



CONCEPTOS Y APLICACIONES DE BIG DATA

Carrera: *Especialización en Inteligencia de Datos orientada a Big Data*

Carga Horaria: 64 Hs.

Año 2018

Profesor a Cargo: *Dr. Wald oHasperué,
Ing. María José Basgall*

OBJETIVO

- Presentar al alumno las tecnologías actuales del diseño e implementación de aplicaciones en Big Data profundizando los conocimientos presentados con experimentación con trabajos en máquina.
- Estudiar los conceptos y fundamentos de Big Data.
- Analizar los principales problemas en las aplicaciones de Big Data
- Estudiar los frameworks actuales para el desarrollo de soluciones en Big Data
- Resolver problemas de Big Data utilizando arquitecturas Cloud.
- Herramientas: Hadoop Mapreduce, Spark, HiveQL

MODALIDAD DE EVALUACION

Los alumnos deberán tener más del 70% de asistencia a clase y aprobar los diferentes trabajos prácticos. Dichos trabajos deberán ser entregados acorde a un cronograma de fechas indicado al inicio del curso.

Al completar el curso se definirá un trabajo final experimental individual que deberá ser entregado dentro de los 60 días posteriores a la finalización el curso.



PROGRAMA

- A. Fundamentos de Big Data
 - Definición y dimensiones en Big Data.
 - Aplicaciones de Big Data.
 - Modelos de datos y modelos de procesamiento en Big Data
 - Etica, seguridad, privacidad en Big Data.
 - Casos de uso. IoT

- B. Modelos y frameworks para Big Data:
 - Ecosistema Hadoop.
 - Paradigma MapReduce.
 - Frameworks de alto nivel

- C. Diseño de aplicaciones en batch
 - Objetivos y alcance. Tipo de aplicaciones.
 - Cascading, HiveQL, Pig.
 - Apache Spark

- D. Procesamiento de stream data
 - Aplicaciones de stream data
 - Tipos de streams
 - Spark Streaming.

- E. Aplicaciones de Big Data sobre Cloud
 - Algoritmos de tratamiento de información textual.
 - Algoritmos de búsqueda inteligente.
 - Algoritmos de tratamiento de imágenes.
 - Estudio de casos.

BIBLIOGRAFIA

- Judith Hurwitz, Alan Nugent, Dr. Fern Halper and Marcia Kaufman (2013). Big Data for dummies. John Wiley & Sons, Inc. ISBN 978-1-118-50422-2.
- Charu C. Aggarwal (2007). Data streams: models and algorithms. Springer US. ISBN 978-0-387-28759-1.
- Soumendra Mohanty, Madhu Jagadeesh and Harsha Srivatsa (2013). Big Data Imperatives: Enterprise Big Data Warehouse, BI Implementations and Analytics. Apress. ISBN 978-1430248729.
- James A. Scott (2015). Getting Started with Apache Spark. MapR Technologies, Inc.



- Tom White (2011). Hadoop: The Definitive Guide. Hadoop: The Definitive Guide. ISBN 978-1-449-38973-4. 2011.
- O'Reilly Media, Inc. (2016). Big Data Now. O'Reilly Media, Inc. ISBN 9781449356712.
- Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman (2011). Mining of Massive Datasets. Cambridge University Press. ISBN 978-1107015357.
- Kord Davis with Doug Patterson (2012). Ethics of Big Data. O'Reilly Media, Inc. ISBN 978-1-449-31179-7.
- Wei Fan and Albert. (2012). Bifet Mining Big Data: Current Status, and Forecast to the Future. SIGKDD Explorations Vol. 14(2) 1-5.
- Xindong Wu, Xingquan Zhu, Gong-Qing Wu and Wei Ding. (2014). Data Mining with Big Data. IEEE transactions on knowledge and data engineering, Vol. 26 (1): 97-107.
- Bifet, A. (2013). Mining big data in real time. Informatica, 37(1).
- Gubbi, J. et al. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), 1645-1660.
- Fan, W., & Bifet, A. (2013). Mining big data: current status, and forecast to the future. ACM SIGKDD Explorations Newsletter, 14(2), 1-5.
- Amatriain, X. (2013). Mining large streams of user data for personalized recommendations. ACM SIGKDD Explorations Newsletter, 14(2), 37-48.
- Leskovec, J. et al. (2014). Mining of massive datasets. Cambridge University Press.
- Laney, D. (2001). 3-D Data Management: Controlling Data Volume, Velocity and Variety, META Group Original Research Note.
- Shanmuganathan, S. (2014). From data mining and knowledge discovery to big data analytics and knowledge extraction for applications in science.
- Kakhani, M. K. et al. (2013). Research Issues in Big Data Analytics. International Journal of Application or Innovation in Engineering & Management (IJAIEM), 2(8).
- Kaisler, S. et al. (2013). Big data: Issues and challenges moving forward. In System Sciences (HICSS), 46th Hawaii International Conference on (pp. 995-1004). IEEE.
- Saha, B., & Srivastava, D. (2014). Data quality: The other face of big data. In Data Engineering (ICDE), 2014 IEEE 30th International Conference on (pp. 1294-1297).
- Chalmers, S. et al. (2013). Big Data-State of the Art.
- Bajpayee, R. et al. (2015). Big Data: A Brief investigation on NoSQL Databases.
- Akay, A. et al. (2015). Network-Based Modeling and Intelligent Data Mining of Social Media for Improving Care. Biomedical and Health Informatics, IEEE Journal.
- Selvaperumal, P., & Suruliandi, A. (2014). A short message classification algorithm for tweet classification. In Recent Trends in Information Technology (ICRTIT), 2014 International Conference on (pp. 1-3). IEEE.
- Wang, Z. et al. (2014). Anomaly Detection through Enhanced Sentiment Analysis on Social Media Data. In Cloud Computing Technology and Science (CloudCom), 2014 IEEE 6th International Conference on (pp. 917-922). IEEE.