

**PLANIFICACIÓN DE CURSO**  
Segundo Semestre académico 2024

**I. ACTIVIDAD CURRICULAR Y CARGA HORARIA**

|   |                 |
|---|-----------------|
| Asignatura: Análisis para Ciencia de Datos  | Código: MMD2002 |
| Semestre de la Carrera: 4   |                 |
| Carrera: Ingeniería Civil en Modelamiento Matemático de Datos   |                 |
| Escuela: Ingeniería   |                 |
| Docente: Andrés Zúñiga  |                 |
| Ayudante: Ignacio Jimenez   |                 |
| Horario: Cátedra: Lunes 10:15 – 11:45 hrs, Miércoles 10:15 – 11:45 hrs, Jueves 14:30 – 16:00 hrs.<br>Ayudantía: Viernes 14:30 – 16:00 hrs |                 |

|  |           |
|--|-----------|
| Créditos SCT:                          | 9         |
| Carga horaria semestral <sup>1</sup> : | 243 horas |
| Carga horaria semanal:                 | 17 horas  |

|   |          |
|---|----------|
| Tiempo de trabajo directo semanal:        | 6 horas  |
| Tiempo de trabajo del estudiante semanal: | 11 horas |

**II. RESULTADOS U OBJETIVOS DE APRENDIZAJE ESPERADOS ESTE SEMESTRE**

|    |  |
|----|--|
| 1) | Conocer las dificultades inherentes al trabajo en ciencia de datos y la importancia de los fundamentos matemáticos para poder abordarlas de manera exitosa.  |
| 2) | Aplicar los fundamentos matemáticos de Álgebra Lineal como proyección, mínimos cuadrados, técnicas espectrales y análisis de componentes principales en diversos contextos.  |
| 3) | Entender conceptos y herramientas de espacios de Hilbert. Aplicar estos conceptos en análisis armónico y en teoría de probabilidades.  |
| 4) | Entender los elementos y herramientas de teoría de probabilidad. Aplicar estos conceptos al diseño de algoritmos aleatorizados para datos masivos y de alta dimensión.   |
| 5) | Implementar las herramientas aprendidas dentro de cada unidad con datos reales, incluido el análisis crítico de sus resultados, en aplicaciones que incluyen, por ejemplo: problemas de regresión, reducción de dimensionalidad, clusterización o procesamiento de señales e imágenes. |

<sup>1</sup> Considere que 1 crédito SCT equivale a 27 horas de trabajo total (directo y autónomo) en el semestre.

### III. UNIDADES, CONTENIDOS Y ACTIVIDADES

| <i>UNIDAD TEMÁTICA 1: ÁLGEBRA LINEAL AVANZADA</i> |   |  |   |   |
|---|---|--|---|---|
| Semana  | Contenidos  | Actividades de enseñanza y aprendizaje |   | Actividades de evaluación diagnóstica, formativa y/o sumativa |
|   |   | Tiempo directo                         | Tiempo trabajo autónomo del o la estudiante |   |
| 1   | <i>Espacios Vectoriales, Subespacios, mapas lineales e isomorfismos</i>   | 3 horas                                | 12.5 horas                                  |   |
| 2   | <i>Espacios Vectoriales de Dimensión Finita, conjuntos linealmente independientes y bases<br/>Mínimos cuadrados y aplicaciones a aprendizaje de máquinas.</i>   | 6 horas                                | 11 horas                                    |   |
| 3   | <i>Producto interno: desigualdad de Cauchy-Schwarz, proyección en subespacios.<br/>Teorema espectral</i>  | 6 horas                                | 11 horas                                    |   |
| 4   | <i>Descomposición en valores singulares (representación de transformaciones lineales).<br/>Análisis de componentes principales y aplicaciones.<br/>Matrices dispersas (pre-condicionamiento)</i>  | 6 horas                                | 11 horas                                    |   |
| <i>UNIDAD TEMÁTICA 2: ESPACIOS DE HILBERT</i>     |   |  |   |   |
| 5   | <i>Espacios de Hilbert: definición y propiedades (Funcionales lineales y bilineales, formas cuadráticas, productos internos y normas, nociones de topología, desigualdades de Cauchy-Schwarz y Bessel, ortogonalidad y complementos ortogonales, bases y dimensión, acotamiento de funcionales)</i> | 6 horas                                | 11 horas                                    | Entrega T1  |

|  |  |         |          |                 |
|--|--|---------|----------|-----------------|
| 6  | <i>Integral de Lebesgue:<br/>funciones simples</i>   | 6 horas | 11 horas | Control 1 (CC1) |
| 7  | <i>Funciones integrables: caso<br/>general.<br/>Funciones cuadrado<br/>integrables</i>                               | 6 horas | 11 horas |                 |
| 8  | <i>Convergencia fuerte y débil</i>   | 6 horas | 11 horas |                 |
| 9  | <i>Dualidad (Teorema de Riesz)<br/>y Proyección sobre un<br/>convexo</i>   | 6 horas | 11 horas |                 |
| <b>UNIDAD TEMÁTICA 3: PROBABILIDADES</b> |  |         |          |                 |
| 10                                       | <i>Introducción a la medibilidad<br/>y sigma álgebras.<br/>Esperanza y Varianza</i>                                  | 6 horas | 11 horas |                 |
| 11                                       | <i>Independencia, correlación y<br/>covarianza.<br/>Normales multivariadas.<br/>Convergencia en<br/>probabilidad</i> | 6 horas | 11 horas |                 |
| 12                                       | <i>Ley de los grandes números y<br/>Teorema Central del Límite</i>   | 6 horas | 11 horas |                 |
| 13                                       | <i>Desigualdades de<br/>concentración: Markov,<br/>Tchebyshev, Chernoff<br/>(Hoeffding)</i>                          | 6 horas | 11 horas | Entrega T2      |
| 14                                       | <i>Reducción de<br/>dimensionalidad (Lemma de<br/>Johnson-Lindenstrauss) y<br/>Aplicaciones</i>                      | 6 horas | 11 horas | Control 2 (CC2) |
| 15                                       | <i>Tópicos adicionales en<br/>probabilidades</i>   | 6 horas | 11 horas |                 |

