



Programa de ANÁLISIS DE DATOS EN REDES COMPLEJAS

1. NOMBRE DE LA UNIDAD CURRICULAR

Análisis de datos en Redes Complejas

2. CRÉDITOS

10 créditos

3. OBJETIVOS DE LA UNIDAD CURRICULAR

Una red es un concepto más amplio que un grafo, y en general se utiliza para modelar sistemas reales complejos como las interacciones entre personas, proteínas, neuronas, computadoras, elementos de comunicación, sistemas de transporte, etc. El análisis de redes se clasifica (al igual que con otros datos) en descriptivo, predictivo y prescriptivo. Este curso se focaliza en el análisis descriptivo y predictivo de datos en redes, que incluyen: la visualización de las redes; el cálculo de métricas y medidas sobre grafos; el modelado de su estructura y otras propiedades; y la predicción de características o procesos en nodos, enlaces y en la red completa. En resumen, se focaliza en las técnicas útiles para extraer nuevo conocimiento de estos sistemas.

El objetivo de esta asignatura es brindar una visión general del análisis de redes, describiendo sus fundamentos, técnicas, y herramientas para su aplicación práctica.

La asignatura presenta un equilibrio entre las visiones teórica y práctica. En lo teórico, se incluyen los principales conceptos del análisis descriptivo y predictivo de redes, a saber: ¿Qué es el análisis de datos en redes? Repaso de los principales conceptos de la matemática en teoría de grafos. Cálculo de características de nodos (centralidad, redes egocéntricas,...), enlaces, y estructura (cohesión local y global, componentes, etc.). Modelado generativos de grafos (redes aleatorias, etc.), y modelado predictivo (espacios embebidos, redes neuronales para grafos, etc.). Resolución de los principales problemas predictivos involucrados: predicción de características y enlaces, detección de comunidades, evaluación de procesos dinámicos y difusión en grafos.

En la parte práctica: se brindan razones para analizar redes, y los principales usos de la disciplina. ¿Cómo se analizan las redes?, ¿Qué herramientas se utilizan? Casos de aplicaciones reales.

Al finalizar el curso, el estudiante podrá diseñar y realizar investigación relacionada al análisis en redes, incluyendo: la recolección de datos, el análisis, y la documentación de resultados en formato científico (publicación académica).

4. METODOLOGÍA DE ENSEÑANZA

La enseñanza se realiza fundamentalmente en modalidad de taller+proyecto, o sea, centrada en trabajos de laboratorio y actividades individuales/grupales asistidos por los docentes con el objetivo de realizar un proyecto final. Se dictan 4 horas semanales de clase teórico-prácticas, incluyendo presentaciones teóricas, clases de práctico en computador, y monitoreo/consultas como apoyo a la realización del proyecto final. Cada alumno deberá dedicar adicionalmente un promedio de 2 horas semanales para estudio adicional y 4 horas semanales para la realización del proyecto final.

Resumen de dedicación horaria (150 horas total):

- 26 hrs teórico
- 34 hrs práctico, consultas y evaluación
- 30 hrs estudio domiciliario
- 60 hrs proyecto

En las clases prácticas se requiere la utilización de un computador personal para realizar los ejercicios en clase. Durante las clases y de forma domiciliaria se realizarán diferentes actividades que estimulen la investigación, la lectura de artículos, y la discusión y reflexión sobre las problemáticas presentadas. En todo momento, se buscará involucrar a los participantes en discusiones activas.

El proyecto final de investigación requiere las siguientes actividades por parte del estudiante: a mediados del curso realizar una propuesta por escrito del proyecto (incluyendo la fuente de los datos, y una propuesta de análisis científico sobre los datos); y al finalizar el curso entregar un informe completo del trabajo (datos, código fuente, e informe de resultados) así como una presentación oral del proyecto.

5. TEMARIO

1. Introducción

- Introducción al análisis de redes
(análisis de datos descriptivo/predictivo/prescriptivo, aprendizaje automático supervisado/no supervisado, definiciones en ciencia de redes)
- Motivación y utilidad del análisis de redes
(datos en formato de redes, visualización, uso de las redes en distintas disciplinas, repositorios de datos disponibles, descripción de problemas emblemáticos de análisis predictivo de redes, introducción a los modelos generativos, y librerías y software usuales)
- Matemática de las redes:
 - repaso de teoría de grafos
(representación de grafos, tipos de grafos, redes de varios modos, caminos, componentes, conectividad, etc., opcional algoritmos y complejidad en redes)
 - [opcional] repaso en inferencia estadística
(probabilidad, estadística e inferencia, estimaciones puntuales, intervalos de confianza, test de hipótesis)
- *Prácticos opcionales en teoría de grafos e inferencia estadística*
- *Capítulos: [SAND] Cap. 1, 2 (1-48). [SANDR] Cap. 1, 2 (1-28). [NE] Cap. 1, 2, 3, 4, 5, 6 (1-167), opcionales 9,10,11 (273-392)*

