

Practica 2 Aprendizaje Automático

Nombre: José Arcos Aneas

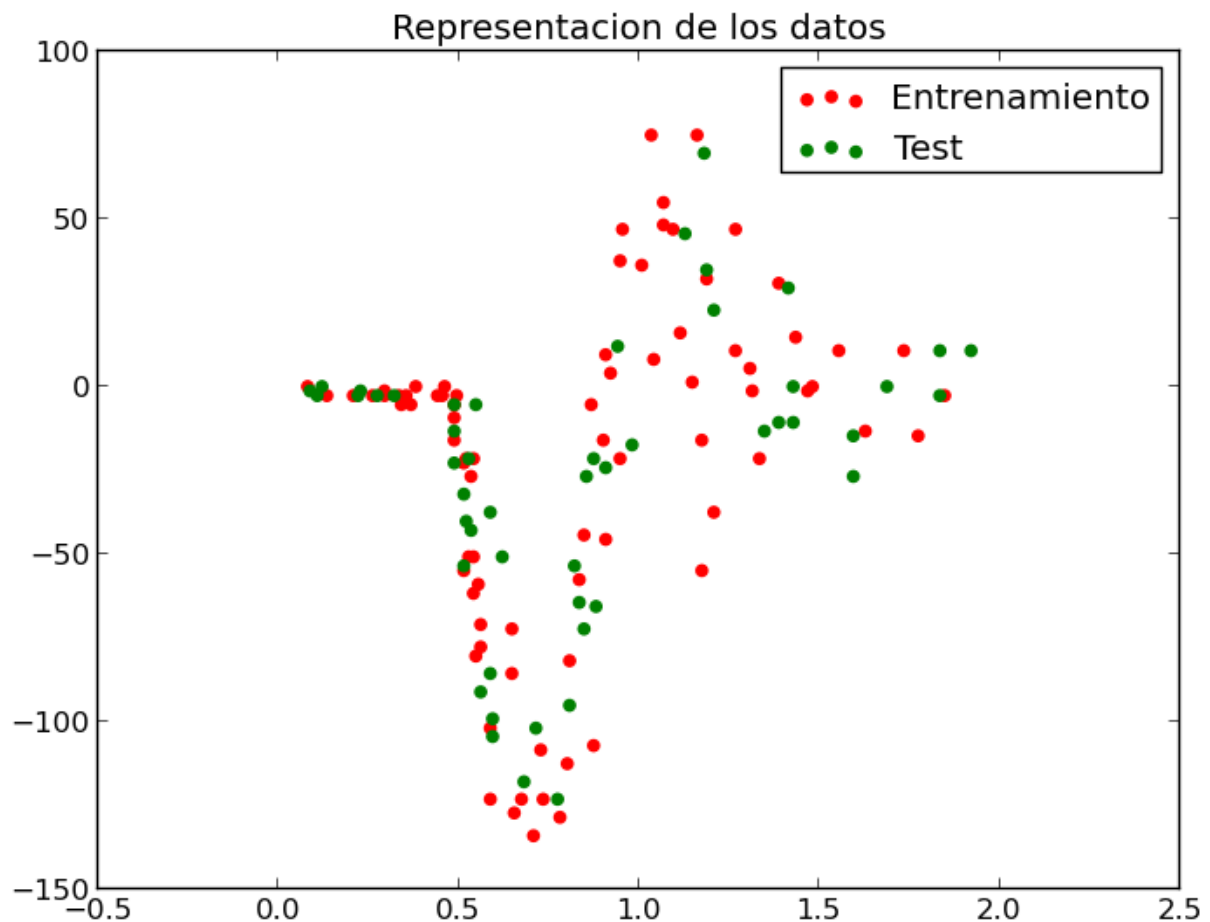
D.N.I: 74740565-H

Correo electrónico: joseaa@correo.ugr.es

Regresión Lineal

La siguiente figura muestra una representación de los datos de test y training obtenidos de los ficheros proporcionados en decsai.

