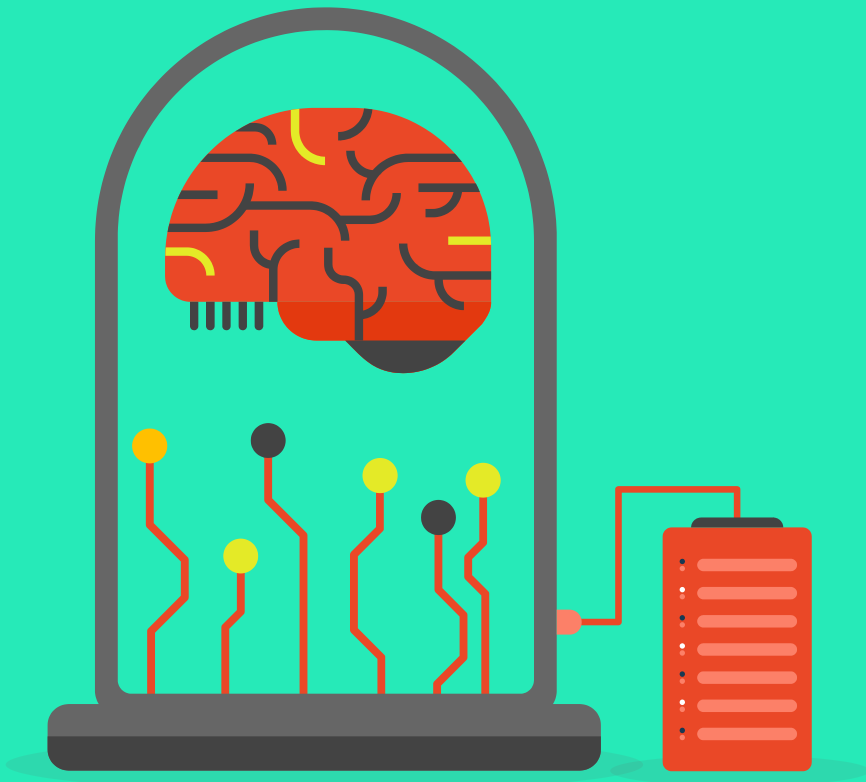


7. Regresión con Vectores de Soporte



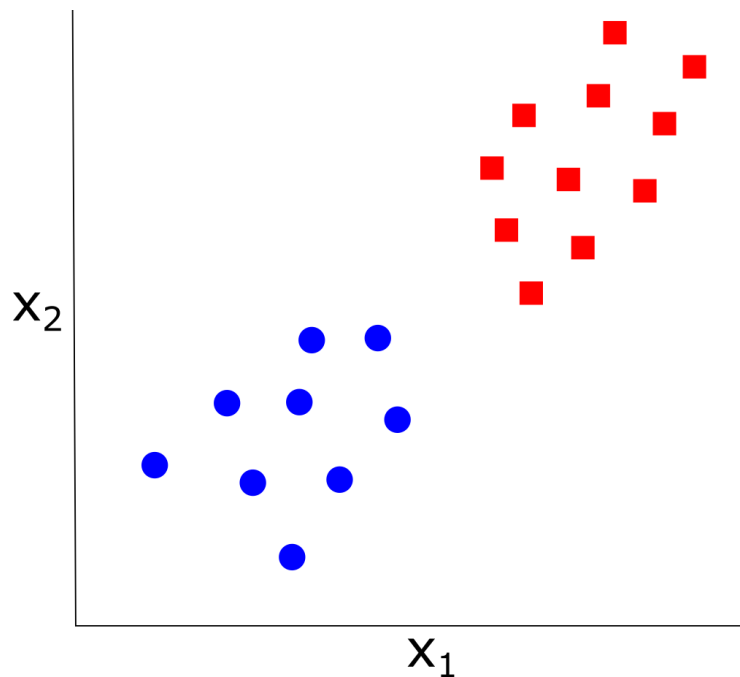
Máquinas de Soporte Vectorial

Las Máquinas de Soporte Vectorial (SVM, por sus siglas en inglés), son un tipo de modelo de Machine Learning capaces de ejecutar tareas de clasificación, regresión y detección de *outliers*. En esta sección exploraremos los conceptos alrededor de la Regresión con Vectores de Soporte (SVR, por sus siglas en inglés), pero a fin de entender su funcionamiento e implementación, es necesario partir de las SVM como modelos de clasificación (más adelante veremos a detalle el problema de la clasificación en el Machine Learning y cómo se implementan los distintos algoritmos).

A fin de exponer de manera breve el problema de la clasificación y en qué consiste, veamos un ejemplo gráfico de clasificación lineal.

