Online / DiHiC Perú

Programa de Especialización Analítica

ADVANCED DATA SCIENCE

BROCHURE 2023

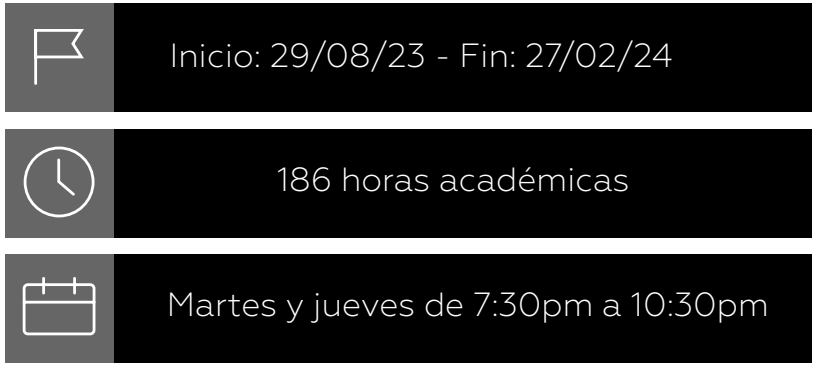
#AprendeDesdeCasa
#AprendeConLosPioneros

PRESENTACIÓN

Actualmente son varios los factores que vienen produciendo cambios en la búsqueda de talento analítico por parte de las empresas, siendo los principales la abundancia de datos, el incremento en la competitividad profesional y el surgimiento de una nueva gama de herramientas de recolección, organización y análisis de los datos. En esa línea, se apuesta por la ciencia de datos, el cloud computing y el big data como disciplinas que pueden dar solución a problemas de negocios en diferentes sectores: banca, retail, consumo masivo, etc.

Por ello, presentamos el **PEA Advanced Data Science**, el cual te permitirá abordar y dar solución a problemáticas de alto grado de complejidad, a partir del uso de técnicas avanzadas Machine Learning, Deep Learning, Big Data, entre otros.

Pre-requisitos:
• Conocimiento en programación con Python y base de datos estructuradas & no estructuradas. • Contar como mínimo con 2 años de experiencia en áreas de Business Analytics, Inteligencia de Negocios, CRM, Big Data o afines.



OBJETIVOS

OBJETIVO GENERAL

Obtener un alto dominio en Ciencia de Datos orientado a resolver problemáticas de mediano y alto grado de dificultad para diversas industrias, tanto on-premise como en entornos cloud. Además, tendrás la capacidad de aportar, desde la perspectiva analítica, a los objetivos y resultados del negocio.

OBJETIVOS ESPECÍFICOS

- Emplear las principales herramientas estadísticas y tecnológicas de vanguardia aplicadas a la ciencia de datos.
- Aplicar métodos supervisados, no supervisados y series de tiempo según casos de estudio planteados.
- Emplear y construir algoritmos aplicados al procesamiento de lenguaje natural (NLP) y aprendizaje profundo (DL).
- Realizar la puesta en producción de productos de machine learning entornos on-premise y on-cloud, así como conocer y aplicar otras técnicas en tendencia.

DIRIGIDO A

Profesionales que se desarrollen y tengan experiencia comprobada en áreas de Business Intelligence, CRM, Tecnologías de Información, Analytics y áreas relacionadas al análisis de datos.

COMPETENCIAS

- Resolver un problema de negocio utilizando las herramientas y técnicas avanzadas de la ciencia de datos.
- Utilizar las principales herramientas estadísticas y tecnológicas de vanguardia aplicadas a la ciencia de datos.
- Identificar el método más adecuado (supervisados, no supervisados y series de tiempo) según problemas de negocio.
- Construir algoritmos aplicados al procesamiento de lenguaje natural (NLP) y aprendizaje profundo (DL).
- Poner en producción de productos de machine learning entornos on-premise y on-cloud, así como conocer y aplicar otras técnicas en tendencia.

CARACTERÍSTICAS

Clases en Vivo

El 100% de las clases que se desarrollan en el programa son en vivo.

Asesoría Académica

Resuelve tus dudas con el asistente académico en línea.

Plataforma E-Learning

Accede en cualquier momento a materiales complementarios: lecturas, videos, tutoriales, clases grabadas y más.

Aprende haciendo

Desarrolla casos con datos reales, incluso puedes proponer casos de tu propio sector.

Proyecto integrador

Pondrás a prueba tus ideas para convertirlas en soluciones analíticas.

Soporte técnico

Asistencia técnica permanente y acceso a máquinas virtuales de ser necesario.

