

فرآیند یادگیری در یادگیری ماشین

دوره پایتون و یادگیری ماشین



</Monolearn>

پارامترهای یادگیری (در یادگیری نظارت شده)

- بخش اصلی در فرآیند یادگیری، تکرار است. (که این تکرارها در واقع همان سعی و خطاها هستند.)
- هر یک از مدل ها یا الگوریتم ها در یادگیری ماشین، تعدادی پارامتر اختیاری یا اجباری دارند که در فرآیند یادگیری مدل استفاده می شوند.
- الگوریتم این پارامترها را آنقدر دستکاری می کند تا به بهترین خروجی (بالاترین دقت) و همچنین کمترین میزان خطا برسد.

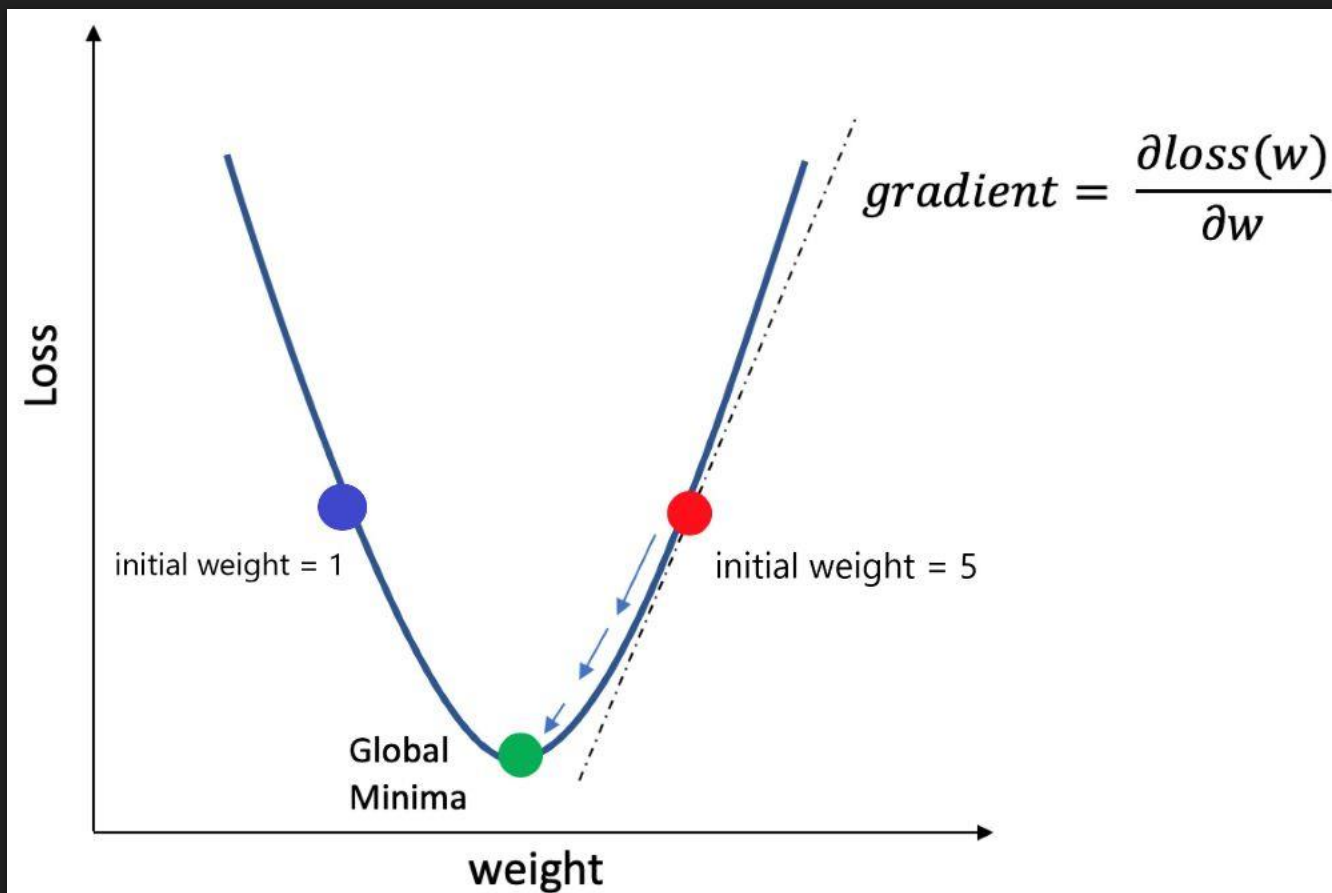
به شکل دقیق تر

- هنگام یادگیری، مدل یادگیری ماشین تلاش می کند تا شرایط یک تابع هدف یا Object function را برآورده کند.
- در عین حال، تلاش می کند تا مقدار Loss یا Error که از طریق تابعی تحت عنوان Loss function محاسبه می شود را کمینه یا مینیمم کند.
- همچنین، هر از پارامترهای ورودی ممکن است در یک مقدار W تحت عنوان Weight یا B تحت عنوان Bias ضرب شوند که مدل یادگیری ماشین سعی می کند تا این وزن ها و بایاس ها را یاد بگیرد.

