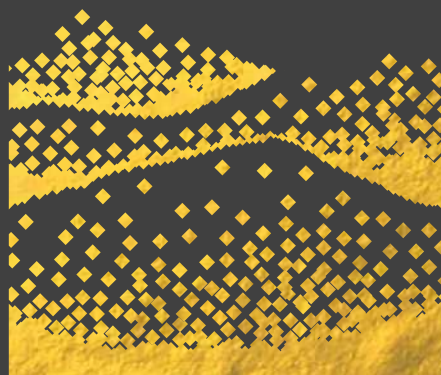


BOURBAKI

COLEGIO DE MATEMÁTICAS

Track de BBVA :
Procesamiento del Lenguaje Natural
(sintaxis y semántica)





Objetivos

A. Enseñar un manejo de Python que permita resolver problemas de negocio utilizando datos que contengan imágenes, texto, características estructuradas, series de tiempo así como audio y vídeo.

B. Familiarizar al estudiante con el lenguaje matemático utilizado en la ciencia de datos así como su utilidad para resolver problemas concretos incluyendo las mejores prácticas en la comunicación de los resultados con los usuarios finales.

C. Construir un amplio portafolio de casos de uso para que los analistas graduados sean capaces de proponer soluciones tanto innovadoras como sólidas con el estado del arte industrial y académico.

D. Evaluar a los estudiantes en las mejores metodologías y técnicas durante manejo de los datos y el modelado matemático, tanto para un análisis descriptivo, predictivo e inclusive prescriptivo.

Perfil estudiantil

Los estudiantes pueden estar dedicados a cualquiera de las siguientes áreas: Ciencia de Datos, Business Intelligence, Ingeniería de datos o Desarrollo de Software.

- **Conocimientos previos sobre Python:** no es necesario que los estudiantes estén familiarizados con este lenguaje de programación pues todos recibirán un curso propedéutico con el objetivo de hacer al curso auto-contenido.
- **Conocimientos previos sobre Matemáticas:** durante los cursos se reforzarán aquellos conceptos matemáticos que no se hayan estudiado durante los estudios formales de los estudiantes. El énfasis de los cursos es en las aplicaciones y no en los aspectos académicos.



Retos, evaluación y acompañamiento

- 1.** Cada semana se realizará un test con 3 preguntas para revisar el aprendizaje del estudiante. Las respuestas se revisarán de manera grupal y además los estudiantes recibirán feedback personalizado.
- 2.** Al finalizar cada bloque se asignará un reto con un caso de uso semejante al implementado, los estudiantes trabajarán en equipos de tres para resolverlo y al finalizar se presentarán sus hallazgos y retroalimentación a todo el grupo.
- 3.** Tanto los profesores como los ayudantes acompañarán a los estudiantes a lo largo del curso y estarán comprometidos en resolver dudas por medio de un canal de comunicación directa. También habrán sesiones de trabajo en grupos de tres para ayudarles a los estudiantes a resolver el reto planteado.



Track BBVA: Ciencia de Datos & IA



Rudimentos de Machine Learning

(overfitting & underfitting)

Machine Learning. Regresiones lineales y árboles. Clasificaciones Multi-clase. Clusterización y reducción de la dimensión. Percepción simple y modelos lineales. Árboles en regresión. Boosting y random Forest. K-means. t-SNE.



Procesamiento de Lenguaje Natural

(semántica y sintaxis)

Clasificación de Textos. Topic Modeling. Encajes y Redes Recurrentes. Grandes Modelos del Lenguaje. Bolsa de palabras y Tf-idf. Clustering v.s. Topic Modeling. Redes Recurrentes. BERT v.s. CHATGPT. RAG. Destilación de modelos.



Profundización de Machine Learning

(variables latentes y secuenciales)

Análisis de Supervivencia. Redes Neuronales. Sistemas de Recomendación. Datos Temporales. Aprendizaje por Refuerzo. Random Survival Forest. Missing valúes. Collaborative filtering. Prophet y sus comparación con ARIMA y LSTM. Q-Learning.

