



TEMARIOS ASIGNATURAS

CURSO 2022/2023

MÁSTER UNIVERSITARIO EN
CIENCIAS SOCIALES COMPUTACIONALES

MATERIA 1

FUNDAMENTOS DE LA CIENCIA SOCIAL COMPUTACIONAL

| | |
|---|--|
| <ol style="list-style-type: none">1. Introducción<ul style="list-style-type: none">○ ¿Qué es la ciencia social computacional (CSS)?○ El paradigma de la CSS○ Primeros ejemplos○ La sociedad como sistema complejo adaptativo○ Principales áreas de la CSS2. Big data<ul style="list-style-type: none">○ Extracción automática de información y data mining○ Técnicas de análisis○ Ejemplos3. Redes sociales<ul style="list-style-type: none">○ Redes complejas: definiciones básicas○ Análisis cuantitativo de redes y software○ Ejemplos4. Complejidad social<ul style="list-style-type: none">○ Fundamentos y características○ Indicadores cuantitativos○ Leyes de la complejidad social5. Modelos y simulaciones<ul style="list-style-type: none">○ Construcción de modelos○ El propósito de las simulaciones○ Software básico: NetLogo○ Ejemplos | <ol style="list-style-type: none">1. Introduction<ul style="list-style-type: none">○ What is computational social science (CSS)?○ The paradigm of CSS○ First examples○ Society as a complex adaptive system○ Main areas of CSS2. Big data<ul style="list-style-type: none">○ Automatic information extraction and data mining○ Analysis techniques○ Examples3. Social networks<ul style="list-style-type: none">○ Complex networks: basic definitions○ Quantitative network analysis and software○ Examples4. Social complexity<ul style="list-style-type: none">○ Fundamentals and characteristics○ Quantitative indicators○ Laws of social complexity5. Models and simulations<ul style="list-style-type: none">○ Model construction○ The purpose of simulations○ Basic software: NetLogo○ Examples |
|---|--|

TEORÍAS DEL COMPORTAMIENTO EN LAS CIENCIAS SOCIALES

| | |
|--|--|
| <ol style="list-style-type: none"> 1. Explicación sociológica y mecanismos sociales <ul style="list-style-type: none"> ○ La explicación intencional y sus problemas ○ Interpretación vs. explicación ○ Modelos clásicos: conductismo, culturalismo, teoría de roles, estructuralismo, interaccionismo, teoría del habitus 2. Teoría de la elección racional y racionalidad de la conducta <ul style="list-style-type: none"> ○ Interacción estratégica y teoría de juegos ○ Acción colectiva y cooperación social ○ Dilemas sociales y sus soluciones 3. Teoría DBO <ul style="list-style-type: none"> ○ Efectos de composición y resultados sociales no intencionados ○ El “barco de Coleman”: el individualismo metodológico y el problema micro-macro ○ El debate sobre la “sociología analítica” 4. Ciencia social conductual <ul style="list-style-type: none"> ○ El “enriquecimiento conductual” de la elección racional ○ Racionalidad limitada, sesgos cognitivos y heurísticas sociales ○ Motivaciones prosociales, reciprocidad y altruismo 5. Influencia social, contagio, imitación y difusión <ul style="list-style-type: none"> ○ Confianza social. ○ Teoría de la señalización y conducta aparentemente irracional 6. Normas sociales y convenciones <ul style="list-style-type: none"> ○ Normas subóptimas e ignorancia plural ○ Falsificación de preferencias ○ Poder, status y jerarquías de status ○ Desigualdad 7. Biología vs. socialización y cultura como factores explicativos <ul style="list-style-type: none"> ○ Ciencia social evolucionaria y teoría evolutiva de la cultura ○ Las bases biológicas de la conducta social | <ol style="list-style-type: none"> 1. Sociological explanation and social mechanisms <ul style="list-style-type: none"> ○ Intentional explanation and its problems ○ Interpretation vs explanation ○ Classical models: behaviorism, culturalism, role theory, structuralism, interactionism, habitus theory 2. Rational choice theory and rationality <ul style="list-style-type: none"> ○ Strategic interaction and game theory ○ Collective action and social cooperation ○ Social dilemmas and their solutions 3. DBO theory <ul style="list-style-type: none"> ○ Composition effects and unintended social outcomes ○ The ‘Coleman’s boat’: methodological individualism and the micro-macro problem ○ The debate on ‘analytical sociology’ 4. Behavioural social science <ul style="list-style-type: none"> ○ The ‘behavioural enrichment’ of rational choice ○ Limited rationality, cognitive biases, and social heuristics ○ Pro-social motivations, reciprocity, and altruism 5. Social influence, contagion, imitation, and diffusion <ul style="list-style-type: none"> ○ Trust and trustworthiness ○ Signalling theory and apparently irrational behaviour 6. Social norms and conventions <ul style="list-style-type: none"> ○ Suboptimal norms and pluralistic ignorance ○ Falsification of preferences ○ Power, status and status hierarchies ○ Inequality 7. Biology vs. socialization and culture as explanatory factors <ul style="list-style-type: none"> ○ The debate on nature vs. nurture ○ Evolutionary social science and cultural theory. Biological basis of social behaviour |
|--|--|

