



ITESO, Universidad
Jesuita de Guadalajara

Ciencia de Datos e
Inteligencia de Negocios

CDIN

Profesor: Dr. Gaddiel Desirena López

gdesirena@iteso.mx



Presentación

- La ciencia de datos es un campo interdisciplinario que involucra a los procesos y sistemas para extraer conocimiento o un mejor entendimiento de grandes volúmenes de datos en sus diferentes formas y formatos.
- Por otra parte, la inteligencia empresarial, inteligencia de negocios o BI (del inglés *business intelligence*), es el conjunto de estrategias, aplicaciones, datos, productos, tecnologías y arquitectura técnica, los cuales están enfocados a la administración y creación de conocimiento a través del análisis de los datos existentes en una organización o empresa.



Propósito General

- Analizar bases de datos mediante el uso de algoritmos de aprendizaje máquina para la toma de decisiones en fenómenos financieros.



A background image showing a person's hands typing on a laptop keyboard. The person is wearing a plaid shirt. The laptop screen displays a blurred image of a person. The overall scene is dimly lit, with a warm, orange-toned light source visible in the upper left corner, creating a soft glow over the scene.

Propósitos Transversales

- Colaborar con los compañeros de clase.
- Generar el hábito de resolver problemas con herramientas computacionales.
- Usar de manera efectiva y eficaz las tecnologías de información para representar e interpretar los conceptos en diferentes formas: numérica, geométrica, algebraica y algorítmica.
- Implementar en Python los conocimientos adquiridos en este curso.

Actitudes y Valores

Se espera que al cursar esta asignatura se desarrolle responsabilidad ante la actividad académica, manifiesta en al menos los siguientes aspectos:

- Participación activa, con compromiso, perseverancia y actitud positiva.
- El cumplimiento de las normas de disciplina establecidas.
- El cumplimiento en tiempo y forma de las actividades que se encomienden como trabajo independiente.
- El desarrollo de espíritu crítico y autocrítico (constructivo) en el análisis del desempeño propio y de los compañeros.



