

TRABAJO OBLIGATORIO:

La mejora de un método conocido

Aprendizaje no supervisado

Félix José Fuentes Hurtado

9 de marzo de 2021

Máster en
Inteligencia Artificial



1. Descripción del trabajo

A lo largo de estas semanas veremos diferentes métodos para aprendizaje no supervisado. Concretamente, en las primeras sesiones de la asignatura estudiaremos numerosos algoritmos de agrupamiento o *clustering*. En todos los casos, las versiones de los algoritmos que veremos en clase serán siempre las básicas u originalmente propuestas.

El objetivo de este trabajo es encontrar y estudiar una versión de alguno de los algoritmos estudiados que se haya propuesto para *mejorar* el algoritmo original. Una vez seleccionado, se estudiará esa nueva versión del algoritmo y se compararán ambos (el original y la versión mejorada).

2. Procedimiento

Para llevar a cabo esta actividad, se deberá elegir uno de los métodos estudiados en clase. Se buscará¹ en la literatura científica relacionada un artículo que proponga una mejora al método seleccionado. La mejora puede ser, por ejemplo, una mejor inicialización del algoritmo, la supresión-sustitución de un parámetro, etc.

Una vez elegido un algoritmo y encontrado un artículo que proponga una mejora para el mismo, se deberán programar ambos algoritmos (el básico² y la versión *mejorada*).

Se compararán los resultados de ambos algoritmos en al menos 2 datasets a elegir, calculando métricas extrínsecas (en uno de ellos) e intrínsecas.³

Se analizarán los resultados y se prestará atención a la contribución de la mejora del algoritmo extendido. ¿Se aprecia en los resultados la mejora que defienden los autores en el artículo estudiado? Se elaborará un informe documentando todo el proceso y con un análisis exhaustivo de los resultados.

3. Entrega

Se entregará un informe completo de todo el proceso, donde se expliquen ambos métodos (básico y mejorado), las pruebas realizadas y el análisis de los resultados.

Junto al informe, se entregará el código fuente (preferiblemente, Jupyter Notebook)

El informe y el código se presentarán mediante el Campus Virtual (no se aceptarán entregas vía email) antes de las **23.59 del día 11 de abril de 2021**.

¹Por ejemplo, buscar en <http://scholar.google.com>

²En clase veremos el código de los diferentes algoritmos estudiados

³Los que se usarán en clase están disponibles en: <https://github.com/flifuehu/unsupervisedlearning>

4. Rúbrica de evaluación

Criterio	No aceptable (0 %)	Mejorable (33 %)	Aceptable (66 %)	Perfecto(100 %)
Presentación de las técnicas (4pts)	Las técnicas no se presentan	Sólo se presenta la técnica nueva	Se presentan las dos técnicas pero existen errores en la explicación	Las dos técnicas (básica y mejorada) se presentan correctamente
Experimentación y análisis (3pt)	No se realiza una comparación	No se tiene en cuenta en la comparación y el análisis la supuesta mejora del algoritmo	Se realiza la comparación sobre pocos conjuntos de datos (solo 1)	Se realiza la comparación sobre al menos 2 conjuntos de datos
Evaluación (2pt)	No se usan criterios válidos para evaluar los métodos	Se usa una única métrica de evaluación (intrínseca o extrínseca)	Se usan métricas de eval. intrínsecas y extrínsecas pero no se analizan los resultados	Se usan varias métricas de evaluación y se analizan los resultados
Código (1pt)	El código no funciona al primer intento	El código es funcional pero no está completo	El código es funcional pero no se diferencian las etapas del análisis	El código es funcional, claro y completo