



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA SINTÉTICO

UNIDAD ACADÉMICA: UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS COAHUILA (UPIIC), ESCUELA SUPERIOR DE CÓMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS TLAXCALA (UPIIT)	
PROGRAMA ACADÉMICO: Licenciatura en Ciencias de Datos	
UNIDAD DE APRENDIZAJE: Estadística avanzada	SEMESTRE: VI, VII

PROPÓSITO DE LA UNIDAD DE APRENDIZAJE:
 Construye modelos estadísticos a partir de la regresión múltiple, los modelos lineales generalizados, los mínimos cuadrados penalizados y la estadística en altas dimensiones.

CONTENIDOS:	I. Aprendizaje estadístico en altas dimensiones II. Modelos estadísticos de regresión múltiple III. Mínimos cuadrados penalizados IV. Modelos lineales generalizados y verosimilitud penalizada V. Inferencia estadística en altas dimensiones
--------------------	--

ORIENTACIÓN DIDÁCTICA:	Métodos de enseñanza	Estrategias de aprendizaje	
	a) Inductivo	X	a) Estudio de casos
	b) Deductivo	X	b) Aprendizaje basado en problemas
	c) Analógico		c) Aprendizaje orientado proyectos

EVALUACIÓN Y ACREDITACIÓN:	Diagnóstica	X	Saberes Previamente Adquiridos	X
	Solución de casos		Organizadores gráficos	
	Problemas resueltos	X	Problemarios	
	Reporte de proyectos		Reporte de seminarios	
	Reportes de indagación	X	Otras evidencias a evaluar: conclusión de discusión, uso de software.	
	Reportes de prácticas	X		
Evaluación escrita	X			

	Autor(es)	Año	Título del documento	Editorial / ISBN
BIBLIOGRAFÍA BÁSICA:	Wahmann, P. & van de Veer, S.	2011	<i>Statistics for High-Dimensional Data. Methods, Theory and Applications</i>	Springer / 9783642201912
	Efron, B. & Hastie, T.	2016	<i>Computer Age Statistical Inference: Algorithms, Evidence and Data Science</i>	Cambridge University Press / 9781107149000
	Fan, J., L. Runze, Zhang, C-H, Zou, H.	2020	<i>Statistical Foundations of Data Science</i>	CRC Press / 9781466510845
	Giraud, C.	2014	<i>Introduction to High-Dimensional Statistics</i>	Chapman and Hall / 9781482237948
	James, G; Witten, D; Hastie, T. & Tibshirani, R.	2017	<i>An Introduction to Statistical Learning: With Applications in R</i>	Springer / 9781461471370



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA DE ESTUDIOS

UNIDAD DE APRENDIZAJE: Estadística avanzada

HOJA 2 DE 7

UNIDAD ACADÉMICA: UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS COAHUILA (UPIIC), ESCUELA SUPERIOR DE CÓMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS TLAXCALA (UPIIT)

PROGRAMA ACADÉMICO: Licenciatura en Ciencia de Datos

SEMESTRE: VI, VII	ÁREA DE FORMACIÓN: Profesional	MODALIDAD: Escolarizada
-----------------------------	--	-----------------------------------

TIPO DE UNIDAD DE APRENDIZAJE:
Teórica-Práctica/Optativa

VIGENTE A PARTIR DE: Agosto 2022/ Enero 2023	CRÉDITOS:	
	Tepic: 7.5	SATCA: 6.3

INTENCIÓN EDUCATIVA

Esta unidad contribuye al perfil de egreso de la Licenciatura en Ciencia de Datos con las habilidades de análisis y aplicación de herramientas y modelos estadísticos avanzadas para la Ciencia de datos. Asimismo, se desarrollan habilidades transversales como la toma de decisiones basada en estadística, la responsabilidad y el trabajo en equipo con un alto sentido ético.

La presente unidad se relaciona de manera antecedente con Estadística, Programación para ciencia de datos, Aprendizaje máquina e inteligencia artificial y Matemáticas avanzadas para ciencia de datos, de manera lateral con Modelado predictivo y Análisis de series de tiempo y consecuente con Temas selectos de aprendizaje profundo, Temas selectos de inteligencia artificial y Simulación básica.

PROPÓSITO DE LA UNIDAD DE APRENDIZAJE

Construye modelos estadísticos a partir de la regresión múltiple, los modelos lineales generalizados, los mínimos cuadrados penalizados y la estadística en altas dimensiones.

