

## ANEXO I PROPOSTA DE PROXECTOS DE INVESTIGACIÓN STEMbach

<b>Coordinación STEMbach na Facultade/Escola UVigo:</b>	
Nome: Emilio Rolán Álvarez	
Enderezo electrónico: rolan@uvigo.es	Teléfono: 986812578
<b>Dirección do proxecto Uvigo:</b>	
Nome: Estela Sánchez Rodríguez	
Enderezo electrónico: esanchez@uvigo.es	Teléfono: 986812493
<b>Co-dirección do proxecto UVigo:</b>	
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 aprendizaxe automática aplicadas á bioloxía**

### Resumo

A aprendizaxe automática (machine learning) é un conxunto de técnicas e algoritmos que permiten a un sistema informático aprender de maneira autónoma a partir de datos e mellorar o seu rendemento nunha tarefa específica, como pode ser a clasificación. Podemos pensar, por exemplo, en clasificar a diferentes especies a partir de características morfolóxicas, clasificar hábitats a partir de variables ambientais ou tratar de predicir o tempo de supervivencia ou a resposta a un tratamento farmacolóxico dun paciente partindo dun conxunto de individuos que xa están clasificados.

### Obxectivo

O principal obxectivo é utilizar técnicas de intelixencia artificial para clasificar un conxunto de mostras. A análise biolóxica, posterior á clasificación proporcionada por estas técnicas, será fundamental para interpretar adecuadamente os resultados obtidos. Neste proxecto os alumnos comprenderán as principais diferenzas entre as técnicas non supervisadas e as supervisadas, entenderán o que son as mostras de adestramento, e iniciaranse co software libre estatístico R para aplicar as técnicas a un problema biolóxico aplicado.

## Plan de traballo

- Fase 1 (6 horas): Explicación das técnicas metodolóxicas.
- Fase 2 (6 horas): Manexo do programa de software libre R.
- Fase 3 (2 horas): Selección e depuración dos datos experimentais.
- Fase 4 (6 horas): Análise de datos e conclusións

Algunhas das reunións poden celebrarse no propio IES, outras serán na Universidade e algunha pode ser en liña.

Traballo virtual: Os directores comprométense a revisar durante todo o período o traballo desenvolvido, ademais de axudarlles na memoria final e na preparación da defensa.