CUESTIONES SOCIALES Y ÉTICAS DEL BIG DATA Y LA IA

| | |
|---|---|
| <p>1. Ética en las Ciencias Sociales Computacionales</p> <ul style="list-style-type: none"> ○ ¿Qué es la ética? La ética como vida buena y como paz social ○ La libertad como constitutiva radical del ser humano y como fenómeno observable en las ciencias sociales ○ Responsabilidad personal ante la influencia psicológica del entorno ○ La racionalidad de la ética: ética pública y ética privada. ¿Se puede razonar sobre lo bueno y lo malo? ○ Ingeniería social: responsabilidad ante las transformaciones de la sociedad <p>2. Efectos en el mundo real y consecuencias no deseadas de los sistemas algorítmicos</p> <ul style="list-style-type: none"> ○ Ética algorítmica programada y ética algorítmica aprendida ○ Cómo valorar la predicción de consecuencias en las decisiones éticas (humanas y algorítmicas) ○ ¿Deberíamos tener miedo a la inteligencia artificial? <p>3. Capitalismo de vigilancia</p> <ul style="list-style-type: none"> ○ Qui prodest? ¿A quién benefician mis búsquedas? ○ Empresas tecnológicas: sostenibilidad económica y creación de adicciones ○ Cómo lograr que la tecnología no se vuelva contra nosotros <p>4. Legislación y protección de datos personales y privacidad</p> <ul style="list-style-type: none"> ○ El modelo europeo frente al norteamericano y el chino ○ ¿Algo que ocultar? Qué son datos sensibles, y por qué son sensibles ○ Por qué no basta solo con el compromiso ético de las grandes corporaciones, y por qué no basta solo con la legislación y el control externo | <p>1. Ethics in Computational Social Sciences</p> <ul style="list-style-type: none"> ○ What is ethics? Ethics as good life and as social peace. ○ Freedom as a radical constitutive of the human being and as an observable phenomenon in the social sciences ○ Personal responsibility before the psychological influence of the environment ○ The rationality of ethics: public ethics and private ethics. Is it possible to reason about right and wrong? ○ Social engineering: responsibility before the transformations of society <p>2. Real-world effects and unintended consequences of algorithmic systems</p> <ul style="list-style-type: none"> ○ Programmed algorithmic ethics and learned algorithmic ethics ○ How to assess the predictability of consequences in ethical decisions (human and algorithmic) ○ Should we be afraid of artificial intelligence? <p>3. Surveillance capitalism</p> <ul style="list-style-type: none"> ○ Qui prodest? Who benefits from my searches? ○ Technology companies: economic sustainability and creation of addictions ○ How to ensure that technology does not turn against us <p>4. Legislation and protection of personal data and privacy</p> <ul style="list-style-type: none"> ○ The European model versus the North American and Chinese models ○ Something to hide? What is sensitive data, and why is it sensitive ○ Why the ethical commitment of large corporations alone is not enough, and why legislation and external control alone is not enough |
|---|---|

