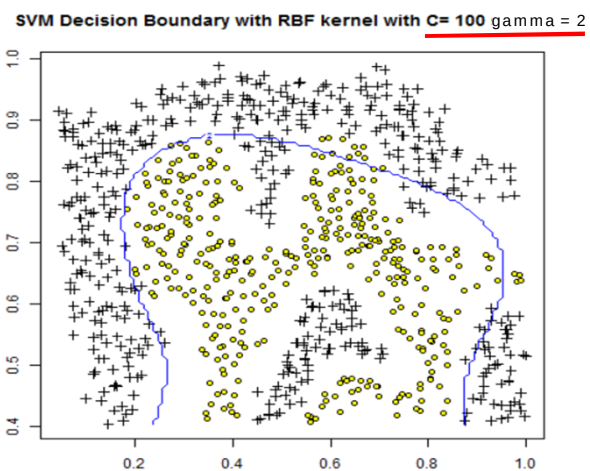
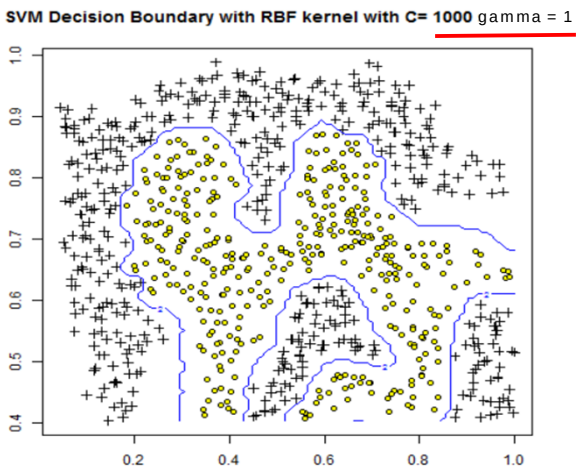
Gaussian RBF Kernel

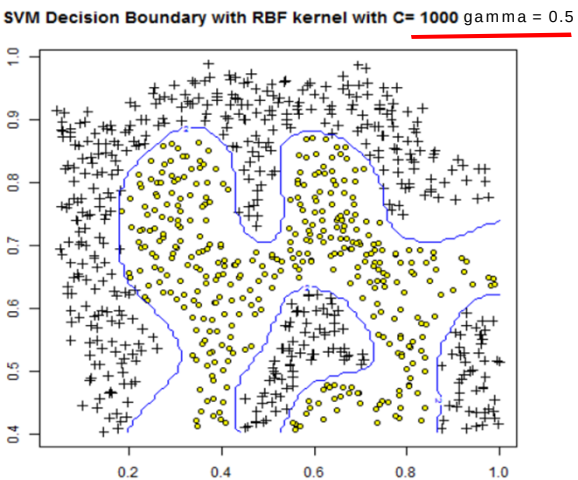
Observe the **change in Decision Boundary** for **different values** of **C** and **gamma**