<p align="center">TIEMPOS ASIGNADOS</p> <p>HORAS TEORÍA/SEMANA: 3.0</p> <p>HORAS PRÁCTICA/SEMANA: 1.5</p> <p>HORAS TEORÍA/SEMESTRE: 54.0</p> <p>HORAS PRÁCTICA/SEMESTRE: 27.0</p> <p>HORAS APRENDIZAJE AUTÓNOMO: 24.0</p> <p>HORAS TOTALES/SEMESTRE: 81.0</p>	<p align="center">UNIDAD DE APRENDIZAJE DISEÑADA POR:</p> <p align="center">Comisión de Diseño del Programa Académico</p>	<p align="center">AUTORIZADO Y VALIDADO POR:</p> <div style="background-color: black; height: 150px; width: 100%;"></div> <p align="center">Director de Educación Superior</p>
	<p align="center">APROBADO POR:</p> <p align="center">Comisión de Programas Académicos del H. Consejo General Consultivo del IPN</p> <p align="center">27/06/2022</p>	




INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Estadística avanzada

HOJA 3 DE 7

UNIDAD TEMÁTICA I Aprendizaje estadístico en altas dimensiones	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Explica las problemáticas de la inferencia estadística en altas dimensiones, a partir del impacto de la dimensionalidad y el análisis de componentes principales.	1.1. Datos con alta dimensionalidad 1.1.1. Contextos donde surge el aprendizaje estadístico de altas dimensiones. 1.1.2. Impacto de la dimensionalidad y la problemática en la inferencia estadística clásica 1.1.3. Objetivos principales del aprendizaje estadístico en altas dimensiones	1.5	1.5	1.5
	1.2. Análisis de Componentes Principales (PCA) en altas dimensiones 1.2.1. Perspectiva de máxima varianza para PCA 1.2.2. Perspectiva de proyección para PCA 1.2.3. Cálculo de vectores propios y aproximaciones de rango menor 1.2.4. El método de PCA en altas dimensiones 1.2.5. Perspectiva de variables latentes para PCA	4.5	3.0	1.5
Subtotal		6.0	4.5	3.0

UNIDAD TEMÁTICA II Modelos estadísticos de regresión múltiple	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Examina la regresión a partir de modelos de expansión y regresión en cordillera.	2.1 Regresión lineal múltiple 2.1.1 Teorema de Gauss Markov. 2.1.2 Pruebas estadísticas en regresión lineal múltiple 2.1.3 Mínimos cuadrados con pesos	3.0	1.5	1.5
	2.2 Construcción de modelos con base de expansión 2.2.1 Regresión polinomial 2.2.2 Regresión con splines 2.2.3 Covariables múltiples	3.0	1.5	1.5
	2.3 Regresión en cordillera 2.3.1 Compromiso sesgo-varianza 2.3.2 Mínimos cuadrados penalizado l2 2.3.3 Interpretación bayesiana 2.3.4 Regresión en cordillera con núcleo	3.0	1.5	1.5
	2.4 Validación cruzada 2.4.1 Dejando fuera una observación 2.4.2 Generalizada	3.0	1.5	1.5
Subtotal		12.0	6.0	6.0



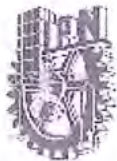
INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