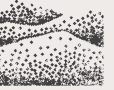


Machine Learning & AI

(Tensores y simulación)

Redes convolucionales. Redes Recurrentes. Datos Multi-Modales. Aprendizaje por refuerzo profundo. Graph Neural Networks. Modelos Seq2seq. Encajes de texto: Doc2vec. Modelos Transformer.





Track BBVA: Data Science & AI

I. Rudimentos de ML: underfitting v.s. overfitting. (Principiante)

II. Profundización en ML: variables latentes & secuenciales. (Intermedio)

III. Procesamiento del Lenguaje Natural: sintaxis y semántica. (Intermedio)

IV. Machine Learning & AI: tensores y simulación. (Avanzado)



Estructura

BBVA tendrá acceso a las grabaciones Estructura por módulo (11 horas).

Semanas de aprendizaje

A. Planteamiento del caso de uso y exploración de los datos.

B. Procesamiento, modelos matemáticos, algoritmo y métricas.

C. Implementación del caso de uso en Python.

Semanas de Reto

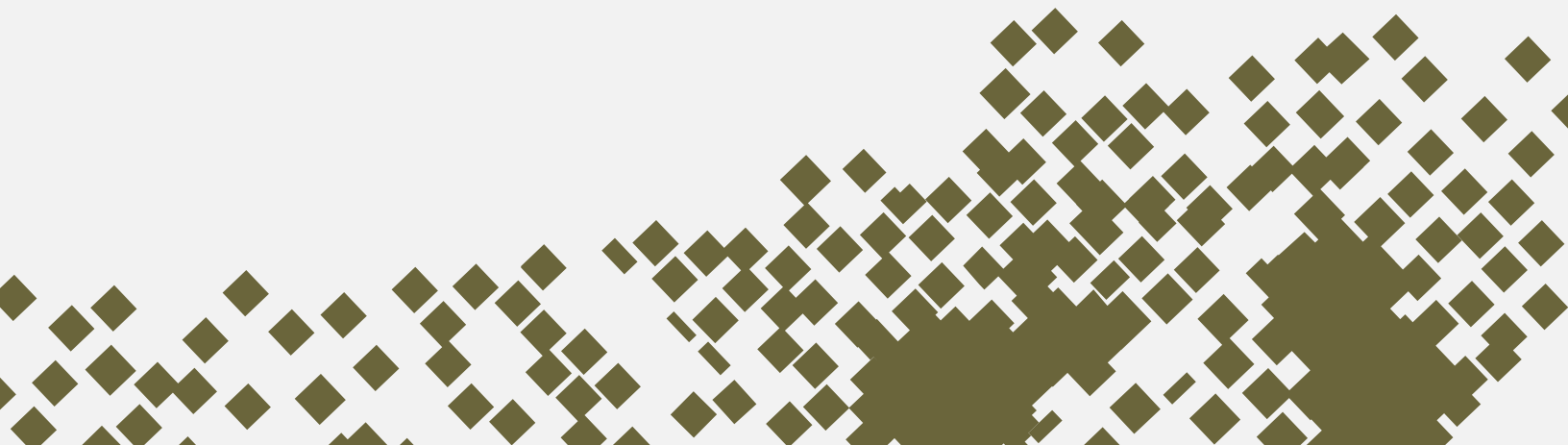
A. Trabajo en el reto.

B. Dudas & Knowledge Check.

C. Presentación del reto en equipos



Procesamiento del lenguaje natural (intermedio): **sintaxis y semántica**





Temario

Módulo I. Análisis de sentimientos

1. Bolsa de palabras
2. Tf-idf & ley de *Zipf*
3. Clasificación y ranking
4. Análisis de sentimientos

Módulo III. Redes densas y embeddings

1. Introducción a las redes densas
2. Encajes de palabras y semántica
3. *Word2vec*
4. Aplicaciones de *embeddings*
5. *Doc2vec*

Módulo V. Grandes modelos del lenguaje

1. Soft fine-tuning con ChatGPT
2. Capas de Atención y *transformer*
3. BERT & GPT
4. RAG y destilación

Módulo II. Topic Modeling

1. Modelos generativos
2. Medidas de similitud y maldición de la dimensión
3. Clustering v.s. Topic Modeling
4. Latent Dirichlet Allocation
5. *Exchangability*

