

Curso extracurricular: Riesgo de Crédito con Aplicación de *Big Data* y *Data Science*

Profesor: Santiago Marcos

Es Licenciado en Economía por la Universidad de Buenos Aires (UBA). Se encuentra realizando una tesis para la Maestría en Generación y Análisis de información Estadística de la UNTREF y ha completado la Diplomatura en Deep Learning en la Universidad Tecnológica Nacional (UTN). Actualmente se desempeña como docente en la UNTREF en cursos de postgrado sobre *Big Data*. Durante los últimos quince años ha trabajado en empresas de software, consumo masivo, consultoría y servicios financieros como científico de datos y desarrollador de modelos estadísticos.

Presentación y objetivo del curso

La ciencia de datos y la estadística ofrecen herramientas que se pueden poner al servicio del desarrollo del conocimiento en distintas disciplinas. Las entidades financieras disponen de forma directa e indirecta de una gran cantidad de datos de sus clientes reales y potenciales que les permiten realizar inferencias de precisión sobre su comportamiento futuro.

El propósito del presente curso se centra en formar profesionales en conocimientos de riesgo crediticio y en la explotación de datos que permitan construir inferencias de comportamiento sobre los clientes. El curso comienza con nociones iniciales y, desde allí, avanza a cuestiones más detalladas. Durante las clases prácticas se utilizará Python. Para el trabajo práctico los alumnos pueden utilizar Python u otro software que sea de su preferencia.

Destinatarios

Estudiantes de grado, posgrado, técnicos, profesionales e investigadores que deseen adquirir conocimientos teóricos y prácticos sobre el riesgo de crédito modelizado para el desarrollo de su carrera vinculado con entidades financieras.

No se requieren conocimientos previos de programación. Resulta deseable una noción general de estadística.

Contenidos

Teoría – Conceptos Se desarrollarán conceptos sobre el funcionamiento del riesgo de crédito minorista y los usos de la estadística aplicada a éste.	Práctica – Python Se trabajará con bases de datos de fuentes públicas (kaggle) con el objetivo de desarrollar un modelo predictivo de riesgo de impago. (default)	
<ul style="list-style-type: none"> · Introducción · Funcionamiento de una entidad financiera minorista · Principales productos. Cerrados y Revolventes. Con y sin Garantía · El rol del riesgo de crédito. · Apetito de riesgo 	<ul style="list-style-type: none"> · Importación de datos de fuentes externas · Manipulación de dataframes 	Clase I
<ul style="list-style-type: none"> · Originación · Bureau de crédito · Central de deudores · Información al momento del alta · Políticas de originación 	<ul style="list-style-type: none"> · Agregación de Datos · Filtros · Manipulación de campos · Cruces de tablas (merge) · Derivación de nuevas variables · Variables Dummy 	Clase II
<ul style="list-style-type: none"> · Comportamiento · Ampliación de productos · Modificaciones de líneas · Reportes de monitoreo de carteras. Vintage – Corte transversal · Reportes de monitoreo de Scores crediticios. Performance Estabilidad. 	<ul style="list-style-type: none"> · Imputación de Variables · Armado de muestras de desarrollo y validación · Estimación con algoritmo de Regresión Logística · Estimación con algoritmo de Regresión Lineal 	Clase III
<ul style="list-style-type: none"> · Morosidad y cobranzas · Estrategias de cobranza · Productos de remediación 	<ul style="list-style-type: none"> · Estimación con algoritmo de Random Forest (ensamble) · Estimación con algoritmo de lightgbm · Métricas de bondad de ajuste 	Clase IV

<ul style="list-style-type: none"> · Estados de cartera (dentro y fuera de balance) · Valuación de carteras 	<ul style="list-style-type: none"> · Estimación con algoritmo de Random Forest (ensamble) (cont.) · Estimación con algoritmo de lightgbm (cont.) 	Clase V
<ul style="list-style-type: none"> · Repaso conceptual 	<ul style="list-style-type: none"> · Puesta en común de los trabajos finales 	Clase VI

Bibliografía principal

- o Material de la cátedra.
- o A Whirlwind Tour of Python – Jake Vander Plas
<https://jakevdp.github.io/WhirlwindTourOfPython/index.html>
- o Python Data Science Handbook – Jake Vander Plas
<https://jakevdp.github.io/PythonDataScienceHandbook/index.html>
- o Hechos y Estadísticas - M. J. Moroney – EUDEBA
- o Data Science - John D. Kelleher & Brendan Tierney - MIT PRESS ESSENTIAL o <https://www.kaggle.com/>
- o Libro vivo de Ciencia de Datos - Pablo Casas
<https://librovivodecienciadedatos.ai/>

Requisitos para la cursada y aprobación

- Cumplir con el 80% de asistencia a los encuentros sincrónicos. La cursada se organiza en un temario secuencial. El mismo se desarrollará durante 6 clases sincrónicas de dos horas y media de duración.
- Exponer y entregar un trabajo práctico final utilizando Python o el software que prefieran. Los alumnos tendrán un espacio para trabajar en grupos (o de forma individual si así lo prefieren). Se dará seguimiento al progreso que vayan realizando por parte del docente.

Organización del curso

Modalidad: online sincrónica y asincrónica

Días y horario: encuentros virtuales los días miércoles de 18.30 a 21.00 horas

Fecha de inicio: 02/08/2023

Fecha de finalización: 06/09/23

Cantidad de clases: 6

Total de horas: 24