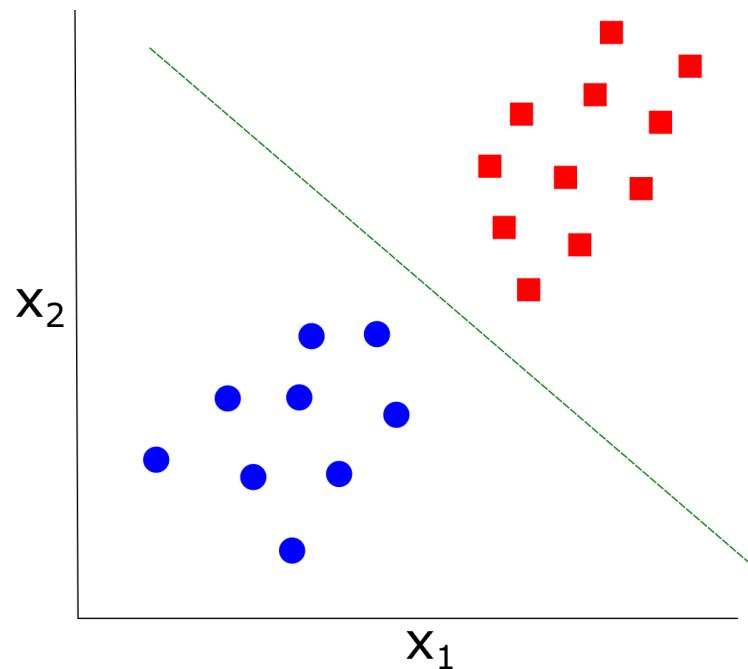
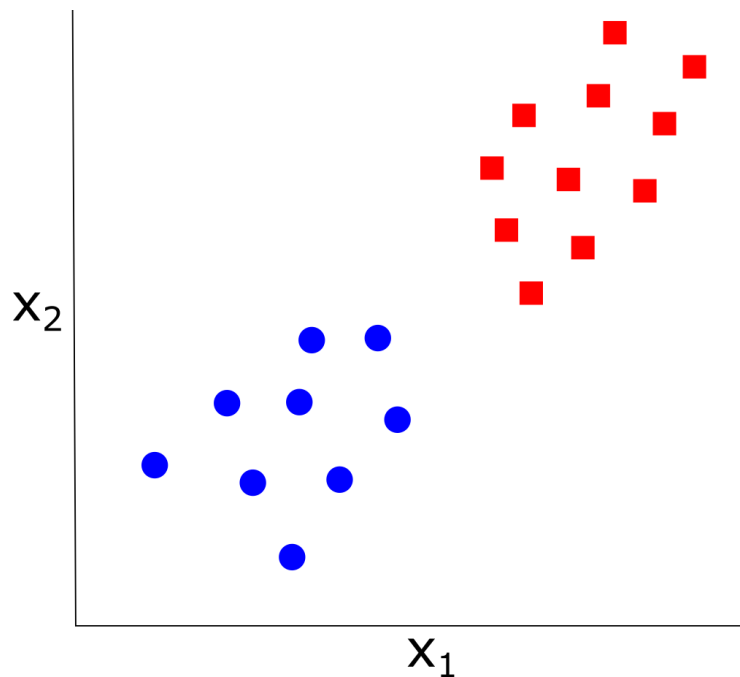
Máquinas de Soporte Vectorial

Supongamos que tenemos un conjunto de datos distribuidos en el espacio de la manera como se observa en la primera figura.



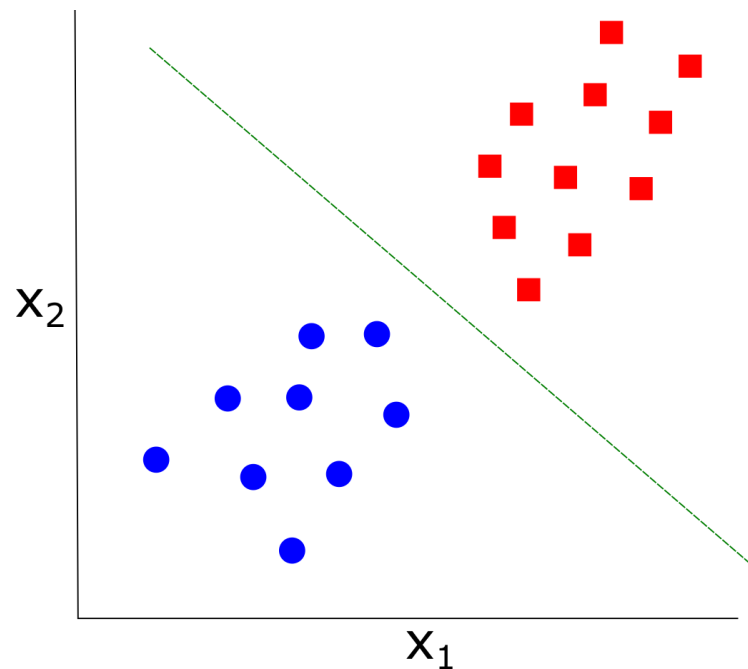
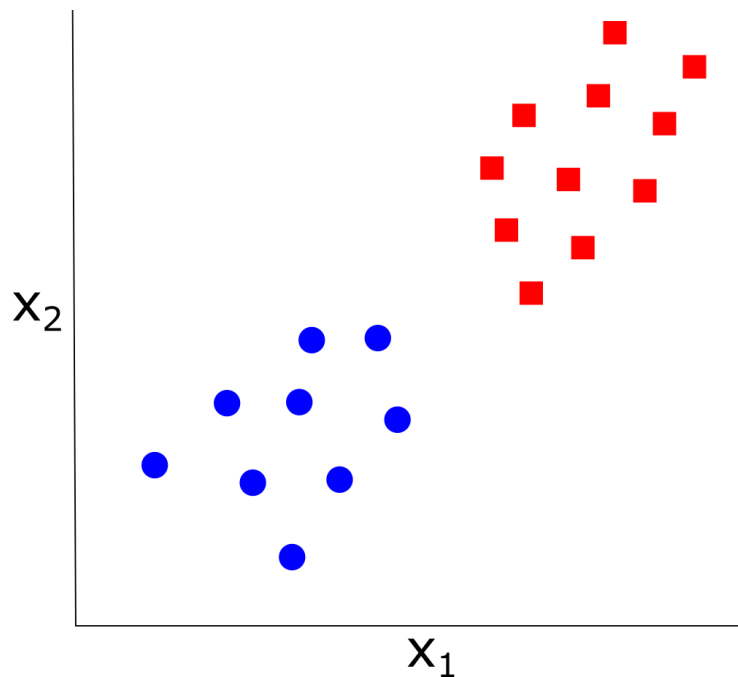
Máquinas de Soporte Vectorial

A partir de la figura, resulta claro que solo bastaría una línea recta que separe las clases para determinar los datos que pertenecen a cada una de ellas.