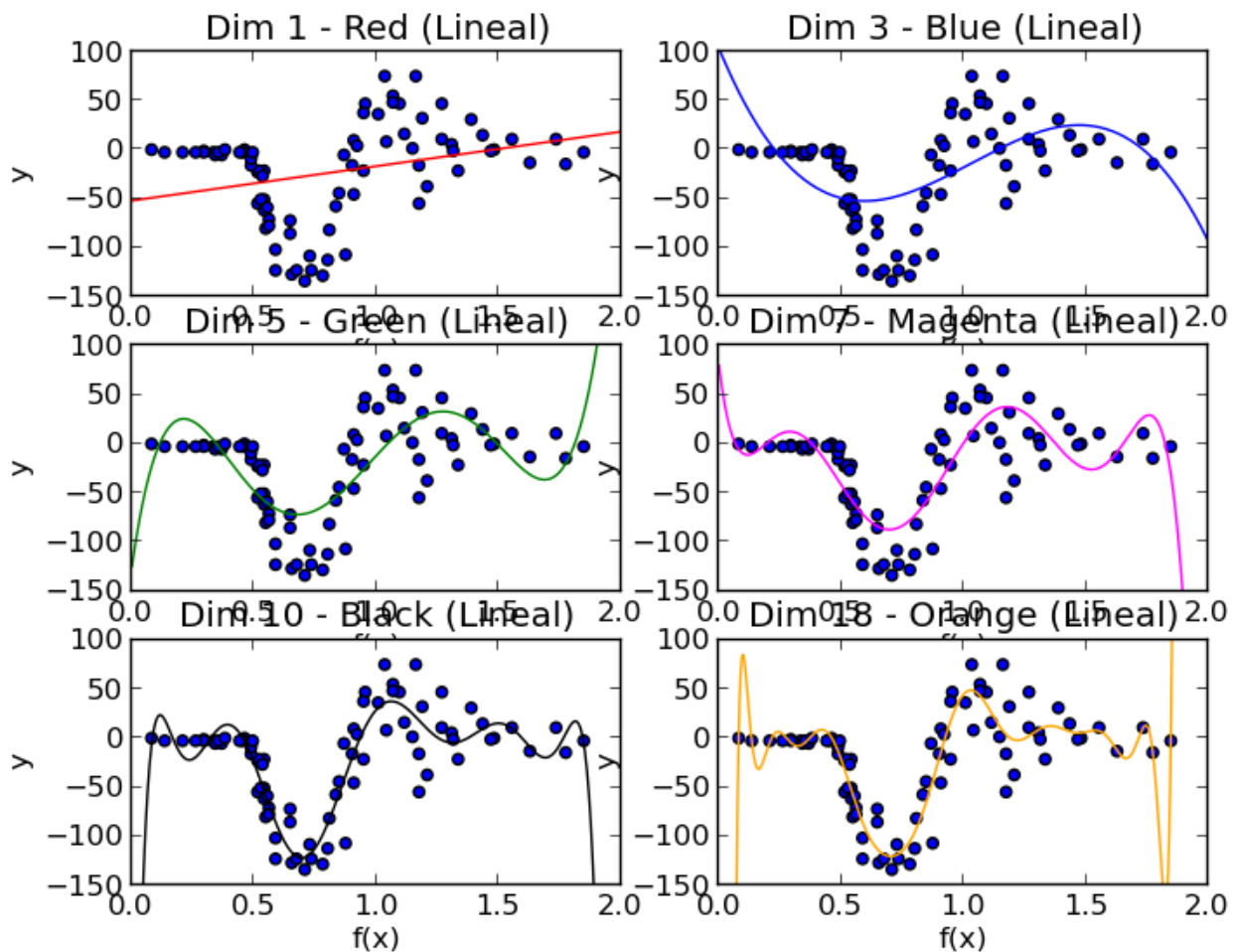
Esta representación nos permite detectar la relación entre dichos valores. Tras la representación se observa que los datos son buenos para clasificación ya que los datos de training son una representación de los datos de test.



Ajuste de regresión lineal.

Se realiza un ajuste variando grado del polinomio para distinguir el grado que mas se ajuste mas a los datos de entrada. El aumento hará que el ajuste sea mejor, pero sin llegar a aumentar demasiado el grado ya que podemos llegar a una situación de sobre ajuste de los datos.

Los polinomios de grado D1,D3 están muy mal ajustados al modelo. Los de dimensión 5 y 7 se van acercando considerablemente y una de dimensión 10 ajusta casi a la perfección los datos. En cambio si seguimos aumentando la dimensión llegamos a una situación de sobre-ajuste como puede observarse en la gráfica de la figura de dimensión 18.



Ajuste de regresión ridge.

Se valora el valor de alpha en el modelo ridge. En los polinomios de grado bajo no observamos una mejora considerable cambiando el valor de alpha.

Para un polinomio de grado 3 y un valor de alpha grande de 10 o 20 se observa una gráfica mas relajada.