گرادیان کاهشی یا Gradient Descent

- یکی از الگوریتم هایی است که مدل های یادگیری ماشین بسیاری استفاده می شود.
- این الگوریتم برای پیدا کردن مقدار کمینه یا Minimum تابع Loss یا Loss function استفاده می شود.
- این الگوریتم یک الگوریتم بهینه سازی (از نوع تکرار شونده) است.
- این الگوریتم از مرتبه ی زمانی $O(kn^2)$ است که K تعداد ویژگی ها یا Feature ها و n تعداد کل داده ها است.
- به شکل دقیق تر، گرادیان کاهشی سعی در پیدا کردن کمینه موضعی یا Local minimum روی یک تابع مشتق پذیر را دارد.

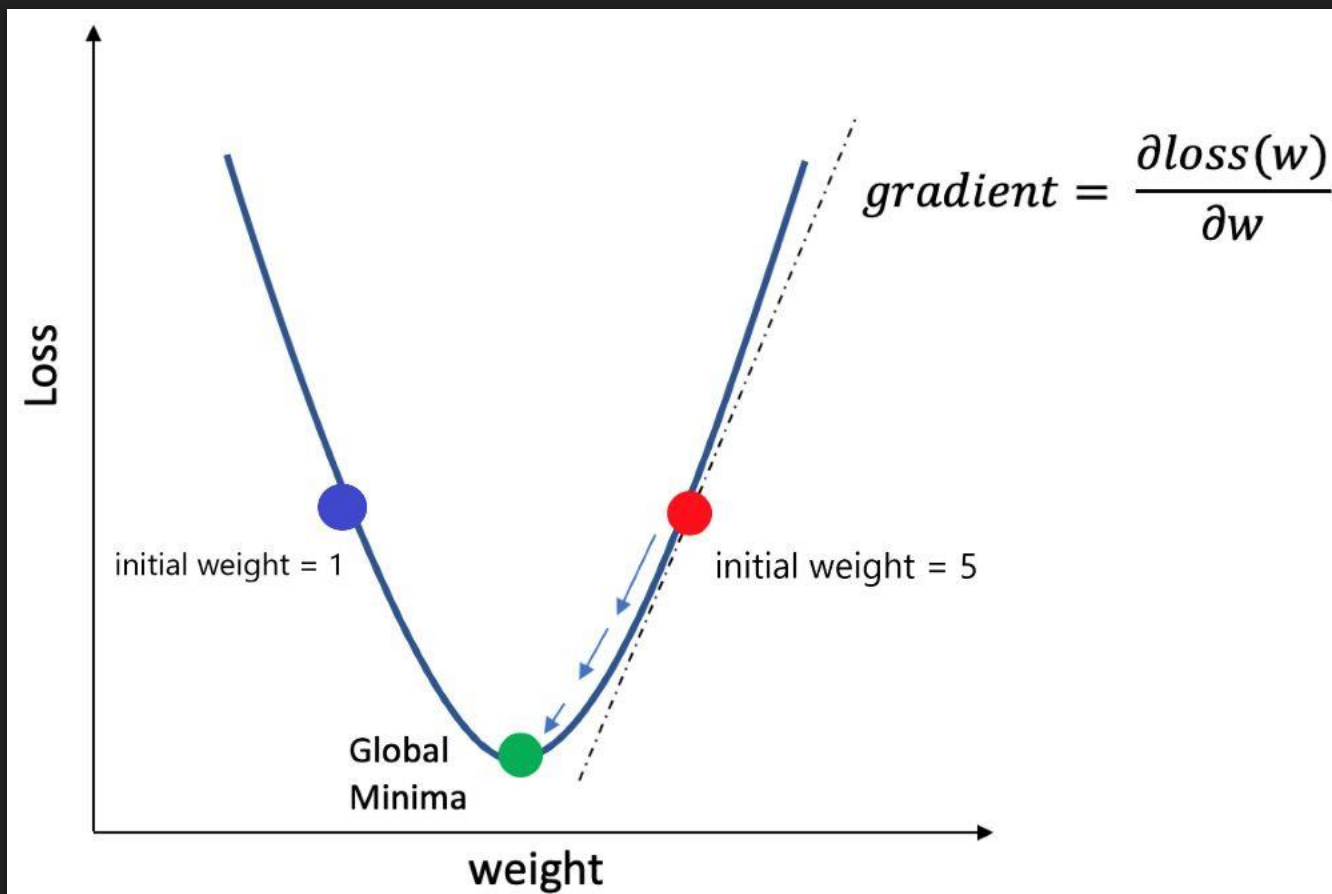
روش کار گرادیان کاهشی یا Gradient Descent



$$\text{new weight} = \text{initial weight} - d \text{ loss}/dw$$

- اگر مقدار W جدید، از مقدار W کمینه، کمتر باشد، یعنی سمت چپ W کمینه باشیم، شیب خط یا مشتق نسبت به W منفی است و این یعنی اینکه new weight مقدار بیشتری خواهد شد و در نتیجه از سمت چپ به سمت W کمینه حرکت می کنیم.
- هنگامی که W جدید، از مقدار کمینه W بیشتر باشد، و در سمت راست W کمینه باشیم، مشتق عبارت بالا مثبت است و این یعنی اینکه مقدار new weight مقدار کمتری خواهد شد و در نتیجه از سمت راست به سمت W کمینه حرکت می کنیم.