#### IV. CONDICIONES Y POLÍTICAS DE EVALUACIÓN

##### Metodología Docente

La metodología docente de trabajo será activo-participativa, mediante cátedras y sesiones de resolución de problemas. La modalidad de la asignatura será presencial, mientras la situación sanitaria lo permita.

##### Evaluaciones

Las evaluaciones permitirán que los y las estudiantes demuestren los resultados de aprendizaje alcanzados en los distintos momentos del proceso de enseñanza. La evaluación se realizará mediante 2 Controles de Cátedra (CC), 2 Tareas (T) y un Examen (EX). Los CC y el EX tendrán una duración de al menos 3 horas cada uno, y serán evaluados de forma presencial.

Tabla 1. Calendario de evaluaciones

| <b>Evaluación</b> | <b>Fecha</b>   |
|-------------------|--|
| T1                | <i>Entrega</i><br>Semana 6<br>(lunes, 23 sept., 23:59hrs)            |
| CC1               | <b><i>Presencial</i></b><br>Semana 7<br>(jueves, 03 oct., 14:00hrs)  |
| T2                | <i>Entrega</i><br>Semana 13<br>(lunes, 11 nov., 23:59hrs)            |
| CC2               | <b><i>Presencial</i></b><br>Semana 14<br>(jueves, 21 nov., 14:00hrs) |
| Examen            | <b><i>Presencial</i></b><br>Semana 16 o 17<br>(por definir por Elng) |

El promedio simple de los Controles de Cátedra (CC) conforma la Nota de Presentación al Examen (NP):

$$NP = (CC1 + CC2 + CC3)/3$$

en donde

$$CC3 = 50\% T1 + 50\% T2.$$

La Nota de Cátedra (NC) está compuesta por el promedio de los Controles de Cátedra (CC) y el Examen (EX) con una ponderación de:

$$NC = 60\% NP + 40\% EX.$$

### **APROBACIÓN**

La condición de aprobación del curso pasa por aprobar la Nota de Cátedra:

$$NC \geq 4.0$$

La Nota Final (NF) del curso, que se asigna al Acta de Notas, corresponde a la Nota de Cátedra:

$$NF = NC.$$

### **INASISTENCIAS A EVALUACIONES**

Toda inasistencia a un Control de Cátedra será calificada con la nota mínima 1.0, si no se justifica. **El Examen (EX) reemplaza automáticamente la menor nota de Controles (CC).** Sin embargo, el reemplazo de la nota del Examen por la nota más baja de CC no aplicará en el caso de una nota mínima asignada producto de una infracción a la integridad académica (por ejemplo, copia).

### **EXENCIÓN AL EXAMEN, Y RECUPERATIVO**

Cualquier estudiante cuya nota NP sea igual o superior a 5.5, está exento de rendir el Examen. Los estudiantes exentos recibirán como nota de Examen (EX) su nota de presentación (NP).

Los estudiantes exentos puede rendir de igual forma el Examen si lo desean, en cuyo caso su EX se considerará sólo si mejora su calificación final del curso.

Si posterior al Examen no se ha aprobado el curso, tendrán derecho a rendir un Examen Recuperativo (ER) sólo los estudiantes que tengan una NC igual a 3.7, 3.8 o 3.9. La aprobación del ER dará como resultado que la NC=4.0 (y no la nota final=4.0)

## **V. BIBLIOGRAFÍA Y RECURSOS OBLIGATORIOS**

- Sheldon Axler, "Linear Algebra Done Right", Third Edition, 2015.
- Lokenath Debnath and Piotr Mikusiński, "Hilbert Spaces with Applications", 2005.
- Alan Karr, "Probability", 1993.

## VI. BIBLIOGRAFÍA Y RECURSOS COMPLEMENTARIOS

- Avrim Blum, John Hopcroft, Ravindran Kannan, "Foundations of Data Science", 2018.
- Paul Halmos, "Introduction to Hilbert Space and the Theory of Spectral Multiplicity", 2013.
- Haim Brezis, "Functional Analysis, Sobolev Spaces and Partial Differential Equations", 2011.
- Pedro Morettin, "Wavelets in Functional Data Analysis", 2017.
- Jean-Pierre Florens, "Econometric Modeling and Inference", 2012.
- Yung Tong, "The Multivariate Normal Distribution", 1990
- Sanjoy Dasgupta, Anupam Gupta, "An elementary proof of a theorem of Johnson and Lindenstrauss". *Random Structures & Algorithms*, 22 (1): 60–65, 2003.