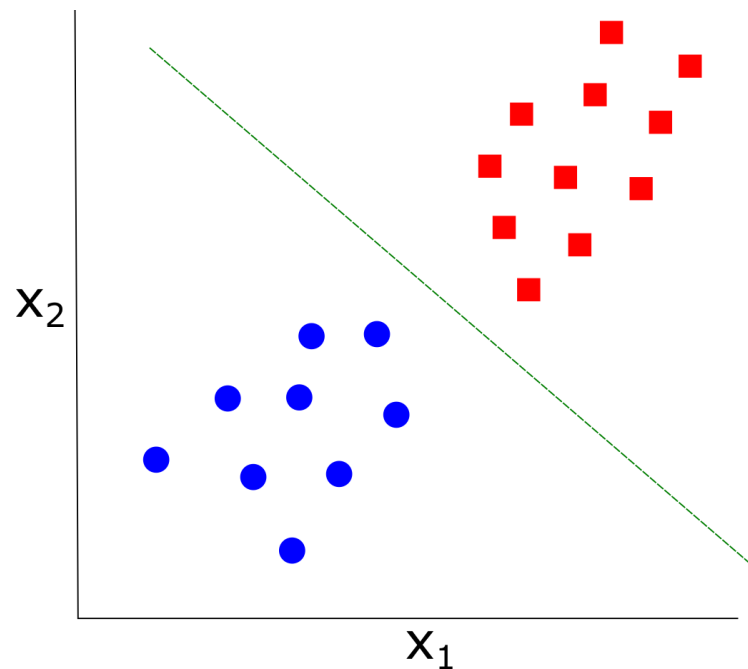
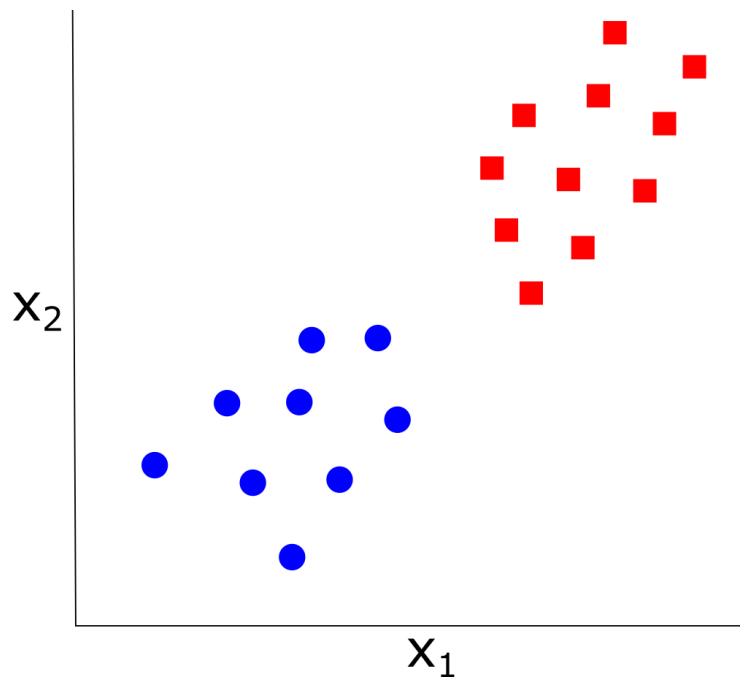
Máquinas de Soporte Vectorial

Una vez obtenida esta línea, podríamos afirmar que cualquier dato a la derecha de ella, pertenece a la clase roja, y cualquiera a la izquierda, pertenece a la clase azul.



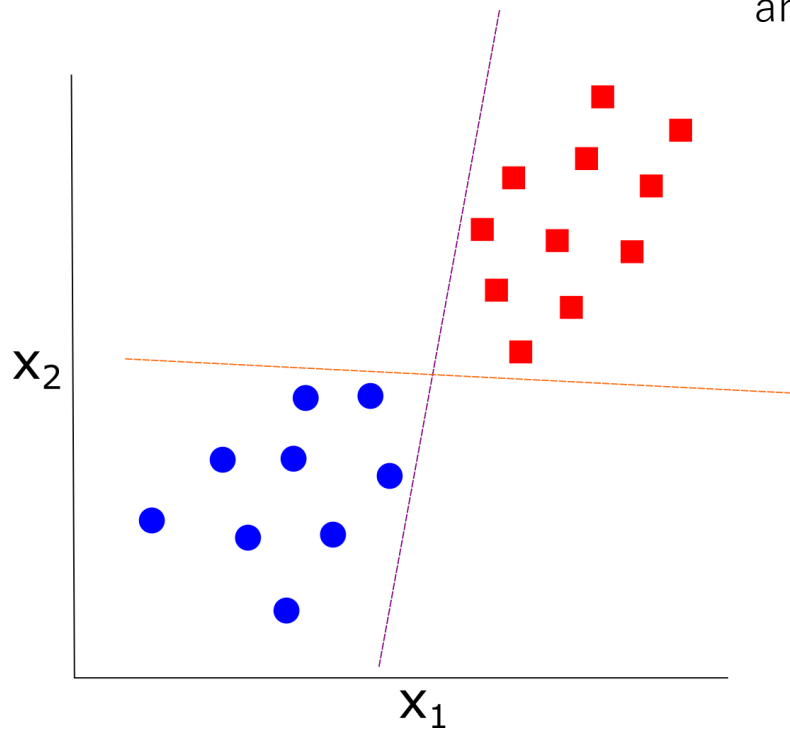
Máquinas de Soporte Vectorial

A este tipo de problemas de clasificación se le llama **Linealmente Separable**, y a la línea que separa ambas clases se conoce como **Borde de Decisión**.