روش کارگرادیان کاهشی یا Gradient Descent



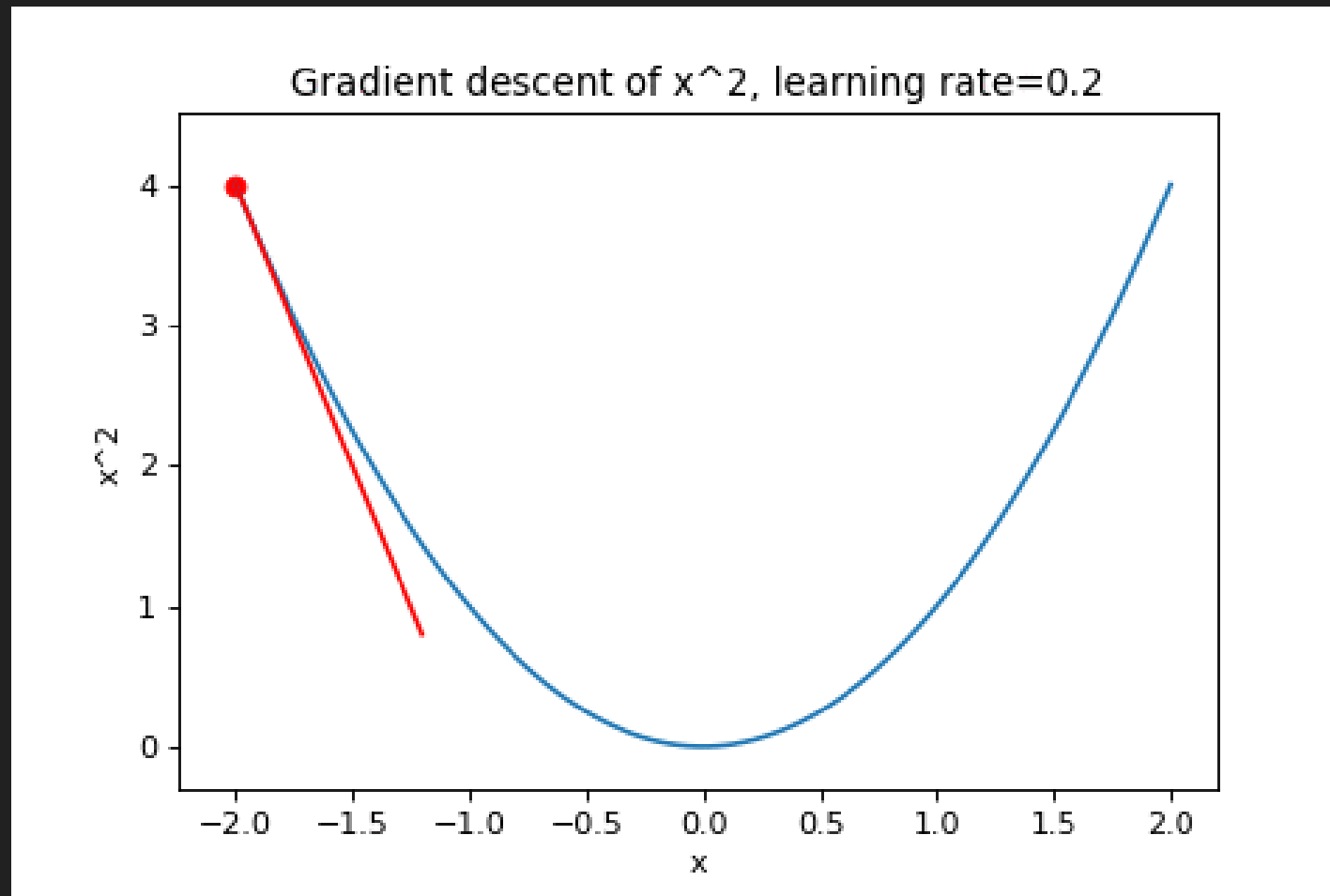
$$\text{new weight} = \text{initial weight} - d \text{ loss}/dw$$

○ در واقع برای بهبود سرعت این فرایند بهتر است از ضربی به اسم learning rate یا نرخ یادگیری استفاده کنیم.

