

Programa



CURSO:CIENCIA DE DATOS RESPONSABLE

TRADUCCIÓN:RESPONSIBLE DATA SCIENCE

SIGLA:IMT2240

CRÉDITOS:10

MÓDULOS:02

CARÁCTER:MÍNIMO

TIPO:CÁTEDRA Y LABORATORIO

CALIFICACIÓN:ESTÁNDAR

DISCIPLINA:CIENCIA DE DATOS

PALABRAS CLAVE:TRANSPARENCIA, INTERPRETABILIDAD, PRIVACIDAD, SESGO, CALIDAD

NIVEL FORMATIVO:PREGRADO

I.DESCRIPCIÓN DEL CURSO

En este curso los estudiantes aprenderán tanto los fundamentos éticos como las herramientas y metodologías básicas que cimentan el uso responsable de los datos en distintos contextos. Se combinará una serie de lecturas con laboratorios, tareas, y proyectos prácticos que permitan a los estudiantes entender tanto la crítica como la solución a problemas relacionados con transparencia, interpretabilidad, privacidad, y sesgo, entre otros.

II.RESULTADOS DE APRENDIZAJE

- 1.Analizar las falencias éticas que puede presentar un proceso de análisis de datos identificando los componentes y técnicas de las ciencias sociales que se pueden utilizar para paliarlas.
- 2.Utilizar herramientas matemáticas y computacionales que garanticen un uso responsable de los datos.
- 3.Construir modelos que faciliten la interpretabilidad de los resultados obtenidos para poder justificar éticamente las decisiones tomadas.
- 4.Evaluar la factibilidad de las soluciones propuestas en base a los códigos vigentes para asegurar su pertinencia legal.

III.CONTENIDOS

- 1.Introducción
 - 1.1.Ejemplos de temas éticos en ciencia de datos
 - 1.2.Imparcialidad y diversidad, imposibilidad teórica
 - 1.3.Transparencia y gobernanza
 - 1.4.Privacidad
- 2.Ciclo de vida de los datos
 - 2.1.Perfiles de datos
 - 2.2.Limpieza de datos
 - 2.3.Calidad de datos vs imparcialidad y transparencia
- 3.Imparcialidad y diversidad

- 3.1. Definiciones de imparcialidad
- 3.2. La necesidad de una mirada socio-tecnológica
- 3.3. Diversidad, regla de Rooney

- 4. Transparencia e interpretabilidad
- 4.1. Necesidad de la transparencia e interpretabilidad
- 4.2. Explicaciones locales
- 4.3. Influencia causal

- 5. Privacidad y protección de datos
- 5.1. Transferencia responsable de datos
- 5.2. Anonimización y sus límites
- 5.3. Privacidad diferencial

- 6. Códigos legales y de conducta
- 6.1. Principios éticos y normativas legales
- 6.2. Regulaciones sobre protección general de datos

IV. ESTRATEGIAS METODOLÓGICAS

- Cátedras
- Estudio de casos
- Discusión en equipos
- Laboratorios
- Expertos invitados

V. ESTRATEGIAS EVALUATIVAS

- Laboratorios: 30%
- Tareas: 30%
- Proyecto final: 40%

VI. BIBLIOGRAFÍA

Mínimo

Julia Stoyanovich, Armanda Lewis. Teaching Responsible Data Science: Charting New Pedagogical Territory. CoRR abs/1912.10564 (2019)

Complementarias

Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. (ProPublica). <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Danah boyd. 2010. Privacy and Publicity in the Context of Big Data. <https://www.danah.org/papers/talks/2010/WWW2010.html>

Meredith Broussard. 2018. Artificial Unintelligence: How Computers Misunderstand the World. MIT Press, Cambridge, Massachusetts.

Alexandra Chouldechova. 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. CoRR abs/1703.00056 (2017). arXiv:1703.00056 <http://arxiv.org/abs/1703.00056>

Tschantz. 2018. Discrimination in Online Personalization: A Multidisciplinary Inquiry. In Conference on Fairness, Accountability and Transparency, FAT 2018, 23-24 February 2018, New York, NY, USA. 20-34. <http://proceedings.mlr.press/v81/datta18a.html>

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL / SEPTIEMBRE 2020