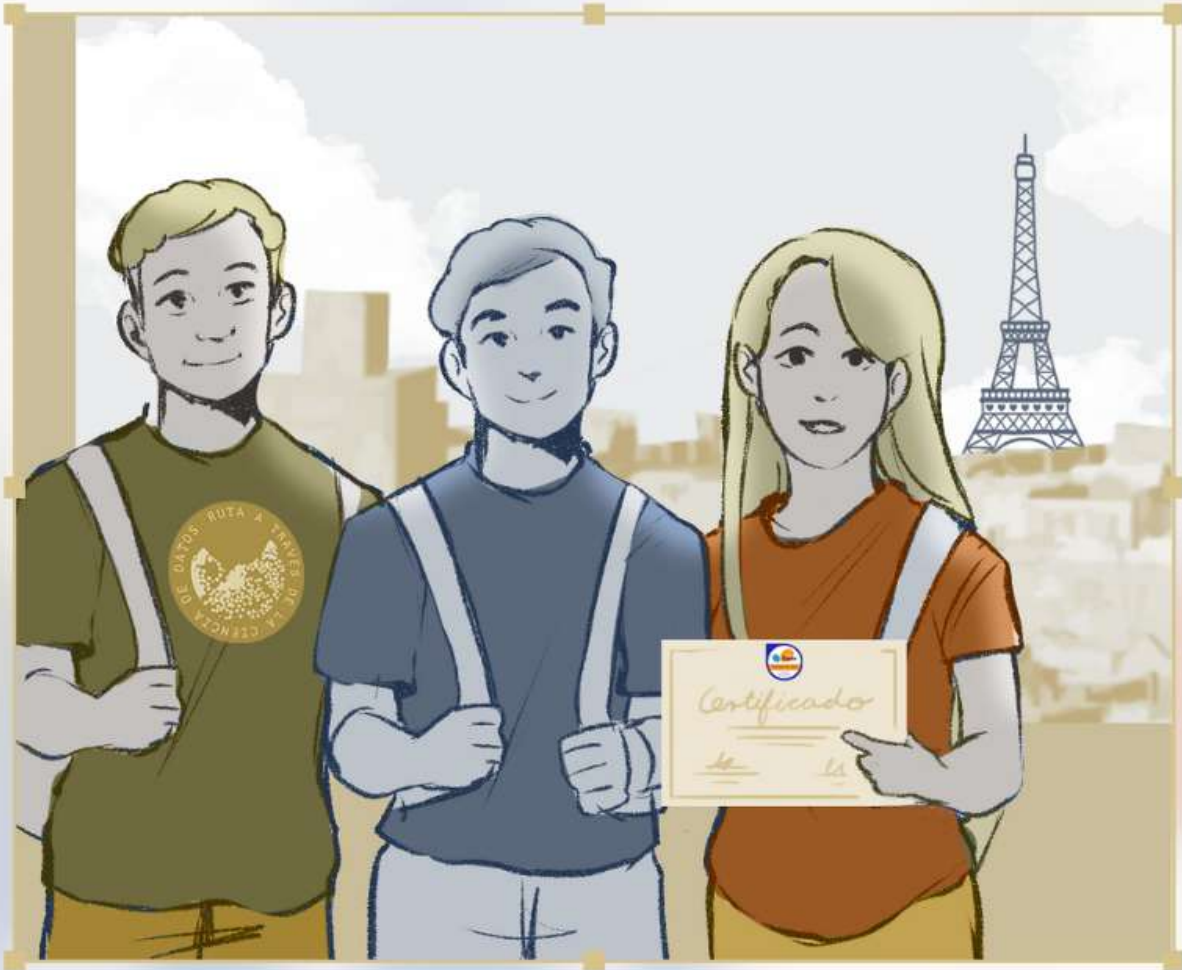
Módulo IV. Redes recurrentes y *Seq2seq*

1. Redes recurrentes y datos secuenciales
2. Redes *LSTM* y la memoria
3. Modelos *Seq2seq*
4. Resúmenes ejecutivos
5. Atención de *Bahdanau* (vainilla)



