



MINERÍA DE SERIES TEMPORALES Y FLUJO DE DATOS

MÁSTER EN CIENCIA DE DATOS E INGENIERIA DE
COMPUTADORES

Trabajo autónomo II: Flujo de Datos

Autores

José Ángel Díaz García
joseangeldiazg02@correo.ugr.es



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

Granada, Abril de 2018

Índice general

1. Introducción	4
1.1. Problema a resolver	4
1.2. Objetivos	5
1.3. Organización del trabajo	6
2. Práctica	7
3. Teoría	8
4. Conclusión	9

Índice de figuras

1.1. Interfaz del software moa.	5
---	---

Índice de tablas

Capítulo 1

Introducción

En este documento encontramos el resultado final alcanzado durante el estudio del apartado de **minería de flujos de datos**, enmarcado dentro de la asignatura de ‘Minería de Series Temporales y Flujos de Datos’ del máster en Ciencia de Datos de la Universidad de Granada. En este primer capítulo, veremos una introducción al problema a resolver así como a los objetivos a alcanzar con esta práctica y la organización del trabajo.

1.1. Problema a resolver

El problema a resolver en esta práctica se centrará en la resolución de ciertos problemas de clasificación con flujos de datos, de manera tanto estática como dinámica así como con cambio o sin cambio de contexto. Para resolver estos problemas, se propone el uso del software MOA [1] [2] el cual usaremos con la línea de comandos o con la interfaz gráfica que podemos ver en la figura 1.1.

Los problemas a resolver serán:

1. Entrenamiento offline (estacionario) y evaluación posterior.
2. Entrenamiento online.
3. Entrenamiento online en datos con concept drift.

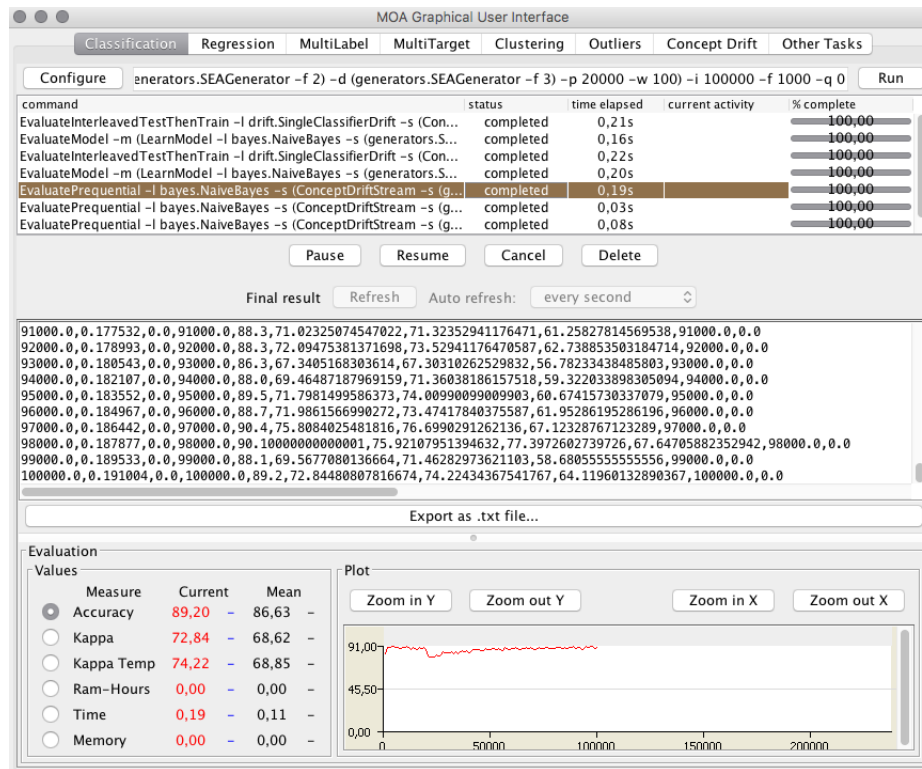


Figura 1.1: Interfaz del software moa.

4. Entrenamiento online en datos con concept drift, incluyendo mecanismos para olvidar instancias pasadas.
5. Entrenamiento online en datos con concept drift, incluyendo mecanismos para reinicializar modelos tras la detección de cambios de concepto.

Posteriormente a la solución de estos problemas, se propone una discusión teórica de los resultados, por lo que encontraremos la práctica dividida en dos apartados, por un lado el teórico (capítulo 3) y por otro el práctico (2).

1.2. Objetivos

Los objetivos de esta práctica serán:

- Asentar y comprender la materia teórica de la minería de flujo de datos vista durante el transcurso de la asignatura.
- Comprender el uso y formas de utilización del software MOA.
- Asentar el conocimiento sobre test estadísticos par comparación de modelos.
- Elaboración de una memoria donde se recojan todos los resultados de manera apropiada.

1.3. Organización del trabajo

La organización del presente documento, se centra en detallar cada uno de los pasos seguidos durante el estudio y resolución del problema planteado en esta introducción, tras la cual tendremos el contenido práctico en el capítulo 2 y el cual representa el grueso de esta memoria. Tras este capítulo encontramos el capítulo 3 donde desde un punto de vista teórico analizamos el concepto de **clasificación** y **concept drift** en minería de flujo de datos. Finalizaremos la memoria con las conclusiones obtenidas en el transcurso de finalización de la misma en el capítulo 4.

Capítulo 2

Práctica

Capítulo 3

Teoría

Capítulo 4

Conclusión

Bibliografía

- [1] Albert Bifet, Geoff Holmes, Richard Kirkby, Bernhard Pfahringer (2010); MOA: Massive Online Analysis; *Journal of Machine Learning Research* 11: 1601-1604
- [2] Moa Web Site <https://moa.cms.waikato.ac.nz>