

Instrucciones para la evaluación de la práctica

Modelos bayesianos

231102

Cada alumno debe entregar dos ficheros, –un informe en formato pdf y un código ejecutable en el lenguaje de programación R–, etiquetados con su nombre de usuario de la Universidad de Oviedo (por ejemplo, `uo123456.pdf`, `uo123456.R`, respectivamente) a través del Campus Virtual antes de la finalización de la última sesión de prácticas de laboratorio (días 11 y 12 de diciembre).

Informe. El informe contará con una extensión máxima de 4 páginas y debería contener al menos las siguientes partes:

- Título, nombre completo, documento nacional de identidad, usuario de uniovi y fecha al inicio de la primera página (obligatorio).
- Introducción (10%).
- Metodología, explicación y justificación. Al menos una aproximación ha de incluir un enfoque bayesiano (obligatorio) (25%).
- Resultados (5%), conclusiones (15%) y referencias (5%).
- Aspectos formales del documento: redacción global, estética, sintaxis, ortografía, maquetación, claridad, LLM, etc. (20%).

Código. El fichero de programación solo debe incluir como código ejecutable una única función de R, llamada como el usuario de uniovi. Consta con dos argumentos de entrada, la base de datos `train` y la base de datos `test`, ambas con la misma estructura, trabaja en mononú-

cleo (no en paralelo) con un tiempo de ejecución inferior a 10 minutos y devuelve el error cuadrático medio.

- Título, nombre completo, documento nacional de identidad, usuario y fecha al inicio del fichero, codificado en UTF-8 (obligatorio).
- Para trabajar, divida la base de datos original en dos partes y emplee una de ellas como entrenamiento y la otra como validación. Para entrenar el sistema no se puede utilizar la variable respuesta de la base de datos test. Solo se puede utilizar esta variable para estimar el error cuadrático medio de la predicción.
- Claridad y comentarios explicativos del código (10%).
- Bondad del ajuste mediante el error cuadrático medio (10%).

Véase el ejemplo que se propone (`uo123456.R`) y la función de chequeo que se emplea para validarla (`chequeofuncion.R`). No emplee código que no entienda o que no sepa lo que hace. Puede ser código malicioso y conlleva consecuencias graves.

Autoría. Referencie si resulta adecuado. Procure que gran parte de su trabajo sea propio y evite que otras personas se aprovechen de su esfuerzo o repliquen su investigación. Es normal que al principio se compartan ideas, pero una vez enfocado el estudio, cada alumno debería realizar una aproximación propia y original (en el sentido de no ser copia) diferente del resto de sus compañeros¹. El programa Turnitin se emplea para detectar posibles plagios y del informe en pdf se deberá poder extraer el texto. Se aplicará un factor de corrección en la calificación de informes similares en cuanto a planteamiento, contenido o resultados.

Defensa. Quien lo desee podrá presentar su trabajo al profesor durante la última clase de laboratorio con una duración máxima de tres minutos.

¹Una actividad desarrollada en clase (el análisis de los suspensos universitarios) mostró que cada persona enfoca la resolución del problema desde su propio punto de vista. Apena que de forma recurrente aparezcan escritos muy similares o parafraseados cuando cada alumno es único e irreplicable.

Calificación. El porcentaje entre paréntesis hace referencia al peso de cada apartado en la evaluación de la práctica de laboratorio. Se seguirá un criterio de comparación entre los informes evaluados.

Investigación. En un hospital de salud mental se han analizado los tiempos de respuesta en milisegundos de 17 personas, de las cuales 11 no presentan síntomas de esquizofrenia (tipo=0) y 6 están diagnosticadas de esquizofrenia (tipo=1). Para cada paciente se ha repetido 30 veces la misma prueba. Se desea modelar el tiempo de respuesta.

```
> datos <- readRDS("esquizofrenico.rds")
> dim(datos)
[1] 510  4
> head(datos)
  tipo paciente intento tiempo
1    0         11      1    240
2    0         11      2    292
3    0         11      3    350
4    0         11      4    254
5    0         11      5    396
6    0         11      6    430
```

Referencia. La base de datos propuesta se ha inspirado en el experimento sobre la esquizofrenia presentado por Belin, T. R., y Rubin, D. B. (1990). «Analysis of a finite mixture model with variance components.» *Proceedings of the American Statistical Association, Social Statistics Section*, 211–215.