



UNIVERSIDAD DE CASTILLA - LA MANCHA

GUÍA DOCENTE

1. DATOS GENERALES

Asignatura: BIOINFORMÁTICA Y BIG DATA	Código: 60622
Tipología: OBLIGATORIA	Créditos ECTS: 6
Grado: 402 - GRADO EN BIOTECNOLOGÍA	Curso académico: 2020-21
Centro: 601 - ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRONÓMICA Y DE MONTES Y BIOTECNOLOGÍA	Grupo(s): 10
Curso: 3	Duración: Primer cuatrimestre
Lengua principal de impartición: Español	Segunda lengua:
Uso docente de otras lenguas: Inglés	English Friendly: S
Página web: campusvirtual.uclm.es	Bilingüe: N

Profesor: LUIS DE LA OSSA JIMENEZ - Grupo(s): 10				
Edificio/Despacho	Departamento	Teléfono	Correo electrónico	Horario de tutoría
ESII / 0.A.12	SISTEMAS INFORMÁTICOS	2413	luis.delaossa@uclm.es	Consultar https://www.esiiab.uclm.es/tutorias.php

2. REQUISITOS PREVIOS

Haber cursado y superado la asignatura "Estadística y métodos computacionales".

3. JUSTIFICACIÓN EN EL PLAN DE ESTUDIOS, RELACIÓN CON OTRAS ASIGNATURAS Y CON LA PROFESIÓN

En esta asignatura se proporcionará a los alumnos los conocimientos necesarios para resolver algunos problemas biológicos mediante métodos computacionales. Por otra parte, se abordará el desarrollo de aplicaciones para análisis de datos biológicos y para la resolución de problemas derivados del uso de volúmenes masivos de datos. En este sentido, también se hará una introducción a las técnicas de aprendizaje automático. Por último, se estudiarán los principios básicos de tratamiento de imagen digital, y las problemáticas y técnicas de procesamiento de imagen más comunes en biología.

4. COMPETENCIAS DE LA TITULACIÓN QUE LA ASIGNATURA CONTRIBUYE A ALCANZAR

Competencias propias de la asignatura

Código	Descripción
CB01	Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio
CB02	Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio
CB03	Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética
CB04	Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado
CB05	Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía
CE14	Conocer el manejo de bases de datos biológicos, bioquímicos y genéticos.
CG01	Capacidad de organización y planificación.
CG02	Capacidad de análisis y síntesis.
CG03	Capacidad para trabajar en equipos multidisciplinares de forma colaborativa y con responsabilidad compartida.
CT01	Conocer una segunda lengua extranjera.
CT02	Conocer y aplicar las Tecnologías de la Información y la Comunicación (TIC).
CT03	Utilizar una correcta comunicación oral y escrita.
CT04	Conocer el compromiso ético y la deontología profesional.

5. OBJETIVOS O RESULTADOS DE APRENDIZAJE ESPERADOS

Resultados de aprendizaje propios de la asignatura

Descripción

Entender las bases de las ciencias de la computación e informática, saber manejarse en el entorno de los principales sistemas operativos para el desarrollo de operaciones básicas y desarrollar programas sencillos de aplicación en Biotecnología en un lenguaje de alto nivel.

Entender los fundamentos de la Programación Orientada a Objeto.

Entender y manejar estructuras de datos.

Manejar e integrar el software existente para el análisis de secuencias biológicas.

Manejar expresiones regulares para la búsqueda de patrones.

Adquirir la capacidad de leer y escribir ficheros de textos.

Saber diseñar experimentos y ajustar los datos obtenidos por regresión lineal y no lineal con herramientas informáticas.

Saber recuperar y aprovechar la información biotecnológica disponible relacionada con las secuencias biológicas, las estructuras de las biomoléculas, la genómica y la proteómica.

Utilizar los métodos para la entrada y salida de datos.

Conocer los fundamentos de los métodos de mejora y aprovechamiento de esas señales y saber aplicar los métodos elementales mediante herramientas informáticas.

Conocer los fundamentos de los principales métodos de tratamiento de secuencias biológicas y saber aplicarlos mediante herramientas informáticas.

Conocer los principios de adquisición de imágenes y otras señales en el contexto biotecnológico y las causas de su degradación.

Conocer métodos para representación gráfica de los resultados.

Representar datos y realizar representaciones de los mismos.

