



Universidad de los Andes
Ingeniería de Sistemas y Computación
IS53301 – Inteligencia de Negocios - sección 2 y 3
Semestre: 2025-10
Profesores:
S1. Manuel Porras - mf.porras@uniandes.edu.co
S2. Haydemar María Nuñez C. - h.nunez@uniandes.edu.co
S3. María del Pilar Villamil G.- mavillam@uniandes.edu.co
S4. Juan Pablo Reyey - jp.reyes39@uniandes.edu.co
S5. Diego Escobar - df.escobar@uniandes.edu.co

Horario:
S1. Lu.Ju. 8:00- 9:20
S2. Ma.Ju. 9:30-10:50
S3. Ma.Ju. 11:00-12:20
S4. Lu.Mi 11:00-12:20
S5. Lu.Mi 5:00- 6:30

Semana	Día	Fecha	Tema	Bibliografía	Preguntas a resolver	Actividades (a entregar)
Módulo 1: Machine learning						
1	Ma	21-ene	Introducción: analítica, historia y ejemplos de proyectos de aprendizaje automático y de bodegas de datos.	[10] Cap. 1. Core Data and Analytics Concepts [13] Cap. 1. So What? Creating Value with Data Science	¿Cuáles son los objetivos de una solución de analítica? ¿Qué características tienen las aplicaciones de analítica? ¿Qué relación existe entre sistemas transaccionales, IN y analítica? ¿Qué caracteriza a un científico de datos y a un ingeniero de datos?	
	Ju	23-ene	Generalidades de aprendizaje automático- tareas de aprendizaje	[1] Cap. 1 The machine learning Landscape, [12]	¿Qué es aprendizaje automático? ¿Qué tipos y tareas de aprendizaje hay en aprendizaje automático? ¿Qué pasos tiene la metodología ASUM-DM?	
2	Ma	28-ene	Metodología para proyectos de aprendizaje automático. Entendimiento de datos: Perfilamiento de datos y análisis de calidad de datos.	[1] Apendix A. machine learning Project Checklist [2] Cap. 4 Building Good Training Datasets – Data Preprocessing. Section Identifying missing values in tabular data pág. 108 * Infografías sobre entendimiento y calidad de datos	¿Qué significa entender los datos? ¿Qué es calidad y perfilamiento de datos? ¿Qué es JupyterLab?	
	Ju	30-ene	Entendimiento y Preparación de datos. Ingeniería de características. Procesamiento de texto	[3] Cap. 4. Feature Engineering for Text Representation (Hasta Advanced Feature Engineering Models, no incluido) * Infografía de ASUM-DM	¿Qué significa preparar los datos? ¿Qué es ingeniería de características? ¿Cuál es el proceso a seguir para preparar textos con el fin de aplicar técnicas de aprendizaje automático (e.g., clasificación, regresión)?	Enunciado proyecto 1
3	Ma	4-feb	Procesamiento de texto e inicio proyecto 1	Enunciado proyecto 1	¿Qué se debe hacer en el proyecto 1? ¿Cuál es la relación con lo que se está trabajando en clase?	
	Ju	6-feb	Tarea de regresión. Formulación de la tarea, revisión de regresión lineal como técnica. Inicia el Laboratorio de regresión.	[1] Cap. 2. End-to-End aprendizaje automático Project. Cap. 4. Training Models. Linear regression.	¿Qué características deben tener las variables de entrada y la variable objetivo para aplicar algoritmos de regresión lineal? ¿Qué métricas se utilizan para evaluar la calidad de un modelo de regresión?	
4	Ma	11-feb	Análisis del resultado de la tarea de regresión. Automatización del proceso de creación del modelo analítico.	[4] Cap. 4 Regression and Prediction - Interpreting the Regression Equation	¿Cómo se interpreta el resultado de una regresión lineal? ¿Qué son los pipelines?	