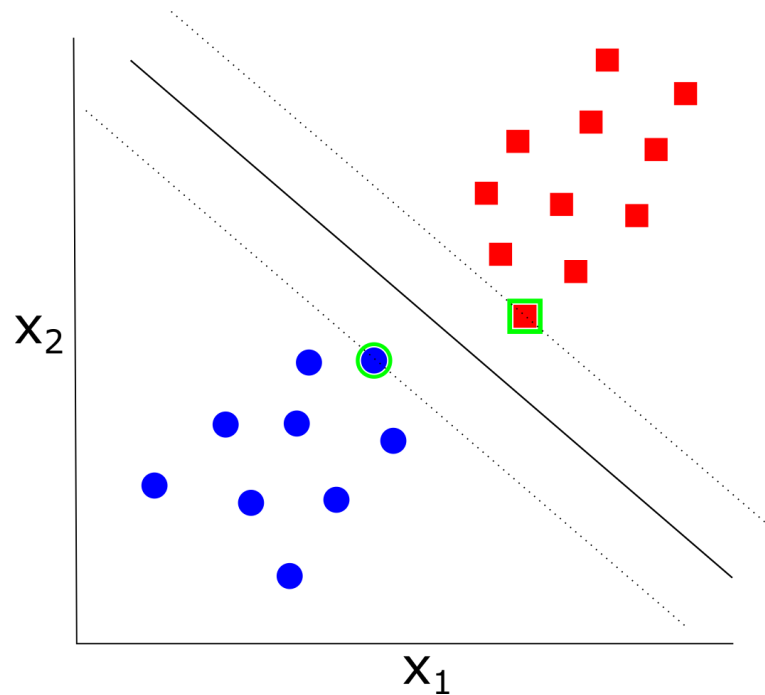
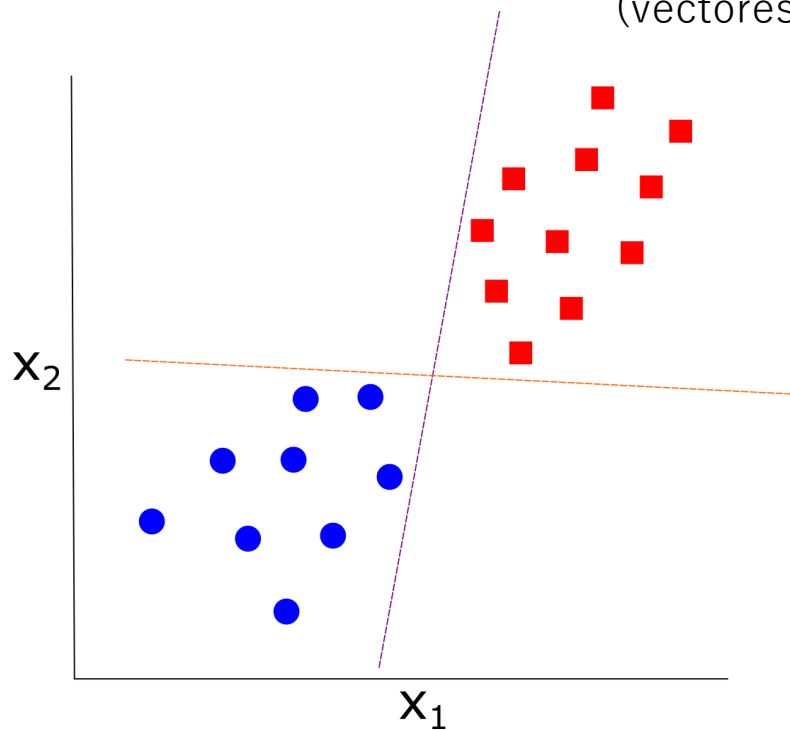
Máquinas de Soporte Vectorial

Ahora bien, técnicamente, cualquiera de los bordes de decisión de la figura de la izquierda es capaz de separar las clases de manera lineal. ¿Pero cuál es el mejor de ambos?



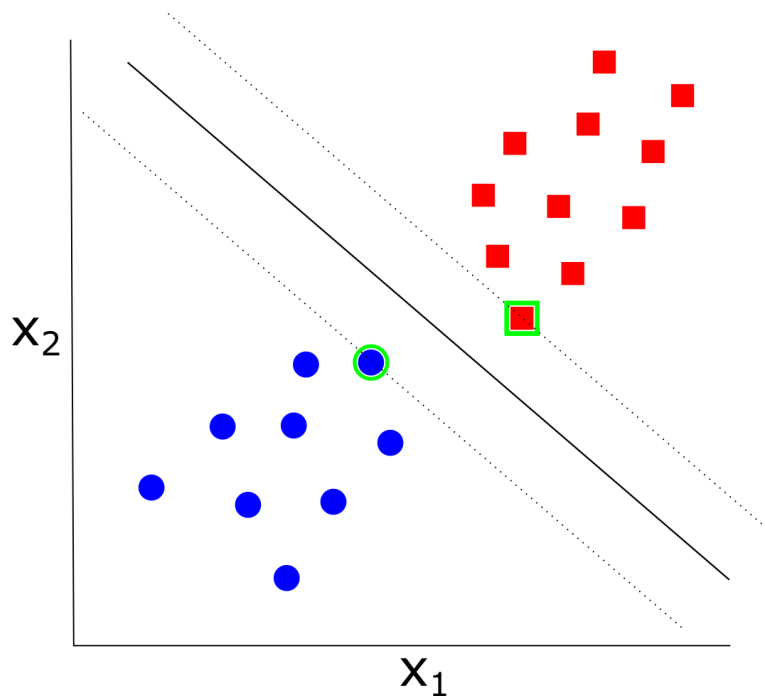
Máquinas de Soporte Vectorial

Las SVM basan su funcionamiento en encontrar el borde de decisión que, acotado por un *margen*, se encuentra más alejado de las clases de entrenamiento más cercanas (vectores de soporte).