MATERIA 2

DISEÑO DE INVESTIGACIÓN PARA LAS CIENCIAS SOCIALES

| | |
|--|---|
| <ol style="list-style-type: none"> 1. Introducción <ul style="list-style-type: none"> ○ La importancia del diseño de investigación en ciencias sociales ○ Qué significa responder a una pregunta con evidencia empírica ○ Tipos de investigación empírica 2. Preguntas de investigación <ul style="list-style-type: none"> ○ Cómo encontrarlas y cómo trabajar con ellas ○ Descripción y explicación ○ Variables y relaciones entre variables 3. Tipos de diseño de investigación y de información empírica <ul style="list-style-type: none"> ○ Estudios de caso, estudios comparados, estudios cuantitativos con muestras grandes ○ Ventajas, desventajas y complementariedades ○ Tipos de evidencia empírica ○ Unidad de análisis y variabilidad ○ Problemas de medición 4. Causalidad <ul style="list-style-type: none"> ○ Predicción, correlación, probabilidad ○ Efectos causales y mecanismos causales ○ Explicaciones basados en mecanismos ○ Niveles de explicación: macro, meso, micro ○ Grafos dirigidos acíclicos 5. Problemas en inferencia causal <ul style="list-style-type: none"> ○ Relaciones espurias, sesgo de selección, collider bias, etc. ○ Problemas de inferencia interna y externa ○ Falacia ecológica, difusión, validez externa, etc. 6. Diseño de investigación para identificar relaciones causales <ul style="list-style-type: none"> ○ Ideal del método experimental en ciencias naturales ○ Contrafácticos y cómo aproximarse a ellos ○ Ventajas y desventajas del diseño de inferencia causal 7. Introducción a la lógica de la metodología de inferencia causal <ul style="list-style-type: none"> ○ Entendiendo las técnicas más comunes: experimentos, difference-in-differences, RDD, matching, etc. | <ol style="list-style-type: none"> 1. Introduction <ul style="list-style-type: none"> ○ Importance of research design in social sciences ○ What it means to answer a question with empirical evidence ○ Types of empirical research 2. Research questions <ul style="list-style-type: none"> ○ How to find them and how to work with them ○ Description and explanation ○ Variables and relationships between variables 3. Types of research design and empirical information <ul style="list-style-type: none"> ○ Case studies, comparative studies, quantitative studies with large samples ○ Advantages, disadvantages and complementarities ○ Types of empirical evidence ○ Unit of analysis and variability ○ Measurement problems 4. Causality <ul style="list-style-type: none"> ○ Prediction, correlation, probability ○ Causal effects and causal mechanisms ○ Mechanism-based explanations ○ Levels of explanation: macro, meso, micro ○ Directed Acyclic Graphs 5. Problems in causal inference <ul style="list-style-type: none"> ○ Confounding, selection bias, collider bias, etc. ○ Problems of internal and external inference ○ Ecological fallacy, diffusion, external validity, etc. 6. Research design to identify causal relationships <ul style="list-style-type: none"> ○ Ideal of experimental method in natural sciences ○ Counterfactuals and how to approach them ○ Advantages and disadvantages of causal inference design 7. Introduction to the logic of causal inference methodology <ul style="list-style-type: none"> ○ Understanding the most common techniques: experiments, difference-in-differences, RDD, matching, etc. |
|--|---|

METODOLOGÍA DE ENCUESTAS I

| | |
|---|--|
| <ol style="list-style-type: none"> 1. Introducción a la investigación con encuestas 2. Las encuestas y el proceso de investigación 3. Medición <ul style="list-style-type: none"> ○ Conceptualización de ideas y operacionalización ○ Preguntas pre-test ○ Error de medición ○ Validación y confiabilidad 4. Modos de recogida de datos <ul style="list-style-type: none"> ○ Cara a cara ○ Teléfono / Móvil ○ Web / Paneles en línea / SMS / Otros ○ Métodos mixtos 5. Muestreo y poblaciones <ul style="list-style-type: none"> ○ Introducción a la teoría del muestreo <ul style="list-style-type: none"> ■ Validez ■ Error ○ Tipos de muestras <ul style="list-style-type: none"> ■ Probabilística ■ No probabilístico ■ Poblaciones ocultas ■ Enfoques de gráficos/redes ○ Ajustes posteriores al muestreo <ul style="list-style-type: none"> ■ Estimación de la varianza ■ Ponderación ■ No respuesta | <ol style="list-style-type: none"> 1. Introduction on Survey Research 2. Surveys and the Research Process 3. Measurement <ul style="list-style-type: none"> ○ Conceptualizing Ideas and operationalizing questions ○ Pre-Testing Questions ○ Measurement error ○ Validation and reliability 4. Modes of Data Collection <ul style="list-style-type: none"> ○ Face-to-face ○ Telephone / Mobile ○ Web / Online panels / SMS / Others ○ Mixed methods 5. Sampling and Populations <ul style="list-style-type: none"> ○ Introduction to sampling theory <ul style="list-style-type: none"> ■ Validity ■ Error ○ Types of samples <ul style="list-style-type: none"> ■ Probabilistic ■ Non-probabilistic ■ Hidden-populations ■ Graph/network approaches ○ Post-sampling adjustments <ul style="list-style-type: none"> ■ Variance estimation ■ Weighing ■ Non-response |
|---|--|