2. Análisis descriptivo de datos en redes: medidas y métricas en redes

- Mapeo del grafo de la red y visualización
- *Práctico 1: Software para visualizar redes (layouts, uso de etiquetas y colores).*
- Medidas de centralidad de nodos
(vectores propios, intermediación, cercanía, PageRank, hubs y autoridades, etc.)
- *Práctico 2: Calcular e interpretar medidas de centralidad de nodo en redes reales*
- Cohesión en las redes
(escalas de cohesión, medidas de solapamiento local y global, enunciar el problema de detección de comunidades)
- La estructura de gran escala de las redes (componentes, cliques, k-cores, efecto de pequeño mundo). Distribución del grado de conectividad (distribuciones power-law y scale-free)
- *Práctico 3: Análisis descriptivo y comparativo sobre diversas redes (medidas de centralidad, cohesión y distribución de grado). Aplicar sobre los datos del proyecto final.*
- *Capítulos: [SAND] Cap. 3, 4 (49-122). [SANDR] Cap. 3, 4 (29-67). [NE] Cap. 7, 8 (168-270)*
-

3. Análisis predictivo de datos en redes: predicción de información faltante, modelos generativos y procesos sobre redes

- Técnicas predictivas en redes
(espacios embebidos de nodos, redes neuronales para grafos, algoritmos de detección de comunidades, aplicación a problemas de predicción sobre redes)
- *Práctico 4: Introducción a los espacios embebidos de nodos. Interpretar resultados*

- *Práctico 5: Introducción a las redes neuronales para grafos. Aplicación para predecir información faltante en redes*
- *Práctico 6: Detectar comunidades. Evaluar medidas de cohesión y solapamiento*
- *Práctico 7: Aplicar sobre los datos del proyecto final: espacios embebidos de nodos, y/o redes neuronales para grafos, y/o algoritmos de detección de comunidades*
- Modelos generativos de redes (tradicionales y profundos), estimaciones y muestro
- *Práctico 8: Crear grafos aleatorios, calcular la distribución del grado y de las componentes, calcular camino más corto, etc.*
- Procesos tradicionales sobre redes: robustez y fallas de componentes, difusión de información en redes (epidemias), evolución y dinámica en redes
- *Práctico 9: Evaluar el efecto de mundo pequeño para la difusión de epidemias*
- *Capítulos: [GRL] Cap. 3, 4, 5, 6, 8, 9 (28-74, 102-122), [SAND] Cap. 4.3.3, 5, 6, 7, 8, 9 (102-111,123-331), [SANDR] Cap. 4.4, 5, 6, 7, 8, 9, 10 (59-66, 69-195), [NE] Cap. 12,13,14,15, 16, 17, 18 (397-704)*

4. [opcional] Análisis prescriptivo de datos en redes: optimización combinatoria sobre redes

- Introducción a la problemática y a las metodologías existentes

6. BIBLIOGRAFÍA

Tema	Básica	Complementaria
Introducción	(1,3)	(4,5)
Análisis descriptivo de datos en redes: medidas y métricas en redes	(1,3)	(4,5,7)
Análisis predictivo de datos en redes: predicción de información faltante, modelos generativos y procesos sobre redes	(1,2)	(4,5,7)
Análisis prescriptivo de datos en redes	(2)	(4,6)

6.1 Básica

1. [SAND] Kolaczyk, E.D. "Statistical Analysis of Network Data: Methods and Models". Springer Series in Statistics. Springer New York, 2009. ISBN 9780387881461. (disponible en portal Timbó)
2. [GRL] Hamilton, William L. "Graph Representation Learning". Synthesis Lectures on Artificial Intelligence and Machine Learning, Volume 14, 3. Morgan and Claypool. September, 2020. (pre-publicación disponible on-line: https://www.cs.mcgill.ca/~wlh/grl_book/)
3. [SANDR] Kolaczyk, E.D. and Csardi, G. "Statistical Analysis of Network Data with R". Use R!, Springer New York, 2014. ISBN 9781493909834. (disponible en portal Timbó)

6.2 Complementaria

4. [NE] Mark Newman. Networks: An Introduction. Oxford University Press, Inc. New York, NY, USA 2010. ISBN:0199206651 9780199206650
5. [BA] Albert-László Barabási, Network Science. (disponible online <http://barabasilab.com/networksciencebook/>)
6. [NCM] David Easley and Jon Kleinberg. Networks, Crowds, and Markets, Reasoning About a Highly Connected World. Cornell University, New York. Date Published: July 2010. ISBN: 9780521195331 (disponible online)
7. [ASN] Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). Analyzing Social Networks. SAGE Publications Limited. ISBN: 9781446247419.

