



### Formato para registro de Unidades de aprendizaje 2021

I.- Datos de identificación de la unidad de aprendizaje

<b>Unidad académica:</b>	Centro de Investigación en Computación (CIC)										
<b>Programa académico:</b>	Doctorado en Ciencias de la Computación										
	<input checked="" type="checkbox"/>	Doctorado				Orientación profesional					
		Maestría			<input checked="" type="checkbox"/>	Orientado a la investigación					
		Especialidad				Con la industria					
						Especialidad médica					
<b>Nombre de unidad de aprendizaje:</b>	Sesión de colegio donde se propuso:		Ordinaria 7, 2023			Fecha de propuesta:		26 de julio de 2023			
	<b>Técnicas avanzadas para la clasificación de patrones</b>										
	Clave de la unidad de aprendizaje:		<b>23B8376</b>			Créditos:		5 <i>REP 2017</i>			
	Semanas del semestre		18	Horas a la semana:		4	Horas totales:		72		
<b>Tipo de unidad de aprendizaje:</b>	Obligatoria:		Optativa:		<input checked="" type="checkbox"/>	Observaciones:					
	Semestre:										
	Teórica (%):	100	Práctica (%):			Teórico-prácticas (%):					
<b>Área del conocimiento:</b>	Ingeniería y Ciencias Fisicomatemáticas		<input checked="" type="checkbox"/>	Ciencias Sociales y Administrativas			Ciencias Médico Biológicas		Interdisciplinario		
<b>Modalidad no escolarizada:</b>	No escolarizada			Nombre de la Plataforma:							
	Mixta			Presencial (%):			En plataforma (%):				
<b>Horas establecidas en el programa de estudios:</b>	Presenciales (si procede) (horas x semana)						En plataforma (horas x semana):				



### Formato para registro de Unidades de aprendizaje 2021

#### I. Aprendizajes que el estudiante deberá demostrar al finalizar

Conocimientos	Habilidades y destrezas	Actitudes y valores
<ul style="list-style-type: none"> <li>Algoritmos de técnicas avanzadas para la clasificación de patrones.</li> </ul>	<ul style="list-style-type: none"> <li>Implementar, depurar y aplicar algoritmos de técnicas avanzadas para la clasificación de patrones en la solución de problemas de Clasificación Inteligente de Patrones, Cómputo Inteligente, Inteligencia Artificial, Reconocimiento de Patrones, Ciencia de Datos, Minería de Datos y áreas afines.</li> </ul>	<ul style="list-style-type: none"> <li>Independencia</li> <li>Creatividad</li> <li>Trabajo colaborativo</li> <li>Responsabilidad</li> </ul>

#### Resolución que aborda la propuesta con su enfoque disciplinar

En las unidades de aprendizaje relacionadas con Clasificación Inteligente de Patrones, Cómputo Inteligente, Aprendizaje Automático, Inteligencia Artificial, Reconocimiento de Patrones, Ciencia de Datos, Minería de Datos y áreas afines, es típico que las restricciones de tiempo obstaculizan el estudio de ciertos algoritmos de técnicas avanzadas. En esta unidad de aprendizaje se pretende cubrir esos relevantes temas; por ello, se abordarán algoritmos de técnicas avanzadas, tales como: selección de atributos, balanceo de clases en conjuntos de datos, imputación, matriz de confusión, conversión y tests estadísticos, entre otros.

#### II. Proximidad formativa

Áreas multi, inter y transdisciplinarias	Líneas de Generación y Aplicación de Conocimiento	Sectores sociales
<ul style="list-style-type: none"> <li>Ciencias de la Computación</li> <li>Clasificación Inteligente de Patrones</li> <li>Cómputo Inteligente</li> <li>Aprendizaje Automático</li> <li>Inteligencia Artificial</li> <li>Reconocimiento de Patrones</li> <li>Ciencia de Datos</li> <li>Minería de Datos</li> <li>Aprendizaje Profundo</li> <li>Metaheurísticas</li> </ul>	<ul style="list-style-type: none"> <li>Inteligencia Artificial y Cómputo Científico</li> </ul>	<ul style="list-style-type: none"> <li>Los algoritmos de técnicas avanzadas para la clasificación de patrones pueden incidir en cualquier sector social, con énfasis en temas sociales sensibles como la salud, aspectos financieros o aplicaciones en distintas actividades relevantes para el ser humano.</li> </ul>