	Ju	13-feb	Tarea de Clasificación. Algoritmos basados en árboles. Preparación de datos -- transformación de datos categóricos a numéricos.	[1] Cap. 2. End-to-End aprendizaje automático Project - Pipelines [1] Cap. 3. Classification (No incluye performance measures). Cap. 6. Decision Trees (No incluye regression). [1] Cap. 6. Decision Trees	¿Qué diferencia hay entre una tarea de clasificación y una de regresión? ¿Cómo funciona la técnica de árboles de decisión? ¿Qué métodos de selección de atributos existen en algoritmos de árboles de decisión? ¿Cuáles son los criterios de terminación del algoritmo base de un árbol de decisión?¿Qué ventajas y desventajas tiene la técnica de árboles de decisión?	Entrega laboratorio 1 (Sábado de esa semana 20:00)
5	Ma	18-feb	Técnica KNN para tarea de clasificación. Análisis de un modelo de clasificación y sus resultados - métricas para evaluar la calidad de un modelo de clasificación. Revisión resultados laboratorio de regresión.	[2] Cap. 3 A Tour of machine learning Classifiers Using scikit-learn. Secc. K-nearest neighbors – a lazy learning algorithm [1] Cap. 3 Classification. Performance Measures.	¿Cómo funciona la técnica de KNN? ¿Qué métricas se utilizan para evaluar la calidad de un modelo de clasificación? ¿Qué es cross-validation?¿Qué es la matriz de confusión?¿Qué son las métricas de precision y de recall?¿Cómo balancear conjuntos de datos desbalanceados y por qué hacerlo?	
	Ju	20-feb	Análisis de resultados de tarea de clasificación.		¿Cómo interpretar el resultado de un árbol de decisión?	Entrega proyecto 1 etapa 1 (Sábado de esa semana 20:00)
6	Ma	25-feb	Tarea de agrupación (Clustering). Preparación de datos: normalización, análisis de variables. Inicia el Laboratorio de agrupación.	[1] Cap. 9. Unsupervised Learning Techniques (k-means, dbscan)	¿Cómo funciona el algoritmo de k-means? ¿Cuándo se utiliza la función de distancia? ¿Cuáles son las funciones de distancia más utilizadas? ¿Qué diferencia hay entre la tarea de agrupación y la de clasificación?	
	Ju	27-feb	Mecanismos de selección del número de grupos en algoritmos de agrupación parcial. Análisis del resultado de agrupación.	[1] Cap. 9. Unsupervised Learning Techniques. Finding the optimal number of clusters	¿Qué es el método de codo (inercia) y cómo se usa en la selección del K? ¿Qué es, para qué sirve y cómo se calcula la silueta de un clúster?	
7	Ma	4-mar	Simulacro previo y en clase aclaración de dudas y repaso parcial		¿Qué dudas me quedan de los temas tratados? ¿Conozco la rúbrica que van a utilizar para calificar el parcial?	
	Ju	6-mar	Proyecto 1 y Práctica de proceso de despliegue de modelos analíticos		¿Qué actividades específicas hay que realizar en procesos de automatización y despliegue de modelos analíticos?	
