



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

Diplomado en Big Data

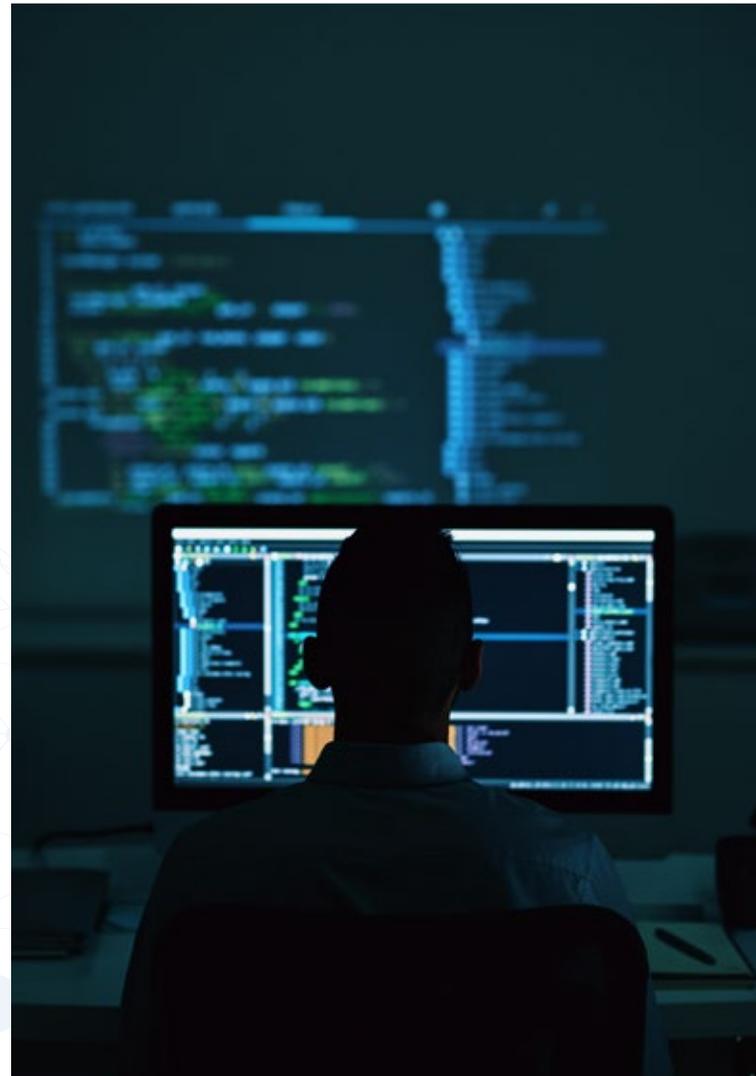
Lidera la transformación digital. Aprende a desarrollar proyectos que te permitirán hacer un uso efectivo de fuentes masivas de datos, tanto estructuradas, como no estructuradas, incluyendo textos, imágenes, voz e interacciones en redes sociales y más.

DIRIGIDO A

- » Ingenieros Civiles Industriales, Civiles Informáticos, Civiles Electrónicos, Civiles Eléctricos.
- » Licenciados en Computación, y otros profesionales licenciados y/o con experiencia en áreas afines, que requieren conocimientos o habilidades en la materia.

OBJETIVOS DE APRENDIZAJE

- » Conocer los principales componentes de la tecnología Big Data.
- » Gestionar proyectos orientados a la introducción de tecnologías en Big Data en la organización.



ESTRUCTURA Y CONTENIDOS DEL PROGRAMA

Curso:

ECOSISTEMA HADOOP

Unidad I: Fundamentos de Hadoop.

- Introducción a Hadoop.
- Hadoop sistema distribuido de archivos: HDFS (Hadoop distributed file system).
- Hadoop distribución de tareas en un cluster: Paradigma Map Reduce.
- Bases de datos no relacionales (NoSQL-Structured Query Language).
- Google Gran Tabla y Hadoop Hbase.

Unidad II: Ecosistema Hadoop.

- Introducción al ecosistema Hadoop de Cloudera.
- Herramientas del ecosistema: Pig, Hive, Impala y Hue.
- Movimiento masivo de datos: Sqoop.
- Datos secuenciales (Streaming): Flume.
- Spark.
- Librería de aprendizaje de máquina: Spark MLlib (Spark Machine learning library).

Curso:

APRENDIZAJE DE MÁQUINA PARA BIG DATA

Unidad I: Conceptos fundamentales de aprendizaje de máquina y big data

- Introducción al aprendizaje de máquina.
- Técnica de vecinos cercanos.
- Redes neuronales de capa oculta.
- Modelos de aprendizaje de máximo margen.
- Optimización y técnicas de descenso de gradiente estocástico.
- Limitaciones de técnicas tradicionales de aprendizaje de máquina para problemas de Big Data.

Unidad II: Técnicas de aprendizaje de máquina para big data

- Técnicas de hashing con colisión local.
- Modelos jerárquicos composicionales y aprendizaje estructural.
- Representaciones distribuidas.
- Redes neuronales convolucionales de aprendizaje profundo.
- Redes neuronales recurrentes.

Curso:

TÉCNICAS DE VISUALIZACIÓN Y SISTEMAS DE RECOMENDACIÓN

Unidad I: Técnicas de visualización

- Fundamentos: Abstracción de datos, marcas y canales, procesamiento humano de la información.
- Representaciones estadísticas y algoritmos.
- Reducción de dimensionalidad.
- Visualización de datos espaciales, temporales y redes sociales.
- Evaluación de sistemas de visualización.
- Herramientas de software y estudio de casos.