### Formato para registro de Unidades de aprendizaje 2021

Estrategia de asociación:  
El estudiante realizará investigaciones documentales en repositorios donde hay conjuntos de datos propios para la investigación científica. Estos repositorios contienen datos relacionados con los sectores sociales en los que incide esta unidad de aprendizaje, haciendo énfasis en temas sociales sensibles como la salud, aspectos financieros o aplicaciones en distintas actividades relevantes para el ser humano. Las experiencias obtenidas en la implementación, la depuración y la aplicación de algoritmos de técnicas avanzadas para la clasificación de patrones se conciliarán y se compartirán con estudiantes y profesionales de otras disciplinas afines, como Cómputo Inteligente, Inteligencia Artificial, Reconocimiento de Patrones, Ciencia de Datos y Minería de Datos, entre otras.

#### III Metodología de enseñanza – aprendizaje

Descripción

Evidencias como proceso de aprendizaje

Evidencias integradoras (resultados que contribuyen al curriculum)	Ponderación



### Formato para registro de Unidades de aprendizaje 2021

#### IV. Descripción de la participación esperada en el estudiante

Receptiva	Resolutiva	Autónoma	Estratégica



## Formato para registro de Unidades de aprendizaje 2021

### Contenido temático

Tema 1. Breve repaso de los conceptos básicos de la clasificación inteligente de patrones	8 horas
1.1 Datasets y complejidad de datos	
1.2 Métodos de validación de modelos	
1.3 Bases conceptuales y algoritmos de clasificación (principales enfoques)	
1.4 Matriz de confusión y medidas de desempeño	
Tema 2. Técnicas avanzadas relacionadas con la matriz de confusión y las medidas de desempeño	8 horas
2.1 Casos extremos y patológicos de sensitivity, specificity, precisión y F1-measure	
2.2 Curvas ROC y AUC	
2.3 Estadística Kappa	
2.4 Ejemplos y aplicaciones	
Tema 3. Técnicas avanzadas para feature selection y object selection	8 horas
3.1 Métodos filter	
3.2 Métodos wrapper y embedded	
3.3 Métodos híbridos	
3.4 Ejemplos y aplicaciones	
Tema 4. Técnicas avanzadas para balanceo de datasets	8 horas
4.1 Métodos de oversampling (ADASYN, ADOMS, ROS, SMOTE, SMOTE-BL, SMOTE-SL)	
4.2 Métodos de undersampling (RUS, CNNTL, CNN, NCL, OSS, TL, SBC)	
4.3 Métodos híbridos (SPIDER, SPIDER2, SMOTE, ENN, SMOTE-TL)	
4.4 Ejemplos y aplicaciones	
Tema 5. Técnicas avanzadas para conversión e imputación	8 horas
5.1 Técnicas avanzadas para conversión	
5.2 Técnicas avanzadas para imputación	
5.3 Ejemplos y aplicaciones	
Tema 6. Técnicas avanzadas para normalización y discretización	8 horas
6.1 Técnicas avanzadas para normalización	
6.2 Técnicas avanzadas para y discretización	
6.3 Ejemplos y aplicaciones	



## Formato para registro de Unidades de aprendizaje 2021

Tema 7. Técnicas avanzadas para detección de outliers	8 horas
7.1 Principales enfoque para la detección de outliers	
7.2 Métodos representativos para la detección de outliers (Isolation Forest, ODIN, Local Outlier Factor, Z-Score, RANSAC)	
7.3 Ejemplos y aplicaciones	
Tema 8. Técnicas avanzadas para atenuar el solapamiento de clases	8 horas
8.1 Concepto de solapamiento de clases y su efecto en el desempeño de los clasificadores	
8.2 Clases linealmente separables	
8.3 Edición de Wilson y otras técnicas avanzadas para atenuar el solapamiento de clases	
8.4 Ejemplos y aplicaciones	
Tema 9. Pruebas estadísticas	8 horas
9.1 Pruebas paramétricas y no paramétricas	
9.2 Pruebas de Friedman, Wilcoxon y poshoc de Holm	
9.3 Ejemplos y aplicaciones	



### Formato para registro de Unidades de aprendizaje 2021

#### V. Secuencia programática

No.	Tema	Objetivo de aprendizaje / competencia específica	Tiempo/Horas/Semanas	

**Tipo de interacción:** ID–Instrucción directa, TC–Trabajo colaborativo, AC–Análisis en campo, RP–Reflexión personal, PE–Presentación expositiva  
*Nota: Replique esta sección las veces que sea necesario para cubrir toda la secuencia programática*

Indicar solo el número de las *Referencias* indizadas en la sección VII de este documento.