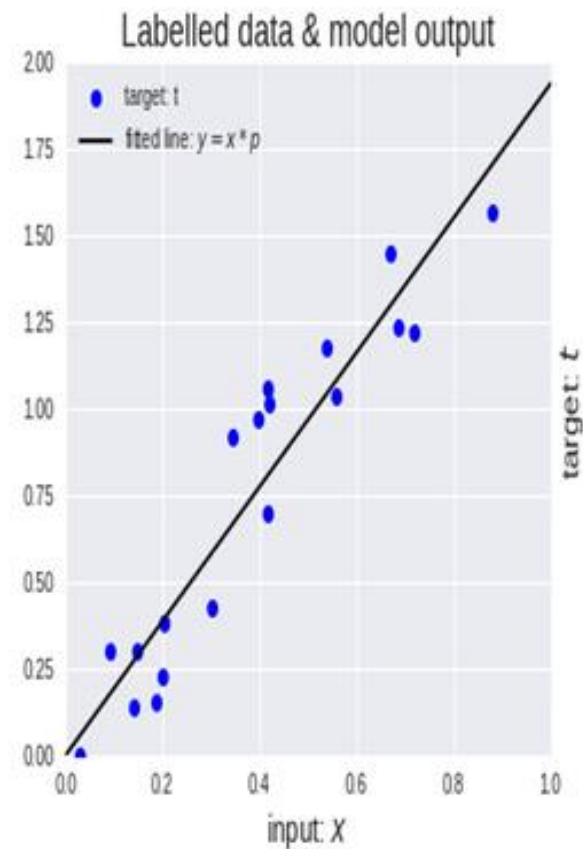
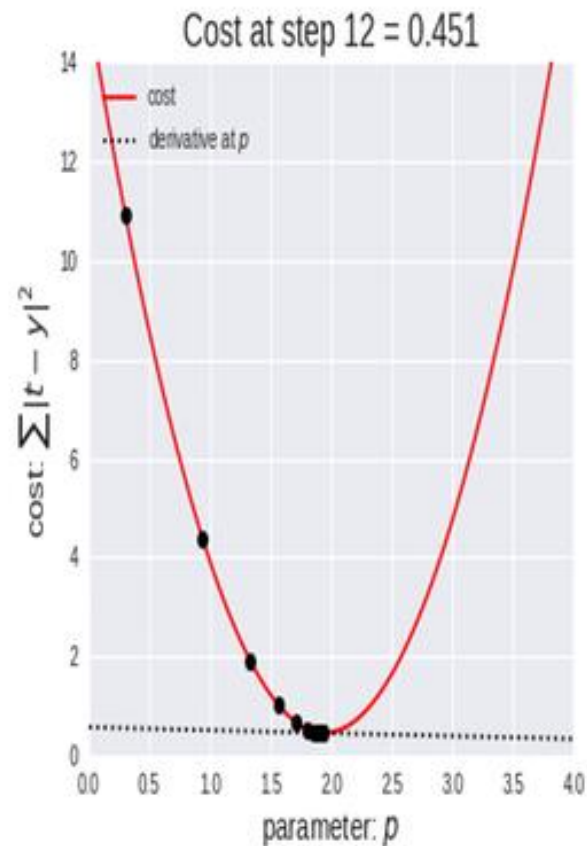
$$w' = w - \alpha * d \text{ loss}/dw$$

Note: alpha is the learning rate.