Unidad II: Sistemas de recomendación

- Datos usados para generar recomendaciones.
- Fuentes de información espacial, temporal y redes sociales.
- Métricas de distancia.
- Preferencias basadas en retroalimentación explícita e implícita.
- Métodos de recomendación: personalizados, colaborativos, basados en contexto o contenido, híbridos.
- Evaluación de sistemas recomendadores.
- Estudio de casos.

Curso:

COMPUTACIÓN DE ALTO RENDIMIENTO Y TEORÍA DE GRAFOS PARA BIG DATA

Unidad I: Computación de alto rendimiento para big data.

- Introducción al procesamiento paralelo de datos.
- Uso de recursos elásticos y procesamiento en la nube.
- Modelos para almacenamiento distribuido de archivos en clusters.
- Modelos para manejo distribuido de memoria en clusters.
- Modelos para ejecución distribuida de proceso en clusters.
- Procesamiento en unidades gráficas (GPUs-graphics processing unit).

Unidad II: Teoría de grafos para big data.

- Conceptos fundamentales de teoría de grafos: grado, isomorfismos, conectividad, centralidad, traversabilidad, búsqueda y ranking, métricas de distancia.
- Base de datos para grafos.
- Grafos como herramienta analítica para big data.
- Grafos para modelamiento y análisis de datos espaciales y redes sociales.
- Grafos para modelamiento y análisis de datos dinámicos.
- Grafos como herramienta de modelamiento y descubrimiento de patrones semánticos.
- Estudio de casos.

Nota: *El orden de los cursos dependerá de la programación que realice la Dirección Académica.*

JEDE DE PROGRAMA

HANS LÖBEL

Ph.D. y Magíster en Ciencias de la Ingeniería, UC. Ingeniero Civil en Ciencia de la Computación, UC. Profesor del Dpto. de Ciencia de la Computación y del Dpto. de Ingeniería de Transporte y Logística de la Escuela de Ingeniería UC.

EQUIPO DOCENTE

NEBIL KAWAS

Data Visualization Engineer @ The Not Company, Profesor del Diplomado de Ingeniería, UC.

SANTIAGO LARRAÍN

Magíster en la Ingeniería, UC. Ingeniero Civil de Industria, UC.

IVÁN LILLO

Ph.D. (c) en Ciencias de la Ingeniería e Ingeniero Civil Electricista, UC. Magíster en Ciencias de la Ingeniería, Profesor Instructor del Departamento de Ciencia de la Computación UC.

HANS LÖBEL

Ph.D. y Magíster en Ciencias de la Ingeniería, UC. Ingeniero Civil en Ciencia de la Computación, UC. Profesor del Dpto. de Ciencia de la Computación y del Dpto. de Ingeniería de Transporte y Logística de la Escuela de Ingeniería UC.

DENIS PARRA

Ph.D. of Philosophy in Information Science, University of Pittsburgh, EE.UU. Ingeniero Civil en Informática de la Universidad Austral de Chile. Profesor Asistente del Departamento de Ciencia de la Computación de Ingeniería, UC.

CRISTIAN RUZ

Ph.D. en Informática, Université de Nice Sophia Antipolis, Francia. Profesor Asistente del Departamento de Ciencia de la Computación UC.

ADRIÁN SOTO

Ph.D. (c) Ciencia de la Ingeniería y Licenciado en Ciencia de la Ingeniería, UC. Áreas de interés: Base de Datos, Data Science y Lenguajes de Programación.

FERNANDO SUÁREZ

Magíster en Ciencias de la Ingeniería, Ingeniero Civil Industrial y Diploma en Tecnologías de Información, UC. Profesor Instructor Adjunto del Departamento de Ciencia de la Computación de la Escuela de Ingeniería UC. Analista de Inversiones Senior, AFP Capital.

REQUISITOS

- » Licenciatura en Ciencias de la Ingeniería o Título profesional Universitario de Ingeniería Civil. Licenciatura o título profesional universitario en una disciplina afín a la Ingeniería.
- » Experiencia laboral de 2 años en el área de tecnologías de información.
- » Experiencia en el uso y/o manejo de sistemas computacionales, como: bases de datos, sistemas ERPs, Data Warehouses, sistemas OLAP, lenguajes de programación, infraestructura HPC o tecnologías afines.
- » Se recomienda buen nivel de comprensión lectora de inglés.

- El programa se inicia con un quorum mínimo de participantes.
- Las salas son asignadas dentro del Campus de Ejecución, NO NECESARIAMENTE es la misma sala todos los días.
- En caso de fuerza mayor, el programa se reserva el derecho a realizar clases por streaming, modificar fechas, lugar y/o profesores.
- Todas las modalidades del programa (dual, presencial y streaming) tienen el mismo valor. Además, recomendamos preguntar las condiciones de cada una al momento de matricularse.



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

DURACIÓN: 144 horas cronológicas

POSTULA EN: programas@ing.puc.cl

**Contáctanos para trabajar contigo
confeccionando el programa
perfecto para tu organización.**

  **+56 9 3353 0870**

www.educacionprofesional.ing.uc.cl

Consulta por descuentos, facilidades de pago y convenios con tu banco para pago en cuotas.