

PROGRAMA DE CURSO ASTROINFORMÁTICA

A. Antecedentes generales del curso:

Departamento	Departamento de Astronomía (DAS)					
Nombre del curso	Astroinformática	Código	AS4501	Créditos	6	
Nombre del curso en inglés	<i>Astroinformatics</i>					
Horas semanales	Docencia	3	Auxiliares	1,5	Trabajo personal	5,5
Carácter del curso	Obligatorio	X		Electivo		
Requisitos	FI3104: Métodos numéricos para ciencias e ingeniería, MA3403: Probabilidades y Estadística					

B. Propósito del curso:

El curso tiene como propósito que los y las estudiantes formulen soluciones algorítmicas a problemas de análisis de datos astronómicos. Para ello, utilizan herramientas estadísticas de inferencia frecuentista y Bayesiana, tanto paramétricas como no paramétricas, o métodos de aprendizaje computacional para procesar e interpretar datos experimentales u observados, considerando modelos existentes.

Los y las estudiantes resuelven problemas de análisis de datos astronómicos mediante el procesamiento de datos masivos y el uso de modelos estadísticos. Asimismo, investigan y analizan una pregunta de investigación en el contexto de un proyecto, para proponer una solución algorítmica aplicada a un problema, cuyos resultados expone en forma oral.

El curso tributa a las siguientes competencias específicas (CE) y genéricas (CG):

CE4: Caracterizar estadísticamente el significado de las mediciones o simulaciones astronómicas.

CE6: Utilizar y escribir programas computacionales para procesar o visualizar datos, con el fin de analizar y comunicar resultados astronómicos.

CE7: Buscar, acceder y utilizar archivos de datos masivos, para generar información de fenómenos astronómicos con herramientas estadísticas.

CG1: Comunicación académica y profesional

Leer de manera comprensiva, analítica y crítica en español. Asimismo, expresar de forma eficaz, clara, precisa e informada sus ideas, opiniones e indagaciones, adecuándose a diversas situaciones comunicativas académicas y profesionales, tanto en lo oral como en lo escrito.

CG2: Comunicación en inglés

Leer y escuchar de manera comprensiva en inglés una variedad de textos e informaciones sobre temas concretos o abstractos, comunicando experiencias y opiniones, adecuándose a diferentes contextos y a las características de la audiencia.

CG3: Compromiso ético

Actuar de manera responsable y honesta, dando cuenta en forma crítica de sus propias acciones y sus consecuencias, en el marco del respeto hacia la dignidad de las personas y el cuidado del medio social, cultural y natural.

CG4: Trabajo en equipo

Trabajar en equipo, de forma estratégica y colaborativa, en diversas actividades formativas, a partir de la autogestión de sí mismo y de la relación con el otro, interactuando con los demás en diversos roles: de líder, colaborador u otros, según requerimientos u objetivos del trabajo, sin discriminar por género u otra razón.

C. Resultados de aprendizaje:

Competencias específicas	Resultados de aprendizaje
CE4	RA1: Utiliza herramientas estadísticas de inferencia frecuentista y Bayesiana, paramétricas y no paramétricas, para interpretar datos.
CE6	RA2: Aplica métodos de aprendizaje computacional, considerando datos astronómicos masivos y/o multidimensionales (series de tiempo, catálogos), para inferir propiedades físicas de un fenómeno.
CE7	RA3: Recolecta y procesa datos astronómicos masivos, considerando principios de computación de alto rendimiento y reducción del volumen de datos para interpretarlos en el contexto de una hipótesis de trabajo o de un análisis exploratorio.
	RA4: Formula soluciones algorítmicas a problemas de análisis de datos astronómicos mediante el procesamiento de datos masivos y el uso de modelos estadísticos.
Competencias genéricas	Resultados de aprendizaje
CG1	RA5: Comunica, en forma oral a sus pares, los resultados de una solución algorítmica aplicada a un problema, considerando, en su escrito o presentación, criterios de claridad, precisión científica y adaptación del mensaje a la audiencia.

CG2	RA6: Lee en inglés textos, artículos y documentación de librerías para extraer e interpretar información de datos masivos en el contexto de la astroinformática.
CG3	RA7: Realiza tareas y ejercicios reportando adecuadamente, según reglas de citado las fuentes de información utilizadas.
CG4	RA8: Trabaja en equipo en un marco de respeto por las ideas del otro, actuando con diligencia y honestidad para el cumplimiento de la actividad.