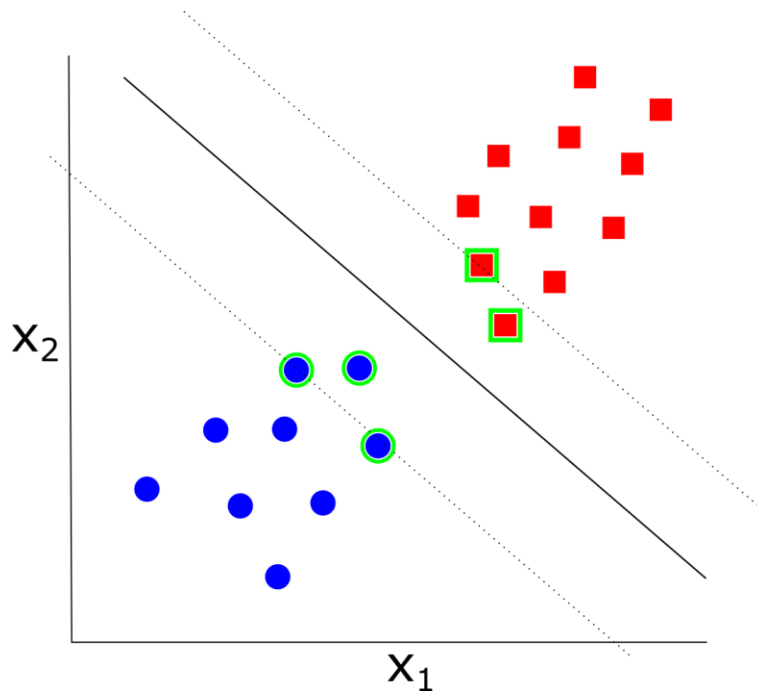
Máquinas de Soporte Vectorial

Incorporar más datos de entrenamiento no afectará el borde de decisión, ya que el margen está siempre determinado por los más cercanos, es decir, los vectores de soporte. A este tipo de clasificación se le conoce como de **Margen Duro**.



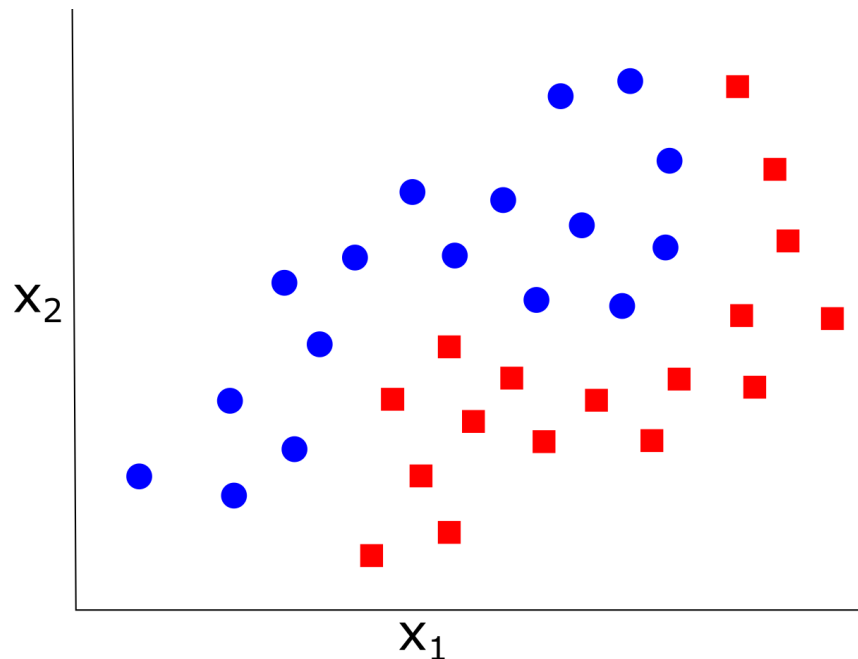
Máquinas de Soporte Vectorial

Si relajamos la restricción sobre la cantidad de vectores de soporte que permitimos entrar en el margen, tenemos entonces clasificación de **Margen Suave**. Esta restricción es un parámetro variable del algoritmo y ayuda a regular el modelo.