8	Ma	11-mar	Parcial 1- Etapa 1 aprendizaje automático		Etapla 1 del parcial, con un conjunto de datos y en python	
	Ju	13-mar	Parcial 1- Etapa 2 aprendizaje automático		Etapla 2 del parcial, parte conceptual aplicada y refinamiento de la parte práctica.	Entrega laboratorio 2 (Sábado de esa semana 20:00)
SR	Ma	18-mar	Semana de receso			
	Ju	20-mar				
Módulo 2: Bodegas de datos y modelos multidimensionales						
9	Ma	25-mar	Proyectos de inteligencia de negocios -- Arquitectura de aplicaciones, Metodología, requerimientos analíticos.	[5] Cap. 2 Data Management Architectures for Analytics hasta: Building data lakes to tame the variety and volume of big data (incluido)	¿Qué diferencias existen entre estos proyectos y los de analítica basados en aprendizaje automático? ¿Qué componentes conforman una arquitectura de una solución de BI?	Se publica el enunciado del proyecto 2
	Ju	27-mar	Generalidades del modelado multidimensional.	[6] Cap. 3. Retail Sales. (hasta pág. 79 o el contenido de transaction FacTables en el ebook)	¿Qué diferencias existen entre un modelo normalizado orientado a transacciones y uno multidimensional?	28 de marzo último día para reportar el 30%
	Vi	28-mar	Charla de temas actuales y tendencias en analítica & BI			Entrega proyecto 1 etapa 2 (Sábado de esa semana 20:00)
10	Ma	1-abr	Modelado multidimensional. Caso retail.	[6] Cap. 3 . Retail Sales. (hasta pág. 79 o el contenido de transaction FacTables en el ebook)	¿Qué son dimensiones y tablas de hecho? ¿Cómo es el proceso para construir modelos multidimensionales?	Revisar los videos de modelado multidimensional
	Ju	3-abr	Proyecto 2: Entender el contexto. Revisar el enunciado, los datos, revisar proyectos similares para plantear requerimientos analíticos iniciales.	[7] Cap. 3. Collecting the Requirements	¿Cómo abordar una entrevista con un cliente para identificar requerimientos analíticos?	
	Vi	4-abr	Apertura del segundo proyecto			
11	Ma	8-abr	Modelar dimensiones: introducir el manejo de historia de atributos de dimensiones en modelos multidimensionales. Caso de aprovisionamiento- adquisición.	[6] Cap. 5. Procurement o Lectura del curso: historia de atributos de dimensiones	¿Qué tipos propone el grupo kimball para el manejo de historia en las dimensiones? ¿Qué características tienen? ¿Qué es una matriz de bus?	
	Ju	10-abr	Análisis descriptivo y el uso de tableros de control Proyecto 2: Diseñar el boceto de los tableros de control de análisis descriptivo	[9] Cap. 1. Dashboards y Cap 5. Dashboard Design	¿Qué características tiene un tablero de control? ¿Qué gráficos debo utilizar para transmitir mejor un mensaje basado en datos? ¿Qué me dicen los datos? ¿Qué decisiones puedo tomar con los datos?	11 de abril último día de retiro de materias
SS	Ma	15-abr	Semana Santa			
	Ju	17-abr				
12	Ma	22-abr	Proyecto 2: Refinar requerimientos analíticos, revisar bocetos y modelos multidimensionales	[11] Cap. 2. choosing an effective visual y Cap. 7. Lessons in storytelling.	¿A quién están dirigidos los tableros de control? ¿Qué historia van a contar?	
	Ju	24-abr	Laboratorio de proceso ETL	[8] Cap. 2: Understanding the ETL Process and Data Pipelines	¿Qué actividades específicas hay que realizar en procesos de ETL? ¿Qué diferencias existen entre ETL y ELT?¿Qué estrategias se pueden proponer para cargar varias fuentes que reporten información de dimensiones o de tablas de hecho? ¿Qué riesgos existen si los datos a cargar no cumplen con lo definido por los dueños de los procesos? ¿Existen mecanismos para corregir y reportar problemas de calidad de datos?	
13	Ma	29-abr	Proyecto 2: ajustar bocetos de tableros y trabajar en el diseño del proceso de ETL	[11] Cap 8. Pulling it all together.	¿A quién están dirigidos los tableros de control? ¿Qué historia van a contar?	
	Ju	1-may	Festivo: Modelado multidimensional. Ejercicios	[6] Cap. 3. Retail sales (A partir de la pág. 79)	¿Qué es una llave subrogada?¿Qué es una dimensión degenerada?¿Qué es un esquema copo de nieve- snowflake?¿Qué es un outtrigger?	Entrega laboratorio 3 (Sábado de esa semana 20:00)
14	Ma	6-may	Modelado multidimensional. Caso inventario.	[6] Cap. 4 Inventory (pág. 114-115, 119-122)	¿Qué tipos de tablas de hecho existen?	