Saber aplicar herramientas básicas del análisis numérico para la resolución de problemas biológicos, químicos, bioquímicos y biotecnológicos.

Resultados adicionales

Conocer el aprendizaje automático y las posibilidades que ofrece en la resolución de problemas en el campo de la biotecnología.

6. TEMARIO

Tema 1: Introducción a la bioinformática

Tema 2: Procesamiento de datos con Python

Tema 2.1 Introducción al entorno Python y a Jupyter notebook

Tema 2.2 Programación con Python

Tema 2.3 Procesamiento de datos con Pandas

Tema 2.4 Expresiones regulares

Tema 3: Exploración y visualización de datos

Tema 3.1 Principios de visualización

Tema 3.2 Visualización con matplotlib/seaborn

Tema 3.3 Otras herramientas para visualización

Tema 4: Análisis de datos y modelos predictivos con statsmodels

Tema 4.1 Estadística descriptiva

Tema 4.2 Test de hipótesis

Tema 4.3 Regresión lineal y logística

Tema 4.4 Clustering y PCA

Tema 4.5 Otras técnicas de aprendizaje automático.

Tema 5: Procesamiento de imágenes en biología

Tema 5.1 Introducción al tratamiento digital de imágenes

Tema 5.2 Operaciones básicas de procesamiento de imágenes

Tema 5.3 Análisis de imágenes

Tema 5.4 Herramientas

Tema 6: Análisis de secuencias genéticas y proteínas.

Tema 6.1 Repositorios y acceso programático

Tema 6.2 Similitud entre secuencias genéticas. Alineamiento.

Tema 6.3 Predicción de estructuras secundaria (ARN) y terciaria (Proteína).

7. ACTIVIDADES O BLOQUES DE ACTIVIDAD Y METODOLOGÍA

Actividad formativa	Metodología	Competencias relacionadas (para títulos anteriores a RD 822/2021)	ECTS	Horas	Ev	Ob	Descripción
Enseñanza presencial (Teoría) [PRESENCIAL]	Método expositivo/Lección magistral	CB02 CE14 CT02	1	25	N	-	Debido a que la asignatura es de carácter instrumental, parte de las clases de teoría se dedicará a la explicación de conceptos básicos, y otra parte a seminarios sobre las distintas herramientas.
Enseñanza presencial (Prácticas) [PRESENCIAL]	Resolución de ejercicios y problemas	CB02 CB03 CB05 CE14 CT02	0.8	20	N	-	Estas clases se dedicarán a la resolución explicación de ejercicios y casos prácticos.
Enseñanza presencial (Prácticas) [PRESENCIAL]	Prácticas	CB01 CB02 CB03 CB04 CB05 CE14 CG01 CG03 CT01 CT02	0.4	10	N	-	Presentación, explicación y resolución de dudas sobre los trabajos prácticos.
Elaboración de memorias de Prácticas [AUTÓNOMA]	Trabajo dirigido o tutorizado	CB01 CB02 CB03 CB04 CB05 CE14 CG01 CG02 CG03 CT01 CT02 CT03 CT04	1	25	S	S	Fruto de cada práctica, se ha de elaborar una memoria con la descripción del proceso y el análisis de los resultados del trabajo.
Pruebas de progreso [PRESENCIAL]	Pruebas de evaluación		0.16	4	S	S	Se llevarán a cabo dos pruebas de progreso a lo largo del curso. Ambas son recuperables en el examen final.
Tutorías de grupo [PRESENCIAL]	Tutorías grupales	CB04 CT04	0.04	1	N	-	Se planifica una sesión para hacer balance de la marcha del curso a mitad del cuatrimestre.
Estudio o preparación de pruebas [AUTÓNOMA]	Trabajo autónomo		2.6	65	N	-	
Total:			6	150			
Créditos totales de trabajo presencial: 2.4			Horas totales de trabajo presencial: 60				
Créditos totales de trabajo autónomo: 3.6			Horas totales de trabajo autónomo: 90				

Ev: Actividad formativa evaluable

Ob: Actividad formativa de superación obligatoria (Será imprescindible su superación tanto en evaluación continua como no continua)

