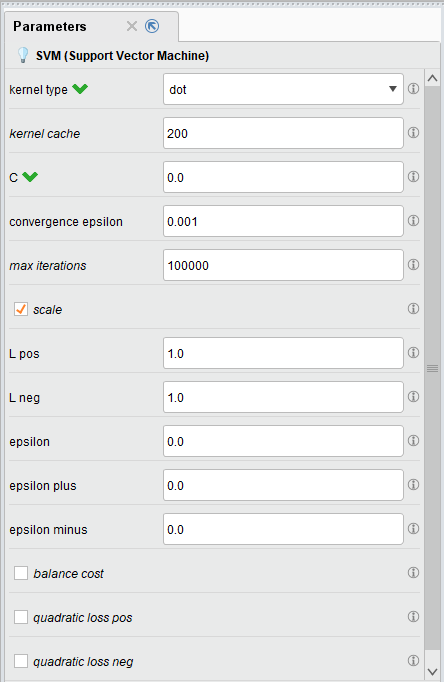
|  |
| --- |
| **Introducción a los Métodos de Aprendizaje Automático** |
| **Ejercicios domiciliarios**  **UT4**  **Algoritmos No Lineales (Parte 1)** |
| **Autor:** Gerardo Fernández - CI: 2858230-7  **Docente:** Ernesto Ocampo |
| **04/09/2021** |
|  |

**Trabajo de Aplicación 7**

**Support Vector Machines**

**Parámetros por defecto**

****

Tipos de kernel

El tipo de función del kernel se selecciona a través de este parámetro. Se admiten los siguientes tipos de kernel: punto, radial, polinomial, neural, anova, epachnenikov, combinación gaussiana, multicuadric

punto

El kernel de puntos está definido por k (x, y) = x \* y, es decir, es el producto interno de x e y.

radial:

El kernel radial está definido por exp (-g || x-y || ^ 2) donde g es la gamma, está especificado por el parámetro gamma del kernel. El parámetro ajustable gamma juega un papel importante en el rendimiento del kernel y debe ajustarse cuidadosamente al problema en cuestión.

Polynomial

El kernel del polinomio se define por k (x, y) = (x \* y + 1) ^ d donde d es el grado del polinomio y está especificado por el parámetro de grado del núcleo. Los núcleos polinomiales son adecuados para problemas en los que todos los datos de entrenamiento están normalizados.

neuronal

El núcleo neuronal se define por una red neuronal de dos capas tanh (a x \* y + b) donde a es alfa y b es la constante de intersección. Estos parámetros se pueden ajustar utilizando los parámetros kernel a y kernel b. Un valor común para alfa es 1 / N, donde N es la dimensión de datos. Tenga en cuenta que no todas las opciones de ayb conducen a una función de kernel válida.

Anova

El núcleo de anova se define por la suma de exp (-g (x-y)) por elevado a la potencia d, donde g es gamma y d es grado. gamma y grado se ajustan por los parámetros de kernel gamma y kernel degree, respectivamente.

epachnenikov

El núcleo epachnenikov es lafunción (3/4) (1-u2) para u entre -1 y 1 y cero para u fuera de ese rango. Tiene dos parámetros ajustables kernel sigma1 y kernel degree.

combinación gaussiana

Este es el núcleo de combinación gaussiana. Tiene parámetros ajustables kernel sigma1, kernel sigma2 y kernel sigma3.

multiquadric

El núcleo multicuadratico se define por la raíz cuadrada de || x-y || ^ 2 + c ^ 2. Tiene parámetros ajustables kernel sigma1 y kernel sigma shift.v

L pos (opcional)

Un factor para la constante de complejidad de SVM para ejemplos positivos. Este parámetro es parte de la función de pérdida.

L neg (opcional)

Un factor para la constante de complejidad SVM para ejemplos negativos. Este parámetro es parte de la función de pérdida.