	Ju	8-may	Proyecto 2: Avanzar en implementación de los tableros de control y de ETL		¿A quién está dirigida la presentación? ¿Qué hallazgos tengo al utilizar los tableros de control construidos? ¿Qué vocabulario y temas presentar al comunicar el resultado de un proyecto de BI?	
15	Ma	13-may	Reflexión de ética en los proyectos centrados en datos		¿Quiero proponer un enunciado para el parcial?	
	Ju	15-may	Proyecto 2: Recomendaciones para la presentación final. Simulacro del parcial 2		¿Qué dudas me quedan en temas de modelado multidimensional?	
	Vi	16-may	Presentación de proyecto 2 (2:00 - 4:00 pm)			Entrega del proyecto 2 1:00 pm.
16	Ma	20-may	Reflexión del proyecto 2. Ejercicios de modelado multidimensional			
	Ju	22-may	Preparación parcial 2 y Cierre del curso.		¿Qué dudas me quedan en temas de modelado? ¿Conozco los errores típicos?	
Parcial 2 - Fecha programada por registro en la semana del 26 de mayo						

Bibliografía disponible en la biblioteca de la Universidad	
[1] GÉRON, Aurélien. "Hands-On aprendizaje automático with Scikit-Learn, Keras, and TensorFlow", 3rd Edition. O'Reilly 2022	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/1g0omta/alma991005483874707681
[2] RASCHKA, Sebastian, MIRJALILI, Vahid. "Python aprendizaje automático". Packt publishing. Second edition. 2017	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/18hc8d9/cdi_proquest_ebookcentral_FBC6005547
[3] SARKAR, Dipanjan. "Text Analytics with Python: A Practitioner's Guide to Natural Language Processing". Apress. 2019	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/1g0omta/alma991005253792907681
[4] BRUCE, Peter, BRUCE, Andrew. "Practical statistics for data scientists: 50 essential concepts". O'Reilly Media, Inc. 2017	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/1g0omta/alma991005388400707681
[5] EAGAR, Gareth. "Data Engineering with AWS". Packt Publishing. 2021.	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/18hc8d9/cdi_safari_books_v2_9781800560413
[6] KIMBALL, Ralph, ROSS, Margy. "The Data Warehouse Toolkit: the definitive guide to dimensional modeling". Third Edition. John Wiley & Sons, Inc. 2013.	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/1g0omta/alma991005385820707681
[7] KIMBALL, Ralph, REEVES, Laura, ROSS, Margy, THORNTHWAITE, Warren. "The Data Warehouse Lifecycle Toolkit". John Wiley & Sons, Inc., 2008.	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/1g0omta/alma991005272374707681
[8] PANDEY, Brij K, SCHOOF, Emily R. "Building ETL Pipelines with Python". 2023.	https://learning-oreilly.com.ezproxy.uniandes.edu.co/library/view/building-etl-pipelines/9781804615256/819123_02.xhtml#_idParaDest-45
[9] KHAN, Arshad. "Visual Analytics for Dashboards: A Step-by-Step Guide to Principles and Practical Techniques". Apress 2024.	https://learning-oreilly.com.ezproxy.uniandes.edu.co/library/view/visual-analytics-for/9798868801198/
[10] BHADRESH, Shiyal. Beginning Azure synapse analytics : transition from data warehouse to data lakehouse. Apress. 2021.	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/1g0omta/alma991005390001007681
[11] NUSSBAUMER KAFILIC, Cole. "Storytelling with Data: A Data Visualization Guide for Business Professionals". Wiley, 2015.	https://uniandes.primo.exlibrisgroup.com/permalink/57U_UDLA/1g0omta/alma991005388510207681
[12] "CRISP-ML(Q). The ML Lifecycle Process"	https://ml-ops.org/content/crisp-ml
[13] VAUGHAN, Daniel. "Data Science: The Hard Parts". O'Reilly Media, Inc. 2023	https://learning-oreilly.com.ezproxy.uniandes.edu.co/library/view/data-science-the/9781098146467/