8. CRITERIOS DE EVALUACIÓN Y VALORACIONES

Sistema de evaluación	Evaluación continua	Evaluación no continua*	Descripción
Pruebas de progreso	50.00%	0.00%	Las pruebas de progreso consisten en un examen escrito y en una entrevista en la que se plantearán ejercicios prácticos. Se realizan dos pruebas, que serán recuperables en el examen final.
Elaboración de memorias de prácticas	15.00%	15.00%	La resolución de cada uno de los casos prácticos debe ser reflejada en una memoria, que será evaluada de manera independiente.
Prueba final	0.00%	50.00%	La prueba final solamente será obligatoria para aquellos que no hayan superado las pruebas de progreso.
Práctico	35.00%	35.00%	A lo largo se proponen 5 problemas o casos prácticos, relativos a cada uno de los temas (excepto el uno).
Total:	100.00%	100.00%	

* En **Evaluación no continua** se deben definir los porcentajes de evaluación según lo dispuesto en el art. 4 del Reglamento de Evaluación del Estudiante de la UCLM, que establece que debe facilitarse a los estudiantes que no puedan asistir regularmente a las actividades formativas presenciales la superación de la asignatura, teniendo derecho (art. 12.2) a ser calificado globalmente, en 2 convocatorias anuales por asignatura, una ordinaria y otra extraordinaria (evaluándose el 100% de las competencias).

Criterios de evaluación de la convocatoria ordinaria:

Evaluación continua:

Es obligatorio obtener una puntuación igual o mayor que 5 sobre 10 en:

- 1) Cada una de las pruebas de progreso.
- 2) Cada una de las prácticas.
- 3) La suma total (que incluye la resolución de problemas o casos).

Los alumnos que no superen las pruebas de progreso, pueden presentarse a la prueba final.

Los alumnos que no superen alguna de las prácticas, pueden volver a entregarlas para su evaluación en la convocatoria ordinaria.

Las prácticas y conjuntos de problemas contienen una parte obligatoria, que se puntúa sobre 10, y una optativa.

Se propondrán fechas de entrega cada práctica y conjunto de problemas. No se penaliza la entrega fuera de plazo. Sin embargo, una vez superada la fecha de entrega correspondiente, es obligatorio hacer la parte optativa, y se puntuará sobre 10 la entrega completa.

Evaluación no continua:

Es obligatorio obtener una puntuación igual o mayor que 5 sobre 10 en:

- 1) La prueba final.
- 2) Cada una de las prácticas.
- 3) La suma total (que incluye la resolución de problemas o casos).

Están exentos de hacer la prueba final los alumnos que hayan superado las pruebas de progreso.

Las prácticas y conjuntos de problemas contienen una parte obligatoria, que se puntúa sobre 10, y una optativa.

Se propondrán fechas de entrega cada práctica y conjunto de problemas. No se penaliza la entrega fuera de plazo. Sin embargo, de cara a la evaluación no continua, es obligatorio hacer la parte optativa, y se puntuará sobre 10 la entrega completa.

Particularidades de la convocatoria extraordinaria:

Es obligatorio obtener una puntuación igual o mayor que 5 sobre 10 en:

- 1) La prueba final.
- 2) Cada una de las prácticas.
- 3) La suma total (que incluye la resolución de problemas o casos).

Están exentos de hacer la prueba los alumnos que hayan superado las pruebas de progreso o el examen final en la convocatoria ordinaria.

Las prácticas y conjuntos de problemas contienen una parte obligatoria, que se puntúa sobre 10, y una optativa.

Se propondrán fechas de entrega cada práctica y conjunto de problemas. No se penaliza la entrega fuera de plazo. Sin embargo, de cara a la evaluación no continua, es obligatorio hacer la parte optativa, y se puntuará sobre 10 la entrega completa.

Particularidades de la convocatoria especial de finalización:

Idéntica a la convocatoria extraordinaria. Las prácticas corresponderán a las realizadas el curso anterior.