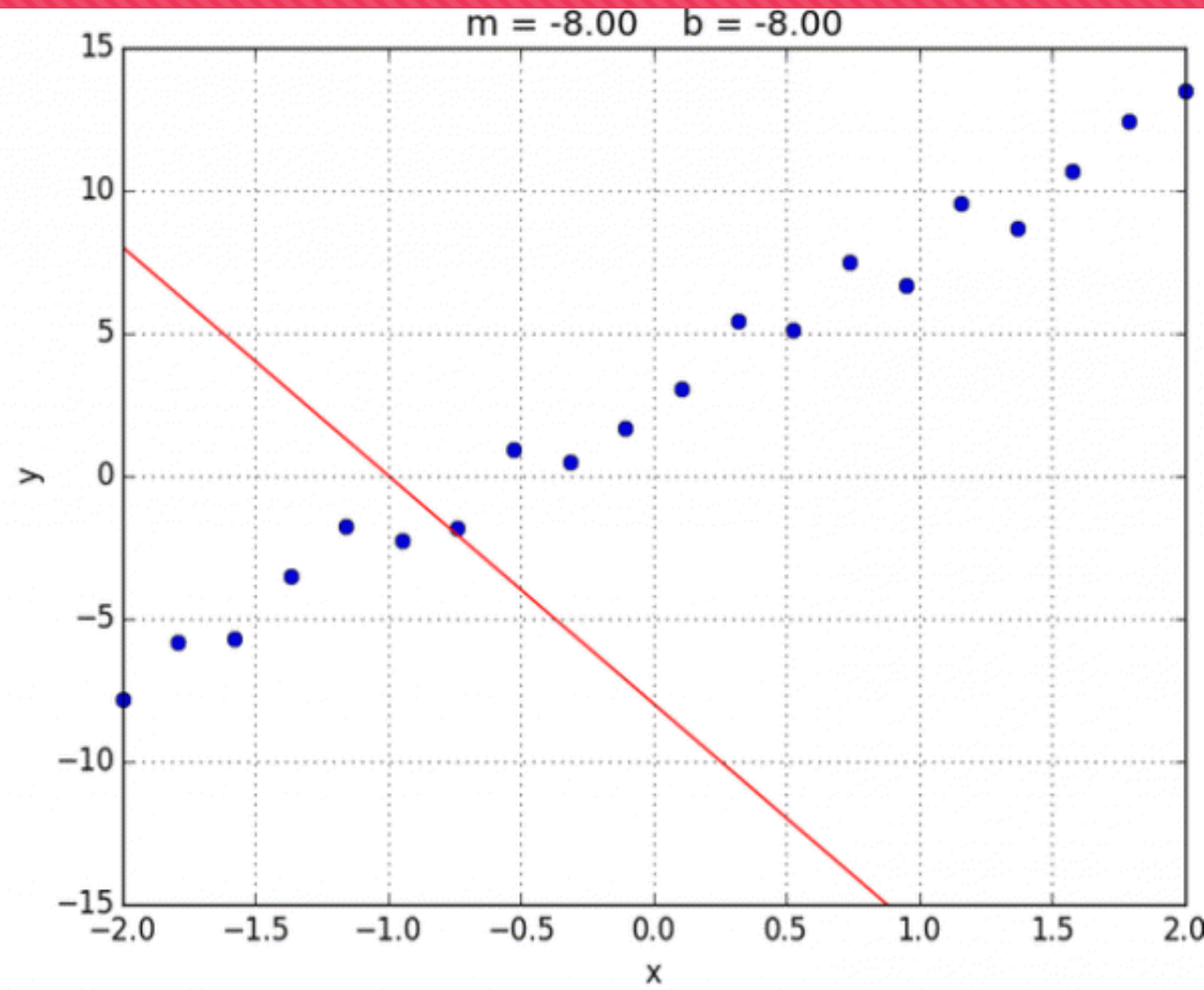
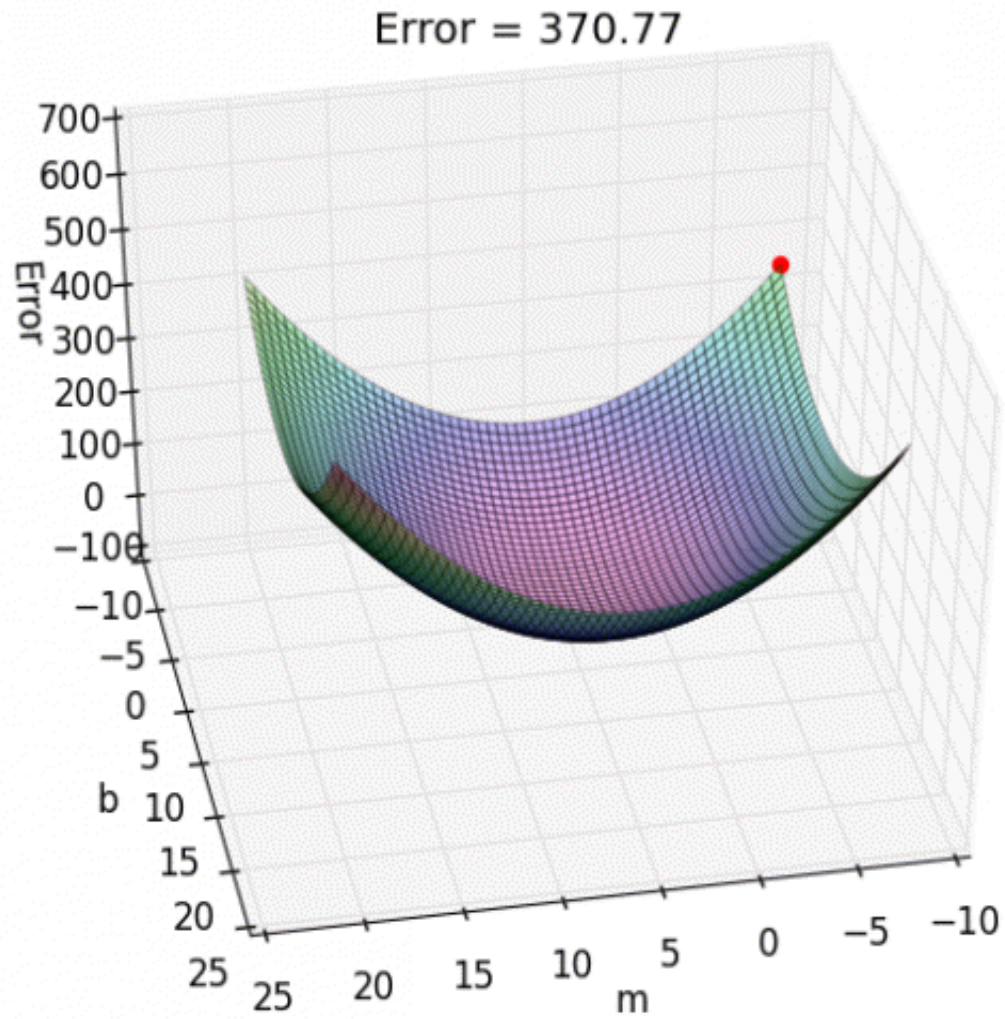
روش کارگرادیان کاهشی یا Gradient Descent