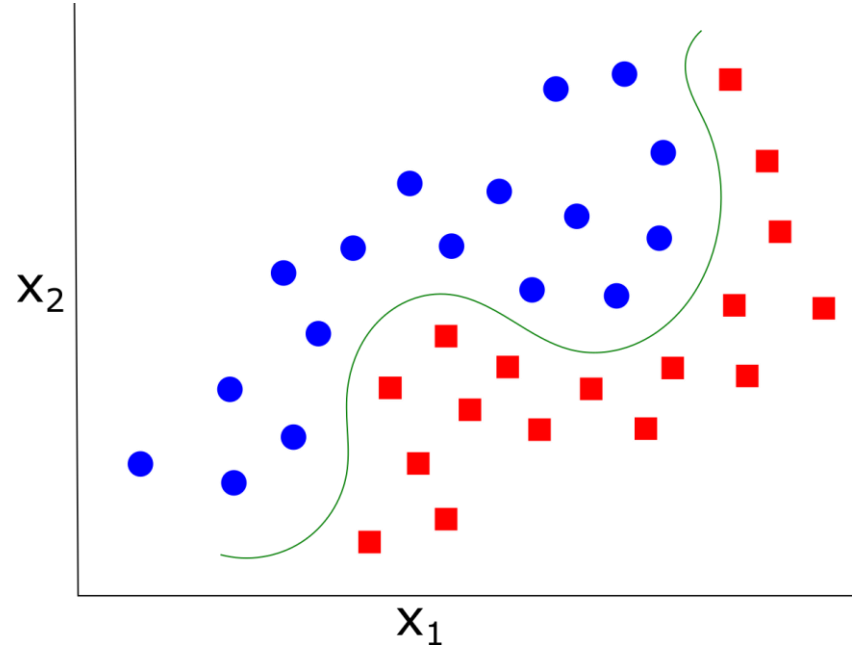
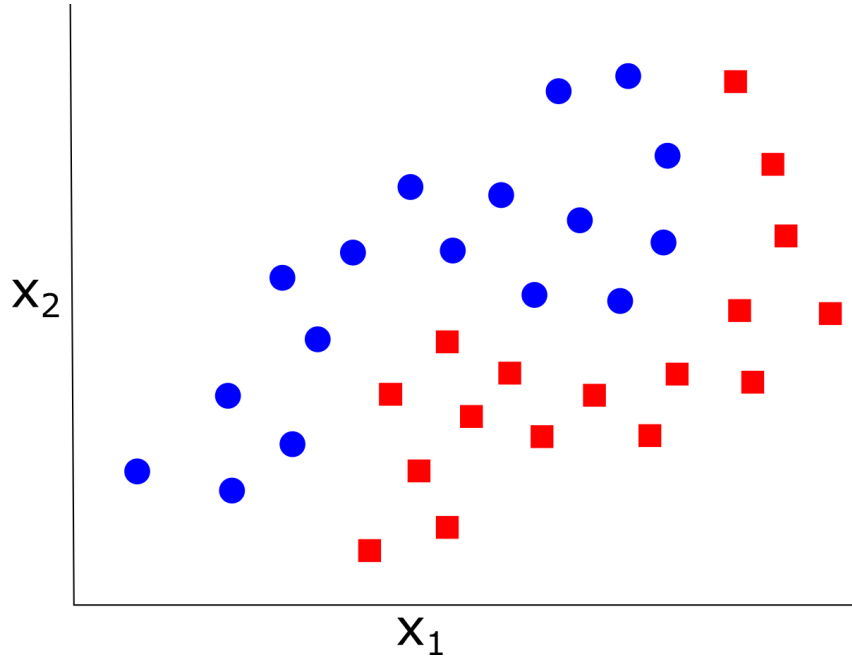
Máquinas de Soporte Vectorial

Ahora bien, ¿qué ocurre en el caso de problemas **Linealmente No Separables**? En el caso de la figura, no es posible dibujar una línea recta que separe las clases.



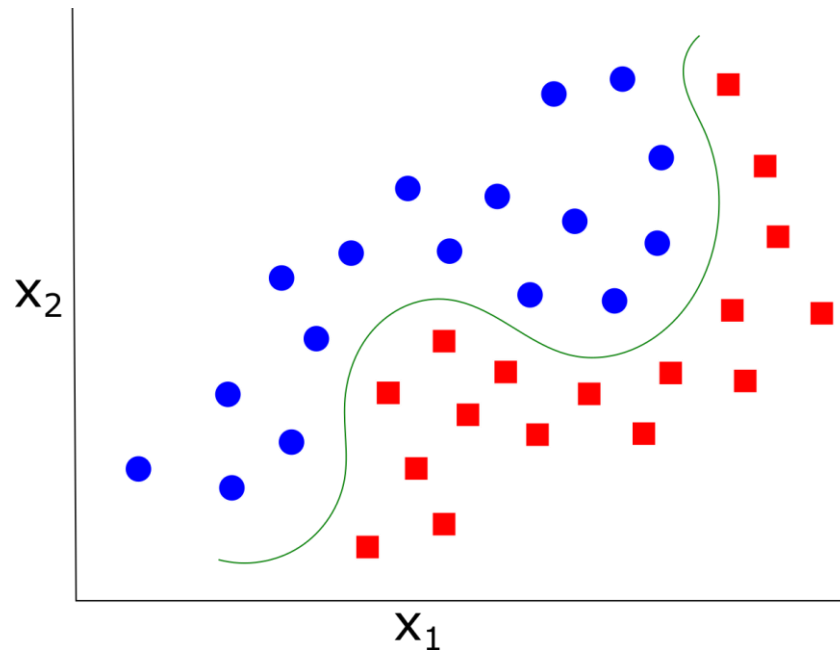
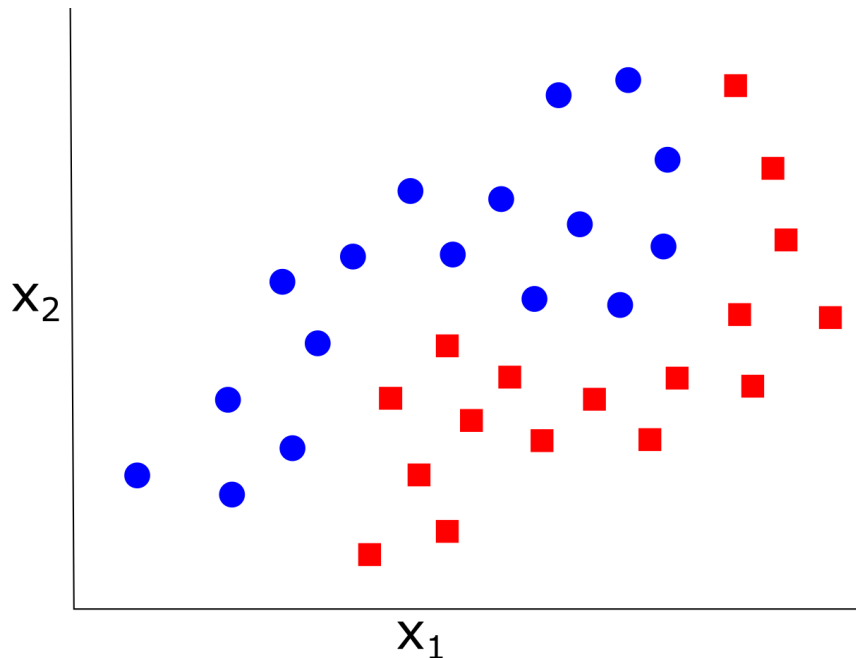
Máquinas de Soporte Vectorial

Pero sí es posible dibujar una curva que lo haga.