NIDAD DE APRENDIZAJE: Estadística avanzada

HOJA 4 DE 7

UNIDAD TEMÁTICA III Mínimos cuadrados penalizados	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Analiza la selección de variables a partir de mínimos cuadrados penalizados y Lasso.	3.1. Criterio clásico de selección de variables	4.5	1.5	1.5
	3.1.1. Selección de subconjunto			
	3.1.2. Relación con regresión penalizada			
	3.1.3. Selección de parámetros de regularización			
	3.2. Mínimos cuadrados penalizados de plegado cóncavo	4.5	1.5	1.5
	3.2.1. Diseños ortonormales			
	3.2.2. Funciones de penalidad			
	3.2.3. Umbrales			
	3.2.4. Propiedades de riesgo			
	3.3. Lasso y regularización L1	4.5	3.0	1.5
3.3.1. Garrote				
3.3.2. Lasso modelo clásico				
3.3.3. Lasso modelo adaptivo				
3.3.4. Red elástica				
3.3.5. Selector Dantzing				
Subtotal		13.5	6.0	4.5
UNIDAD TEMÁTICA IV Modelos lineales generalizados y verosimilitud penalizada	CONTENIDO	HORAS CON DOCENTE		HRS AA
UNIDAD DE COMPETENCIA Aplica modelos lineales generalizados a partir de algoritmos y propiedades.	4.1. Modelos lineales generalizados	3.0	1.5	1.0
	4.1.1. Familia exponencial			
	4.1.2. Elementos de modelos lineales generalizados			
	4.1.3. Máxima verosimilitud			
	4.1.4. Cálculo de mínimos cuadrados repesados iterativo			
	4.1.5. Análisis de residuales			
	4.2. Aplicaciones y modelos	3.0	1.5	1.0
	4.2.1. Modelos Bernoulli y binomial			
	4.2.2. Modelos para respuestas de conteo			
	4.2.3. Modelos para respuestas continuas no negativas			
	4.2.4. Modelos con errores normales			
	4.3. Algoritmos	3.0	3.0	1.0
	4.3.1. Aproximación local cuadrática			
	4.3.2. Aproximación local lineal			
	4.3.3. Mayorización y minimización adaptiva iterativa local			
	4.4. Propiedades de muestreo en dimensión baja	3.0		1.5
	4.4.1. Condiciones de regularidad			
4.4.2. Propiedad Oracle				
4.4.3. Propiedades de muestreo con dimensión divergente				
4.5. Propiedades de muestreo en ultra dimensiones	1.5			
4.5.1. Estimador Lasso penalizado y propiedad de riesgo				
4.5.2. Propiedad Oracle fuerte				
Subtotal		13.5	6.0	6.0



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



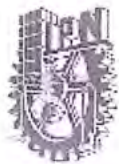
UNIDAD DE APRENDIZAJE: Estadística avanzada

HOJA 5 DE 7

UNIDAD TEMÁTICA V Inferencia estadística en altas dimensiones	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Realiza inferencia estadística en altas dimensiones a partir de la regresión lineal, modelos lineales generalizados y modelos gráficos gaussianos.	5.1 Inferencia en regresión lineal	3.0	1.5	1.5
	5.1.1 Sesgo desvanecido de estimadores de regresión			
	5.1.2 Elección de pesos			
	5.1.3 Inferencia para el nivel de ruido			
	5.2 Inferencia en modelos lineales generalizados	3.0	1.5	1.5
	5.2.1 Lasso desespesificado			
	5.2.2 Estimador de la función de score de correlacionada			
	5.3 Modelos gráficos gaussianos	3.0	1.5	1.5
	5.3.1 Inferencia vía mínimos cuadrados penalizado			
	5.3.2 Tamaño de muestra en regresión y modelos gráficos			
Subtotal		9.0	4.5	4.5

ESTRATEGIAS DE APRENDIZAJE	EVALUACIÓN DE LOS APRENDIZAJES
Estrategias de Aprendizaje Basado en Problemas El estudiante desarrollará las siguientes actividades: <ol style="list-style-type: none"> 1. Indagación documental acerca de distribuciones de probabilidad a partir de datos 2. Resolución de problemas prácticos relacionados con los temas vistos en las unidades temáticas 3. Discusión grupal dirigida de preguntas estratégicas preparadas por el profesor que ayuden al alumno a construir los conocimientos con base a la teoría 4. Uso de software de acceso libre para realizar análisis estadístico de datos 5. Realización de prácticas 	Evaluación diagnóstica. Portafolio de evidencias: <ol style="list-style-type: none"> 1. Reporte de indagación 2. Problemas resueltos 3. Conclusión de discusión 4. Solución de problemas con uso de software 5. Reporte de prácticas 6. Evaluación escrita

RELACIÓN DE PRÁCTICAS			
PRÁCTICA No.	NOMBRE DE LA PRÁCTICA	UNIDADES TEMÁTICAS	LUGAR DE REALIZACIÓN
1	Alta dimensionalidad en ciencia, tecnología y negocios	I	Sala de cómputo
2	Base de dígitos MINIST y reconstrucción con PCA	I	
3	Generación de nuevos datos usando variables latentes	I	
4	Transformaciones Box-Cox de la variable respuesta	II	
5	Pruebas estadísticas en regresión lineal múltiple	II	
6	Algoritmos numéricos para mínimos cuadrados penalizados	III	
7	Regularización de parámetros para mínimos cuadrados penalizados	III	
8	Selección de variables vía verosimilitud penalizada	IV	
9	Calibración de parámetros de selección	IV	
10	Comparación numérica entre pruebas estadísticas en modelos lineales generalizados	V	
11	Aplicación de prueba de verosimilitud parcial penalizada en un análisis empírico	V	
		TOTAL DE HORAS: 27.0	