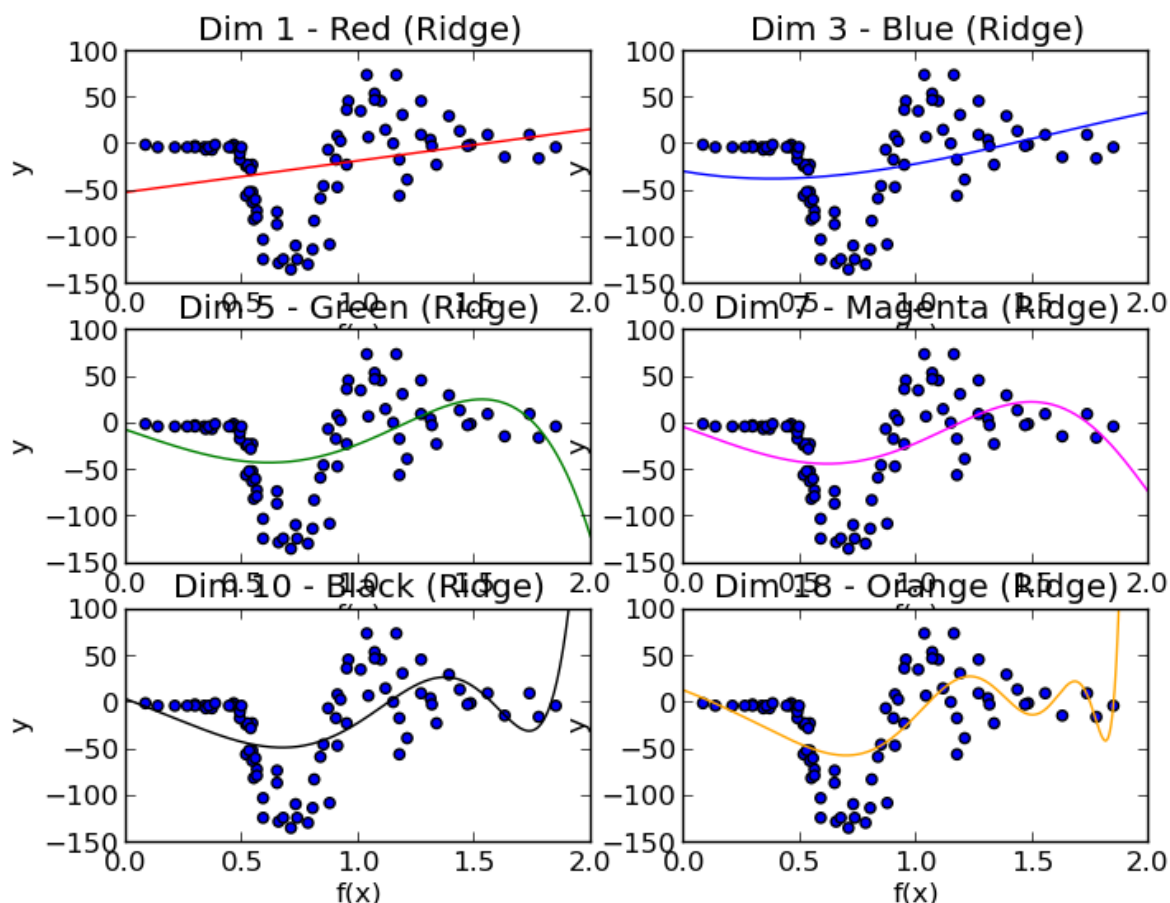
En polinomios de grado 5,7,10 en cambio para valores pequeños de α como 0,5 se observa cierta mejora pero para valores grandes los resultados no mejoran mucho la situación.

Por ultimo en una gráfica que presente un sobre-ajuste podemos ver que para valores pequeños del α como 0.1 la gráfica se suaviza aunque con valores grandes como 20 el modelo se aleja mucho de lo esperado.

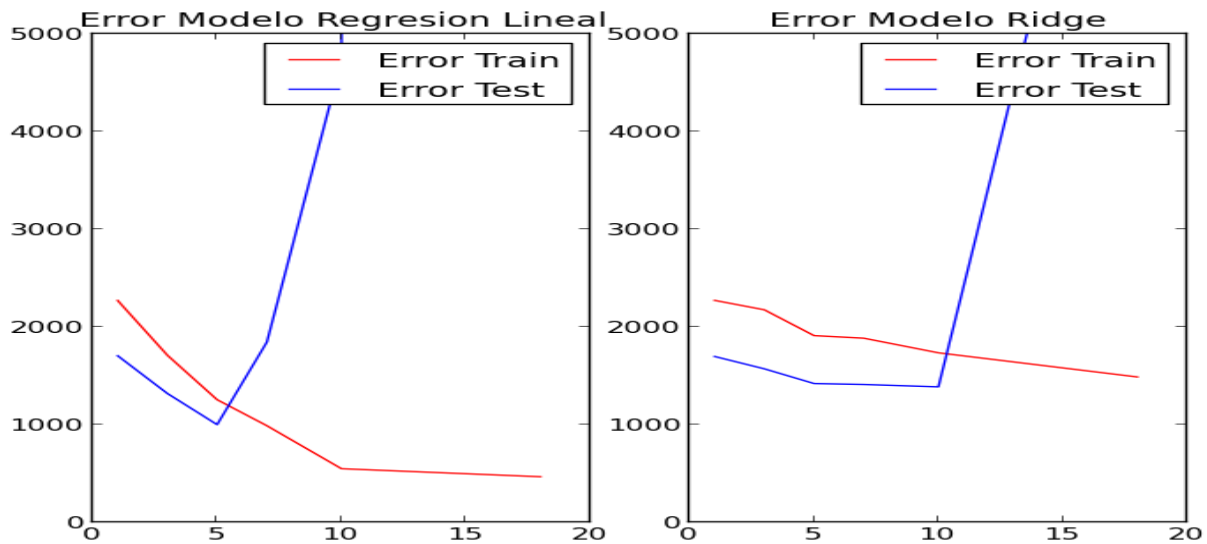
La conclusión a la que se puede llegar es que en una gráfica que presente sobre-ajuste valores pequeños de α pueden mejorar mientras valores grandes no ayudan.