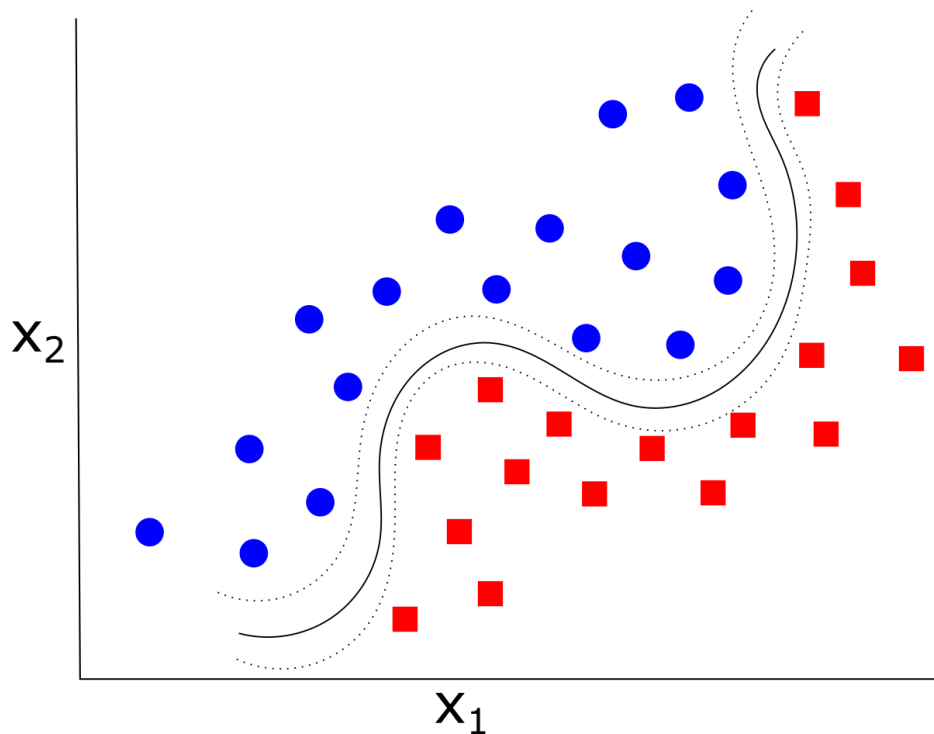
Máquinas de Soporte Vectorial

En este sentido, la clasificación con problemas no lineales con SVM se logra al incorporar funciones no lineales (polinomios, por ejemplo) llamados **kernels**.



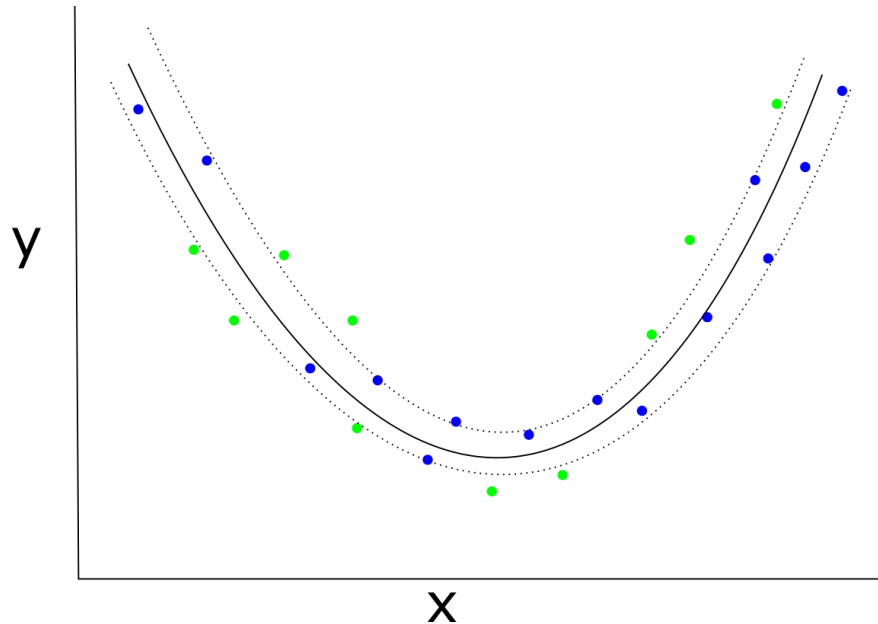
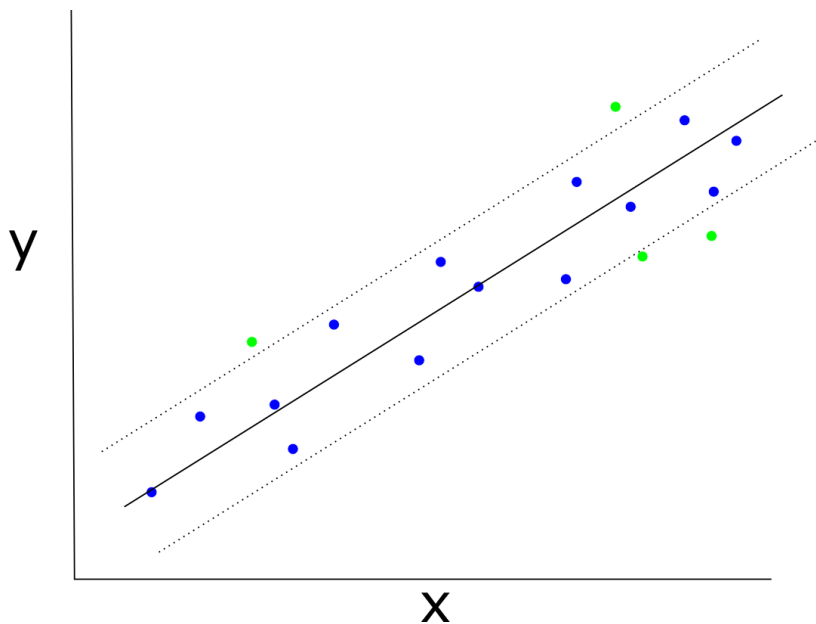
Máquinas de Soporte Vectorial

Del mismo modo que en el caso lineal, el algoritmo busca maximizar el margen alrededor del borde de decisión a partir de los vectores de soporte.