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Estadística avanzada

HOJA 6 DE 7

Bibliografía																		
Tipo	Autor(es)	Año	Título del documento	Editorial	Documento													
					Libro	Antología	Otros	Texto	Simulador	Imagen	Tutorial	Video	Presentación	Diccionario	Otro			
C	Agresti, A.	2015	<i>Foundations of Linear and Generalized Linear Models</i>	Wiley / 9781118730034	X													
B	Bühlmann, P. & van de Geer, S.	2011	<i>Statistics for High-Dimensional Data. Methods, Theory and Applications</i>	Springer / 9783642201912	X													
C	Deisenroth, M. P., Faisal, A. A. & Ong, C. S.	2020	<i>Mathematics for Machine Learning</i>	Cambridge University Press / 9781108470049	X													
B	Efron, B. & Hastie, T.	2016	<i>Computer Age Statistical Inference: Algorithms, Evidence and Data Science</i>	Cambridge University Press / 9781107149892	X													
B	Fan, J., L. Runze, Zhang, C-H, Zou, H.	2020	<i>Statistical Foundations of Data Science</i>	CRC Press/ 9781466510845	X													
B	Giraud, C.	2014	<i>Introduction to High-Dimensional Statistics</i>	Chapman and Hall / 9781482237948	X													
B	James, G; Witten, D; Hastie, T. & Tibshirani, R.	2017	<i>An Introduction to Statistical Learning: With Applications in R</i>	Springer / 9781461471370	X													
C	McCullagh, P. & Nelder, J.A.	1989*	<i>Generalized Linear Models</i>	CRC Press / 9780412317606	X													
Autor, año, título y Dirección Electrónica																		
THE MNIST DATABASE of handwritten digits, Y. LeCun, C. Cortes, C.J.C. Burges. Recuperado el 30 de mayo de 2021 de: http://yann.lecun.com/exdb/mnist/					X													
How to develop LASSO Regression Models in Python. Recuperado el 30 de mayo de 2021 de: https://machinelearningmastery.com/lasso-regression-with-python/																		X
Interactive Chaos: Making things simple. Recuperado el 30 de mayo de 2021 de: https://interactivechaos.com/es/manual/tutorial-de-machine-learning/presentacion																		
Wolfram: Machine Learning. Recuperado el 30 de mayo de 2021 de: https://www.wolfram.com/featureset/machine-learning/?source=nav																		
UCI Machine Learning Repository, recuperado el 30 de mayo de 2021 de: https://archive.ics.uci.edu/ml/index.php																		
Wolfram Alpha LLC. (2019). Wolfram Alpha computational intelligence. England. Recuperado el 3 de noviembre de 2019, de: www.wolframalpha.com																		

*Bibliografía clásica



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Estadística avanzada

HOJA 7 DE 7

PERFIL DOCENTE: Ingeniería, Licenciatura y/o Maestría en Ciencias Físico Matemáticas o afines

EXPERIENCIA PROFESIONAL	CONOCIMIENTOS	HABILIDADES DIDÁCTICAS	ACTITUDES
Preferentemente 2 años en el área de ciencias básicas relacionadas con matemáticas o ingeniería. Mínima de dos años de docencia a Nivel Superior.	En estadística a nivel superior y uso de softwares para análisis estadístico de datos. Del Modelo Educativo Institucional.	Discursivas Cognoscitivas Metodológicas De conducción del grupo Para evaluar Coordinar grupos de aprendizaje Organizar equipos de aprendizaje Planificación de enseñanza Manejo de estrategias didácticas centradas en el aprendizaje Manejo de TIC en enseñanza y para el aprendizaje	Responsabilidad Tolerancia Honestidad Respeto Paciencia Disciplina Constancia

ELABORÓ

Dr. Víctor Manuel Pérez Abreu
Carrión
Profesor Coordinador