D. Unidades temáticas:

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
1	RA1, RA2, RA6	Estadística para astronomía	2,5 semanas
Contenidos		Indicador de logro	
1.1. Distribuciones relevantes para astronomía. 1.2. Estadística paramétrica y no paramétrica. 1.3. Inferencia clásica y Bayesiana en espacios multidimensionales, incluyendo Markov Chain Monte Carlo.		El/la estudiante: 1. Analiza e interpreta tipos de distribuciones estadísticas a partir de ejemplos reales. 2. Mide variables físicas, con intervalos de confianza, a partir de datos observacionales, considerando un modelo estadístico. 3. Cuantifica el alcance de la interpretación de los datos observacionales de forma paramétrica o no paramétrica.	
Bibliografía de la unidad		[1] Cap. 3, 4, 5. [2] Cap. 2, 3, 4, 5. [3] Cap. 2, 3.	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
2	RA1, RA4, RA6, RA7	Datos astronómicos	1,5 semanas
Contenidos		Indicador de logro	
2.1. Incertezas: Fuentes de error. 2.2. Datos: Adquisición de datos astronómicos. 2.2.1. Principios básicos de bases de datos. 2.2.2. Aplicaciones de bases de datos astronómicas.		El/la estudiante: 1. Identifica y caracteriza fuentes de error típicas en datos astronómicos. 2. Identifica fuentes de datos astronómicos, accediendo y recogiendo muestras representativas para la resolución de problemas astronómicos. 3. Determina, en un modelo estadístico (por ejemplo, distribuciones normales o de <i>Poisson</i>), las componentes determinísticas y aleatorias que describen los datos. 4. Formula soluciones algorítmicas para la interpretación de datos astronómicos.	

	5. Maneja, de manera responsable, las fuentes de información utilizada, citándolas cuando corresponda.
Bibliografía de la unidad	[2] Cap. 1, Apéndice C, D. [3] Cap. 7.

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
3	RA2, RA3, RA4, RA6, RA8	Aprendizaje computacional	3,5 semanas
Contenidos		Indicador de logro	
3.1. Estimación de densidad. 3.2. Métricas para clasificación: exactitud y precisión, falsos positivos y negativos, F1-score, curva ROC, matriz de confusión, desbalanceo. 3.3. Métodos de clasificación supervisados y no supervisados. 3.4. Regresión.		El/la estudiante: <ol style="list-style-type: none"> 1. Construye una muestra para métodos supervisados, utilizando bases de datos aplicables a un problema. 2. Elabora un algoritmo para la resolución de un problema. 3. Ejecuta métodos de validación cruzados, determinando la aplicabilidad del algoritmo desarrollado a un conjunto de datos en particular. 4. Aplica algoritmos de agrupamiento, estimación de densidad, clasificación o regresión para el análisis de grandes volúmenes de datos, considerando la familia de algoritmos supervisados o no supervisados. 5. Compara métodos de clasificación o regresión utilizando métricas asociadas, considerando las limitaciones y ventajas de cada uno. 	
Bibliografía de la unidad		[1] Cap. 6, 7, 8, 9.	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
4	RA1, RA2, RA4, RA6	Herramientas de análisis temporal y espacial	2,5 semanas
Contenidos		Indicador de logro	
4.1. Variabilidad y escalas de tiempo en astronomía. 4.2. Procesos estocásticos, autocorrelación y modelos autoregresivos. 4.3. Variabilidad periódica: transformadas de Fourier, <i>wavelets</i> y periodogramas. 4.4. Análisis espacial, correlación espacial y <i>teselaciones</i> . 4.5. Interpolación temporal y espacial mediante procesos gaussianos.		El/la estudiante: <ol style="list-style-type: none"> 1. Clasifica tipos de variabilidad de forma fenomenológica o basada en modelos físicos. 2. Selecciona modelos autorregresivos, analizando fenómenos astrofísicos según su naturaleza estocástica. 3. Calcula periodogramas con datos a partir de series de tiempo astronómicas para la determinación de períodos. 4. Aplica métodos de análisis de correlación espacial, considerando parámetros correspondientes a los datos. 5. Utiliza procesos gaussianos para interpolar datos temporales o espaciales. 	
Bibliografía de la unidad		[1] Cap. 10, Apéndice E. [2] Cap. 11.	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
5	RA2, RA3, RA6, RA8	Computación de alto rendimiento	1,5 semanas
Contenidos		Indicador de logro	
5.1. Principios básicos de computación de alto rendimiento (<i>High performance computing</i>). 5.2. Distribución de tareas y flujos de trabajo (<i>pipelines</i>). 5.3. Procesamiento de datos masivos.		El/la estudiante: <ol style="list-style-type: none"> 1. Aplica herramientas de computación de alto rendimiento ejecutando tareas básicas para el procesamiento de datos masivos de forma distribuida. 2. Explica la estructura de tareas y flujos de trabajo (<i>pipelines</i>), justificando una serie de pasos lógicos para ejecutar un algoritmo. 3. Trabaja en equipo en un ejercicio aplicado, actuando con honestidad y diligencia. 	
Bibliografía de la unidad		[1] Cap. 2 [5] Cap. 1, 2, 6 y tutoriales.	