En gráficas poco ajustadas podemos observar que un valor alto de α ayuda a mejorar el modelo.

Y para valores que se acercan al modelo podemos observar que valores pequeños ayudan a que mejore el modelo en cambio valores grandes hacen que este empeore.



Podemos llegar a la conclusión de que un valor que ayuda a que la gráfica mejore es 0.6. El modelo ridge es mas robusto y combinado correctamente el α y el valor el grado del polinomio se pueden suavizar el porcentaje de errores.



Valores de alpha para lasso

Se realiza el ejercicio apartado b buscando los para los distintos valores de alpha cual es el que menor error tiene.

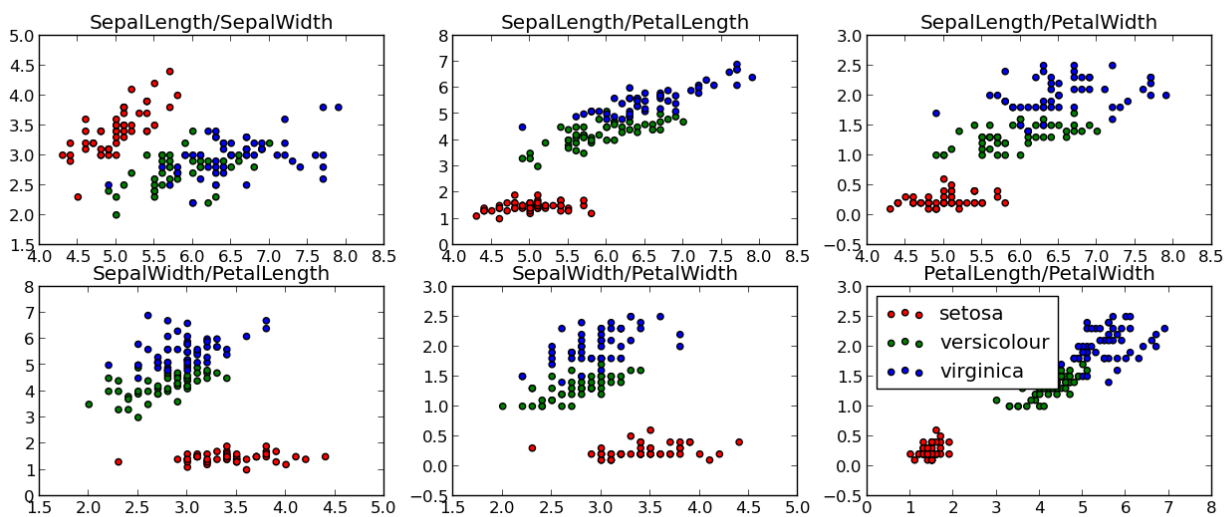
Alpha=0.1 = 25.4628167966
 Alpha=0.2 = 25.5665128166
 Alpha=0.3 = 25.6983327783
 Alpha=0.4 = 25.9050638141
 Alpha=0.5 = 26.201227152
 Alpha=0.6 = 26.5870228113
 Alpha=0.7 = 27.0622158855
 Alpha=0.8 = 27.6269139024
 Alpha=0.9 = 28.2812519386

Tras diferentes comprobaciones. El apartado no incluye la validación cruzada. Pero tras varias ejecuciones de la función se observa que el menor valor corresponde al valor de alpha para 0,1. Al aumentar el alpha lo hace también el error.

Perceptrón

El siguiente apartado nos ayuda a conocer que parejas de variables son mejores para clasificar estos valores de entrada.

Se representan las diferentes gráficas empleando parejas de variables.



Tras probar con los diferentes conjuntos de parejas notamos que todos son capaces de clasificar a la perfección los datos del conjunto 1, ya sabemos que es debido a que son las que mayor variación presentan en relación a los demás conjuntos. Los conjuntos que mejor clasifican son $\{2,3\}$ y $\{0,3\}$ aunque no son linealmente separables. Los demás conjuntos presentan mayor error en las fronteras. Con esto ya saber que el atributo que mejor diferencia a las clases 1 y 2 es el 3.

Dibujamos los hiperplanos para las parejas $\{2,3\}$ y $\{0,3\}$ que son las parejas del guión de practicas (1,2 y 2,4).

El modelo de perceptrón consigue ajusta el modelo para la clase 1 pero las clases 2 y 3 no son separables. La clase 2 no puede ser separada de las clases 1 y 3. En el caso del conjunto de variables 2,3 la separación mejora hasta llegar a separarlos casi en su totalidad, pero no se encuentra ningún hiperplano que los separe en su totalidad. El hiperplano se encuentra demasiado cerca de la clase 2 lo causa errores en la clasificación.

