
PROGRAMA DE ESTUDIOS: BIG DATA

KOICA

IGU HANDONG GLOBAL
UNIVERSITY



Programa de Big Data

Conocimientos y competencias

- **Análisis y visualización de datos:** Desarrollar habilidades en el análisis y visualización de grandes conjuntos de datos para derivar ideas significativas.
- **Aprendizaje automático y modelos predictivos:** Adquirir experiencia en la creación e implementación de modelos de aprendizaje automático para el análisis predictivo.
- **Ingeniería de datos:** Aprenda a procesar, limpiar y transformar datos brutos en un formato adecuado para el análisis.
- **Tecnologías de Big Data:** Explora herramientas y marcos para trabajar con big data, incluyendo computación distribuida y almacenamiento.
- **Comunicación de los resultados de los datos:** Mejorar la capacidad de comunicar eficazmente los resultados de los datos a las partes interesadas, tanto técnicas como no técnicas.

Detalles del curso

1. Programación en R (40 horas, Básico)

El curso "Programación en R" es un programa meticulosamente diseñado que sirve como puerta de entrada esencial para las personas que aspiran a profundizar en los reinos del análisis estadístico y la visualización de datos utilizando R, una de las herramientas más potentes en la ciencia de datos. Este curso está estructurado para atender a un público amplio, desde principiantes sin experiencia previa en programación hasta profesionales que buscan aumentar sus conocimientos de análisis de datos. A lo largo de ocho semanas, los participantes se embarcarán en un viaje a través de los fundamentos de la programación en R, avanzando desde la sintaxis básica y los tipos de datos hasta la manipulación de datos complejos y las técnicas de visualización avanzadas.

Los participantes adquirirán experiencia práctica en el uso de RStudio, el principal entorno de desarrollo integrado para R, lo que les permitirá escribir, probar y depurar su código de manera eficiente. El plan de estudios está cuidadosamente intercalado con tareas basadas en proyectos que desafían a los estudiantes a aplicar su aprendizaje para analizar conjuntos de datos, derivar ideas y presentar sus hallazgos a través de visualizaciones convincentes.

El curso aprovecha una combinación de modos de impartición síncronos y asíncronos para adaptarse a diversos estilos de aprendizaje y horarios. Las sesiones en directo facilitan la interacción en tiempo real con los instructores y los compañeros, mientras que las clases y otros materiales grabados son accesibles para el estudio a ritmo propio.

Temas: Tipos y estructuras de datos, Manipulación de datos, Estadística básica con R, Visualización de datos, Manipulación avanzada de datos, Conceptos de programación.

2. Introducción a Big Data (40 horas, Básico)

Este curso ofrece una inmersión profunda en los conceptos fundamentales de los macrodatos, incluida su definición, características y los retos que plantean.

El plan de estudios está estructurado para ofrecer una visión global de las tecnologías y metodologías de big data. Empezando por los fundamentos del almacenamiento y la gestión de datos, los estudiantes aprenderán sobre marcos de computación distribuida como Hadoop y Spark, que permiten un procesamiento eficiente de grandes conjuntos de datos. El curso cubrirá aspectos clave del análisis de datos, incluyendo minería de datos, algoritmos de aprendizaje automático y métodos estadísticos adaptados a entornos de big data.

Una característica significativa de este curso es su enfoque en las herramientas y plataformas que son

fundamentales en el ecosistema de grandes datos. Los estudiantes adquirirán destreza en el uso de bases de datos NoSQL para el almacenamiento escalable, Apache Spark para el procesamiento de big data y herramientas de visualización de datos para comunicar ideas.

Más allá de los conocimientos técnicos, el curso aborda las consideraciones éticas y de privacidad inherentes al análisis de grandes volúmenes de datos. Los estudiantes participarán en debates y estudios de casos que ponen de relieve las implicaciones sociales de la recopilación y el análisis de datos, incluidas las cuestiones de seguridad de los datos, los derechos de privacidad y el uso ético de los datos. Este enfoque holístico garantiza que los graduados no sólo sean técnicamente competentes, sino también éticamente conscientes, preparados para navegar por las complejidades de la era digital con integridad.

Temas: Tecnologías de Big Data, análisis y minería de datos, almacenamiento de Big Data, aprendizaje automático con Big Data, herramientas de análisis de Big Data, consideraciones éticas y tendencias futuras

3. Manipulación y visualización de datos (40 horas, nivel intermedio)

A través de una experiencia de aprendizaje inmersiva con Python en su núcleo, los participantes se involucrarán en el proceso integral de limpieza de datos, manipulación y creación de visualizaciones convincentes. Aprovechando las potentes bibliotecas de Python como pandas, Matplotlib, Seaborn y Plotly, los estudiantes navegarán a través de escenarios de datos del mundo real, preparándolos para los desafíos de la industria de la ciencia de datos.