MALLA CURRICULAR

HERRAMIENTAS ANALÍTICAS

- Statistics for Data Science
  (16 horas académicas)
- Web Scraping & API's
  (12 horas académicas)
- Proyectos aplicados de Data Science
  (4 horas académicas)

ADVANCED MACHINE LEARNING

- Métodos Supervisados
  (24 horas académicas)
- Métodos No supervisados
  (12 horas académicas)
- Time Series Forecasting
  (12 horas académicas)

NLP & DEEP LEARNING

- Natural Language Processing
  (12 horas académicas)
- Deep Learning
  (24 horas académicas)

DATA SCIENCE TRENDS

- Big Data Analytics
  (16 horas académicas)
- Machine Learning Engineering
  (20 horas académicas)
- MLOps
  (5 horas académicas)
- Grafos
  (5 horas académicas)

HERRAMIENTAS ANALÍTICAS

1

Statistics for Data Science

Fundamentos Estadísticos necesarios para la Ciencia de Datos y Machine Learning.

- Estadística descriptiva.
- Estadística inferencial.
- Pruebas de Hipótesis.
- Data Wrangling.
- Feature engineering.
- Casos de uso.

2

Web Scraping & API's

Identificar la importancia de los datos externos para nuestra empresa y generar soluciones eficientes.

- Web Crawling y Web Scraping.
- Consecuencias del escraper. Bloqueo de páginas web.
- Métodos de extracción de datos. XPATH & CSS Selector.
- Web-scraping estático – LXML.
- Web-scraping dinámico. Librería Selenium.
- Inicialización del driver con Python.
- Extracción de datos de portales de datos abiertos.
- API de redes sociales (Twitter, Spotify, Youtube, etc).

3

Proyectos aplicados de Data Science

Conocer casos de uso de diversas industrias y emplearlos como referencia para sus proyectos.

- Presentación de casos de uso de diversas industrias.

  -Sector Financiero.
  -Sector Retail.
  -Sector Seguros.
  -Sector Telecomunicaciones.
  -Sector Educación
  -Sector Gobierno.

ADVANCED MACHINE LEARNING

4

Métodos Supervisados

Aplicar técnicas de análisis predictivo avanzado a través de la construcción de modelos.

- Partición y balanceo de datos.
- Métricas de evaluación.
- Validación de modelos.
- Modelos de regresión.
- Modelos de clasificación.
- Ajuste de hiper parámetros.

5

Métodos No supervisados

Técnicas de agrupamiento automático (clustering) para segmentar y encontrar insights para el negocio.

- Clustering jerárquico.
- Clustering k-means.
- Clustering PAM.
- Clustering CLARA.
- Clustering DBSCAN.
- Clustering RFM.
- Análisis de componentes principales.
- Análisis de componentes múltiples.
- Reglas de Asociación.

6

Time Series Forecasting

Algoritmos orientados a series de tiempo y desarrollar capacidades de análisis y procesos de forecasting.

- Características de una serie de tiempo.
- Proceso estacionario. Raíces unitarias
- Métodos de pre procesamiento para series temporales.
- Técnicas de detección de patrones.
- Métodos de suavización exponencial.
- Pronósticos con dominio en el tiempo. Modelo Box – Jenkins.
- Modelos: AR, MA, ARIMA, SARIMA.
- Métodos multivariados para análisis temporal.

DATA SCIENCE TRENDS

aws

7

Big Data Analytics

Big Data y sus principales servicios de cloud para realizar actividades con grandes volúmenes de datos.

- Introducción al Big Data y Cloud Computing.
- Ingesta de datos (tiempo real, near real time, migración BD).
- Almacenamiento de datos (creación de buckets, tablas, inserción, ingesta, creación BD).
- Procesamiento de datos.
- Análisis de datos.
- Visualización de datos.
- Componentes de seguridad en entornos Cloud Computing.

8

Machine Learning Engineering

Comprender y aplicar sistemas on-cloud (AWS) para poner en producción modelos de machine learning.

- Sistemas de Machine Learning.
- Machine Learning Pipeline: escribiendo código para producción.
- Sirviendo un modelo a través de Rest API.
- Herramientas ci/cd para data science.
- Ejecutando aplicaciones en contenedores.
- Desplegando en IaaS (Aws ECS).
- Airflow.
- Despliegue en serverless.

9

MLOps (Workshop)

Comprender los fundamentos del despliegue y su flujo del ciclo de vida de los modelos de Machine Learning.

- Importancia de MLOps en Machine Learning
- Requerimientos para poder hacer MLOps.
- Flujo de vida de un modelo en producción.
- Registro y versionamiento de los modelos.
- Casos aplicados con MLOps.

10

Grafos (Workshop)

Comprender los fundamentos de los grafos, su escalabilidad y su aplicación en proyecto IA.