Con los datos de accuracy, precision y recall podemos comprobar que conjunto de variables es mejor para realizar la clasificación.

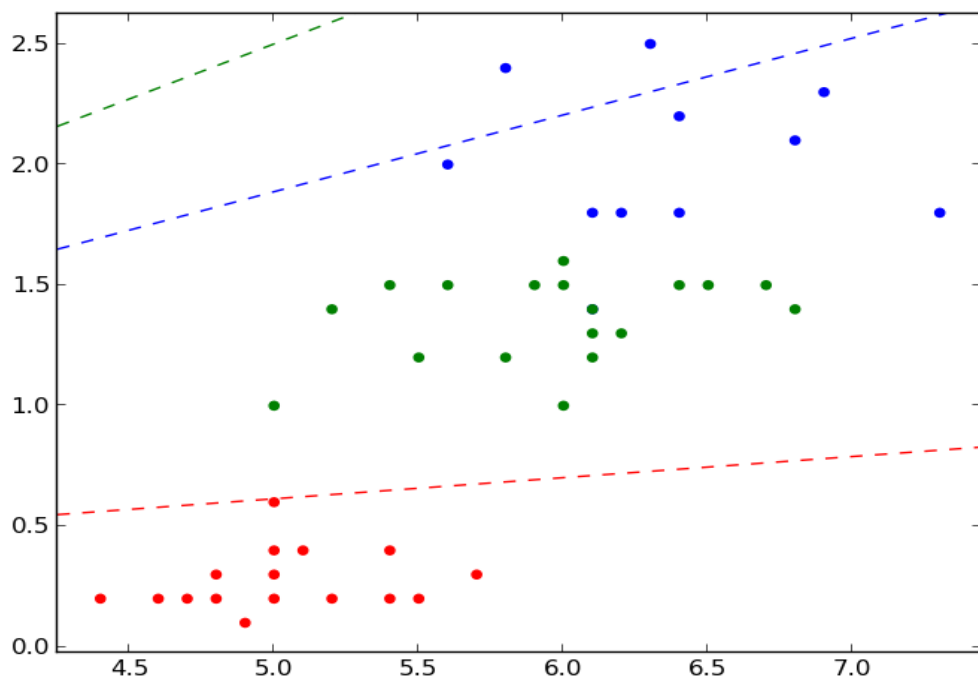
Accuracy nos indica como de bueno es el ajuste. Para el clasificador $\{0,3\}$ podemos observar que el accuracy es menor y por tanto clasifica peor que el clasificador $\{2,3\}$. Para el recall que indica cuantos elementos se han perdido de cada clase podemos ver que para el primer clasificador el numero aumenta en las 3 clases, cambio en el segundo clasificador los valores mejoran considerablemente, aunque es de tener en cuenta que los valores de la clase 3 siguen perdiéndose en su gran mayoría aumentando el error del clasificador.

La precisión que indica cuantos elementos se han clasificado en su clase para la primera pareja se observa que los datos de la clase 1 se clasifican bien en su totalidad, en cambio para otras clases dos y tres la clasificación no es tan precisa aunque cabe destacar que la clase 3 tiene mejor precision que la clase 2.

Para la segunda pareja de variables las clases 1 y 3 se clasifican perfectamente pero la precisión en la clase dos sigue siendo del mismo porcentaje.

Podemos llegar a la conclusión de que el segundo clasificador mejora al primero, ya que clasifica mejor la clase 3, aunque ambos presenten errores graves al clasificar la clase 2.

Hiperplano perceptrón pareja $\{0,3\}$ y datos recall, accuracy y precision.

