







MACHINE LEARNING AND DATA SCIENCE

PRESENTACIÓN GENERAL
PROGRAMA DE FORMACIÓN

GENERALIDADES

- La ciencia de datos (o *data science* - DS en Inglés) aborda el análisis y explotación del creciente volumen de datos.
- Los métodos de aprendizaje computacional (ML) permiten extraer información y conocimiento, en forma de patrones, modelos predictivos y descriptivos, de forma automática.
- Estas tecnologías son parte integral de la 4ta revolución industrial, la cual se refiere al cambio dramático que está experimentando la sociedad y la industria gracias a la adopción de diferentes tecnologías.

| | |
|--|--|
|  | Inicio Desde Abril 09 de 2019 <small>Consultar individualmente más adelante.</small> |
|  | Martes a Jueves 6:00pm a 9:00pm |
|  | 39 horas presenciales por módulo |
|  | Universidad Nacional de Colombia Sede Bogotá |

PROGRAMA FORMATIVO MLDS

MLDS-1 Introducción al Análisis y visualización de datos con Python – **Abr 9 y May 16**

MLDS-2 Introducción Aprendizaje Computacional Machine learning – **May 21 y Jun 20**

MLDS-3 Procesamiento de grandes volúmenes de info (Big data) – **Jun 25 y Jul 23**

MLDS-4 Desarrollo de Sistemas Analíticos en la nube – **2019-2**

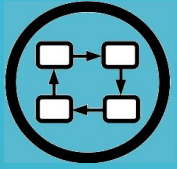
MLDS-5 Procesamiento y entendimiento de lenguaje natural – **2019-2**

MLDS-6 Deep learning – **2019-2**



OBJETIVO PRINCIPAL

- ✓ Formar a los asistentes en las competencias y habilidades técnicas que les permitan aplicar herramientas conceptuales y tecnológicas del análisis de datos y el aprendizaje computacional a la solución de problemas reales.



METODOLOGÍA

- a) Exposición magistral
- b) Talleres de trabajo en equipo
- c) Sustentación y argumentación de proyectos

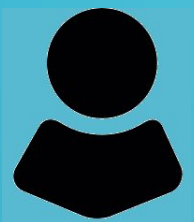
30%: Presentación de temas

70%: Talleres guiados



CERTIFICACIÓN

- ✓ Este curso ofrece certificación expedida por la Facultad de Ingeniería de la Universidad Nacional de Colombia Sede Bogotá si el estudiante cursa con éxito el 80% o más del tiempo efectivo del curso/diplomado/programa.
- ✓ También es necesario aprobar con una nota mínima los exámenes o talleres en el desarrollo de la metodología.
- ✓ Los usuarios que los cursen (tres o más módulos dentro del semestre en el que son ofertados) obtendrá el Certificado Diplomado, adicional al certificado individual de cada módulo.
- ✓ El programa de formación contará con nuevos módulos que se irán abriendo y ofertando al público.

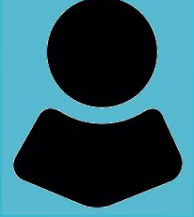


PERFIL DEL ASPIRANTE

El programa de formación en MLDS está dirigido a todas las personas que desde sus roles tienen responsabilidad o incidencia en los procesos de análisis de datos y desarrollo de sistemas inteligentes y analíticos: gerentes de informática, analistas, desarrolladores y responsables de área.

Igualmente, el programa de formación en MLDS también está dirigido a aquellos profesionales en áreas tales como la ingeniería, ciencias naturales, economía, matemáticas y estadística, interesados en diversas aplicaciones del aprendizaje computacional y el desarrollo de sistemas inteligentes y analíticos.

Dado el enfoque práctico del programa de formación en MLDS se requiere que los estudiantes tengan familiaridad previa con conceptos de programación.



PRESENTACIONES

Fabio A. González, PhD

INGENIERO DE SISTEMAS Y COMPUTACIÓN

- Ingeniero de Sistemas, Universidad Nacional de Colombia (1993); MSc. en Matemáticas, Universidad Nacional de Colombia [1998]; Ph.D. & M.Sc. in Computer Science, The University of Memphis (2003); Profesor visitante University of Louisville, USA (2009); Profesor visitante Ecole Nationale Supérieure de Techniques Avancées (ENSTAParisTech), Francia (2010); Profesor visitante University of Houston, USA (2017).
- Profesor Titular del Departamento de Ingeniería de Sistemas e Industrial. Coordinador del programa de Doctorado en Ingeniería de Sistemas y Computación, Universidad Nacional de Colombia, Bogotá. Ampla experiencia en investigación en aprendizaje computacional, minería de datos, recuperación de información, procesamiento de lenguaje natural y visión por computador. Participación en diversos proyectos de investigación internacionales y nacionales. Más de 100 publicaciones científicas en revistas, libros y conferencias internacionales. Lidera el grupo de investigación MindLab (Machine Learning Perception and Discovery Lab, <http://mindlaboratory.org>)

MÓDULO MLDS-01

INTRODUCCIÓN AL ANÁLISIS Y VISUALIZACIÓN DE DATOS CON PYTHON

El objetivo de este módulo es entrenar al estudiante en el uso del lenguaje de programación Python como herramienta para el análisis y visualización de datos. El módulo abordará los elementos básicos del lenguaje, los ambientes de programación que soportan su uso interactivo y las bibliotecas especializadas para el análisis y visualización de datos (**39 horas - Teórico/Práctico**).

CONTENIDO MLDS-01 - I

• **Semana 1:**

- Introducción
 - Qué es el análisis de datos?
 - Configuración entorno de trabajo: Colaboratory - Jupyter notebook environment
- Introducción a Python
- Librería numérica de Python: NumPy

• **Semana 2**

- Análisis de datos con Pandas
- Visualización de datos con Pandas

CONTENIDO MLDS-01 - IV

• **Semana 3:**

- Principios de visualización de información
- Visualización de datos con Matplotlib
- Visualización de datos con Seaborn

• **Semana 4:**

- Visualización interactiva de datos con Plotly y Cufflinks
- Visualizaciones interactivas web con Bokeh
- Visualización de datos complejos: datos geográficos: mapas cloropléticos

• **Semana 5:**

- Presentación proyecto final del módulo y cierre (3 horas)

BIBLIOGRAFÍA MLDS-01

1. Wes McKinney. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media, Inc., 2012.
2. Tamara Munzner. Visualization Analysis and Design. A K Peters Visualization Series, CRC Press, 2014.
3. Igor Milovanovic. Python Data Visualization Cookbook. Packt Publishing Ltd, 2013.
4. UCLA Statistical Computing Resources. Disponible en:
<http://wiki.stat.ucla.edu/socr/index.php/EBook>