METODOLOGÍA DE ENCUESTAS II

| | |
|--|--|
| <ol style="list-style-type: none"> 1. Análisis de datos de encuesta <ul style="list-style-type: none"> ○ Operacionalización y preparación de variables ○ Análisis con datos de sección cruzada ○ Análisis con datos panel ○ Análisis con datos jerárquicos ○ Ejemplos prácticos 2. Predicción con datos de encuesta <ul style="list-style-type: none"> ○ Predicción versus explicación ○ Evaluación del modelo predictivo 3. Tratamiento de los casos perdidos <ul style="list-style-type: none"> ○ Análisis de datos completos ○ Sustitución por la media/mediana ○ Imputación múltiple 4. Presentación de datos y resultados 5. Ejemplos prácticos | <ol style="list-style-type: none"> 1. Survey data analysis <ul style="list-style-type: none"> ○ Operationalization and data cleaning ○ Cross-sectional data analysis ○ Panel data analysis ○ Multilevel analysis 2. The use of survey data in forecasting <ul style="list-style-type: none"> ○ Predicting versus explaining ○ Performance evaluación 3. Dealing with missing data <ul style="list-style-type: none"> ○ Complete case analysis ○ Mean/median substitution ○ Multiple imputation 4. Data reporting 5. Worked examples |
|--|--|

ANÁLISIS DE REDES SOCIALES

| | |
|--|---|
| <ol style="list-style-type: none"> 1. Introducción, definiciones y principales conceptos <ul style="list-style-type: none"> ○ Datos "enredados". Procesos sociales que puede representarse como redes ○ Definición de red y redes aleatorias 2. Medidas para el análisis relacional de datos de redes sociales <ul style="list-style-type: none"> ○ Análisis estadístico de redes sociales ○ Medidas locales: grado, transitividad, asortatividad ○ Centralidad en redes sociales ○ Estructura de comunidades y de mundo pequeño ○ Importancia de los enlaces fuertes y débiles 3. Estructura y comportamiento de las redes <ul style="list-style-type: none"> ○ Procesos de contagio social. Homofilia vs. influencia ○ Formación de enlaces, dinámica de redes ○ Polarización y coordinación en redes sociales ○ Resiliencia de redes sociales 4. Visualización de redes sociales <ul style="list-style-type: none"> ○ Introducción a paquetes de visualización de redes sociales ○ Taller de visualización de redes sociales en Twitter 5. Ejemplos prácticos <ul style="list-style-type: none"> ○ Análisis de conversaciones en redes sociales ○ Modelos de propagación de epidemias en redes ○ Marketing en redes sociales ○ Segregación en redes sociales y zonas urbanas | <ol style="list-style-type: none"> 1. Introduction, definitions and main concepts <ul style="list-style-type: none"> ○ "Linked" data. Social processes that can be represented as networks ○ Definition of network and random networks 2. Measures for relational analysis of social network data <ul style="list-style-type: none"> ○ Statistical analysis of social networks ○ Local measures: degree, transitivity, asortativity ○ Centrality in social networks ○ Structure of communities and small world phenomena ○ Importance of strong and weak links 3. Structure and behavior of networks <ul style="list-style-type: none"> ○ Social contagion processes. Homophilia vs. influence ○ Formation of links, network dynamics ○ Polarization and coordination in social networks ○ Resilience of social networks 4. Visualization of social networks <ul style="list-style-type: none"> ○ Introduction to social media display packages ○ Workshop on visualization of social networks on Twitter 5. Practical examples <ul style="list-style-type: none"> ○ Analysis of conversations in social networks ○ Models of epidemic propagation in networks ○ Marketing in social networks ○ Segregation in social networks and urban areas |
|--|---|

