

Laboratorio de Diseño y Optimización de Operaciones

Disciplina asociada: Ingeniería Industrial

Escuela: Ingeniería y Ciencias

Departamento Académico: Ingeniería Industrial

Programas académicos: 8 IIS11

Requisitos: Estar cursando al menos el periodo 8

Equivalencia: IN3017

Intención del curso en el contexto general del plan de estudios:

Curso de nivel avanzado en ingeniería industrial que integra los conocimientos adquiridos por el alumno acerca de *ingeniería industrial clásica y creación de valor, con un enfoque en el área de Ciencia de Datos*.

Requiere conocimientos previos de estadística, algoritmos de predicción (promedios móviles), administración de la producción e inventarios, actividades que agregan valor, mapeo de procesos y, es deseable, tener conocimientos básicos en programación.

Como resultado del aprendizaje el alumno será capaz de aplicar los conocimientos adquiridos durante su carrera en el diseño, análisis y mejoramiento de procesos productivos, junto con algunas herramientas y conceptos de Ciencia de Datos, por medio de una experiencia de laboratorio.

Objetivo general de la Unidad de Formación:

Al finalizar el curso el alumno será capaz de seleccionar el conjunto de herramientas más adecuadas para analizar y mejorar un sistema, herramientas tanto del área de Ingeniería Industrial como de Ciencia de Datos.

Contenido temático del curso:

1. Introducción
 - 1.1. Relación: Ingeniería Industrial y Ciencia de Datos
Ingeniero Industrial con esteroides (atacar problemas con más variedad de herramientas).
2. Conceptos Básicos: Ciencia de Datos
 - 2.1. Situación actual: Crecimiento en los datos (0.5% de la información se usa y se analiza)
 - 2.2. ¿Qué es Ciencia de Datos?
 - Data Science ≠ Big Data.
 - Todo empieza con una pregunta.
 - Tareas (Clasificación, predicción, etc.)
 - Tipos de datos
 - 2.3. Perfil de un científico de datos
 - Diferencia: Data Scientist, Data engineer, Data analyst, Software engineer.

- Perfil (matemáticas, negocio, programación, estadística, aprendizaje de máquina)
- Importancia del trabajo en equipo
- 2.4. Pasos para realizar un proyecto de Ciencia de Datos
 - ¿Qué es un buen proyecto de Ciencia de datos?
 - CRISP-DM - Mapeo de procesos
 - Triste realidad: 80% es limpieza de datos, 20% es lo demás y quejarse de la limpieza de datos.
 - Importancia del EDA y de la visualización de los datos
- 2.5. Entregables de un proyecto de Ciencia de Datos
- 2.6. ¿Cómo saber si el proyecto tuvo éxito?
- 3. Conceptos básicos: Aprendizaje de Máquina
 - 3.1. ¿Qué es aprendizaje de máquina?
 - 3.2. Aprendizaje supervisado y no supervisado
 - Ejemplo de modelos (principalmente modelos para predicción: Promedios Móviles, Árboles de Decisión, Bosques Aleatorios)
 - 3.3. Estadística tradicional vs. Aprendizaje de máquina
 - 3.4. Importancia de las Métricas
- 4. Descripción de Herramientas
 - 4.1. Versiones de código: Github (crear cuenta en github y explicar funcionalidad)
 - 4.2. Lenguajes de programación: R y Python
 - R (mini curso - enfocado a visualización)
 - Python (mini curso - enfocado a modelado) *si hay tiempo
 - 4.3. Almacenamiento y procesamiento en la nube (descripción, ventajas, *mini ejemplo en IBM Cloud)
 - 4.4. Reportes y presentaciones (Jupyter Notebook y Rmarkdown)
- 5. Storytelling
 - 5.1. Presentación de soluciones y propuestas

- Todos estos temas enfocados a la elaboración de un proyecto a lo largo del curso.

Objetivos específicos de aprendizaje por tema:

1. Comprender que la perspectiva y herramientas que da el área de ciencia de datos tienen gran relación con las tareas que realiza y problemas que abarca un ingeniero industrial.
2. Comprender los conceptos básicos que se utilizan como científico de datos y aplicarlos para el entregable final del proyecto del curso.
3. Comprender los conceptos básicos relacionados a aprendizaje de máquina y su relación (contraste) con la estadística tradicional, para que el alumno pueda proponer y evaluar diversos modelos, tanto tradicionales como de aprendizaje de máquina, como solución a una problemática dada como proyecto del curso.
4. Aplicar (de forma básica) algunas las herramientas propuestas para desarrollar un proyecto de ciencia de datos y mostrar su utilidad y alcance ante diversos escenarios y problemas.
5. Presentar de forma oral y escrita los resultados.

Metodología de enseñanza y actividades de aprendizaje:

1. Se recomienda que los alumnos realicen prácticas en grupos e investigaciones individuales.
2. Exposiciones por parte del instructor.
3. Asignación de tareas como trabajo previo a la realización de las prácticas.
4. Entrega de reporte de actividades, avances y acuerdos
5. Presentación de los resultados alcanzados ante el instructor y sus compañeros.

Tiempo estimado de cada tema:

Presentación (Primer día de clases)	1.5 horas
Tema 1	1.5 horas
Tema 2	9 horas
Primer parcial (Examen teórico - Entrega Primer Avance de Proyecto)	
Tema 3	9 horas
Tema 4	19 horas
Segundo Parcial (Examen teórico - Entrega Segundo Avance de Proyecto) - A mitad del tema 4	
Tema 5	2 horas
Evaluación (3 hrs presentación final / 1.5 hrs primer parcial / 1.5 hrs segundo parcial)	6 horas
Total de horas	48 horas

Criterios de evaluación sugerida:

Para la evaluación del aprendizaje de los alumnos se cuenta con procedimientos y criterios que permiten dar seguimiento y evaluar los resultados del proceso de aprendizaje. Los procedimientos y la ponderación de cada uno de ellos son los siguientes

Primer Examen Parcial	15 %
Primer Avance de Proyecto	5 %
Segundo Examen Parcial	15 %
Segundo Avance de Proyecto	5 %
Actividades en clase / Tareas	10 %
Asistencia	5 %
Proyecto Final	30 %
Presentación Final	15 %
Total	100 %

Bibliografía sugerida:

Matloff, N. S. (2011). *The art of R programming: A tour of statistical software design*.

Patterson, Kerry. (Eds.) (2012) *Crucial conversations: tools for talking when stakes are high*

Knaflitz, C. N. (2015). *Storytelling with data: a data visualization guide for business professionals*.

Material de apoyo:**Perfil del Profesor:**

(143501) Maestría en Ingeniería Industrial ; (142701) Maestría en Ingeniería de Sistemas ; (143501) Doctorado en Ingeniería Industrial; (142701) Doctorado en Ingeniería de Sistemas CIP:143501, 142701

Idioma en que se imparte la materia: Español