

**DENOMINACIÓN DE LA ASIGNATURA O UNIDAD DE APRENDIZAJE****ANALÍTICA DE DATOS**

<b>CICLO ESCOLAR(SEMESTRE):</b>	5
<b>ÁREA</b>	ÁREA MAYOR
<b>TIPO</b>	OBLIGATORIA
<b>CLAVE DE LA ASIGNATURA:</b>	24247
<b>SIGLA DE LA ASIGNATURA:</b>	II055
<b>HORAS CON DOCENTE:</b>	4
<b>HORAS INDEPENDIENTES:</b>	2
<b>CRÉDITOS:</b>	6
<b>CARACTERIZACIÓN:</b>	Teórica-Práctica
<b>INSTALACIONES:</b>	LABORATORIO: LABORATORIO DE CÓMPUTO
<b>COORDINACIÓN</b>	INGENIERÍA INDUSTRIAL
<b>PRERREQUISITOS</b>	

**FINES DEL APRENDIZAJE O FORMACIÓN**

- Distinguir los fundamentos estadísticos de la analítica de datos mediante su tratamiento y las técnicas apropiadas, para su procesamiento.
- Aplicar las técnicas más usadas en la modelación predictiva para el estudio de patrones en bases de datos y los métodos, para evaluar su desempeño e implementación.
- Aplicar las técnicas más importantes en modelación descriptiva para el estudio de la estructura de un conjunto de datos, así como de las técnicas de asociación y secuencia surgidos en distintas industrias.
- Identificar el método adecuado para el estudio de conjuntos de datos y su evolución a lo largo del tiempo, para la estimación de eventos futuros.

**CONTENIDO TEMÁTICO (TEMAS Y SUBTEMAS)****1.-Preparación de datos.**

- 1.1 Conceptos elementales de estadística multivariada.
- 1.2 Tratamiento estadístico de los datos: datos ausentes, atípicos, imputación.
- 1.3 Reducción de dimensionalidad: PCA, MDS y t-SNE.
- 1.4 Reducción de datos y transformación.
- 1.5 Visualización de datos.
- 1.6 Ejercicios.

**2.-Modelos predictivos.**

- 2.1 Modelos lineales.
- 2.2 Árboles de decisión.
- 2.3 Redes neuronales artificiales.
- 2.4 Máquinas de soporte vectorial.
- 2.5 Naive Bayes y k vecinos más cercanos.
- 2.6 Métodos de ensamble.

**3.-Validación y desempeño de modelos.**

- 3.1 Introducción a los métodos de validación.
- 3.2 Estabilidad del modelo.
- 3.3 Validación cruzada y robustez.
- 3.4 Regularización y selección de modelos.

### 3.5 Ejercicios.

#### 4.-Modelos descriptivos.

##### 4.1 Conceptos elementales.

##### 4.2 Análisis de conglomerados jerárquicos.

##### 4.3 K-medias y algoritmos relacionados.

##### 4.4 Reglas de asociación y secuencia.

##### 4.5 Casos de aplicación.

#### 5.-Modelos de series de tiempo y pronósticos.

##### 5.1 Definición, tipos y utilidad.

##### 5.2 Evaluación de pronósticos.

##### 5.3 Métodos para series estacionales.

##### 5.4 Métodos basados en la tendencia.

##### 5.5 Introducción a los modelos de Box-Jenkins.

### ACTIVIDADES DE APRENDIZAJE BAJO LA CONDUCCIÓN DE UN ACADÉMICO

- Realización de estudios de caso de analítica de datos y series de tiempo y pronósticos.
- Resolución de ejercicios de analítica de datos y series de tiempo y pronósticos.
- Resolución de problemas con conjuntos de datos teóricos y reales.
- Realización de proyectos que incluyan la aplicación de técnicas de analítica de datos y/o series de tiempo y pronósticos.

### ACTIVIDADES DE APRENDIZAJE INDEPENDIENTES

- Realización de ejercicios de modelación descriptiva, predictiva y de series de tiempo y pronósticos.
- Resolución de problemas de modelación descriptiva, predictiva y de series de tiempo y pronósticos.
- Realización de proyectos que incluyan el pre-procesamiento de datos y algún tipo de modelación estadística o de series de tiempos o de pronósticos.
- Elaboración de proyectos finales de analítica de datos o de series de tiempo o pronósticos.

### CRITERIOS DE EVALUACIÓN

INSTRUMENTO	PORCENTAJE
	La suma de los porcentajes debe representar el 100%
1. Examen parcial 1 (Modelación descriptiva).	20 %
2. Examen parcial 2 (Pre-procesamiento de datos).	20 %
3. Resolución de ejercicios y problemas.	20 %
4. Examen final (Modelación predictiva).	20 %
5. Proyecto final.	20 %
<b>TOTAL:</b>	<b>100%</b>

### MODALIDADES TECNOLÓGICAS E INFORMÁTICAS

**NO APLICA** debido a que el programa es modalidad escolarizada.

### BIBLIOGRAFÍA

1. Daza, A. (2016). *Data mining: Minería de datos*. Perú: Macro.
2. García-Herrero, J. , Molina, J. , Berlanga, J. y et al. . (2018). *Ciencia de datos: técnicas analíticas y aprendizaje estadístico*. España: Alfaomega.
3. Montgomery, D. , Jennings, C. y Kulahci, M. (2016). *Introduction to Time Series Analysis and Forecasting*. USA: Wiley.
4. Tan, P. , Steinbach, M. y Kumar, V. (2018). *Introduction to Data Mining*. USA: Pearson.
5. VanderPlas, J. (2017). *Python Data Science Handbook: Essential Tools for Working with Data*. USA: O'Reilly Media.