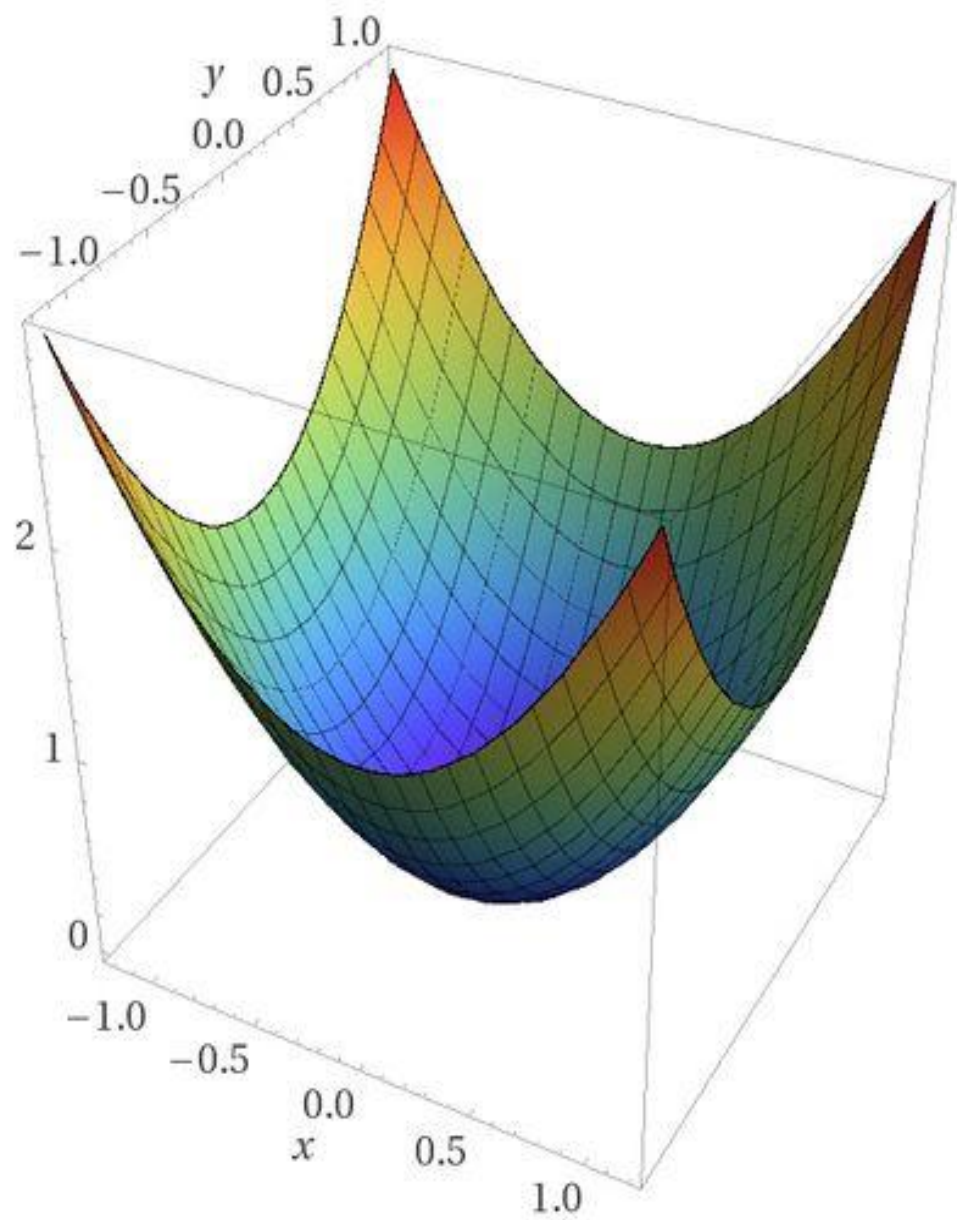


روش کارگرادیان کاهشی یا Gradient Descent

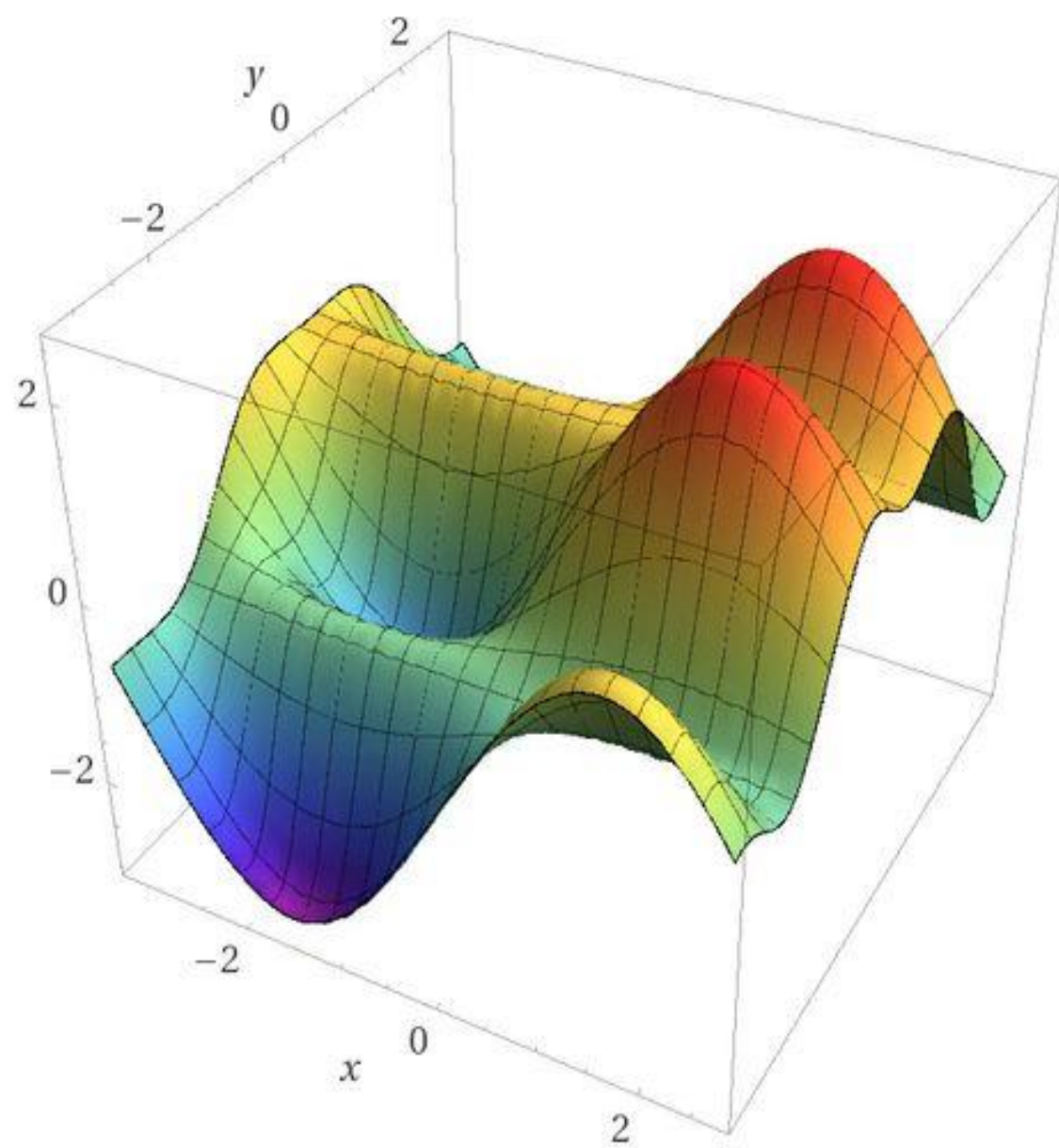


