



## 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 complejos como las interacciones entre personas, proteínas, neuronas, computadoras, elementos de comunicación, sistemas de transporte, etc. El análisis de una red incluye el modelado de su estructura, la visualización, y el cálculo de medidas, todos muy útiles para extraer nuevo conocimiento de estos sistemas.

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

La unidad curricular presenta un equilibrio entre las visiones teórica y práctica. En lo teórico, se incluyen los principales conceptos del análisis de redes, a saber: ¿Qué es el análisis de datos? Repaso de los principales conceptos de la matemática en teoría de grafos. Modelado de grafos (redes aleatorias). Análisis desde los nodos (centralidad, redes egocéntricas,...). Análisis desde los enlaces (detección de comunidades). Dinámica 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 de datos en redes, incluyendo: la recolección y organización de datos, el análisis científico de esos datos, 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):

- 32 hrs teórico
- 28 hrs práctico
- 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); al final del curso entregar un informe completo del trabajo (datos, código fuente y informe de resultados) así como una presentación oral del proyecto.

#### 5. TEMARIO

##### 1. Motivación

- Utilidad del análisis de redes (análisis de datos, aprendizaje automático, datos en formato de redes, visualización)
- El uso de las redes en distintas disciplinas (comunicación e información, biología, física, sociología)
- La representación de grafos (nodos, enlace, matriz de adyacencia, grado de un nodo, redes de uno y dos modos)
- Práctico: Introducción al software para analizar redes. Analizar mi red social online (recolectar datos de Internet LinkedIn/Facebook/Twitter, visualizar resultados)
- Capítulos: [NE] Cap. 1,2,3,4,5 (1-107)

##### 2. Matemática de las redes: teoría de grafos e inferencia estadística

- Grafos ponderados, hiper-grafos, grafos bipartitos, árboles. grafos planares, etc. Caminos, componentes, conectividad, caminantes aleatorios, etc.

- Repaso de probabilidad e inferencia estadística clásica.
- Práctico: Software para visualizar redes (uso de etiquetas y colores)
- Capítulos: [NE] Cap. 6 (107-167)

### 3. Medidas en redes

- Medidas de centralidad de nodos (vectores propios, intermediación, cercanía, Page-Rank, hubs y autoridades, etc.)
- Detección de comunidades (componentes, cliques, k-cores, etc.)
- Práctico1: Calcular e interpretar medidas de centralidad de nodo en redes reales (LinkedIn/Facebook/Twitter, etc.)
- Práctico2: Calcular e interpretar solapamiento y separación de comunidades en redes reales (colaboración científica, blogs políticos, ingredientes de cocina, etc.)
- Capítulos: [NE] Cap. 7 (168-235)

### 4. Modelos de redes

- La estructura de gran escala de las redes (componentes, efecto de pequeño mundo)
- Distribución del grado de conectividad (distribuciones power-law y scale-free)
- Grafos aleatorios y aleatorios generalizados
- Creación de redes
- Práctico: Mediante software especializado (NetLogo, Gephi, R), crear grafos aleatorios, calcular la distribución del grado y de las componentes, calcular camino más corto, etc.
- Capítulos: [NE] Cap. 8,12,13,14 (235-272, 397-588)

### 5. Procesos sobre redes

- Fallas de componentes
- Difusión de información en redes (epidemias)
- Dinámica en redes
- Práctico: Evaluar el efecto de mundo pequeño para la difusión de epidemias
- Capítulos: [NE] Cap. 16,17,18 (589-705)

### 6. Tendencias en el análisis de redes

- Redes de gran escala
- Algoritmos y software para el cómputo en redes de gran escala
- Práctico1: Software para redes de gran escala. Análisis de una red de gran escala
- Práctico2: Leer y discutir trabajos académicos recientes en el área, o aplicaciones novedosas.

## 6. BIBLIOGRAFÍA

Tema	Básica	Complementaria
Motivación	(1,2)	(3,4)
Matemática de las redes	(1,2)	(3,4)
Medidas en redes	(1,2)	(3,4,6)
Modelos de redes	(1,2)	(3,4,6)
Procesos sobre redes	(1,2)	(3,4)
Tendencias en el análisis de redes	(1)	(5)

### 6.1 Básica

- [NE] Mark Newman. Networks: An Introduction. Oxford University Press, Inc. New York, NY, USA 2010. ISBN:0199206651 9780199206650
- [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

- [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)
- [BA] Albert-László Barabási, Network Science. (disponible online <http://barabasilab.com/networksciencebook/>)
- [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)
- [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:** Análisis de algoritmos, Algoritmos y estructura de datos, 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. El foco de las últimas 4 semanas es el proyecto final. 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.

Semana 1	Motivación (4 hs de clase)
Semana 2	Matemática de las redes (4 hs de clase)
Semana 3	Matemática de las redes (4 hs de clase)
Semana 4	Medidas en redes (4 hs de clase)
Semana 5	Medidas en redes (4 hs de clase)
Semana 6	Medidas en redes (4 hs de clase)
Semana 7	Medidas en redes (2 hs de clase). Proyecto final (2 hs de clase)
Semana 8	Modelos de redes (4 hs de clase)
Semana 9	Modelos de redes (4 hs de clase)
Semana 10	Modelos de redes (4 hs de clase)
Semana 11	Modelos de redes (4 hs de clase)
Semana 12	Procesos sobre redes (4 hs de clase)
Semana 13	Procesos sobre redes (4 hs de clase)
Semana 14	Tendencias en el análisis de redes (2 hs de clase). Proyecto final (2 hs de clase)
Semana 15	Proyecto final (4 hs de clase)

**A3) MODALIDAD DEL CURSO Y PROCEDIMIENTO DE EVALUACIÓN**

Se realizará un proyecto de investigación en la temática de la unidad curricular. Dependiendo de la currícula de estudiantes en cada edición y de las restricciones de cada carrera, este proyecto podrá ser realizado en grupos.

No hay examen final.

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

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

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 asistencia a clase; la lectura previa de los temas a cubrir en cada clase; y la frecuencia/calidad de las contribuciones en las clases teóricas y prácticas. No se descarta la utilización de una evaluación escrita en algunas clases sobre la lectura pautada.

Además, durante la ejecución de la asignatura se presentarán ejercicios domiciliarios, que son fuertemente recomendados para el seguimiento del curso, pero no serán requeridos ni evaluados.

#### **A4) CALIDAD DE LIBRE**

Esta asignatura 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**

Para el Curso: Curso aprobado de Introducción a la Investigación de Operaciones  
Para el 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**

Para el Curso: Las previas definidas en forma general para todas las electivas técnicas, no tiene previas específicas.

Para el 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. 170.**

**Fecha** 5.12.17 **Exp.** 060120-001189-16