9. SECUENCIA DE TRABAJO, CALENDARIO, HITOS IMPORTANTES E INVERSIÓN TEMPORAL

No asignables a temas

Horas	Suma horas
Pruebas de progreso [PRESENCIAL][Pruebas de evaluación]	4
Tutorías de grupo [PRESENCIAL][Tutorías grupales]	1
Estudio o preparación de pruebas [AUTÓNOMA][Trabajo autónomo]	65
Comentarios generales sobre la planificación: Esta planificación es ORIENTATIVA, pudiendo variar a lo largo del periodo lectivo en función de las necesidades docentes, festividades, o por cualquier otra causa imprevista. La planificación semanal de la asignatura podrá encontrarse de forma detallada y actualizada en la plataforma Campus Virtual. La planificación temporal podrá verse modificada ante causas imprevistas.	
Tema 1 (de 6): Introducción a la bioinformática	
Actividades formativas	Horas
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	1
Periodo temporal: Semana 1	

Tema 2 (de 6): Procesamiento de datos con Python	
Actividades formativas	Horas
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	6
Enseñanza presencial (Prácticas) [PRESENCIAL][Resolución de ejercicios y problemas]	4
Enseñanza presencial (Prácticas) [PRESENCIAL][Prácticas]	2
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo dirigido o tutorizado]	5
Periodo temporal: Semanas 1, 2, 3 y 4	
Tema 3 (de 6): Exploración y visualización de datos	
Actividades formativas	Horas
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	2
Enseñanza presencial (Prácticas) [PRESENCIAL][Resolución de ejercicios y problemas]	4
Enseñanza presencial (Prácticas) [PRESENCIAL][Prácticas]	2
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo dirigido o tutorizado]	5
Periodo temporal: Semanas 5 y 6	
Tema 4 (de 6): Análisis de datos y modelos predictivos con statsmodels	
Actividades formativas	Horas
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	6
Enseñanza presencial (Prácticas) [PRESENCIAL][Resolución de ejercicios y problemas]	4
Enseñanza presencial (Prácticas) [PRESENCIAL][Prácticas]	2
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo dirigido o tutorizado]	5
Periodo temporal: Semanas 7, 8 y 9	
Tema 5 (de 6): Procesamiento de imágenes en biología	
Actividades formativas	Horas
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	5
Enseñanza presencial (Prácticas) [PRESENCIAL][Resolución de ejercicios y problemas]	4
Enseñanza presencial (Prácticas) [PRESENCIAL][Prácticas]	2
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo dirigido o tutorizado]	5
Periodo temporal: Semanas 10, 11 y 12	
Tema 6 (de 6): Análisis de secuencias genéticas y proteínas.	
Actividades formativas	Horas
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	5
Enseñanza presencial (Prácticas) [PRESENCIAL][Resolución de ejercicios y problemas]	4
Enseñanza presencial (Prácticas) [PRESENCIAL][Prácticas]	2
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo dirigido o tutorizado]	5
Periodo temporal: Semanas 13, 14 y 15	
Actividad global	
Actividades formativas	Suma horas
Enseñanza presencial (Prácticas) [PRESENCIAL][Resolución de ejercicios y problemas]	20
Enseñanza presencial (Teoría) [PRESENCIAL][Método expositivo/Lección magistral]	25
Enseñanza presencial (Prácticas) [PRESENCIAL][Prácticas]	10
Estudio o preparación de pruebas [AUTÓNOMA][Trabajo autónomo]	65
Pruebas de progreso [PRESENCIAL][Pruebas de evaluación]	4
Tutorías de grupo [PRESENCIAL][Tutorías grupales]	1
Elaboración de memorias de Prácticas [AUTÓNOMA][Trabajo dirigido o tutorizado]	25
Total horas: 150	

10. BIBLIOGRAFÍA, RECURSOS						
Autor/es	Título/Enlace Web	Editorial	Población	ISBN	Año	Descripción
	Documentación Matplotlib http://matplotlib.org/					
Jake VanderPlas	Python Data Science Handbook https://jakevdp.github.io/PythonDataScienceHandbook/	O'Reilly Media, Inc.		9781491912058	2016	
Ravishankar Chityala	Image Processing and Acquisition using Python	Chapman & Hall/CRC		978-1466583757		
Tim J. Stevens, Wayne Boucher	Python Programming for Biology: Bioinformatics and Beyond Documentación de Pandas http://pandas.pydata.org/ Introducción a la programación con Python https://www.u-cursos.cl/ingenieria/2011/2/CC3501/1/material_docente/bajar?id_material=381752	Cambridge University Press		978-0521720090	2015	
William W. Cohen	A Computer Scientists Guide to Cell Biology	Springer		978-0-387-48275-0	2007	
Sebastian Bassi	Python for Bioinformatics	(Chapman & Hall/CRC Computational Biology Series)		978-1138035263	2018	