

## Diplomado en Data Science and Machine Learning applied to Financial Markets

Coordinadores académicos: M.F. Yolanda Espinosa Félix, M.F.  
José Jorge Ramírez Olvera y Act. Gerardo Durán Martín

### Objetivo general

En el diplomado se presentarán las herramientas necesarias para el desarrollo de la ciencia de datos en un ámbito financiero. Los participantes utilizarán estas herramientas para analizar, manipular, comprender y representar datos. Además, adquirirán bases teóricas y de programación para comprender los modelos de inteligencia artificial y para aplicar el mejor modelo a un problema dado.

### ¿A quién va dirigido?

A profesionistas del sector financiero del país (bancos, casas de bolsa, sociedades de inversión, afores, casas de cambio, operadores de derivados, socios liquidadores, arrendadoras financieras, empresas de factoraje financiero, compañías de seguros y fianzas, entidades gubernamentales, entidades reguladoras, consultorías, etc.), en las áreas financieras o en las de planeación de empresas no financieras que quieran especializarse en inteligencia artificial aplicada a los mercados financieros.

Los participantes deben poseer nociones de programación y bases sólidas de matemáticas financieras, álgebra lineal, cálculo y estadística. Es indispensable traer una computadora portátil a las sesiones (con procesador i5 o equivalente).

### Módulo I

#### INTRODUCCIÓN A PYTHON CON FINANZAS

##### Objetivo

En este módulo, los participantes conocerán el lenguaje de programación Python por medio de aplicaciones en finanzas. Python es conocido por su uso en aplicaciones en ciencia de datos y aprendizaje de máquina, y su base de usuarios es de las que tiene mayor crecimiento en el mundo.

##### Temario

1. Introducción a la programación
  - 1.1. Programación y aplicaciones
  - 1.2. Lenguajes
  - 1.3. Por qué Python
2. Objetos dentro de Python
  - 2.1. Cadenas, flotantes y enteros
  - 2.2. Variables
  - 2.3. Listas y tuplas
  - 2.4. Diccionarios y conjuntos
3. Instrucciones de control

- 3.1. Booleanos
- 3.2. Condicionales
- 4. Ciclos
  - 4.1. *For*
  - 4.2. *While*
- 5. Funciones
  - 5.1. Motivación
  - 5.2. Parámetros
  - 5.3. Regreso de valores
- 6. Librería estándar
  - 6.1. *Datetime*
  - 6.2. *Collections*
  - 6.3. *Os*
- 7. Usos avanzados
  - 7.1. Funciones lambda
  - 7.2. *Map, Filter, Reduce*
  - 7.3. Listas por comprensión
- 8. Programación orientada a objetos
  - 8.1. Clases
  - 8.2. Estructura de una clase
- 9. Programación funcional
  - 9.1. Funciones lambda
  - 9.2. *List Comprehensions*
  - 9.3. Cálculo de predicados: funciones *any*, *all*
- 10. Primeras aplicaciones financieras en Python
  - 10.1. Extracción de información de Bloomberg con Python
  - 10.2. Matemáticas financieras
    - 10.2.1. Convenciones de tasas de interés
    - 10.2.2. Ecuaciones de valor
    - 10.2.3. Anualidades
    - 10.2.4. Tablas de amortización
  - 10.3. Mercado de dinero
    - 10.3.1. Valuación de bonos
    - 10.3.2. Tasas de rendimiento
    - 10.3.3. Medidas de riesgo
      - 10.3.3.1. Duración
      - 10.3.3.2. Convexidad
    - 10.3.4. Interpolación

## **Módulo 2**

### **ESTADÍSTICA Y PROBABILIDAD CON PYTHON**

#### **Objetivo**

Los participantes se familiarizarán con las principales áreas de conocimiento para fortalecer su capacidad de elaborar modelos avanzados y matemáticamente rigurosos mediante el lenguaje de programación Python.

#### **Temario**

- 1. Álgebra lineal
  - 1.1. Matrices, vectores y escalares

- 1.2. Operaciones con matrices
  - 1.2.1. Suma
  - 1.2.2. Resta
  - 1.2.3. Multiplicación
- 1.3. Determinantes e inversas
- 1.4. Eigenvalores y Eigenvectores
- 1.5. Aplicaciones
- 2. Teoría de la información
  - 2.1. Entropía
  - 2.2. Divergencia KL
  - 2.3. Información mutua
- 3. Probabilidad
  - 3.1. Reglas de la probabilidad
  - 3.2. Distribuciones discretas
    - 3.2.1. Bernoulli
    - 3.2.2. Binomial
    - 3.2.3. Multinomial
  - 3.3. Distribuciones continuas
    - 3.3.1. Normal
    - 3.3.2. Normal multivariada
    - 3.3.3. Distribución  $t$  de Student
    - 3.3.4. Gamma
    - 3.3.5. Beta
    - 3.3.6. Dirichlet
  - 3.4. Esperanza, varianza y covarianza
  - 3.5. El teorema de Bayes
- 4. Estadística inferencial
- 5. Estadística bayesiana
- 6. Aplicaciones a matemáticas de portafolios
- 7. Simulación con Python
  - 7.1. Movimiento browniano y Lemma de Itô
  - 7.2. Montecarlo
  - 7.3. Cholesky
  - 7.4. Montecarlo estructurado
  - 7.5. Poisson

## **Módulo 3**

### **CIENCIA DE DATOS**

#### **Objetivo**

Los participantes examinarán las herramientas y las ideas esenciales de la ciencia de datos como análisis previo a aplicaciones más rigurosas. Se pondrán de relieve las aplicaciones de la ciencia de datos para explorar datos financieros desde una nueva perspectiva.

