

CONCLUSIONES Y RECOMENDACIONES

1. Esta página web brinda al usuario la posibilidad de desplazar uno o más puntos de la muestra para observar la forma de la curva discriminante antes y después de dichos movimientos. Sería interesante que el usuario convierta unos pocos puntos del conjunto de datos en datos extremos (*outliers*) y trate luego de aplicar su conocimiento para explicar la reacción o insensibilidad de un determinado método ante dichos cambios extremos. Disponiendo el conjunto de datos convenientemente el usuario podrá observar que la regresión lineal de grado 1 es sensible ante un dato extremo mientras que la regresión logística de grado 1 no siempre será influenciada por ese dato.
2. Puesto que esta herramienta ofrece al usuario la posibilidad de construir su propia distribución de datos, se recomienda que construya conjuntos de datos tan poco comunes o novedosos como la distribución de círculos concéntricos (*Figura 4.29*) u otros conjuntos donde la tarea de discriminación o de captación de patrones no sea una labor sencilla. Esto,

con el objetivo de palpar la potencia o debilidad de cada método según el caso propuesto. Específicamente, en la discriminación de un conjunto de datos con distribución de círculos concéntricos el método de los kvecinos más cercanos captó perfectamente el patrón de datos utilizando valores de $k=1,5$ y 10 .

3. De los 5 métodos que se han expuesto en el presente trabajo se podría decir que no existe uno que sea el mejor para resolver la tarea de discriminación. Sin embargo, sí existen unos métodos que son más efectivos que otros dependiendo del conjunto al cual se apliquen.
4. Cuando se trata de distribuciones irregulares de datos que requieren una gran expresividad uno de los métodos que debería considerarse es el kvecinos más cercanos que aunque es un método retardado ofrece interesantes posibilidades si se escoge un conveniente valor para k .
5. Al comparar el método anticipativo de árbol de decisión con el método retardado de kvecinos más cercanos el usuario podrá comprobar por sí mismo que el costo computacional (*tiempo de respuesta*) de este último método es considerablemente alto comparado con el primero, especialmente si se escoge un valor alto para k .

6. El aplicativo web permite al usuario visualizar hasta 4 curvas discriminantes aplicadas a la misma muestra en el mismo plano. Se recomienda que el usuario aproveche esta posibilidad comparando visualmente distintos métodos aplicados a una misma situación.

7. A excepción del método de Naive Bayes Kernel, los otros 4 métodos tienen parámetros que dependiendo del valor que se les establezca pueden hacer que el comportamiento de los métodos varíe. Se recomienda probar cada método con diferentes valores de su parámetro correspondiente para observar cuál es el cambio que el parámetro provoca en el método y tratar, si es posible, de encontrar un valor idóneo para este parámetro. En relación a esto, el usuario podrá notar que existen grados de poda (*parámetros de parada*) para los cuales la precisión del árbol de decisión es mejor. Al aplicar el árbol de decisión a una muestra con clases distribuidas exponencialmente el error de prueba bajó de 0.11 a 0.06 cuando su parámetro de parada aumentó de 2% a 8%.

8. Valiéndose de los errores de prueba que se registran en el Tablero de resultados de la página web el usuario podrá llevar a cabo experimentos globales que consisten en la aplicación de todos los métodos con distintos valores de sus parámetros. Particularmente, en un experimento global

aplicado a un conjunto de datos en el cual cada atributo de cada clase se distribuía exponencialmente, el método que obtuvo el menor error de prueba fue el árbol de decisión con el 94% de aciertos.

9. Aunque tanto los métodos de regresión como el árbol de decisión proporcionan un modelo general con el cual discriminar nuevos elementos, el sistema de reglas que produce el árbol de decisión es más comprensible desde el punto de vista humano que los modelos de regresión.
10. Aunque el método de árbol de decisión es comprensible, computacionalmente eficiente y produce en general buenos resultados, adolece de limitada expresividad en ciertos conjuntos de datos y de inestabilidad ante variaciones de la muestra. Específicamente, en 6 muestras tomadas de una población compuesta por 2 clases distribuidas exponencialmente la eficacia de este método al discriminar varió del 87% al 96% al utilizar un parámetro de parada del 16%.
11. El método de Naive Bayes Kernel clasificó correctamente el 91% de los datos de prueba al ser aplicado a una muestra con distribución de Cuadrados traslapados. En este caso Naive Bayes superó en 2% a la regresión polinómica de grado 2. Sin embargo, cuando no se está seguro

de la independencia condicional de los atributos que determinan los elementos de la muestra es aconsejable contrastar este método con otros que no tengan el limitante de la independencia.

12. Al aumentar la complejidad de los modelos de regresión no siempre se logra una mejoría en el error de prueba, antes esta complejidad puede conllevar a que el modelo se sobreajuste a los datos de entrenamiento. Específicamente, al aumentar la complejidad de la regresión polinómica en la discriminación de una muestra con clases distribuidas exponencialmente no se logró ninguna mejoría. Con la regresión de grado 1 se obtuvo el error de prueba de 0.11 mientras que con las regresiones de grado 2 y 3 este error se incrementó ligeramente a 0.12.

13. Aunque los métodos de regresión no son tan expresivos como los kvecinos más próximos, éstos a veces pueden ser muy útiles para captar el patrón de ciertos conjuntos de datos y discriminar correctamente cuando el grado de sus polinomios es idóneo. En particular, en un experimento diseñado para medir la capacidad de cada método para captar el patrón de la distribución de círculos concéntricos, el método de regresión polinomial de grado 4 captó bien el patrón logrando una eficacia del 97%.