



# Universidad Tecnológica de la Mixteca

Clave DGP: 200089

## Maestría en Ciencia de Datos

### PROGRAMA DE ESTUDIOS

#### NOMBRE DE LA ASIGNATURA

**Proyecto Aplicado III**

SEMESTRE	CLAVE DE LA ASIGNATURA	TOTAL DE HORAS
<b>Tercer semestre</b>	<b>371031</b>	<b>40 Mediación docente 65 Estudio independiente</b>

#### OBJETIVO(S) GENERAL(ES) DE LA ASIGNATURA

Obtener resultados a partir de la base de datos aplicando técnicas del análisis multivariante y de datos masivos.

#### TEMAS Y SUBTEMAS

##### 1 Selección de variables

- 1.1 Definición de variables irrelevantes y redundantes.
- 1.2 Métodos de selección

##### 2. Imputación de datos perdidos o faltantes

- 2.1 Imputación simple
- 2.2 Imputación múltiple

##### 3. Diseño del modelo de aprendizaje máquina o profundo

- 3.1 Elección de modelo
- 3.3 Pipeline
- 3.4 Train-Test split

#### ACTIVIDADES DE APRENDIZAJE

Desarrollo práctico de procesos y métodos que apliquen los conocimientos adquiridos respecto al análisis multivariante y de datos masivos.

#### CRITERIOS Y PROCEDIMIENTOS DE EVALUACIÓN Y ACREDITACIÓN

Exámenes parciales y final. Tareas Simulaciones en computadora. Proyectos. Esto tendrá una equivalencia del 100% en la calificación final del semestre

#### BIBLIOGRAFÍA (TIPO, TÍTULO, AUTOR, EDITORIAL Y AÑO)

##### Básica:

1. A First Course in Multivariate Statistics. Bernard Flury. Springer 1997.
2. An Introduction to Multivariate Statistical Analysis (3ra. Ed.). T. W. Anderson. Wiley 2003.
3. Big Data: Principles and Best Practices of Scalable Real-Time Data Systems" por Nathan Marz y James Warren.

**Consulta:**

1. "Learning Spark: Lightning-Fast Data Analytics" por Holden Karau, Andy Konwinski, Patrick Wendell y Matei Zaharia
1. Using Multivariate Statistics (7ma. Ed.) Barbara G. Tabachnick & Linda S. Fidell. Pearson 2021.
2. Ortega, J.M. (2023). Big data, machine learning y data science en python. Editorial. RA-MA, S.A. 669 pp

**PERFIL PROFESIONAL DEL DOCENTE**

Maestría o Doctorado en Computación, Informática, Matemáticas Aplicadas, Estadística, Ciencia de Datos o afines, con conocimientos en el uso de software como Python y R.

**Vo.Bo**

M.T.C.A. ERIK GERMÁN RAMOS PÉREZ  
COORDINADOR DE LA UNIVERSIDAD VIRTUAL

**AUTORIZÓ**

L.I. MARIO ALBERTO MORENO ROCHA  
VICE-RECTOR ACADÉMICO