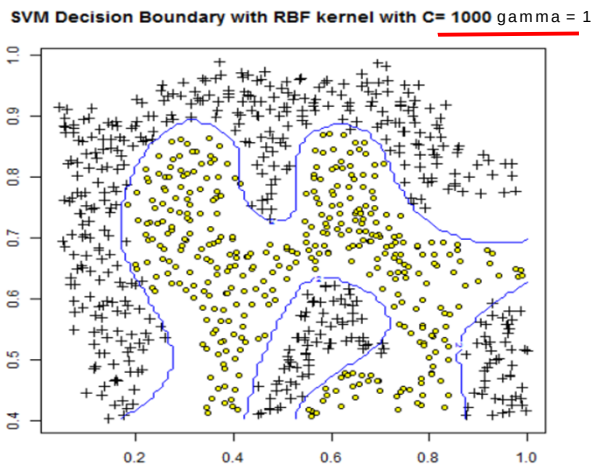




Gaussian RBF Kernel

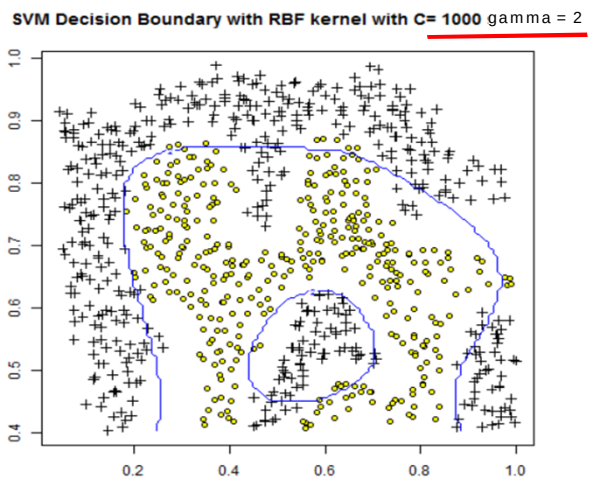
Observe the **change in Decision Boundary** for **different values** of **C** and **gamma**





Gaussian RBF Kernel

Observe the **change in Decision Boundary** for **different values** of **C** and **gamma**



Gaussian RBF Kernel

Appropriate values for **C** and **gamma** are arrived after examining **Confusion Matrix**

Agenda

- What is SVM?
- Ideology behind SVM
- Intuition Development
- Terminologies used in SVM
- How does it work?
- What is Kernel trick?
- Types of kernels
  - Polynomial Kernel
  - Gaussian RBF Kernel
- Support Vector Regression
- Pros and Cons of SVM
- Data preparation for SVM
- Use Case - House Prices

Support Vector Regression (SVR)

As of now everyone is familiar with SVM or Support Vector Machine.

But **SVR** is **different from SVM**.

As the name suggests **SVR** is a **regression algorithm**, so we can use **SVR** for **Continuous values** instead of Classification which is SVC.

Terminologies related to SVR

Kernel

- The function used to map a **lower** dimensional data to **higher** dimensional data.

Hyper Plane

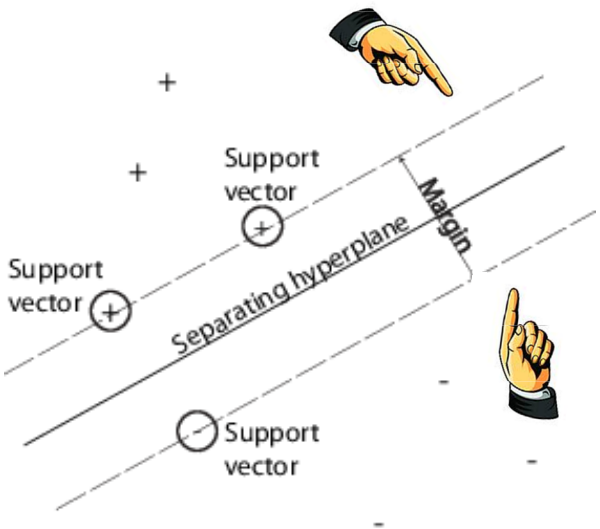
- In **SVM** this is basically the **separation line between the data classes**.
- **In SVR** we are going to **define it as the line** that will help us **predict the *continuous value* or *target value***.

Terminologies related to SVR

Boundary line

- In SVM, there are **two lines** other than Hyper Plane which creates a **margin**.

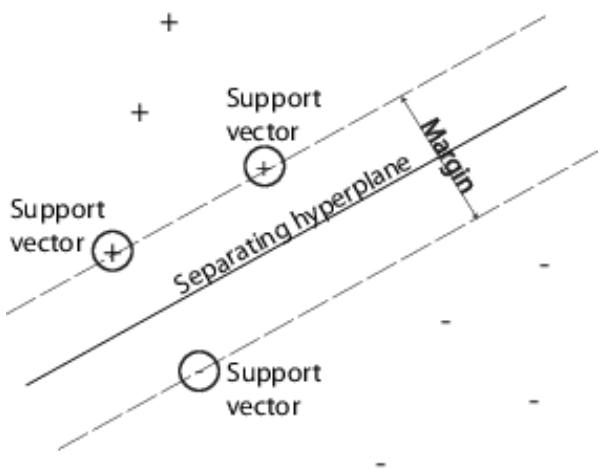
- The **support vectors** can be on the **Boundary lines** or outside it. This boundary line separates the two classes. In SVR the concept is same.