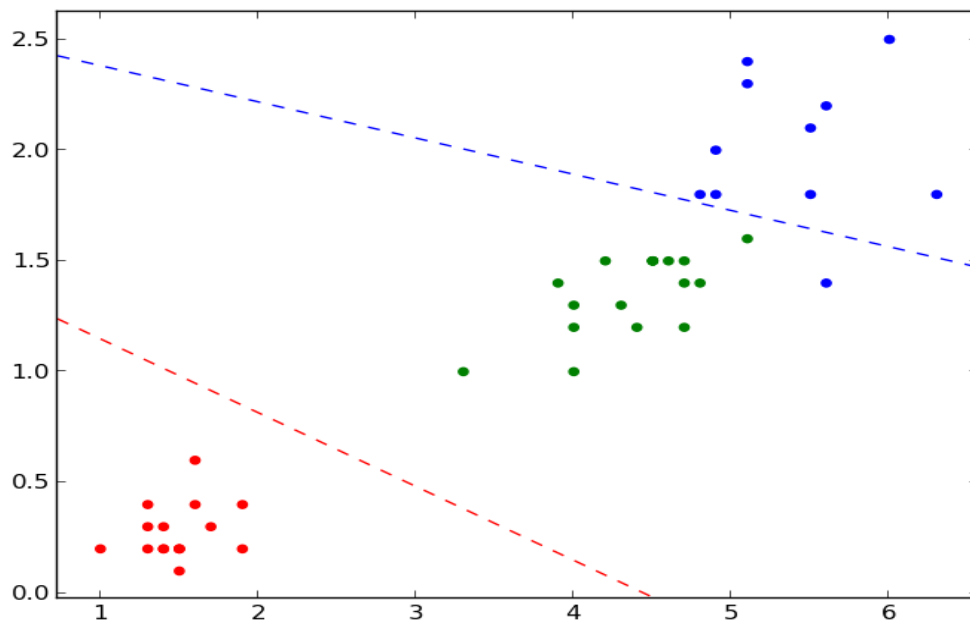


accuracy: 0.619047619048

recall: [1. 0. 0.79487179]

precision por clase: [0.46575342 0. 0.96875]

Hiperplano perceptrón pareja {2,3} y datos recall, acuaracy y precision.



Accuracy: 0.605238095238

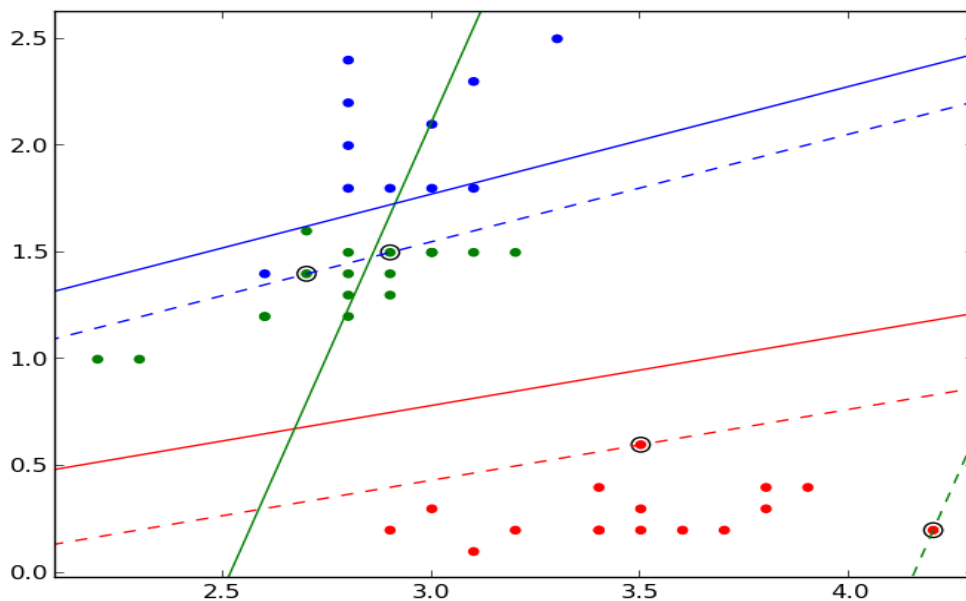
recall: [1. 0. 1.]

precision por clase: [0.72340426 0. 0.67241379]