Dirigido a aspirantes a científicos de datos, profesionales que deseen mejorar sus habilidades en el manejo y la visualización de datos y estudiantes de campos relacionados que deseen aumentar sus capacidades analíticas, este curso ofrece una vía para dominar el manejo y la visualización de datos. Una vez finalizado, los participantes adquirirán una profunda competencia en el uso de Python para la ciencia de datos, listos para hacer frente a complejos desafíos de datos y elevar su carrera en el floreciente campo de la analítica de datos.

Temas: Introducción a la gestión de datos, Técnicas de limpieza de datos, Transformación de datos, Introducción a la visualización de datos, Visualización avanzada de datos

4. Minería de datos (40 horas, Intermedio)

Presenta una exploración exhaustiva de los conceptos, metodologías y herramientas de la minería de datos, centrándose en la extracción de información significativa a partir de vastos conjuntos de datos. El plan de estudios tiende un puente entre el conocimiento teórico y la aplicación práctica, preparando a los estudiantes para afrontar complejos retos de análisis de datos en diversos sectores. Mediante la integración de principios de la informática, la estadística y el aprendizaje automático, el curso ofrece un enfoque multidisciplinar que enriquece la experiencia de aprendizaje y proporciona a los estudiantes un conjunto de habilidades versátiles.

Un aspecto central de la filosofía del curso es el énfasis en el aprendizaje práctico. A través de sesiones de laboratorio dedicadas y un extenso trabajo de proyecto, los estudiantes se involucrarán directamente con problemas de minería de datos del mundo real utilizando herramientas de software avanzadas como scikit-learn de Python y la suite de minería de datos Weka.

El plan de estudios del curso incluye una amplia gama de estudios de casos y proyectos extraídos de diversos ámbitos, como las finanzas, la atención sanitaria, el análisis de las redes sociales, etcétera. Estas aplicaciones del mundo real subrayan la importancia de las técnicas de minería de datos para resolver los retos actuales y permiten a los estudiantes explorar el potencial de la analítica de datos en diferentes

sectores. La inclusión de tendencias y retos actuales garantiza que los estudiantes no sólo aprendan técnicas de minería de datos, sino que también comprendan sus implicaciones y aplicaciones en el mundo moderno.

Las consideraciones éticas en la minería de datos constituyen una parte crucial del contenido del curso. Dado que la privacidad de los datos, la seguridad y el uso ético de la información son cada vez más importantes, el curso aborda estas cuestiones de frente.

Temas: Preprocesamiento de datos, minería de reglas de asociación, clasificación, agrupación, detección de anomalías, temas avanzados

5. Análisis de datos (40 horas, Avanzado)

Los participantes explorarán los entresijos de la recopilación, limpieza, análisis y visualización de datos, dotándolos de las habilidades necesarias para afrontar los retos que plantean los datos en el mundo real. Mediante una combinación de conocimientos teóricos y aplicación práctica, este curso pretende desmitificar el proceso de análisis de datos y poner de manifiesto el poder de la toma de decisiones basada en datos en diversos sectores.

Un elemento central de la filosofía del curso es un enfoque de aprendizaje práctico, en el que los estudiantes se sumergen en proyectos reales de análisis de datos utilizando software estándar del sector como Python, R, SQL, Tableau y PowerBI. Estas sesiones prácticas están diseñadas para salvar la brecha entre los conceptos teóricos y su aplicación, permitiendo a los estudiantes experimentar de primera mano los matices de los flujos de trabajo de análisis de datos. Al trabajar con conjuntos de datos reales y resolver problemas auténticos, los participantes desarrollarán una base sólida en el manejo de datos, el análisis estadístico y la creación de visualizaciones de datos convincentes.

A lo largo del curso, los estudiantes trabajarán individualmente y en equipo en proyectos que reflejan las complejidades y los retos del análisis de datos en contextos profesionales. La culminación del curso es un proyecto final en el que los estudiantes aplican sus conocimientos acumulados para analizar un conjunto de datos, extraer ideas significativas y presentar sus conclusiones de una manera profesional.

Reconociendo las dimensiones éticas del análisis de datos, el curso integra debates sobre la privacidad de los datos, la seguridad y el uso ético de la información. Mediante la exploración de estudios de casos y temas de actualidad en la gobernanza de datos, los participantes aprenderán a navegar por las consideraciones éticas integrales para el papel de un analista de datos en el mundo centrado en los datos de hoy.

Data Analytics está diseñado para un público diverso, incluyendo estudiantes sin experiencia previa en análisis de datos, profesionales que buscan aumentar su conjunto de habilidades, y cualquier persona curiosa sobre el aprovechamiento del poder de los datos para la toma de decisiones. Con su enseñanza experta, opciones de aprendizaje flexibles y un plan de estudios que equilibra la teoría con la práctica, este curso constituye un punto de entrada ideal en el campo del análisis de datos.

Temas: Fundamentos estadísticos para el análisis de datos, Preprocesamiento de datos, Herramientas de análisis de datos, Análisis exploratorio de datos (AED), Visualización de datos, Aprendizaje automático para el análisis de datos, Ética y privacidad en el análisis de datos.