épsilon (opcional)

Este parámetro especifica la constante de insensibilidad. No hay pérdida si la predicción se encuentra muy cerca del valor real. Este parámetro es parte de la función de pérdida.

epsilon plus (opcional)

Este parámetro es parte de la función de pérdida. Especificar épsilon solo para la desviación positiva.

epsilon minus (opcional)

Este parámetro es parte de la función de pérdida. Especificar épsilon solo para la desviación negativa.

costo del saldo (opcional)

Si está marcado, adaptar Cpos y Cneg al tamaño relativo de las clases.

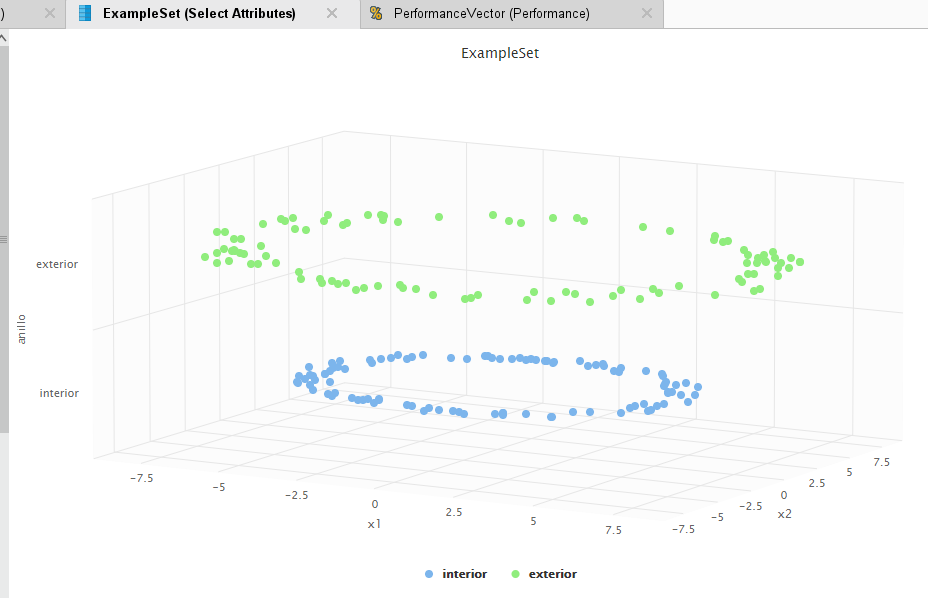
posición de pérdida cuadrática

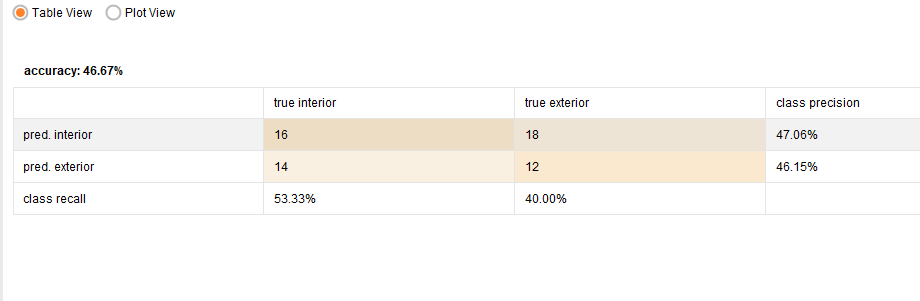
Utilizar la pérdida cuadrática para la desviación positiva. Este parámetro es parte de la función de pérdida.

pérdida cuadrática neg (opcional)