Número	RA al que tributa	Nombre de la unidad	Duración en semanas
6	RA1, RA2, RA3, RA4, RA5, RA6, RA7, RA8	Desarrollo de un proyecto en el contexto de la astroinformática	3,5 semanas
Contenidos		Indicador de logro	
5.1. Trabajo de proyecto en astroinformática. 5.2. Planteamiento de un problema. 5.3. Revisión y análisis de las variables. Revisión de fuentes bibliográficas sobre el tema a investigar. 5.4. Metodología de ejecución del proyecto (por ejemplo, Metodologías ágiles). 5.5. Acciones para una planificación y distribución de tareas a desarrollar. 5.6. Propuesta y desarrollo de investigación. 5.7. Manejo de datos. 5.8. Análisis estadístico e interpretación física.		El/la estudiante: <ol style="list-style-type: none"> 1. Analiza una pregunta de investigación en el contexto de un problema de investigación acotado, considerando datos disponibles y su origen, métodos estadísticos y métodos computacionales. 2. Recoge e Interpreta información de las fuentes o textos consultados como antecedente para respaldar la investigación y su posible solución. 3. Propone una solución algorítmica a un problema de investigación acotado, considerando datos y métodos de procesamiento. 4. Organiza y distribuye con su equipo las tareas para ejecutar el proyecto en un marco de respeto y responsabilidad. 5. Expone oralmente los resultados de la solución algorítmica, considerando un planteamiento claro del tema, una justificación de las metodologías usadas, y el alcance de dicha solución. 	
Bibliografía de la unidad		Página https://agilemanifesto.org/ https://www.nature.com/articles/d41586-019-01184-9 .	

E. Estrategias de enseñanza - aprendizaje:

El curso considera las siguientes estrategias:

- Clases expositivas y apoyadas por computación interactivas (Jupyter notebooks).
- Resolución de problemas.
- Exposiciones
- Trabajo guiado personal y en equipo.

F. Estrategias de evaluación:

El curso considera distintas instancias de evaluación:

- Miniejercicios semanales (cuestionarios). Se busca que los y las estudiantes puedan demostrar aprendizajes a partir de la lectura de diferentes textos y documentos, extrayendo conceptos aplicables al desarrollo del proyecto.
- Tareas.
- Proyecto (se organiza en torno a reuniones periódicas donde se revisan los avances para la exposición final).

De todas formas, al inicio del semestre se informará sobre el tipo de evaluación y la ponderación que se asignará a cada evaluación.

Dentro de los temas relacionados con las lecturas estas versarán sobre aprendizaje computacional, naturaleza de datos astronómicos, distribuciones estadísticas e inferencia, herramientas de análisis temporal y espacial, computación de alto rendimiento, entre otros.

G. Recursos bibliográficos:

Bibliografía obligatoria:

- [1] Željko, I. et al. *Statistics* (2014). *Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data: A Practical Python Guide for the Analysis of Survey Data*. Princeton University Press.
- [2] Feigelson, E.D., and Babu, G.J. (2012). *Modern statistical methods for astronomy: with r applications*. Cambridge University Press.
- [3] Wall, J. V., and Jenkins, C. R. (2012). *Practical Statistics for Astronomers*. Cambridge Observing Handbooks for Research and Astronomers: 2nd edition.

Bibliografía complementaria:

- [4] Hastie, Tibshirani & Friedman (2009). *The Elements of Statistical Learning*, Springer.
- [5] Eijkhout, V., et al (2015). *Introduction to High Performance Scientific Computing*.
- [6] Murphy, K. (2012). *Machine Learning: A Probabilistic Perspective*.

H. Datos generales sobre elaboración y vigencia del programa de curso:

Vigencia desde:	Otoño, 2021
Elaborado por:	Francisco Förster, Valentino González
Validado por:	Validación CTD de Astronomía
Revisado por:	Área de Gestión Curricular