



Facultad de
Ingeniería

coursera

MIAD



Maestría
en Inteligencia
Analítica de Datos

La Universidad de los Andes como institución que almacena, y recolecta datos personales, requiere obtener su autorización para que de manera libre, previa, expresa, voluntaria y debidamente informada, permita a todas las dependencias académicas y/o administrativas, recolectar, recaudar, almacenar, usar, circular, suprimir, procesar, compilar, intercambiar, dar tratamiento, actualizar y disponer de los datos que han sido suministrados y que se han incorporado en distintas bases o bancos de datos, o en repositorios electrónicos de todo tipo con que cuenta la Universidad. Esta información es, y será utilizada en el desarrollo de las funciones propias de la Universidad en su condición de institución de educación superior, de forma directa o a través de terceros, a no ser que usted le manifieste lo contrario de manera directa, expresa, inequívoca y por escrito a la cuenta de correo electrónico dispuesta para tal efecto: habeasdata@uniandes.edu.co – Conozca más aquí: www.uniandes.edu.co/datospersonales

MIAD



Maestría
en Inteligencia
Analítica de Datos

Machine Learning y Procesamiento de Lenguaje Natural

Semana 1: Marzo 27-Marzo 31 de 2023



AGENDA

1. Bienvenida y Aspectos Generales del Curso

1. Sobre el Curso
2. Objetivos del Curso
3. Equipo
4. Sesiones y Cronograma General
 1. Estructura de una semana del curso
 2. Cronograma General
5. Estructura de Evaluación
 1. Peso porcentual de recursos calificables

2. Grupos de Trabajo en el Curso

3. Esquema de las Sesiones Sincrónicas

4. Canales de Apoyo

5. Preguntas y Respuestas

MIAD



Maestría
en Inteligencia
Analítica de Datos

Machine Learning y Procesamiento de Lenguaje Natural

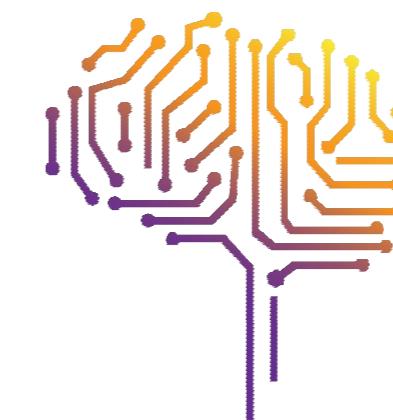
**Maestría
en
Inteligencia
Analítica de
Datos**

01

ASPECTOS GENERALES

Maestría en Inteligencia
Analítica de Datos

Universidad de los Andes





Sobre el Curso



El machine learning es un campo que ha revolucionado el mundo al proveer soluciones efectivas a diferentes problemas a través de algoritmos capaces de aprender patrones en un conjunto particular de datos.



El procesamiento de lenguaje natural utiliza diferentes técnicas y algoritmos para comprender el lenguaje humano, por lo que es una de las áreas de aprendizaje automático más aplicadas.

Específicamente, en las organizaciones, el uso de **machine learning constituye una ventaja competitiva**, ya que permite generar soluciones innovadoras y precisas a diferentes problemáticas. Por lo anterior, en este curso **estudiaremos los conceptos y técnicas fundamentales del machine learning y del procesamiento del lenguaje natural (NLP)**, con un enfoque en problemas reales para su implementación en contextos organizacionales.

Objetivos del Curso



Implementar sistemas productivos de Machine Learning para que puedan ser usados en diversos ambientes organizacionales.

Crear modelos de Machine Learning de acuerdo con las necesidades particulares de una organización.

Implementar modelos de procesamiento de Lenguaje Natural en contextos organizacionales pertinentes.

Seleccionar modelos predictivos, con base en sus implicaciones técnicas, para que puedan ser usados en contextos organizacionales.

