectures, 120 minutes structured activities, and 120 minutes self-study per week for 16 weeks
Creditpoints:	2
Prerequisites course(s):	Statistics
Course outcomes:	After taking this course the students have ability to: CO1. Ability to identify the characteristics of datasets and compare the trivial data and big data for various applications. CO2. Ability to select and implement machine learning techniques and computing environment that are suitable for the applications under consideration. CO3. Ability to solve problems associated with batch learning and online learning, and the big data characteristics such as high dimensionality, dynamically growing data and in particular scalability issues.

	CO4. Ability to understand and apply scaling up machine learning techniques and associated computing techniques and technologies. CO5. Ability to recognize and implement various ways of selecting suitable model parameters for different machine learning techniques. CO6. Ability to integrate machine learning libraries and mathematical and statistical tools with modern technologies like hadoop and mapreduce.

<table>
<tr><td>Content:</td><td>The course materials are organized to cover an overview of the subject matter and four topics in detail: big data analytics, big data computing environment, machine learning techniques, and scaling up machine learning. Therefore, the course materials are divided into five areas:
• Conceptualization and summarization: Representation of data. Modeling of
machine learning techniques. Application of big data computing technologies.
• Trivial data versus Big data: Representation learning. Publicly available datasets.
Scalability and Scaling up techniques. Report writing using Latex.
• Big data computing environment: Modern data analytics technologies like Hadoop and MapReduce. Suitable programming languages like Python, Java and C. Big data friendly machine learning scikit-learn libraries. Software platforms like Matlab or R.
• Machine learning techniques: Three phases of machine learning. types of learning. support vector machine. decision trees and random forests. deep learning.
• Scaling up machine learning: Dimensionality reduction techniques like principal
component analysis and feature hashing. Online processing technique called stochastic gradient descent. Big data machine learning models.</td></tr>
<tr><td>Study/exam achievements:</td><td>Three assignments (20% each = 60% total), project (20%), and final exam (20%)The final mark will be weight as follow:

<table>
<tr><th>No</th><th>CO</th><th>Assessment Object</th><th>Assessment Technique</th><th>Weight</th></tr>
<tr><td>1</td><td>CO1</td><td rowspan="3">Scientific paper</td><td rowspan="3">Assignment</td><td>20%</td></tr>
<tr><td>2</td><td>CO2</td><td>20%</td></tr>
<tr><td>3</td><td>CO3</td><td>20%</td></tr>
<tr><td>4</td><td>CO4</td><td rowspan="2">Project</td><td rowspan="2">Presentation</td><td>20%</td></tr>
<tr><td>5</td><td>CO5</td><td>20%</td></tr>
<tr><td>6</td><td>CO6</td><td>Final exam</td><td>Assignment</td><td>20%</td></tr>
<tr><td colspan="4">Total</td><td>100%</td></tr>
</table></td></tr>
<tr><td>Forms of media:</td><td>LCD Projector, Laptop / Computer, White Board</td></tr>
<tr><td>Literature:</td><td>1. Big Data Analytics: Systems, Algorithms, Applications. 2019. https://doi.org/10.1007/978-981-15-0094-7.
2. Hurwitz, J. (2017). Big data for dummies.
3. Turkington, G. (2013). Hadoop beginner's guide: Learn how to crunch big data to extract meaning from the data avalanche. (Hadoop Beginner's Guide.) Birmingham, UK: Packt Publ.
4. Tonidandel, S., King, E., & Cortina, J. M. (2016). Big data at</td></tr>
</table>

	work: The data science revolution and organizational psychology. 5. Simon, P. (2015). Too big to ignore: The business case for big data. 6. Błażewicz, G., Wydawnictwo Naukowe PWN., & Dadan Translations. (2018). Marketing automation revolution: Using the potential of Big Data. Warszawa: Wydawnictwo Naukowe PWN. 7. In Buyya, R., In Calheiros, R. N., & In Vahid, D. A. (2016). Big data: Principles and paradigms. Cambridge, MA: Morgan Kaufmann is an imprint of Elsevier. 8. Furht, B., & Villanustre, F. (2016). Big Data Technologies and Applications.

PLO and CO mapping

	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8	PLO9	PLO10
CO1					✓					
CO2					✓					
CO3					✓					
CO4					✓					
CO5					✓					
CO6					✓					