UNIVERSITAS MATARAM

(University of Mataram)

FAKULTAS TEKNIK

(Faculty of Engineering)

PROGRAM STUDI TEKNIK INFORMATIKA

(Department of Informatics Engineering)

MODULE HANDBOOK DESCRIPTION

Big Data (K22B51)

Module designation	Big Data
Semester(s) in which the module is taught	*5 / fourth year*
Person responsible for the module	*Dr.Eng I Gde Putu Wirarama Wedashwara Wirawan ST., MT.*
Language	*Indonesian*
Relation to curriculum	*Compulsory*
Teaching methods	*Lectures, Discussions, Project*
Workload (incl. contact hours, self-study hours)	Contact Hours every week, each week of the 16 weeks/semester including Evaluation ● 3 x 50 minutes lecturer/week ● 3 x 60 minutes class exercise/week ● Self Study hours = 120 minutes/week Total workload 340 minutes/week
Credit points	*3 (~ 3,2 ECTS)*
Required and recommended prerequisites for joining the module	Probability and Statistics (D18KK114) Database System (D18KB206)

Module objectives/intended learning outcomes	The main objective of Big Data is to learn about big data concepts and infrastructure, statistical analysis, processing flow and technical implementation of programming in big data environments such as Hadoop and spark. The courses are run on a project basis to train students to work together and be responsible for completing projects in the big data field, such as data crawling, visualisation, and preparation for the entire ETL (Extract Transform Load). Based on these main objectives, the application of big data courses have subject learning outcomes, namely: 1. Able to work together in groups and be responsible for their respective roles to complete big data tasks such as data crawling, visualization, preparation for the entire ETL (Extract Transform Load). 2. Able to explain big data concepts and infrastructure and run big data processing flows conceptually and technically in programming. *3.* Able to configure big data environments such as Hadoop and spark as well as implement programming involving statistical analysis.
Content	Big Data discusses big data concepts and infrastructure, statistical analysis, processing flow and technical implementation of programming in big data environments such as Hadoop and spark. The courses are run on a project basis to train students to work together and be responsible for completing projects in the big data field, such as data crawling, visualisation, and preparation for the entire ETL (Extract Transform Load).
Examination forms	*Assignments, Quiz, Simulation, Project Based Assignments*
Study and examination requirements	*Assignments 25%,* *Quiz 25%,* *Project based assignments 50%*

Reading list	1. Zikopoulos, P., & Eaton, C. (2011). Understanding big data: Analytics for enterprise class hadoop and streaming data. McGraw-Hill Osborne Media. 2. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile networks and applications, 19(2), 171-209. 3. Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. IEEE transactions on knowledge and data engineering, 26(1), 97-107. 4. McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. Harvard business review, 90(10), 60-68. 5. Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics-Challenges in topic discovery, data collection, and data preparation. International journal of information management, 39, 156-168. 6. Cattell, R. (2011). Scalable SQL and NoSQL data stores. Acm Sigmod Record, 39(4), 12-27. 7. Van der Veen, J. S., Van der Waaij, B., & Meijer, R. J. (2012, June). Sensor data storage performance: SQL or NoSQL, physical or virtual. In 2012 IEEE fifth international conference on cloud computing (pp. 431-438). IEEE. 8. Isele, R., Umbrich, J., Bizer, C., & Harth, A. (2010, November). LDspider: An open-source crawling framework for the Web of Linked Data. In Proceedings of the 2010 International Conference on Posters & Demonstrations Track (Vol. 658, pp. 29-32). 9. Cho, J. (2001). Crawling the web: discovery and maintenance of large-scale web data. Computer science. Stanford University. 10. White, T. (2012). Hadoop: The definitive guide. " O'Reilly Media, Inc.". 11. Venner, J. (2009). Pro hadoop. Apress. 12. Rathee, S. (2013, November). Big data and Hadoop with components like Flume, Pig, Hive and Jaql. In International conference on cloud, big data and trust (Vol. 15). 13. Pol, U. R. (2016). Big data analysis: Comparison of hadoop mapreduce, pig and hive. International Journal of Innovative Research in Science, Engineering and Technology, 5(6), 9687-93. 14. Drabas, T., & Lee, D. (2017). Learning PySpark. Packt Publishing Ltd.