

ANEXO I PROPOSTA DE PROXECTOS DE INVESTIGACIÓN STEMbach

Coordinación STEMbach na Facultade/Escola UVigo:	
Nome: Emilio Rolán Álvarez	
Enderezo electrónico: rolan@uvigo.es	Teléfono: 986812578
Dirección do proxecto Uvigo:	
Nome: Estela Sánchez Rodríguez	
Enderezo electrónico: esanchez@uvigo.es	Teléfono: 986812493
Co-dirección do proxecto UVigo:	
Nome: Miguel Ángel Mirás Calvo	
Enderezo electrónico: mmiras@uvigo.es	Teléfono: 986812449

NOTA: Os custos derivados da execución deste proxecto de investigación tales como os desprazamentos do profesorado da UVigo ao centro educativo ou doutras actividades establecidas no plan de traballo, correrán a cargo do centro educativo ao que se asigne este proxecto.

Título

Técnicas de aprendizaje automático aplicadas a la biología

Resumo

El aprendizaje automático (machine learning) es un conjunto de técnicas y algoritmos que permiten a un sistema informático aprender de manera autónoma a partir de datos y mejorar su rendimiento en una tarea específica, como puede ser la clasificación. Podemos pensar, por ejemplo, en clasificar a diferentes especies a partir de características morfológicas, clasificar hábitats a partir de variables medioambientales o tratar de predecir el tiempo de supervivencia o la respuesta a un tratamiento farmacológico de un paciente partiendo de un conjunto de individuos que ya están clasificados.

Obxectivo

El principal objetivo es utilizar técnicas de inteligencia artificial para clasificar un conjunto de muestras. El análisis biológico, posterior a la clasificación proporcionada por estas técnicas, será fundamental para interpretar adecuadamente los resultados obtenidos. En este proyecto los alumnos comprenderán las principales diferencias entre las técnicas no supervisadas y las supervisadas, entenderán lo que son las muestras de entrenamiento, y se iniciarán con el software libre estadístico R para aplicar las técnicas a un problema biológico aplicado.

Plan de trabajo

Fase 1 (6 horas): Explicación de las técnicas metodológicas.

Fase 2 (6 horas): Manejo del programa de software libre R.

Fase 3 (2 horas): Selección y depuración de los datos experimentales.

Fase 4 (6 horas): Análisis de datos y conclusiones

Algunas de las reuniones pueden celebrarse en el propio IES, otras serán en la Universidad y alguna puede ser online.

Trabajo virtual: Los directores se comprometen a revisar durante todo el período el trabajo desarrollado, además de ayudarles en la memoria final y en la preparación de la defensa.