#### **Temario**

- 1. Bases de datos y SQL
  - 1.1. Introducción a las bases de datos con MySQL

- 1.2. Relación entre tablas: llaves primarias y foráneas
- 1.3. Creación de una base de datos
- 1.4. Consultas en una base de datos
- 2. Python científico
  - 2.1. Numpy: arreglos matriciales
  - 2.2. Gráficas científicas con Matplotlib
  - 2.3. Herramientas científicas con Scipy
- 3. Introducción al análisis de datos
  - 3.1. DataFrames con Pandas
  - 3.2. Manipulación de información
    - 3.2.1. Desde una base de datos
    - 3.2.2. CSV
    - 3.2.3. JSON
  - 3.3. Creación de nuevas variables
  - 3.4. Limpieza y validación de información
  - 3.5. *Webscrapping*
    - 3.5.1. Lenguaje html, CSS y el DOM
    - 3.5.2. *Request, Get y Post*
    - 3.5.3. Acopio de información
    - 3.5.4. Limpieza de información
    - 3.5.5. Extracción de información de textos
- 4. Análisis de datos financieros
  - 4.1. Sentimiento de mercado
  - 4.2. Análisis de tendencias en distintos subyacentes
- 5. Análisis y construcción de portafolios
  - 5.1. Medición del desempeño
  - 5.2. Factores explicativos del desempeño
- 6. Análisis del lenguaje natural
  - 6.1. Sentimiento de mercado
    - 6.1.1. Redes sociales
    - 6.1.2. Noticias
    - 6.1.3. Expresiones regulares (regex)
    - 6.1.4. N-Grams
    - 6.1.5. Naive Bayes
- 7. Detección de oportunidades de inversión
  - 7.1. Estrategias de valor relativo
    - 7.1.1. Curvas de tasas de interés
    - 7.1.2. Volatilidad
  - 7.2. Robo-advisor
- 8. *Trading* algorítmico
  - 8.1. Costos de transacción
  - 8.2. Cobertura e impacto en liquidez
  - 8.3. Ejecución de órdenes
  - 8.4. *Trading* de pares
  - 8.5. Reversión a la media

## **Módulo 4**

### **MACHINE LEARNING**

## Objetivo

Los participantes conocerán la teoría, los conceptos y las prácticas del aprendizaje de máquina, un subconjunto del campo de la inteligencia artificial, y entenderán el uso de modelos con librerías de Python. Examinarán las aplicaciones para administración de riesgos de mercado, crédito y contraparte, y estudiarán la generación e implantación de estrategias de *trading*.

## Temario

1. Introducción al aprendizaje de máquina
  - 1.1. ¿Aprendizaje de máquina o inteligencia artificial?
  - 1.2. Nociones básicas
    - 1.2.1. Definición y motivación para métodos de aprendizaje
    - 1.2.2. Modelos supervisados
    - 1.2.3. Modelos no supervisados
  - 1.3. La regresión lineal en el aprendizaje de máquina
  - 1.4. *Overfitting* y *underfitting*
2. Selección y entrenamiento de modelos
  - 2.1. Cross-Validation
  - 2.2. Regularización
    - 2.2.1. L1
    - 2.2.2. L2
    - 2.2.3. Elastic Net
  - 2.3. El método del gradiente descendente
  - 2.4. Normalización
3. Modelos supervisados
  - 3.1. La regresión logística
    - 3.1.1. Análisis de errores
  - 3.2. Árboles de decisión
    - 3.2.1. El algoritmo CART
  - 3.3. Máquinas de soporte vectorial
    - 3.3.1. Motivación
    - 3.3.2. Lagrangianos
    - 3.3.3. *kernel* trick
    - 3.3.4. El algoritmo SMO
4. Ensemble Learning
  - 4.1. Random Forests
  - 4.2. Votos de clasificadores
  - 4.3. Bagging
  - 4.4. Boosting
  - 4.5. Multiestrategia de selección de portafolios
5. Modelos no supervisados
  - 5.1. K-Nearest Neighbors
  - 5.2. K-means
  - 5.3. Cálculo de la densidad de kernel
  - 5.4. Reducción de dimensiones
    - 5.4.1. PCA
    - 5.4.2. T-SNE
    - 5.4.3. Construcción de curvas de tasas de interés ajustadas por colateral
    - 5.4.4. Volatilidades implícitas
6. Aprendizaje de máquina para series de tiempo