روش کارگرادیان کاهشی یا Gradient Descent



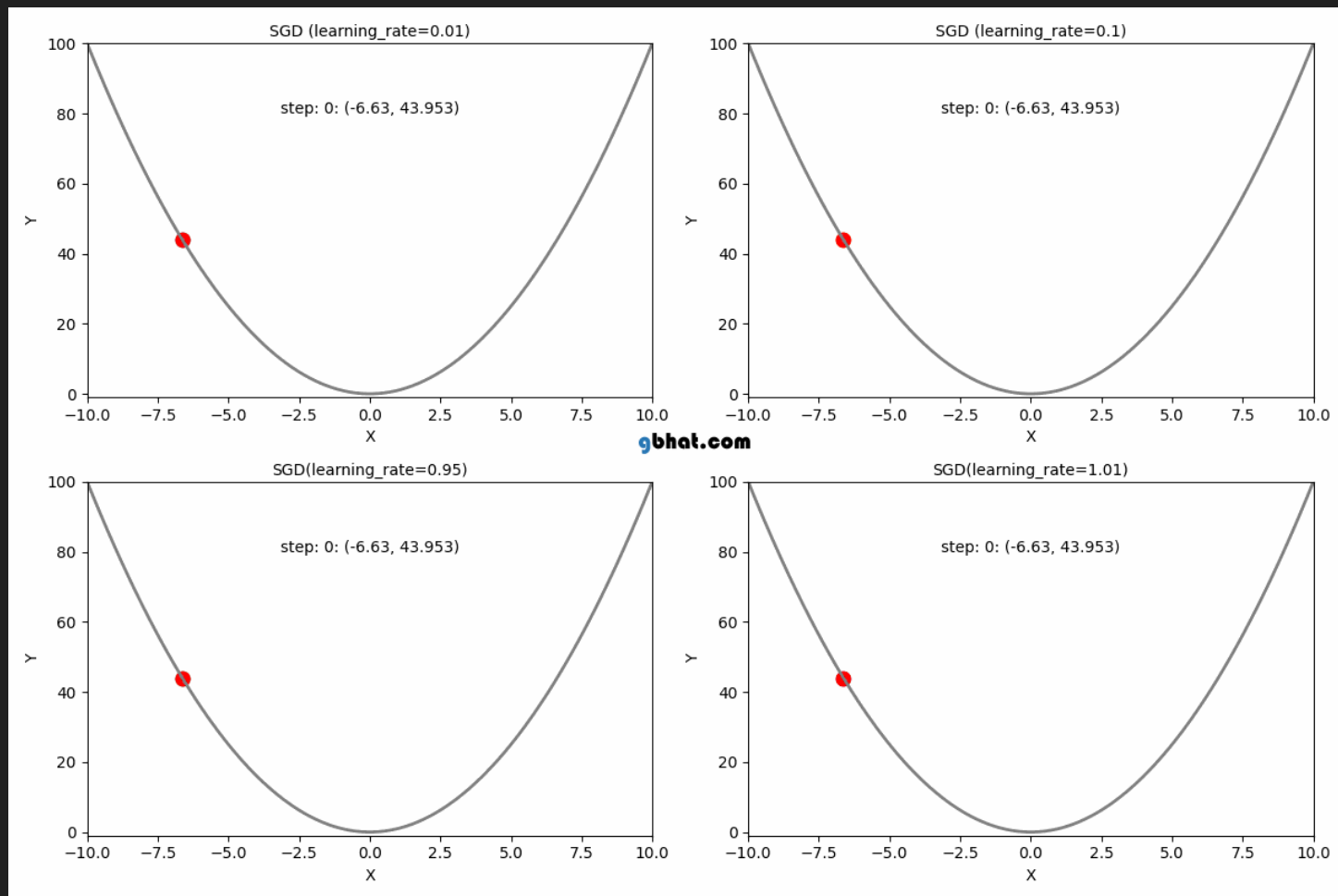


Computed by Wolfram|Alpha