Equipo Docente 2023-12

Profesores



David Zarruk Valencia
Profesor Principal

Tutores



**Jaime
Acevedo**
Tutor líder



**Ana María
Quintero**
Tutor líder



**Alisson
Bernal**



**Nicolás
Danies**



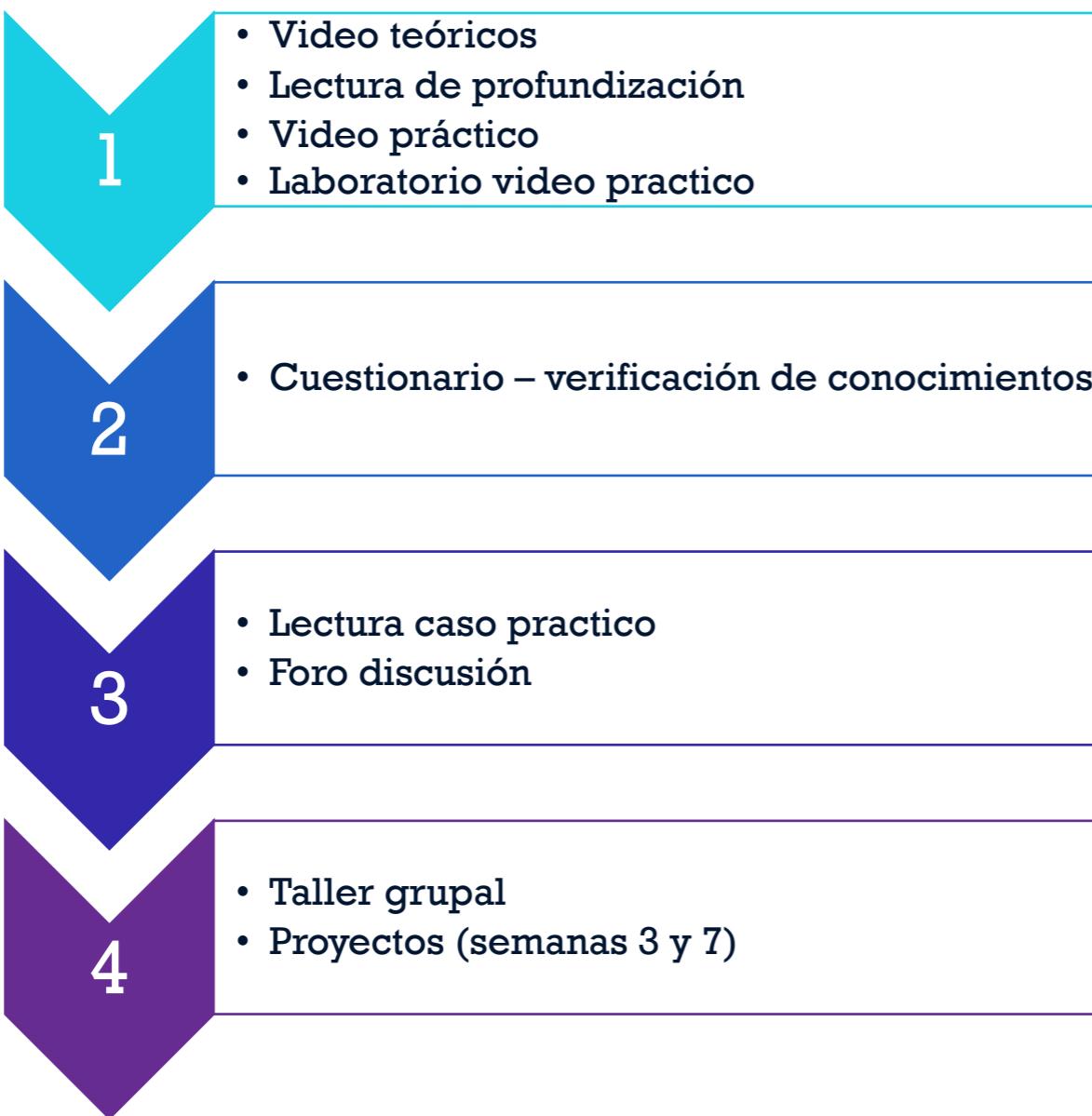
**Juan Esteban
Rojas**



**Felipe
Rueda**

Sesiones y Cronograma Semanal

Esquema general de una semana



Las actividades del curso están diseñadas para comprender las temáticas asociadas a cada una de las semanas, adquirir habilidades para desarrollar modelos en Python e integrar y aplicar los contenidos tratados en la resolución de distintos tipos de problemas predictivos. Cada semana cuenta con videos, lecturas de profundización, tutoriales y cuestionarios, que le permitirá al estudiante prepararse para los talleres individuales y grupales que se han propuesto durante las semanas.