MATERIA 3

PROGRAMACIÓN DE DATOS

| | |
|--|---|
| <ol style="list-style-type: none"> 1. Programación en R <ul style="list-style-type: none"> ○ Tipos de datos y operaciones principales ○ Condicionales and control de flujo ○ Bucles y programación vectorial ○ Funciones ○ Entender errores y depurar el código 2. Manejo de datos <ul style="list-style-type: none"> ○ Importar y exportar datos tabulares ○ Filtrado y reordenación de datos ○ Análisis exploratorio básico ○ Agregación de datos. La familia "apply" ○ Combinar data frames ○ Formatos largo y ancho 3. Ecosistema R <ul style="list-style-type: none"> ○ Paquetes y CRAN ○ Tendencias y referentes 4. Tidyverse <ul style="list-style-type: none"> ○ El operador "Pipe" ○ Paquete tidyr ○ Paquete dplyr 5. Tipos de datos avanzados <ul style="list-style-type: none"> ○ Factores ○ Fechas y horas: paquete lubridate ○ Cadenas de texto: expresiones regulares y paquete stringr | <ol style="list-style-type: none"> 1. R programming <ul style="list-style-type: none"> ○ Basic data types and operations ○ Conditionals and control flow ○ Loops and vectorial programming ○ Functions ○ Understanding errors and debugging 2. Data handling <ul style="list-style-type: none"> ○ Importing and exporting tabular data ○ Data filtering and reordering ○ Basic Exploratory Data Analysis ○ Data aggregation. The apply family ○ Joining data frames ○ Long and wide formats 3. R Ecosystem <ul style="list-style-type: none"> ○ Packages and CRAN ○ Important contributors and future trends 4. Tidyverse <ul style="list-style-type: none"> ○ Pipe operator ○ tidyr package ○ dplyr package 5. Advanced data types <ul style="list-style-type: none"> ○ Factors ○ Dates and times: lubridate package ○ Strings: regular expressions and stringr package |
|--|---|

VISUALIZACIÓN DE DATOS

| | |
|--|---|
| <ol style="list-style-type: none"> 1. Fundamentos de la práctica gráfica <ul style="list-style-type: none"> ○ Por qué hacer gráficas ○ Integridad gráfica ○ Percepción gráfica ○ Principios de representación gráfica 2. La gramática de los gráficos en R <ul style="list-style-type: none"> ○ Construir gráficos capa a capa ○ Guías y escalas ○ Sistemas de coordenadas ○ Facetas ○ Estilos 3. Visualización de datos en R <ul style="list-style-type: none"> ○ Distribución ○ Correlación ○ Ranking ○ Parte de un todo ○ Evolución ○ Mapa ○ Flujo ○ Otras técnicas 4. Comunicación de datos en R <ul style="list-style-type: none"> ○ R Markdown ○ Documentos y notebooks ○ Presentaciones ○ Otros formatos ○ Dashboards | <ol style="list-style-type: none"> 1. Fundamentals of graphical practice <ul style="list-style-type: none"> ○ Why graphics ○ Graphical integrity ○ Graphical perception ○ Principles of graphical representation 2. The grammar of graphs in R <ul style="list-style-type: none"> ○ Building graphs layer by layer ○ Guides and scales ○ Coordinate systems ○ Facets ○ Themes 3. Data visualization in R <ul style="list-style-type: none"> ○ Distribution ○ Correlation ○ Ranking ○ Part of a whole ○ Evolution ○ Maps ○ Flow ○ Other techniques 4. Data communication in R <ul style="list-style-type: none"> ○ R Markdown ○ Documents and notebooks ○ Presentations ○ Other formats ○ Dashboards |
|--|---|