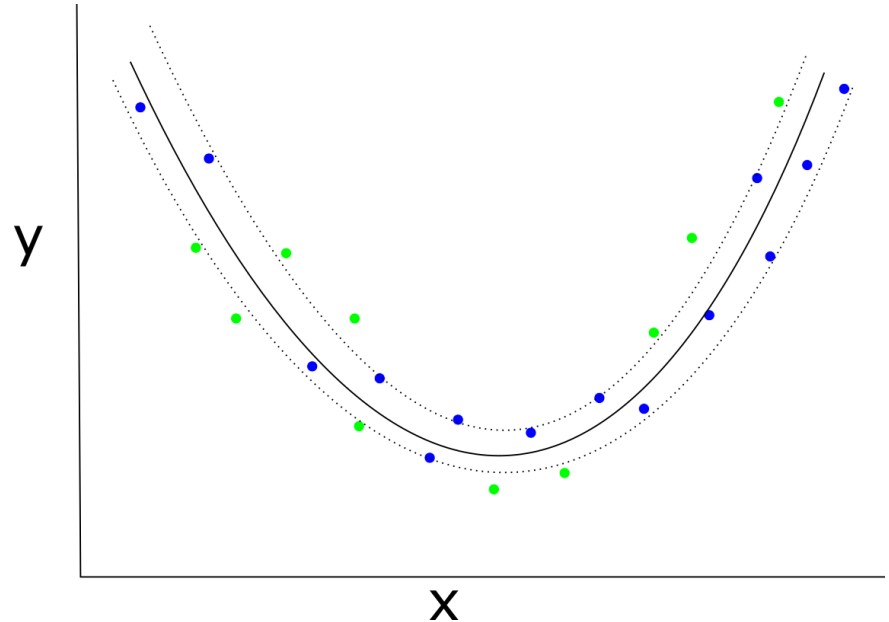
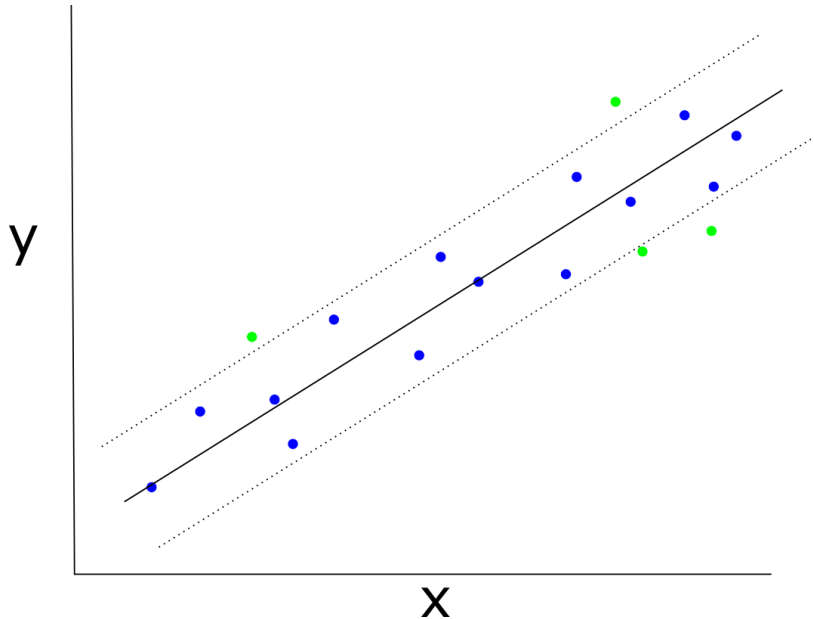
Regresión con Vectores de Soporte

De manera similar, es posible realizar **Regresión** con vectores de soporte. Sin embargo, en el caso de la regresión, el objetivo es el contrario: se trata de encontrar el margen que mayor cantidad de puntos incluya, limitando la cantidad que caigan afuera.



Regresión con Vectores de Soporte

El ancho del margen, y por lo tanto la regularización del modelo, estará regulado por una constante llamada ϵ .



Regresión con Vectores de Soporte

La derivación matemática del cómo se obtienen los bordes de decisión para las SVM y las SVR es bastante compleja, e implica la solución de un problema de optimización con restricciones el cual se resuelve de manera exacta con una técnica llamada Programación Cuadrática. Sin embargo, también es posible obtener las soluciones para los coeficientes del borde de decisión haciendo uso del método del Descenso del Gradiente. Para el caso de un problema linealmente separable, la función de costo tiene la forma:

$$J(\mathbf{w}, b) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^m \max(0, 1 - t^{(i)}(\mathbf{w}^T x^{(i)} + b))$$

En donde el primer término obligará al modelo a tener pesos pequeños y margen grande, y el segundo asegura que el modelo cometa la menor cantidad de violaciones del margen posibles.

Octavo Notebook Práctico

Adicionalmente, es importante mencionar que la Regresión con Vectores de Soporte y las Máquinas de Vectores de Soporte en general, son sensibles a la escala de las variables, por lo que es siempre necesario aplicar el Escalado Estándar (Standard Scaling) antes de llevar a cabo el entrenamiento del modelo.

Ya que conocemos los fundamentos y la intuición detrás de este nuevo tipo de modelo de Machine Learning, veamos el cómo se implementa y se resuelven problemas de regresión con las Máquinas de Soporte Vectorial.

Octavo Notebook Práctico