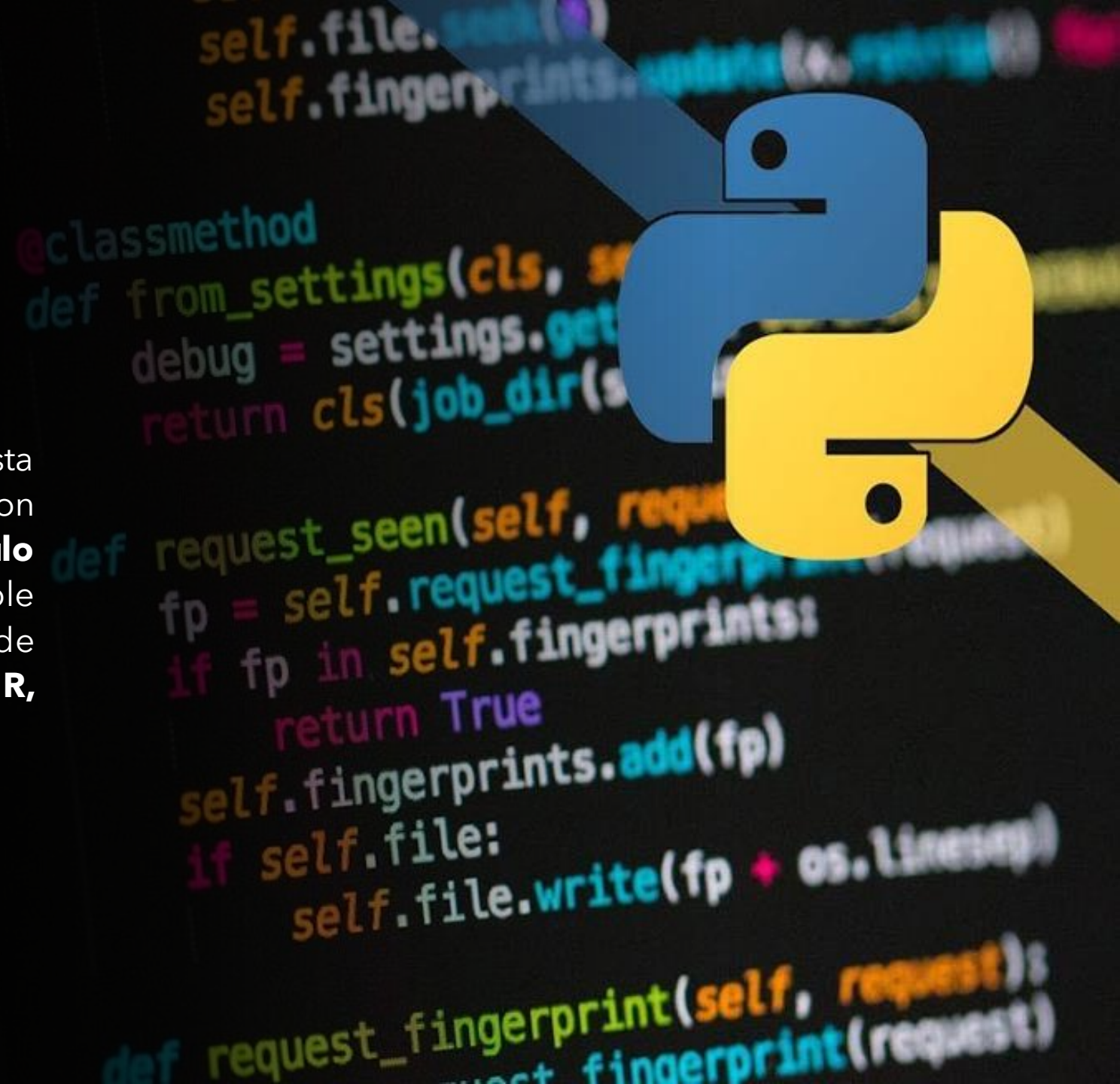
Actitudes y Valores

- El sentido de la ética, evitando, en particular, cometer actos deshonestos en la realización de las actividades evaluativas.
- El desarrollo de la capacidad para identificar características personales al afrontar procesos de aprendizaje y, como consecuencia, para aprender con mayor independencia.
- Diálogo abierto, directo y respetuoso tanto con el profesor como con los compañeros.
- Tolerancia y respeto