7. CONOCIMIENTOS PREVIOS EXIGIDOS Y RECOMENDADOS

7.1 Conocimientos Previos Exigidos:

Probabilidad y Estadística

7.2 Conocimientos Previos Recomendados:

Algoritmos y estructura de datos,
Matemática Discreta,
Aprendizaje Automático,
Investigación Operativa.

ANEXO A
Para todas las Carreras

A1) INSTITUTO

Instituto de Computación

A2) CRONOGRAMA TENTATIVO

Se planifica un cronograma de 15 semanas. Las clases teóricas y prácticas acompañan el curso de principio a fin. En la tercer semana los estudiantes deberán seleccionar un proyecto (propuesto o validado por el docente). Trabajarán en el proyecto hasta el final del curso, donde deberán hacer una presentación oral y entregar un informe escrito. El foco de las últimas 2 semanas es el proyecto final.

Semana 1	Introducción (4 hs de clase)
Semana 2	Introducción (2 hs de clase). Análisis descriptivo de datos en redes (2 hs de clase)
Semana 3	Análisis descriptivo de datos en redes (4 hs de clase)
Semana 4	Análisis descriptivo de datos en redes (4 hs de clase)
Semana 5	Análisis descriptivo de datos en redes (4 hs de clase)
Semana 6	Análisis predictivo de datos en redes (2 hs de clase). Proyecto final (2 hs de clase)
Semana 7	Análisis predictivo de datos en redes (4 hs de clase)
Semana 8	Análisis predictivo de datos en redes (2 hs de clase). Proyecto final (2 hs de clase). Entrega: propuesta proyecto final
Semana 9	Análisis predictivo de datos en redes (4 hs de clase)
Semana 10	Análisis predictivo de datos en redes (4 hs de clase)
Semana 11	Análisis predictivo de datos en redes (4 hs de clase)
Semana 12	Análisis predictivo de datos en redes (4 hs de clase)
Semana 13	Análisis predictivo de datos en redes (4 hs de clase)
Semana 14	Análisis predictivo de datos en redes (2 hs de clase). Análisis prescriptivo de datos en redes (2 hs de clase)
Semana 15	Proyecto final (4 hs de clase). Entrega: reporte proyecto final y presentación oral

A3) MODALIDAD DEL CURSO Y PROCEDIMIENTO DE EVALUACIÓN

Se realizará un proyecto de investigación en la temática de la unidad curricular. Este proyecto podrá ser realizado en grupos dependiendo del alcance del tema elegido para el proyecto y de la curricula de estudiantes en cada edición.

No hay examen final.

Las actividades del curso (y su ponderación para la nota final) consisten en:

Formato Aprobado por resolución N°113 del CFI de fecha 04.07.2017

- (5%) Propuesta escrita de proyecto final
- (15%) Participación en clase
- (20%) Presentación oral del proyecto
- (60%) Informe en formato científico del proyecto de investigación

La unidad curricular se aprueba con el 60% de los puntos, teniendo un mínimo del 60% en cada actividad (propuesta, informe, presentación, y participación).

Para la participación en clase se tomará en consideración: la resolución de los ejercicios de práctico; la asistencia a clase; la lectura previa de los temas a cubrir en cada clase; y la frecuencia/calidad de las contribuciones en las clases.

A4) CALIDAD DE LIBRE

Esta unidad curricular no adhiere a resolución del consejo sobre condición de libre.

A5) CUPOS DE LA UNIDAD CURRICULAR

Cupos mínimos: no hay cupo

Cupos máximos: no hay cupo

ANEXO B para la(s) carrera(s) Ingeniería en Computación (plan 97), Licenciatura en Computación

B1) ÁREA DE FORMACIÓN

Investigación Operativa

B2) UNIDADES CURRICULARES PREVIAS

Curso: Curso aprobado de Introducción a la Investigación de Operaciones

Examen: No aplica

ANEXO B para la(s) carrera(s) Ingeniería en Computación (plan 87)

B1) ÁREA DE FORMACIÓN

No corresponde

B2) UNIDADES CURRICULARES PREVIAS

Curso: Las previas definidas en forma general para todas las electivas técnicas, no tiene previas específicas.
Examen: No aplica

Observación: para el plan 87 de Ingeniería en Computación debe valer como electiva técnica, acreditándose una electiva completa.

APROB RES CONSEJO DE FAC. ING.

18 JUN 29/6/2021 Exp. 060120 - 001188-11