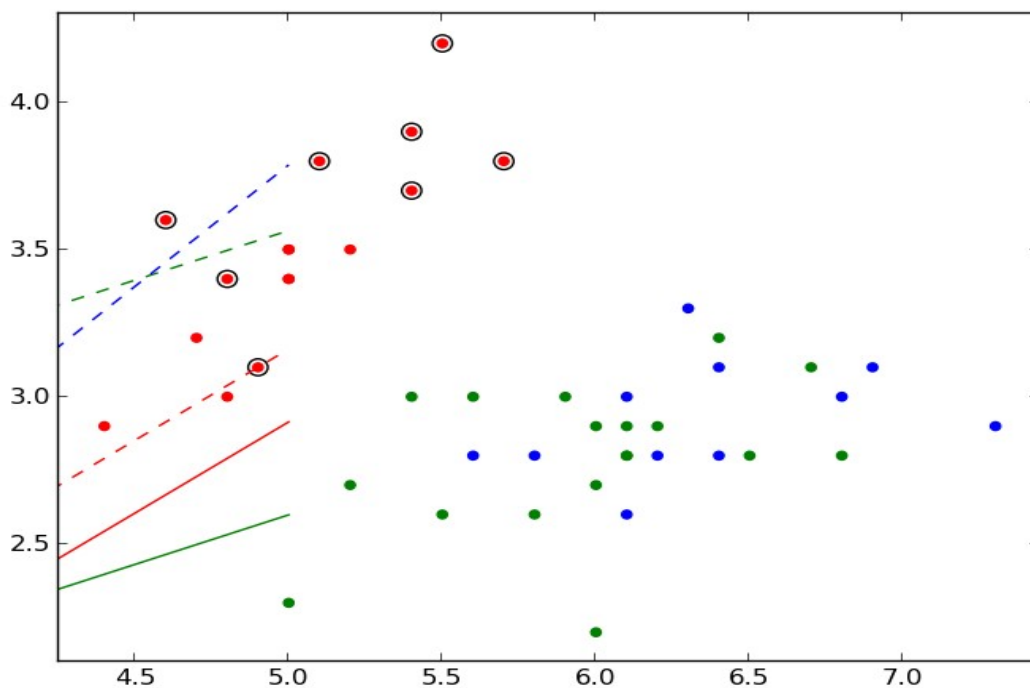
SVM

Para pintar los hiperplanos lo que se hace es calcular la norma w y calcular los puntos a una distancia $1/\|w\|$ del hiperplano. Se emplea la función “decision function” de la clase SVC. Se observa que las soluciones son parecidas a las de perceptron. La pareja $\{0,2\}$ sigue sin delimitar las clases 2 y 3, no es posible alcanzar un hiperplano óptimo. El clasificador con las clases $\{2,3\}$ podemos ver con el hiperplano se consigue el hiperplano óptimo para clase 1 y 2, en el caso de la clase 3 el hiperplano es bastante mejor que en el caso anterior.

Para la pareja $\{1,3\}$ el hiperplano es el siguiente:

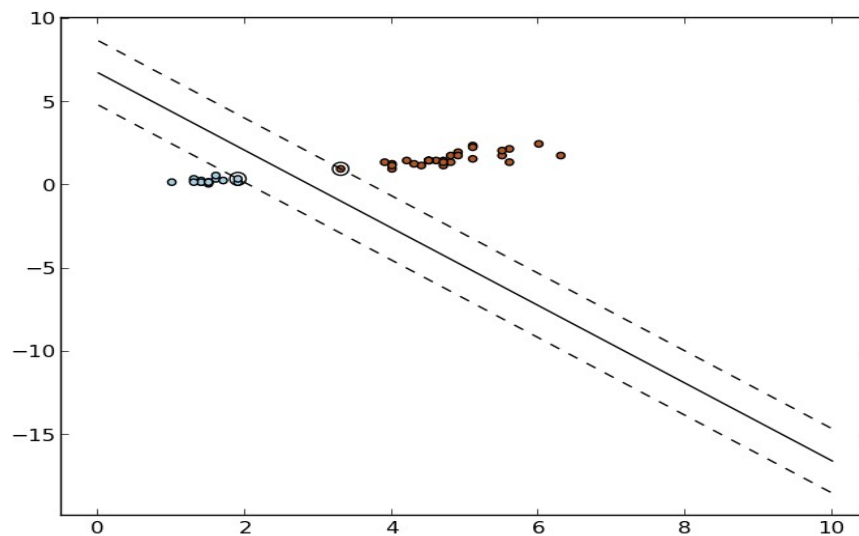


Para la pareja $\{0,1\}$ el hiperplano es el siguiente:



Por alguna razón no me pinta el hiperplano completo cuando incluyo esta pareja de variables y otras mas.

La función declarada en el código fuente como SVMApartadoA utiliza `svm.SVC(kernel='linear')` y el resultado del hiperplano es el siguiente :



Estimación del valor de C

El valor de C determina cuantos valores de error se permiten como máximo.

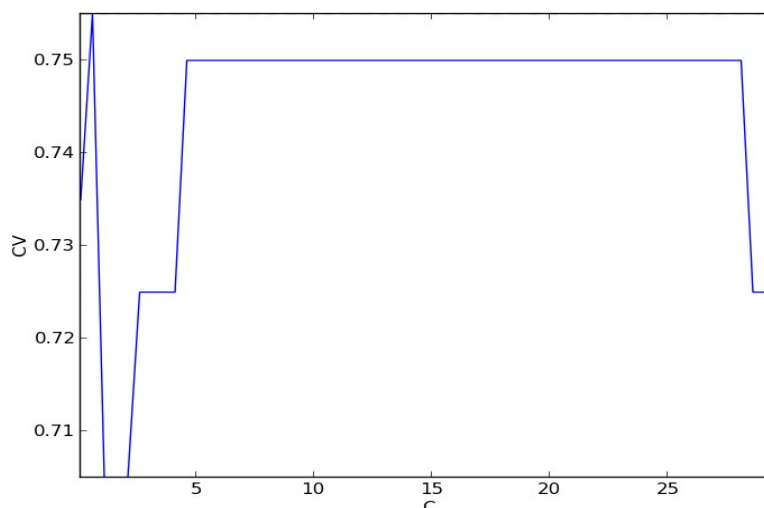
Utilizando validación cruzada CV=10, el valor para la primera pareja seria C=10 la segunda pareja seria C=15.