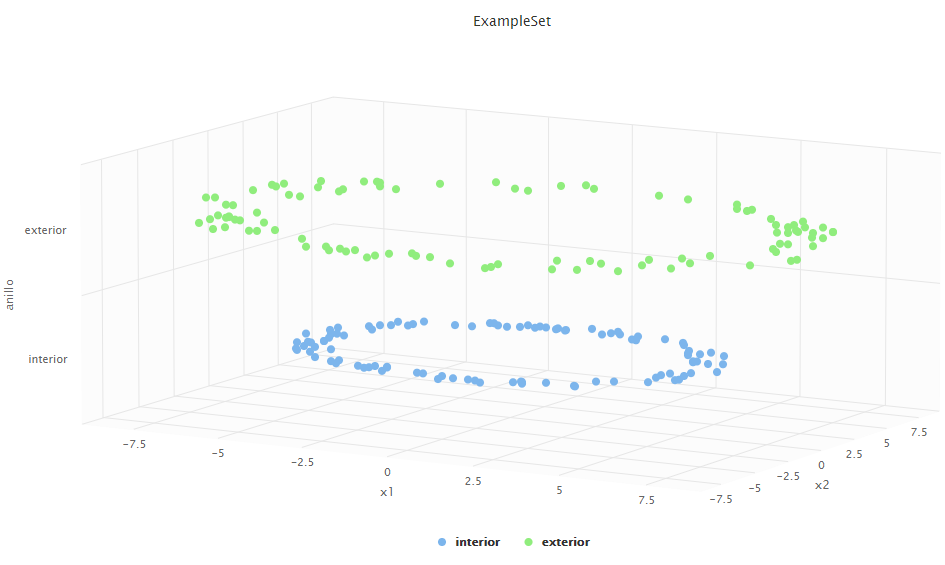
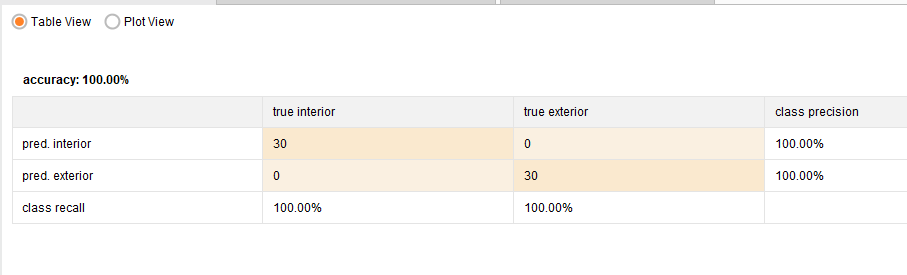
Utilizar la pérdida cuadrática para la desviación negativa. Este parámetro es parte de la función de pérdida.

**Kernel dot**





**Kernel polynomial**



*Constante de complejidad C*

La constante de complejidad de SVM establece la tolerancia para la clasificación errónea, los valores de C más altos permiten límites 'más suaves' y los valores más bajos crean límites 'más duros'. Una constante de complejidad demasiado grande puede provocar un ajuste excesivo, mientras que los valores demasiado pequeños pueden provocar una generalización excesiva.

