

Especialista Universitario/a en PROYECTOS DE ANALÍTICA PREDICTIVA DE DATOS PARA LA INDUSTRIA 4.0

MÓDULO 1: ASPECTOS GENERALES DE LA ANALÍTICA DE DATOS Y SUS APLICACIONES PARA LA INDUSTRIA 4.0 (3 ECTS)

- 1- Presentación del título propio: objetivos y estructura de contenidos
- 2- Industria 4.0
- 3- Procesos industriales & Automatización (visión general)
- 4- Analítica de datos & Big Data
- 5- Contextos de aplicación industrial de la analítica predictiva de datos: Mantenimiento predictivo; Optimización de procesos / eficiencia / calidad de producto.
- 6- Emprendimiento en analítica de datos para la Industria 4.0
- 7- Casos de éxito de empresas industriales

Bibliografía

- Kagermann et al. (2013) Recommendations for implementing the strategic initiative INDUSTRIE 4.0;

http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Material_fuer_Sonderseiten/Industrie_4.0/Final_report_Industrie_4.0_acc_e_ssible.pdf
- E. Mandado et al. (2009) *Autómatas Programables y Sistemas de Automatización*, 2a ed. ISBN: 9788426715753
- M. Niño, A. Illarramendi, (2015). Entendiendo el Big Data: antecedentes, origen y desarrollo posterior. *DYNA New Technologies*, 2(1), [8p.]
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann
- D. Galar, U. Kumar (2017). *eMaintenance: Essential Electronic Tools for Efficiency*.

Academic Press

MÓDULO 2: EL ROL DEL *DIRECTOR DE PROYECTOS* DE ANALÍTICA PREDICTIVA DE DATOS EN LA INDUSTRIA 4.0 (4 ECTS)

- 1- Perfiles principales en los equipos de proyectos de analítica de datos
- 2- Definición del Business Case
- 3- Aspectos metodológicos sobre la dirección y gestión de proyectos:
 - El rol del director de proyecto
 - Metodologías clásicas (basado en PMBoK)
 - Metodologías ágiles
- 4- Metodología y perfiles necesarios para la ejecución de proyectos de analítica predictiva de datos (basado en CRISP-DM)
- 5- Claves específicas de los proyectos para la industria manufacturera: aspectos prácticos derivados de la práctica profesional

Bibliografía

Relacionada con PMBoK:

- <https://www.pmi.org/pmbok-guide-standards>
- <http://store.rmcpjproject.com/pmp-exam-prep-ninth-edition>
- <https://pmi-mad.org/>
- <https://pmi-mad.org/index.php/quienes-somos/branch-pais-vasco>

Relacionada con metodologías ágiles:

- <http://agilemanifesto.org/iso/es/principles.html>
- www.infoq.com/minibooks/kanban-scrum-minibook

- <http://www.lulu.com/shop/henrik-kniberg/scrum-and-xp-from-the-trenches/paperback/product-1673516.html>

Relacionada con CRISP-DM

- CRISP-DM 1.0: Step-by-step data mining guide:
<ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf>

MÓDULO 3: LA INGENIERÍA DE DATOS EN PROYECTOS DE ANALÍTICA DE DATOS EN LA INDUSTRIA 4.0 (4 ECTS)

1- Instrumentación y captura de datos en diferentes ámbitos de monitorización industrial 2-

Infraestructura Big Data en el dominio de la Industria 4.0: entornos on-premise vs cloud

- Introducción a distribuciones Big Data
- Plan de escalabilidad: Dimensionamiento de un cluster
- Entornos on-premise
- Soluciones Cloud

3- Arquitecturas Big Data y Ecosistema de herramientas

- Paradigmas y arquitecturas Big Data en la Industria 4.0
- Ecosistema Hadoop
- Otras herramientas Big Data

4- Modelos de programación distribuidos y bases de datos NoSQL

- Programación paralela y sistemas de ficheros distribuidos
- Hadoop MR
- Spark
- Bases de datos NoSQL

Bibliografía

- MapR → <http://maprdocs.mapr.com/home/>
- Hortonworks → <https://docs.hortonworks.com/>
- Cloudera → <https://www.cloudera.com/documentation.html>
- Big Data: Principles and Best Practices of Scalable Real-time Data Systems, <https://www.manning.com/books/big-data>
- Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems, <http://shop.oreilly.com/product/0636920032175.do>
- Hadoop: The Definitive Guide, <http://shop.oreilly.com/product/0636920033448.do>
- Learning Spark: Lightning-Fast Big Data Analysis, <http://shop.oreilly.com/product/0636920028512.do>
- Getting Started with NoSQL, <http://shop.oreilly.com/product/9781849694988.do>

MÓDULO 4: LA CIENCIA DE DATOS Y LOS MODELOS DE ANALÍTICA PREDICTIVA EN LA INDUSTRIA 4.0 (9 ECTS)

- 1- Conceptos introductorios de analítica de datos e inferencia estadística. Correlación y causalidad.
- 2- Ontología general de técnicas de análisis de datos: clusterización, clasificación/predicción, optimización
- 3- Procedimiento típico de un proceso de analítica de datos y de validación de modelos
- 4- Aspectos prácticos de filtrado, limpieza de datos, rellenado de ceros, preprocesado
- 5- Detección de muestras atípicas (outliers): concepto y técnicas
- 6- Revisión de algoritmos y modelos principales:

- Clusterización o agrupación: taxonomía y técnicas
- Redes neuronales
- Máquinas de vectores soporte
- Árboles de decisión y regresión
- Inferencia y aprendizaje bayesiano
- Mezclas de modelos

7- Visualización de Datos

8- Revisión de implementaciones y plataformas

9- Categorización de aplicaciones industriales:

- Clasificación y reconocimiento de patrones
- Gestión del estado de salud de activos industriales
- CBM y mantenimiento predictivo
- Optimización del ciclo de vida
- Gestión eficiente de la producción y de la distribución logística

Bibliografía

- Aggarwal, C. C. (Ed.). (2014). *Data classification: algorithms and applications*. CRC Press
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825-2830

TRABAJO DE INVESTIGACIÓN (10 ECTS)

El alumnado desarrollará un proyecto final sobre un escenario real de negocio industrial. Partiendo de un caso anonimizado de una empresa de un sector concreto de manufactura y el acceso a un volcado en bruto de datos sobre los que trabajar, el alumnado deberá recorrer diferentes fases de un proyecto de analítica de datos, interactuar con implicados y explorar diferentes tecnologías, para finalmente exponer sus resultados y conclusiones desde un punto de vista tanto técnico como de negocio.