Temáticas Semanales

Módulo	Objetivos
1. Árboles de decisión y ensamblaje	<ul style="list-style-type: none"> • Reconocer las técnicas de uso de árboles de decisión y ensamblaje. • Identificar la relación de cada uno de los parámetros de los modelos de ensamblaje con su ejecución. • Crear e implementar árboles de decisión y ensamblajes de forma manual.
2. Random Forest y XGBoost	<ul style="list-style-type: none"> • Reconocer las características de los modelos de ensamblaje Random Forest y XGBoost y sus respectivos parámetros de ejecución. • Ajustar los parámetros de los modelos Random Forest y XGBoost para lograr su mejor desempeño. • Crear e implementar modelos de ensamblaje Random Forest y XGBoost.
3. Machine Learning como servicio en la nube	<ul style="list-style-type: none"> • Seleccionar la combinación de algoritmos que permita obtener el mejor desempeño productivo de un modelo. • Crear e implementar interfaces de Programación de Aplicaciones (APIs) en servicios de la nube para poder disponibilizar un modelo en un contexto organizacional.
4. Introducción al Procesamiento del Lenguaje Natural (NLP)	<ul style="list-style-type: none"> • Reconocer las características y la utilidad de los modelos de procesamiento de lenguaje natural. • Reconocer e implementar distintas técnicas para el preprocesamiento de textos en modelos NLP.
5. Análisis de sentimientos	<ul style="list-style-type: none"> • Reconocer e implementar distintas técnicas para el preprocesamiento de textos en modelos NLP. • Crear e implementar modelos predictivos basados en datos de textos.
6. Introducción a Redes Neuronales	<ul style="list-style-type: none"> • Reconocer el concepto de red neuronal, sus respectivos tipos y aplicaciones. • Reconocer los parámetros para la implementación de redes neuronales. • Crear e implementar redes neuronales para modelos de clasificación, usando librerías especializadas en Python.
7. NLP usando Redes Neuronales I	<ul style="list-style-type: none"> • Reconocer el concepto de red neuronal, sus respectivos tipos y aplicaciones. • Crear el mejor modelo de predicción de NLP, de acuerdo con el contexto particular de una organización. • Implementar proyectos de NLP que abarquen el proceso de selección de datos, preprocesamiento, modelación, análisis de resultados y disponibilización.
8. NLP usando Redes Neuronales II	<ul style="list-style-type: none"> • Ajustar los parámetros de ejecución de redes neuronales para lograr su mejor desempeño. • Crear el mejor modelo de predicción de NLP, de acuerdo con el contexto particular de una organización. • Implementar proyectos de NLP que abarquen el proceso de selección de datos, preprocesamiento, modelación, análisis de resultados y disponibilización.

Estructura Evaluación del Curso

Cuestionarios semanales (20%, cada uno 2.5%)

- Árboles de decisión y ensamblajes.
- Random Forest y XGboost.
- Machine Learning como servicio en la nube.
- Introducción al procesamiento de lenguaje natural.
- Análisis de sentimientos.
- Introducción a redes neuronales.
- NLP usando redes neuronales.
- Calibración de parámetros en redes neuronales.

Talleres individuales y grupales (25%, cada uno 5%)

- Taller individual: Construcción e implementación de árboles de decisión y métodos de ensamblaje.
- Taller grupal: Construcción e implementación de modelos Bagging, Random Forest y XGBoost.
- Taller grupal: Tokenización de textos
- Taller grupal: Análisis de sentimientos y técnicas de NLP.
- Taller grupal: Redes Neuronales.

Proyectos grupales (50%, cada uno 25%)

- Proyecto 1: Predicción de precios de vehículos usados.
- Proyecto 2: Clasificación de género de películas.

Actividades de coevaluación del trabajo en equipo (5%, cada uno 2.5%)

- Actividad de coevaluación semana 5.
- Actividad de coevaluación semana 8.