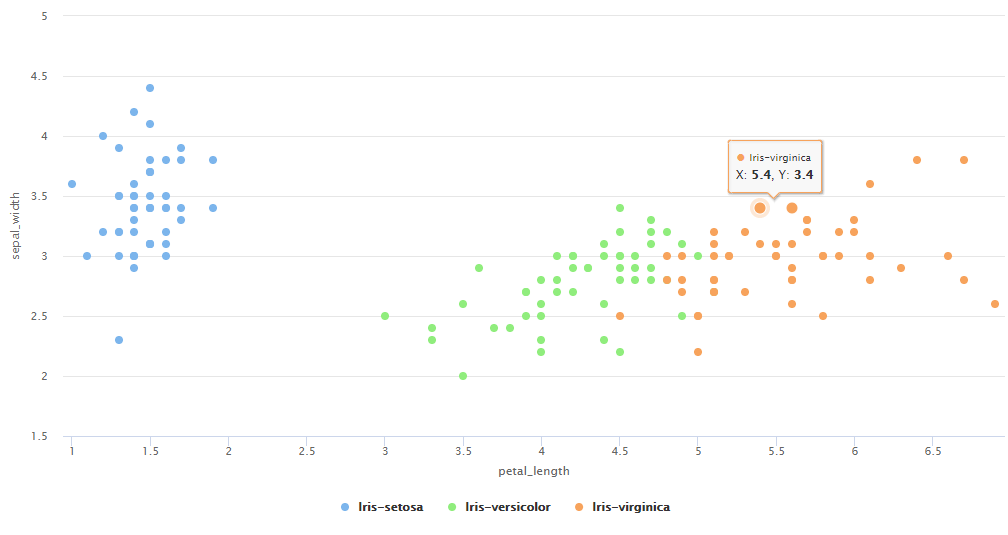
**Trabajo de Aplicación 8**

**k-NN**

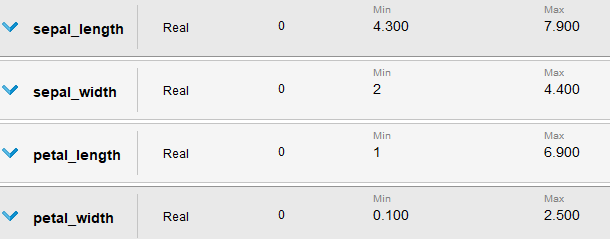
**Ejercicio 1**

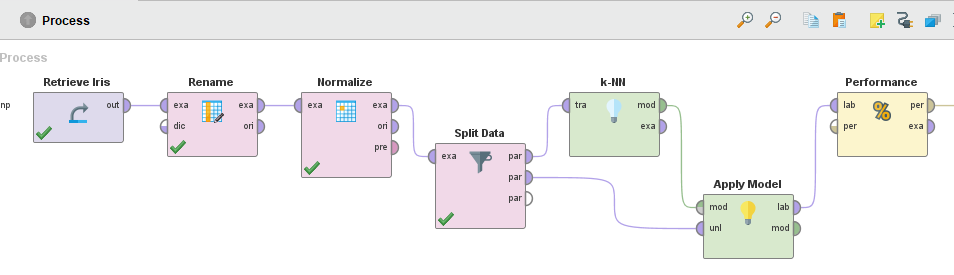
Ver archivo excel [KVecinosMasCercanos\_sol](https://docs.google.com/spreadsheets/d/1iAQ1nxRLuRyws7q-gUBcLLHYO19hcpEsD7srM18qc9A/edit?usp=sharing)

**Ejercicio 2**

Los ejemplos de iris-setosa están mucho más alejados de las otras dos clases, por lo cual es mas fácil clasificar un nuevo ejemplo cerca de esta clase. Sin embargo un punto cerca de las otras dos es más difícil 

Como tarea de preparación previa de datos los mismos deben normalizarse, ya que por ejemplo hay diferencias en la escala utilizada por petal\_width y el resto de los atributos.





voto ponderado

Si se establece este parámetro, los valores de distancia entre los ejemplos también se tienen en cuenta para la predicción. Puede ser útil ponderar las contribuciones de los vecinos, para que los vecinos más cercanos contribuyan más que los más lejanos.

tipos de medida

Este parámetro se utiliza para seleccionar el tipo de medida que se utilizará para encontrar los vecinos más cercanos.

Las siguientes opciones están disponibles:

Medidas mixtas: Las medidas mixtas se utilizan para calcular distancias en el caso de atributos tanto nominales como numéricos.

Medidas nominales: En el caso de solo atributos nominales, se pueden usar diferentes métricas de distancia para calcular distancias en estos atributos nominales.

Medidas numéricas: En el caso de solo atributos numéricos, se pueden usar diferentes métricas de distancia para calcular distancias en estos atributos numéricos.

Divergencias Bregmann: Las divergencias de Bregmann son tipos de medida de "cercanía" más genéricos que no satisfacen la desigualdad o simetría del triángulo.