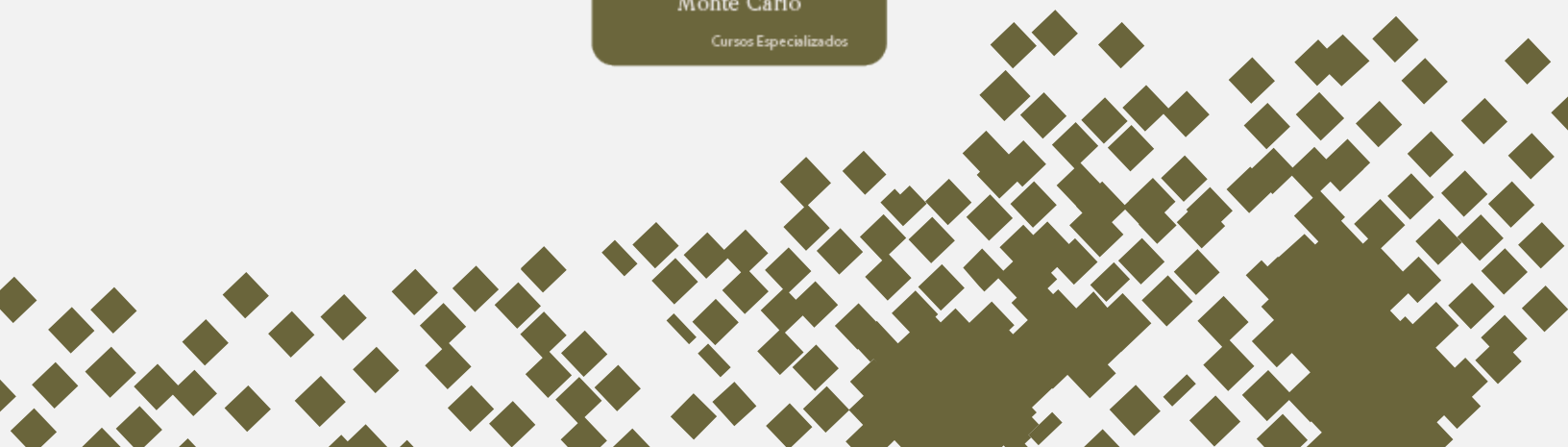
BOURBAKI
COLEGIO DE MATEMÁTICAS

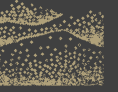
El Colegio Bourbaki ofrecerá la preparación para el examen de certificación en Scikit-Learn



Todos los inscritos al Track de Ciencia de Datos & AI, Track de Finanzas Cuantitativas & AI y el curso del Track para BBVA recibirán gratuitamente el examen de certificación así como una guía que les ayudará a estudiar.

Al finalizar el año, aquellos inscritos en alguno de estos programas que hayan hecho más puntos serán invitados a una Master Class en París del equipo de desarrolladores de Scikit-Learn.





Colegio de Matemáticas Bourbaki

Colegio de Matemáticas Bourbaki es un espacio para el aprendizaje personalizado. Aquí se imparten métodos de enseñanza para acercar a las personas al trasfondo matemático en múltiples fenómenos y procesos. Nuestro objetivo es vincular la academia con aplicaciones de la realidad imperante.

La institución tiene una responsabilidad con la sociedad: procurar siempre el rigor académico en todos sus servicios. Su objetivo es convertir el conocimiento matemático en una inversión sostenible, que genere riqueza.

La especialización es la piedra central de la comprensión de los problemas, el primer paso en la ruta hacia las respuestas que demanda la realidad. La selección de material educativo asertivo y la atención a los detalles correctos aleja a Colegio Bourbaki de los cursos express. Asumimos que con claridad, constancia y transparencia, se construye un camino intelectual sincero.



BOURBAKI

ESCUELA DE MATEMÁTICAS

CURIOSIDAD Y SABER

CDMX, MEXICO

Estaremos encantados de explicar con
detalle sobre nuestros cursos

Escríbenos por mail o WhatsApp

info@colegio-bourbaki.com

+52 56 2141 7850

[www.colegio-
bourbaki.com](http://www.colegio-bourbaki.com)