EXTRACCIÓN DE DATOS

| | |
|--|--|
| <ol style="list-style-type: none"> 1. Introducción al Web Scraping <ul style="list-style-type: none"> ○ ¿Qué es el Web Scraping? ○ Tipos de Web Scraping ○ Formatos de datos: XML y HTML ○ Acceso práctico a XML y HTML ○ Automatización de los programas de Web Scraping ○ Selenium y scraping basado en JavaScript ○ Cuestiones éticas del Web Scraping ○ Ejercicios prácticos 2. APIs de datos <ul style="list-style-type: none"> ○ Qué es una API ○ Fundamentos de la comunicación de la API ○ Introducción al formato JSON ○ Crea tu propia API (y compártela) ○ Arquitectura REST ○ Las API como forma de compartir y obtener datos (de cualquier tipo) ○ Automatización de las solicitudes de la API ○ Hablar con las bases de datos ○ Autenticación y acceso ético a las APIs ○ Ejercicios prácticos 3. Automatización de la adquisición de datos <ul style="list-style-type: none"> ○ ¿Por qué necesitamos la automatización? ○ Acceso a los servidores ○ Tecnologías para la automatización de programas ○ Automatización de trabajos cron ○ Tareas de registro ○ Ejercicios prácticos | <ol style="list-style-type: none"> 1. An introduction to Web Scraping <ul style="list-style-type: none"> ○ What is Web Scraping? ○ Types of Web Scraping ○ Data formats: XML and HTML ○ Practical access to XML and HTML ○ Automation for Web Scraping programs ○ Selenium and JavaScript based scraping ○ Ethical issues with Web Scraping ○ Practical exercises 2. Data APIs <ul style="list-style-type: none"> ○ What is an API ○ Fundamentals of API communication ○ An introduction to the JSON format ○ Create your own API (and share it) ○ REST architecture ○ APIs as a way to share and obtain data (any kind) ○ Automation of API requests ○ Talking with Databases ○ Authentication and ethical access to APIs ○ Practical exercises 3. Automation of Data Acquisition <ul style="list-style-type: none"> ○ Why do we need automation? ○ Accessing servers ○ Technologies for automating programs ○ Automating cron jobs ○ Logging tasks ○ Practical exercises |
|--|--|

MINERÍA DE TEXTO

| | |
|--|---|
| <ol style="list-style-type: none"> 1. Introducción teórica al Procesamiento del Lenguaje Natural <ul style="list-style-type: none"> ○ Breve historia de la lingüística computacional y principales avances ○ Qué es el PLN y su papel en la Inteligencia Artificial ○ Estructura de un pipeline básico de PLN ○ Tareas y aplicaciones más comunes en la industria ○ Importancia actual en la sociedad digital, principales iniciativas 2. Introducción práctica al análisis automático del lenguaje con R <ul style="list-style-type: none"> ○ Importación de texto original, diseño de dataset y creación de estructuras de datos ○ Limpieza de texto, eliminación de stopwords y símbolos, valores ausentes y duplicados ○ Procesos de división y tokenización ○ Análisis básicos: conteo de palabras, extracción de n-gramas, tablas de frecuencias ○ Análisis intermedios: análisis de distintividad, tf-idf, bag of words 3. Introducción al análisis de sentimiento <ul style="list-style-type: none"> ○ Qué es el análisis automático del sentimiento en un texto: la opinión, la emoción y la intención del emisor ○ Casos reales de análisis de sentimiento en la industria y limitaciones ○ Ejercicios prácticos de análisis automático de sentimiento: uso de lexicones y diccionarios, asignación automática de sentimiento, segmentación, nubes de palabras ○ Creación de gráficos e informes de análisis de sentimiento 4. Introducción al modelado de tópicos <ul style="list-style-type: none"> ○ Qué es el modelado de tópicos, usos en la industria ○ Clasificación de textos en categorías: métodos supervisados y no supervisados ○ Ejercicios prácticos de modelado de tópicos: asociación de palabras y tópicos, identificación y caracterización de grupos naturales, términos comunes y solapamiento ○ Creación de gráficos e informes de modelado de tópicos para identificación de ideas representativas 5. Modelos de lenguaje <ul style="list-style-type: none"> ○ Qué son los modelos de lenguaje pre-entrenados y su impacto en el desarrollo del PLN y el aprendizaje automático ○ Usos e implicaciones en la industria y situación actual, principales iniciativas ○ Ejercicios prácticos de uso y evaluación de modelos predictivos básicos con datos en texto | <ol style="list-style-type: none"> 1. Theoretical introduction to Natural Language Processing <ul style="list-style-type: none"> ○ Brief history of computational linguistics and main developments ○ What is Natural Language Processing and its role in Artificial Intelligence ○ Structure of a basic NLP pipeline ○ Most common tasks and applications in the industry ○ Current importance in the digital society, main initiatives 2. Practical introduction to automatic language analysis with R <ul style="list-style-type: none"> ○ Source text import, dataset design and creation of data structures ○ Text cleaning, removal of stopwords and symbols, missing values and duplicates ○ Splitting and tokenization processes ○ Basic analysis: word count, n-gram extraction, frequency tables ○ Intermediate analysis: distinctiveness analysis, tf-idf, bag of words 3. Introduction to sentiment analysis <ul style="list-style-type: none"> ○ What is automatic sentiment analysis in a text: opinion, emotion and intention of the speaker ○ Real-world cases of sentiment analysis in the industry and limitations ○ Practical training on automatic sentiment analysis: use of lexicons and dictionaries, automatic sentiment mapping, segmentation, word clouds ○ Creation of sentiment analysis graphs and reports 4. Introduction to topic modeling <ul style="list-style-type: none"> ○ What is topic modeling, main uses in the industry ○ Classifying text into categories: supervised and unsupervised methods ○ Practical training in topic modelling: word and topic association, natural group identification and characterization, common terms and overlapping ○ Creation of topic modeling graphs and reports for identification of representative ideas 5. Language models <ul style="list-style-type: none"> ○ What are pre-trained language models and their impact on NLP and Machine Learning development ○ Uses and implications in the industry and current status, main initiatives ○ Practical training on the use and evaluation of basic predictive models with text data |
|--|---|