- 6.1. Series de tiempo
- 6.2. Series de tiempo como un problema de aprendizaje de máquina
- 6.3. Validación de modelos de series de tiempo
- 7. Aplicaciones
  - 7.1. Detección de fraudes en un ámbito no supervisado
  - 7.2. Distribución de rendimientos positivos y negativos
  - 7.3. Perfilamiento de inversionistas
  - 7.4. VaR mediante aprendizaje de máquina
  - 7.5. Medición del riesgo de crédito y riesgo de contraparte
    - 7.5.1. Estimación de *Credit Value Adjustment* (CVA)
- 8. Optimización de portafolios
  - 8.1. Modelos dinámicos convexos
  - 8.2. Restricciones de régimen de inversión
  - 8.3. Restricciones no genéricas:
    - 8.3.1. Apalancamiento
    - 8.3.2. Liquidez

## **Módulo 5** **DEEP LEARNING**

### Objetivo

Los participantes se introducirán en la teoría y las ideas de las redes neuronales y sus usos y entenderán el poder de los modelos que se encuentran a la cabeza de importantes empresas de tecnología.

### Temario

- 1. Tensorflow y Keras
- 2. Funciones de activación
  - 2.1. La función sigmoideal
  - 2.2. Tangente hiperbólica
  - 2.3. ReLu
- 3. Las redes *FeedForward*
  - 3.1. El algoritmo del perceptrón
  - 3.2. Redes neuronales de una capa
  - 3.3. Redes neuronales multicapa
  - 3.4. El método de backpropagation
- 4. Métodos de aprendizaje
  - 4.1. Gradiente descendente
  - 4.2. Momentum
  - 4.3. Nesterov
  - 4.4. AdaGrad
  - 4.5. Adam
  - 4.6. Momentum con tasa de aprendizaje sistematizada
- 5. Regularización
  - 5.1. L1 y L2 para una red neuronal
  - 5.2. Batch Norm
  - 5.3. Dropout
- 6. Redes neuronales convolucionales
  - 6.1. Filtros
- 7. Redes neuronales recurrentes

- 7.1. El problema del gradiente divergente
- 7.2. Entrenamiento de una RNN
- 7.3. General Recurrent Unit (GRU)
- 7.4. Long Short Term Memory (LSTM)
- 8. Autoencoders
- 9. Aplicaciones
  - 9.1. Rendimientos esperados
  - 9.2. Portafolios
  - 9.3. Estimación de CVA
  - 9.4. Desarrollos experimentales: el caso de México
    - 9.4.1. Inversión mediante la aproximación de fundamentales
    - 9.4.2. LSTMs para series de tiempo

## **Coordinadores Académicos**

### **Yolanda Espinosa Félix**

Es maestra en Finanzas y licenciada en Economía por el ITAM, y cuenta con el nivel I del programa Chartered Financial Analyst (CFA). En el sector privado se ha desempeñado en casas de bolsa realizando análisis financiero y valuando compañías que cotizan en la Bolsa Mexicana de Valores (BMV), y en el sector público en la Secretaría de Hacienda, área de Crédito Público (emisión de deuda gubernamental), y como asesora del subsecretario de Hacienda. Fue responsable de plantear la estrategia financiera del Fideicomiso de Apoyo al Rescate de Autopistas Concesionadas y participó también en proyectos de innovación e ingeniería financiera. En CitiBanamex/Accival trabajó en el área de Análisis de Inversiones, mientras que en el área de Asset Management ocupó la Gerencia de Administración de Portafolios. Desde 2003 es catedrática de licenciatura, maestría y diplomados en el ITAM, donde imparte cursos de inversiones, instrumentos financieros, finanzas corporativas y administración de riesgos.

### **José Jorge Ramírez Olvera**

Cuenta con más de 15 años de experiencia en el mercado financiero y de labor docente dentro del ITAM a nivel licenciatura, maestría y diplomados (coordinador de los diplomados de Derivados Financieros y Productos Financieros Estructurados); encargado de las mesas de estructuración en distintas instituciones financieras, entre las que destacan: ING Bank, Bank of America – Merrill Lynch y Cantor Structuring (filial de Cantor Fitzgerald). Su especialidad se centra en la cobertura de derivados exóticos, desarrollo de metodologías de valuación de instrumentos financieros y programas de emisión internacionales y locales. Principal contacto con clientes (institucionales y corporativos), intermediarios financieros y autoridades. José Jorge es licenciado en Actuaría, licenciado en Matemáticas Aplicadas y maestro en Finanzas por el ITAM.

## **Gerardo Durán Martín**

Es egresado de la carrera de Actuaría en la Universidad Marista y tiene una especialización en Ciencia de Datos, además de diversas certificaciones en temas de estadística, aprendizaje profundo y aprendizaje de máquina. Actualmente trabaja en Analysic Nabla en el desarrollo de soluciones en el área de la ciencia de datos, donde ha creado modelos de optimización de portafolios, análisis del sentimiento de mercado y pronóstico de mercado. Antes, fue gerente de riesgos en Banco Santander. Ha impartido cursos de aprendizaje de máquina y programación a nivel licenciatura en la Universidad Marista y ha dictado cursos de introducción al aprendizaje de máquina con Python en DataVision y HSBC.