Las entregas se deben realizar el día especificado hasta las 11:59 pm hora COL; las entregas fuera de tiempo sólo se considerarán en casos de fuerza mayor, acorde con el reglamento de la Universidad.

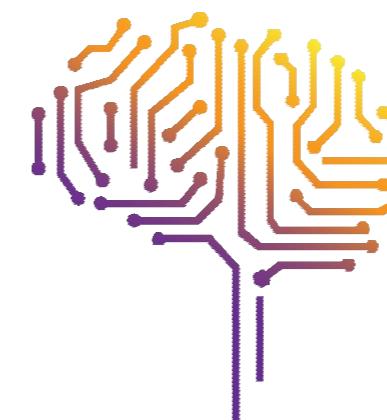
Lo anterior aplica para todas las entregas calificables, aunque la plataforma de Coursera permita el envío de estas fuera de tiempo.

02

GRUPOS DE TRABAJO

Maestría en Inteligencia Analítica de Datos

Universidad de los Andes



Descripción Grupos de Trabajo



¿Para cuáles entregas?



¿Cómo se crearon estos grupos?

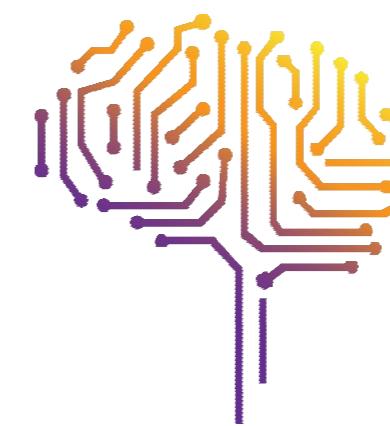


03

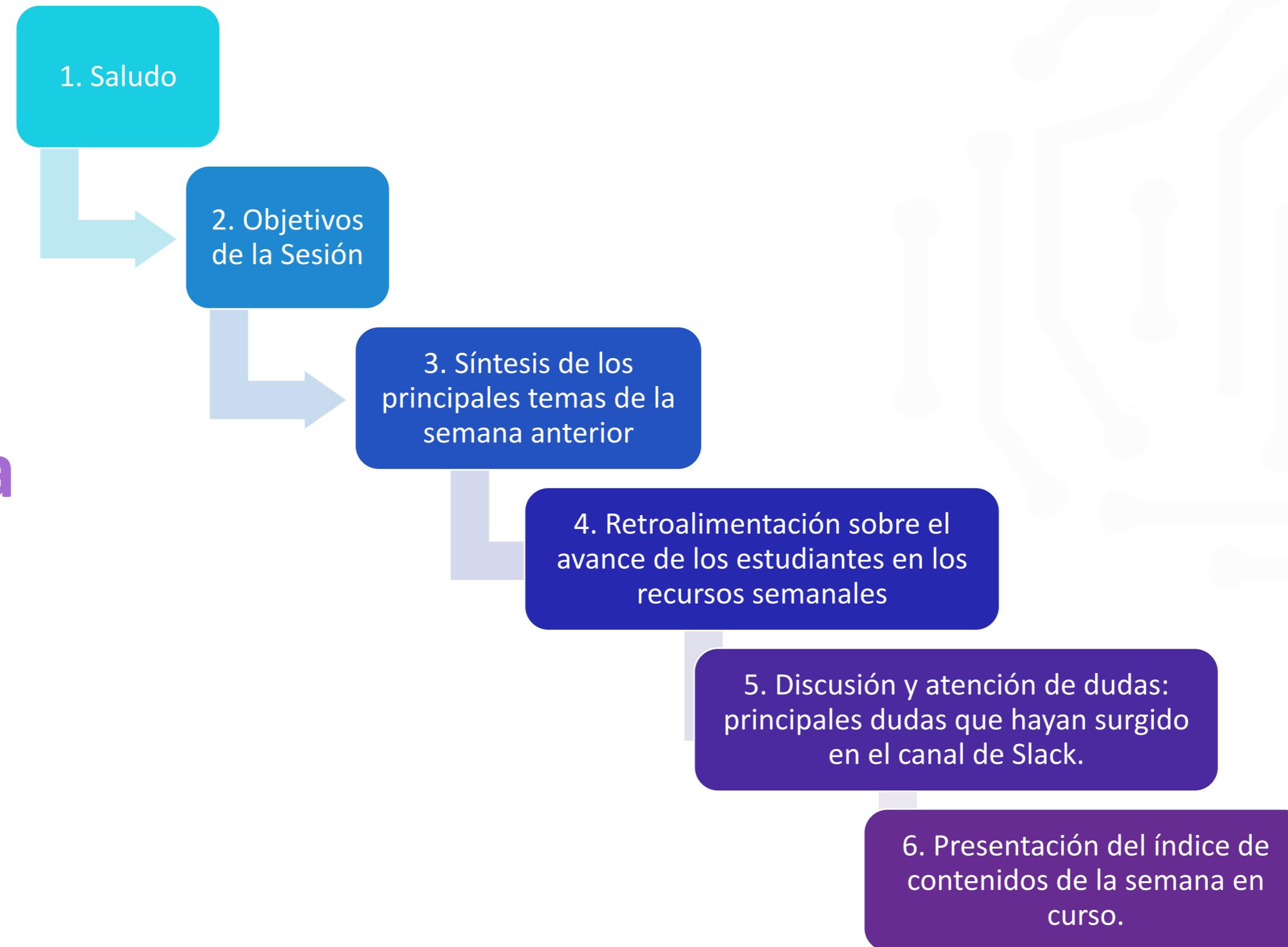
ESQUEMA DE LAS SESIONES SINCRÓNICAS

Maestría en Inteligencia
Analítica de Datos

Universidad de los Andes



Esquema General



Particularidades Contenidos Semana 1

A

Introducción al curso

- Vídeo: Bienvenida al curso
Publicado 4m
- Lectura: Presentación del curso
Publicado 12m
- Lectura: Navegación del curso
Publicado 10m
- Cuadro de aviso de la discusión: Saludar y presentarse
Publicado 10m

B

Actividades individuales prácticas

- Lectura: Tutorial en Github: principales conceptos
Publicado 45m
- Tarea calificada por el personal: Taller: Construcción e implementación .
Publicado 4h 3 objetivos

C