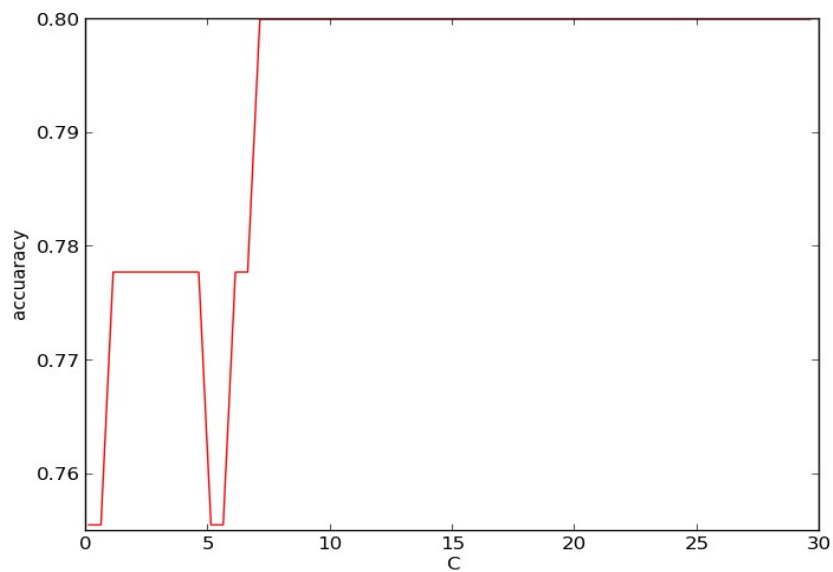
Representación y valoración de los diferentes Kernerls.

Para la pareja {0,3} la clase 1 es linealmente separable. En relación a los kernel el area que ocupa la región de cada clase varia, asi la lineal tiene un área mayor para la clase 1 que la polinomial. Para la clase 3 la mas grande es el kernel rbf y para la clase 2 la mas pequeña.

Para la pareja {2,3} podemos observar que son casi linealmente separables, aunque presena errores en los limites de las clases. Esto confirma que las variables 2 y 3 son las mas características de este ejemplo.

Para este apartado se muestran dos gráficas que representan el valor de C respecto a CV en la primera gráfica y en la segunda el valor de C respecto a Accuaracy.

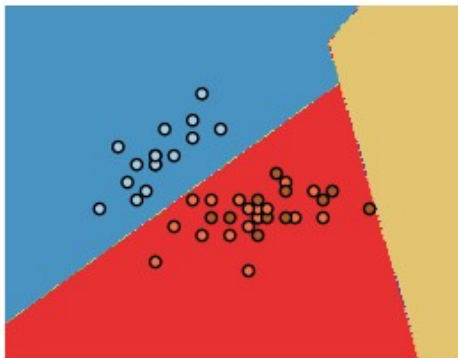




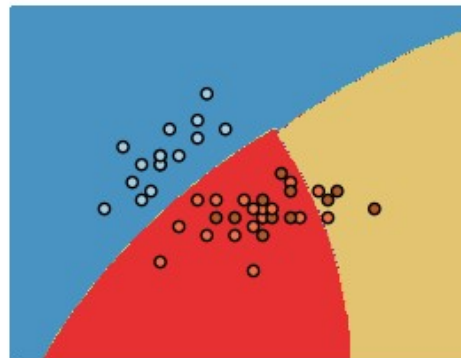
Además se muestra en una gráfica el resultado de los clasificadores utilizando los diferentes kernel que se especifican en el guión de practicas.

Este corresponde a la pareja $\{0,3\}$

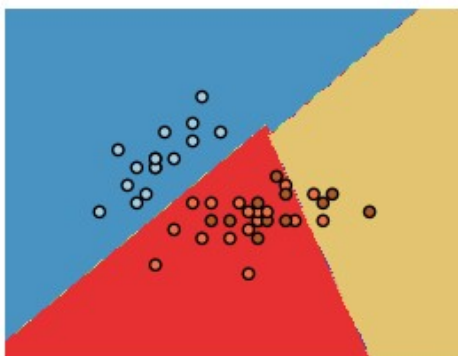
SVC with linear kernel



SVC with RBF kernel

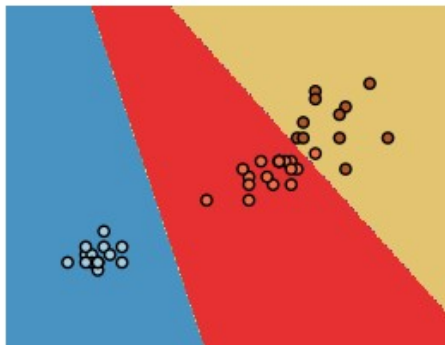


SVC with polynomial kernel

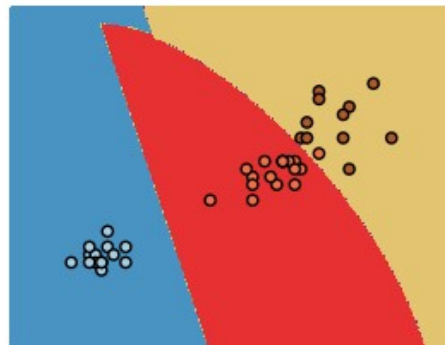


Este corresponde a la pareja {2,3}

SVC with linear kernel



SVC with RBF kernel



SVC with polynomial kernel