MATERIA 4

ESTADÍSTICA Y CIENCIA DE DATOS I

| | |
|---|---|
| <ol style="list-style-type: none"> 1. Estimación paramétrica y no paramétrica 2. Inferencia avanzada 3. Introducción a la modelización avanzada 4. Ejemplos prácticos | <ol style="list-style-type: none"> 1. Parametric and non parametric estimation 2. Advanced Inference 3. Introduction to advanced modelization 4. Empirical examples |
|---|---|

ESTADÍSTICA Y CIENCIA DE DATOS II

| | |
|--|--|
| <ol style="list-style-type: none"> 1. Modelos Lineales Generalizados (GLM) <ul style="list-style-type: none"> ○ Modelos de regresión con variables dependientes e independiente categóricas ○ Modelos para otro tipo de variables dependientes: binarias, con relación de orden, multinomiales, conteos, etc. 2. Modelos Lineales Generalizados Mixtos (GLMM) 3. Ejemplos con datos reales | <ol style="list-style-type: none"> 1. Generalized Linear Models (GLM) <ul style="list-style-type: none"> ○ Regression models for categorical dependent and independent variables ○ Other models for other types of qualitative dependent variables: binary, ordered, multinomial, counting, etc. 2. Generalized Linear Mixed Models (GLMM) 3. Real Life Examples |
|--|--|

MODELADO AVANZADO

| | |
|---|---|
| <ol style="list-style-type: none"> 1. Introducción al Machine Learning <ul style="list-style-type: none"> ○ ¿Entender o predecir? ○ Sesgo vs Varianza ○ Evaluación del rendimiento de los modelos 2. Aprendizaje no-supervisado <ul style="list-style-type: none"> ○ Reducción de dimensión: PCA ○ Clustering: k-means, métodos jerárquicos 3. Aprendizaje supervisado <ul style="list-style-type: none"> ○ Clasificación: aprendizaje estadístico (clasificadores Bayesianos), machine learning (knn, árboles de decisión, random forest, gradient boosting, redes neuronales) ○ Regresión avanzada: selección de modelos, regularización, selección de variables 4. Casos de estudio para todos los temas | <ol style="list-style-type: none"> 1. Introduction to Machine Learning <ul style="list-style-type: none"> ○ To explain or to predict? ○ Bias vs Variance ○ Performance evaluation 2. Unsupervised Learning <ul style="list-style-type: none"> ○ Dimensionality reduction: PCA ○ Clustering: k-means, hierarchical methods 3. Supervised Learning <ul style="list-style-type: none"> ○ Classification: statistical learning (Bayesian classifiers), machine learning (nearest neighbors, decision trees, random forest, gradient boosting, neural networks) ○ Advanced Regression: model selection, regularization tools, feature selection 4. Case Studies for all the topics |
|---|---|

INFERENCIA CAUSAL PARA LAS CIENCIAS SOCIALES

| | |
|--|---|
| <ol style="list-style-type: none">1. ¿Por qué la causalidad? Introducción al marco de salidas potenciales2. La comparación experimental3. Datos observacionales y comparación experimental: matching4. Quasi-experimentos: Diseños de regresión discontinua5. Quasi-experimentos: Variables instrumentales6. Diferencias en diferencias y datos de panel7. Inferencia causal aplicada: Evaluación de políticas | <ol style="list-style-type: none">1. Why causality? Introduction to the potential outcome frameworks2. The experimental benchmark3. Observational data and the experimental benchmark: Matching4. Quasi-experiments: Regression Discontinuity Designs5. Quasi-experiments: Instrumental variables6. Difference-in-differences and panel data7. Causal inference applied: Policy evaluations |
|--|---|

uc3m | Universidad
Carlos III
de Madrid