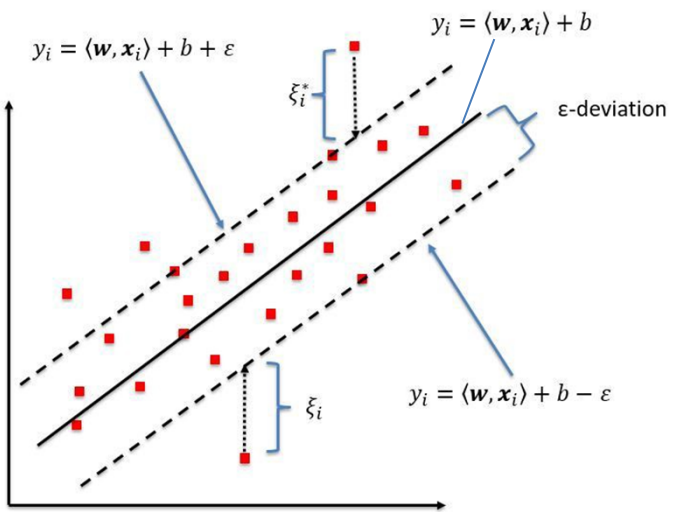
Terminologies related to SVR

Support vectors

- This are the data points which are **closest** to the **boundary**. The distance of the points is **minimum or least**.



SVR



'ξ' (epsilon) is a hyper-parameter

SVR

Basically we are trying to decide a **decision boundary** at **'ξ'** distance from the **original hyper plane** such that data points **closest to the hyper plane** or the **support vectors** are **within that boundary line**.

Agenda

- What is SVM?
- Ideology behind SVM
- Intuition Development
- Terminologies used in SVM
- How does it work?
- What is Kernel trick?
- Types of kernels
  - Polynomial Kernel
  - Gaussian RBF Kernel
- Support Vector Regression
- Pros and Cons of SVM
- Data preparation for SVM
- Use Case - House Prices

Pros and cons of SVM

Pros

- It is really **effective** in the **higher dimension**.
- Effective when the **number of features are more** than **training examples**.
- Best algorithm when **classes are separable**.
- The **hyperplane** is **affected** by only the **support vectors** thus **outliers have less impact**.

Pros and cons of SVM

Cons

- For larger dataset, it requires a **large amount** of time to process.
- **Does not** perform well in case of **overlapped classes**.
- Selecting, appropriately **hyper parameters** of the SVM that will allow for **sufficient generalization performance**.
- Selecting the appropriate **kernel function** can be tricky.

Agenda

- What is SVM?
- Ideology behind SVM
- Intuition Development
- Terminologies used in SVM
- How does it work?
- What is Kernel trick?
- Types of kernels
  - Polynomial Kernel
  - Gaussian RBF Kernel
- Support Vector Regression
- Pros and Cons of SVM
- Data preparation for SVM
- Use Case - House Prices

Preparing data for SVM

Numerical Conversion

- SVM assumes that your inputs are **numerical** instead of **categorical**.
- So you can convert them using one of the most commonly used **one hot encoding** or **label-encoding etc**.

Preparing data for SVM

Binary Conversion

- Since SVM is able to classify only binary data so you would need to convert the **multi-dimensional dataset into binary form** using (**one vs the rest method / one vs one method**) conversion method.

Agenda

- What is SVM?
- Ideology behind SVM
- Intuition Development
- Terminologies used in SVM
- How does it work?
- What is Kernel trick?
- Types of kernels
  - Polynomial Kernel
  - Gaussian RBF Kernel
- Support Vector Regression
- Pros and Cons of SVM
- Data preparation for SVM
- Use Case - House Prices

Use Case : House Prices

- Ask a home buyer to describe their **dream house**, and they probably won't begin with the **height** of the basement ceiling or the **proximity** to an east-west railroad. But this dataset proves that **much more** influences price negotiations than the *number of bedrooms or a white-picket fence*.

Use Case : House Prices

- With **81 explanatory variables** describing (almost) every aspect of **residential homes in Ames, Iowa**, this dataset challenges you to **predict the final price of each home**.