Preparación actividades grupales

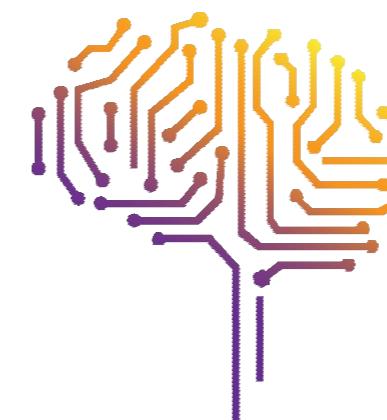
- Lectura: Tres etapas de un proyecto de equipo exitoso
Publicado 10m
- Lectura: Conformación de equipos
Publicado 15m

04

CANALES DE APOYO

Maestría en Inteligencia
Analítica de Datos

Universidad de los Andes



Canales de Comunicación

Slack:

Canal por el que podremos comunicarnos continuamente y con mayor interacción. En este canal estarán los tutores para responder sobre dudas/inquietudes respecto al curso.

Por este canal se podrán formular las dudas o temas que deseen profundizar en la sesión sincrónica siguiente para incluirlas en la agenda.

Correo-Salesforce: solicitudes-miad@uniandes.edu.co

Podrán tramitar casos-solicitudes específicos respecto a calificaciones, circunstancias especiales, etc. Este es un canal oficial de respuesta y seguimiento.





PREGUNTAS

CONTÁCTANOS

PÁGINAS WEB - Ingeniería Industrial:

<https://industrial.uniandes.edu.co/es/programas-academicos/maestrias>

CORREO Maestría: solicitudes-miad@uniandes.edu.co

OTROS ENLACES DE IMPORTANCIA

Matrículas: <https://matriculas.uniandes.edu.co/>

Registro: <http://registro.uniandes.edu.co/>

Bloque Neón: <http://bloqueneon.uniandes.edu.co/>



AnalyticsUA



AnalyticsUA



AnalyticsUA



MIAD