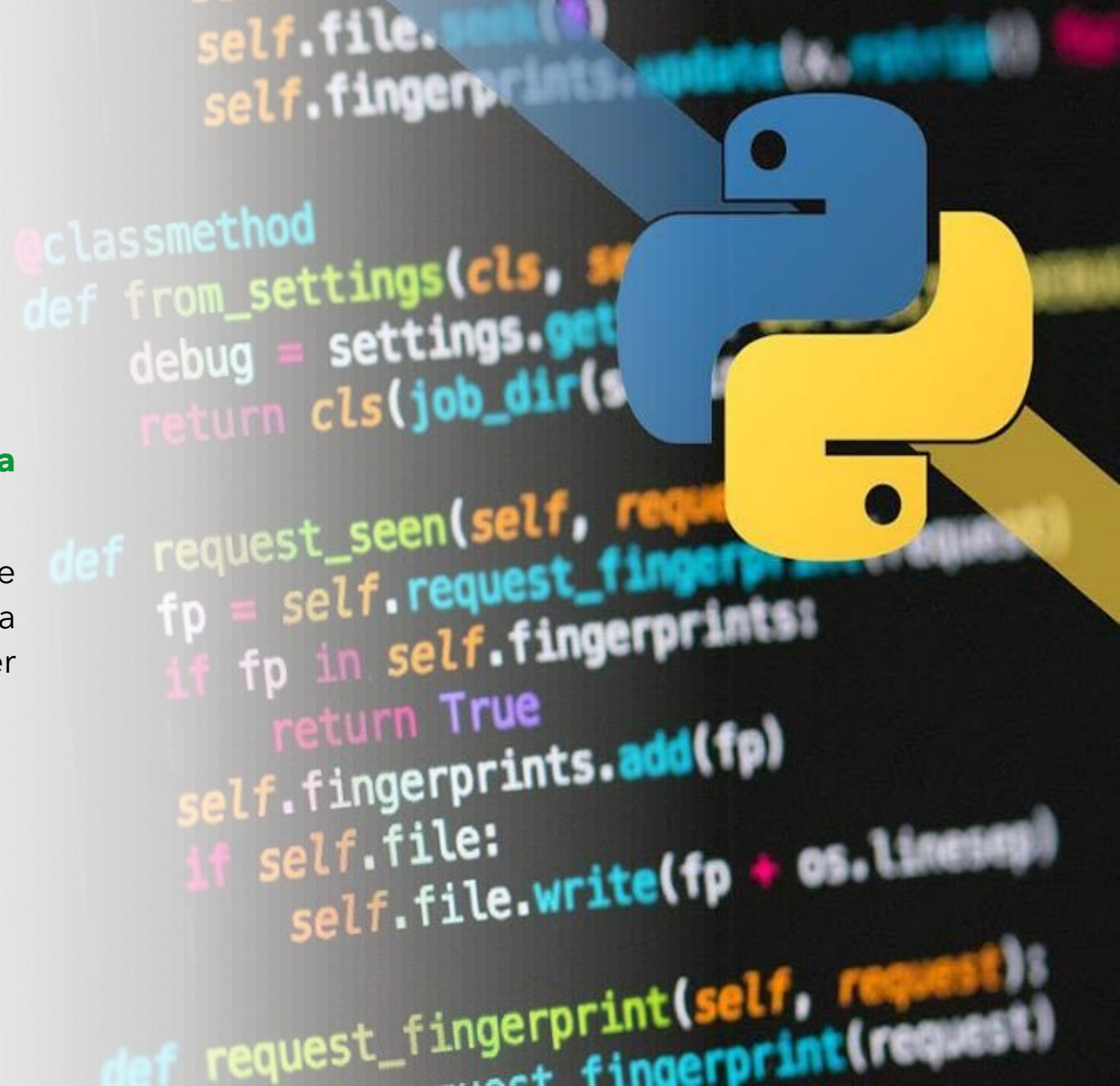
Recursos

- Para tener un desempeño satisfactorio de esta asignatura es indispensable contar con conocimientos de **programación, cálculo diferencial y estadística**. Es deseable manejar algún lenguaje científico de programación, por ejemplo, **MatLab, R, Python**.



Recursos

- **Tener instalado Suite de Anaconda**
(<https://www.anaconda.com/>)
- Anaconda viene con una suite de herramientas gráficas llamada "Anaconda Navigator". Trabajaremos con Spyder y Jupyter Notebook.



The background of the slide features a dark, blurred image of financial tools: a calculator with buttons like '+' and '-' visible on the left, a stack of coins on the right, and a line graph with data points in the center. A small orange horizontal bar is located in the top left corner.

¿Qué se espera al terminar el curso?

- Al finalizar con éxito la asignatura se tendrán las herramientas básicas para la limpieza, visualización de datos provenientes de diversas fuentes de información y el uso de métodos de clasificación para la obtención de conocimiento a partir de datos provenientes de fenómenos financieros como portafolios de inversión y series de tiempo de precios de activos financieros.

Programa del Curso

3 Módulos



1.- Minería de Datos: Datos y su clasificación

Descripción estadística básica de los datos
Visualización de datos
Limpieza de datos
Medición de similitud



2.- Preprocesamiento de datos

Integración de datos
Reducción de datos
Búsqueda de patrones (Clustering)



3.- Aprendizaje Máquina

Análisis de discriminante lineal (LDA)
Regresión Logística
Máquina de vector soporte (SVM)
Introducción a Redes Neuronales

Evaluación del Aprendizaje



Productos

% de Calificación

Proyecto Final

40

3 exámenes parciales

30

Tareas

30

TOTAL:

100%



Políticas y Lineamientos del curso



Clases en línea (Webex o Microsoft Teams).



Evaluación por módulo: examen teórico y práctico.



Se realizarán prácticas durante las sesiones de clase



Todas las prácticas serán realizadas en el software Python (Anaconda).









En sesión de clase no se permite hacer uso de celulares.



El pase de lista se hará en cada clase al inicio de la sesión y la cámara web prendida.



Políticas y Lineamientos del curso

-  Entrega de tareas en CANVAS.
-  No se recibirán tareas tiempo después de la fecha límite.
-  Micrófonos apagados excepto si se les avisa lo contrario.
-  La calificación de los exámenes parciales debe ser aprobatorio para tener derecho a calificación final.
-  La calificación aprobatoria es mayor o igual a 6.0 sin aplicar criterios de redondeo.
-  Se permite como máximo el 20% de inasistencias al curso (6 inasistencias), para mantener el derecho a calificación.

Referencias Bibliográficas

- Data Science for Business: What You Need to Know about Data Mining and Data-Analytic, Provost, F. y Fawcett, T., O'Reilly Media, 2013, ISBN: 978-1-449-37428-0 978-1-449-37429-7
- Foundations of Statistical Algorithms: with Reference to R Packages, Weihs, C. Mersmann, O. y Ligges, U., CRC Press, 2014, ISBN: 978-1-4398-7885-9.
- An Introduction to Machine Learning, Kubat, M., Springer, 2015, ISBN: 978-3-319-20009-5.
- Predictive Analytics with Microsoft Azure Machine Learning; Roger Barga, Valentine Fontama, Wee Hyong Tok, 2015, ISBN: 978-1-4842-1201-1
- Python Business Intelligence Cookbook; R. Dempsey, 2015, ISBN: 978-1-78528-966-8.