#### VI. Habilitadores tecnológicos

Disposiciones	Especificaciones / descripción de efectos
Conectividad	
Habilidades digitales	
Interoperabilidad	
Datos abiertos	
<i>Big Data</i>	
<i>Machine Learning</i>	
Simulación	



### Formato para registro de Unidades de aprendizaje 2021

<a href="#">Realidad aumentada</a>	
<a href="#">Otro...</a>	

#### Conferencias magistrales

1.
2.
3.

#### Notas complementarias


#### VII. Referencias

#### Documentales / electrónicas

4. Flach, P. (2012). Machine learning: the art and science of algorithms that make sense of data. Cambridge university Press.
5. Aggarwal, C. C. (2015). Data Classification: Algorithms and Applications. Data Mining and Knowledge Discovery Series. CRC Press.
6. García, S., Luengo, J., & Herrera, F. (2015). Data preprocessing in data mining. Springer.
7. Kanji, G. K. (2006). 100 statistical tests. Sage.
8. Berrar, D. (2019). Performance measures for binary classification. Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics, 1-3, pp. 546 - 560
9. Wang S., Wu J., Liu P. (2013) An improved performance measure for AUC. Journal of Bionanoscience, 7 (1), pp. 59 - 65
10. Kushagra, Kumar R., Jain S. (2023) Feature Selection for Medical Diagnosis Using Machine Learning: A Review Lecture Notes in Electrical Engineering, 984 LNEE, pp. 179 - 189
11. Islam M.R., Lima A.A., Das S.C., Mridha M.F., Prodeep A.R., Watanobe Y. (2022) A Comprehensive Survey on the Process, Methods, Evaluation, and Challenges of Feature Selection IEEE Access, 10, pp. 99595 - 99632
12. Jiang, Z., Pan, T., Zhang, C., & Yang, J. (2021). A new oversampling method based on the classification contribution degree. Symmetry, 13(2), 194.
13. Omar, B., Rustam, F., Mehmood, A., & Choi, G. S. (2021). Minimizing the overlapping degree to improve class-imbalanced learning under sparse feature selection: application to fraud detection. IEEE Access, 9, 28101-28110.
14. Aggarwal, C. C., & Aggarwal, C. C. (2017). An introduction to outlier analysis (pp. 1-34). Springer International Publishing.





### Formato para registro de Unidades de aprendizaje 2021

15. Adnan F.A., Jamaludin K.R., Wan Muhamad W.Z.A., Miskon S. (2022) A review of the current publication trends on missing data imputation over three decades: direction and future research. <i>Neural Computing and Applications</i> , 34 (21), pp. 18325 - 18340
16. Li Z., Usman M., Tao R., Xia P., Wang C., Chen H., Li B. (2023) A Systematic Survey of Regularization and Normalization in GANs. <i>ACM Computing Surveys</i> , 55 (11), art. no. 232.

#### VIII. Créditos y responsabilas

Responsabilidad	Nombre completo	Clave de nombramiento /No. de empleado
Coordinador (Autor)	Cornelio Yáñez Márquez	14160-EG-19/6
Participante (Coautor)	Amadeo José Argüelles Cruz	14976-EJ-20/6
Participante (Coautor)	Antonio Alarcón Paredes	15782-EA-22
Participante (Coautor)	Ana María Magdalena Saldaña Pérez	16951-EB-22
Participante (Coautor)	Yenny Villuendas Rey	15344-EC-22
Asesor didáctico / Diseñador Instruccional		
Tecnólogo educativo / Comunicólogo		
Corrector de estilo		
Programador multimedia / Diseñador gráfico		
Otro...		



**Formato para registro de Unidades de aprendizaje 2021**

Por la División de Operación y Promoción al Posgrado de la SIP	Por la Subdirección de Diseño y Desarrollo de la DEV
Nombre _____	Nombre _____
FIRMA _____	FIRMA _____

<b>VERIFICACIÓN PARA SU PUESTA EN OPERACIÓN</b>	<b>REVISIÓN TÉCNICO-PEDAGÓGICA PARA LA MODALIDAD</b>
Por la Dirección de Posgrado	Por la Dirección para la Educación Virtual
Nombre _____	Nombre _____
FIRMA _____	FIRMA _____
SELLO DE VALIDACIÓN	