- Concepto de Grafo y su importancia en la Inteligencia Artificial.
- Representación de Grafos y sus elementos.
- Potenciando la analítica con el uso de grafos.
- Escalabilidad de proyectos usando grafos.
- Casos aplicados con Grafos

NATURAL LANGUAGE PROCESSING & DEEP LEARNING

11

Natural Language Processing

Técnicas para el tratamiento de información no estructurada, como textos, comentarios de redes sociales, entre otros.

- Introducción al procesamiento de lenguaje natural.
- Minería de texto. Pasos para ejecutar un proyecto de Text Mining.
- Expresiones regulares. StopWords.
- Calculando la frecuencia de palabras. Nubes de palabras.
- Bigramas y Comparación del uso de palabras.
- Clasificación de textos. Analizando las características del texto.

12

Deep Learning

Fundamentos de las redes neuronales y sus primeras aplicaciones en el procesamiento de información no estructurada.

- Introducción al aprendizaje profundo.
- Introducción de las redes neuronales.
- Propagación forward y backward.
- Funciones de Activación.
- Funciones de pérdida.
- Frameworks Deep learning.
- Mejorando un modelo deep learning.
- Transferencia de aprendizaje.
- Redes convolucionales.
- Redes recurrentes.
- Modelos de secuencia avanzadas.
- Pytorch.

Agile Data Science

Revisión de metodologías ágiles que te permitan organizar el flujo y desarrollo de proyectos de Machine Learning.

Presentación de Proyecto

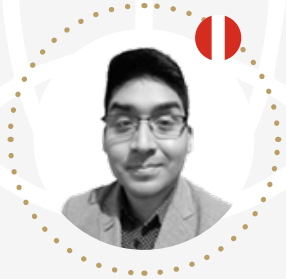
Planificación, desarrollo, modelamiento y ejecución de un proyecto integral de Ciencia de Datos que resuelva un problema de negocio mediante el uso de las técnicas analíticas aprendidas en el programa.

PLANA DOCENTE

ALEX RAYON

Vicerrector de Transformación Digital

University of Deusto



BRIAN ALARCON

Data Science Manager

Rappi

ALEJANDRO CORREA

Chief Artificial Intelligence Officer

Rappi bank

EDWIN ALDANA

Catedrático Conacyt

Cinvestav

LUIS CAJACHAHUA

Principal Data Scientist

teradata.

*En caso de contingencias podría cambiar alguno de los docentes por otro profesional de similar perfil.

ARTURO ROJAS
Product Owner
Data & Analytics
BCP

ROBERTO SAMANIEGO
Subgerente de Créditos
Digitales PYME
Interbank

ANGEL TINTAYA
Senior Data
Engineer
number8

CLAUDIA A. IPARRAGUIRRE
Managing
Director
orekan

LUIS GARAYAR
Gerente Adjunto
Data Scientist
BCP

GIANFRANCO PALOMINO
Head of Data Science
& Analytics
YANBAL

*En caso de contingencias podría cambiar alguno de los docentes por otro profesional de similar perfil.

PROYECTO INTEGRADOR

Desarrollo de un caso
end-to-end de Data Science.

Aplicar los conocimientos impartidos en el PEA en el desarrollo de un proyecto complejo de ciencia de datos.

EVALUACIÓN

+ Evaluaciones parciales: 40%
+ Trabajo final: 60%
+ Nota mínima aprobatoria: 14
+ Asistencia mínima: 80%

CERTIFICACIÓN

CERTIFICADOS PARCIALES:

+ Herramientas analíticas. 32 horas acad.
+ Advanced Machine Learning. 48 horas acad.
+ Natural Processing Language y Deep Learning. 36 horas acad.
+ Data Science Trends. 46 horas acad.

CERTIFICADO FINAL:

+ Especialista en Advanced Data Science
  186 horas académicas.

* Todos los certificados se emitirán de forma digital luego de concluido el programa.

Buscamos liderar la transformación de las empresas.

+14 Años de Experiencia

Desde el año 2009 capacitamos con técnicas de análisis de datos a profesionales de diferentes empresas y sectores.

+100 Expertos en Analítica

Nuestra plana docente ocupa los cargos más importantes en las áreas analíticas de todos los sectores.

+15K Profesionales Capacitados

Nuestros alumnos inscritos pertenecen a las mejores compañías del medio y amplifican con nosotros su red de contactos especializada.

+50 Capacitaciones Especializadas

Contamos con una variedad de líneas temáticas y niveles de especialización.

+300 Empresas Asesoradas

Las empresas top del mercado buscan nuestra asesoría y les brindamos soluciones analíticas ad hoc.

+5 Big Data Analytics Summit

Organizamos el evento más grande del Perú, con los mejores ponentes internacionales.

Formando profesionales mediante la analítica de los datos.

Perú