Computed by Wolfram|Alpha

روش کارگرادیان کاهشی یا Gradient Descent

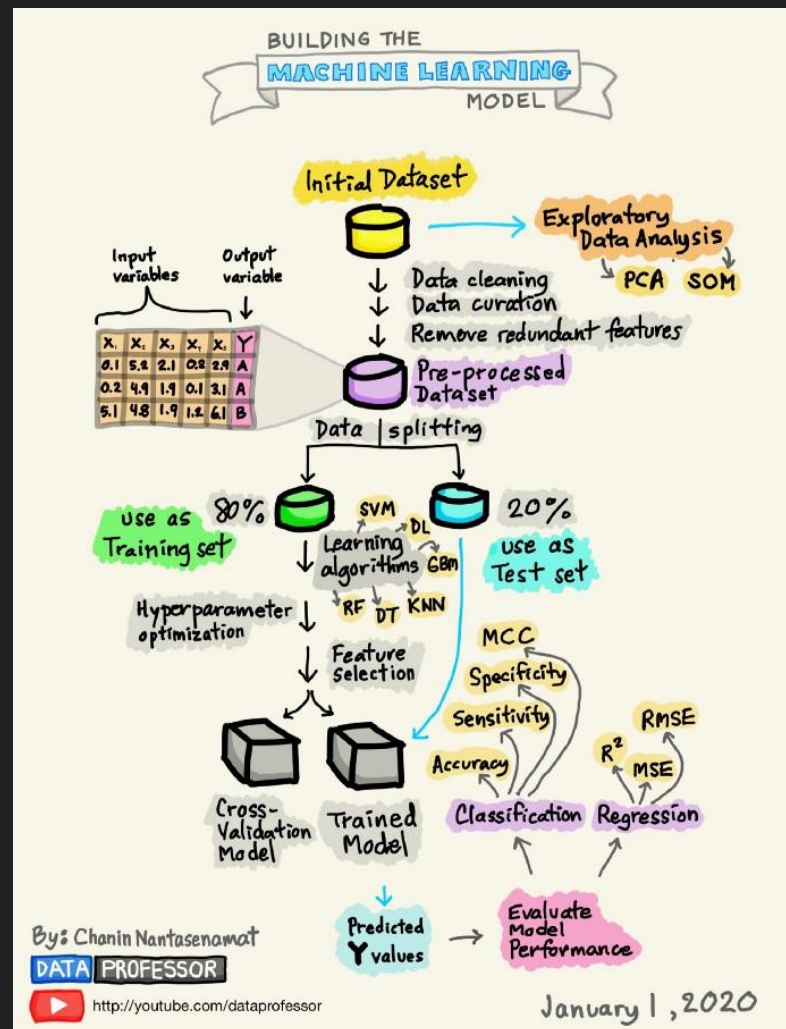


منابع تکمیلی

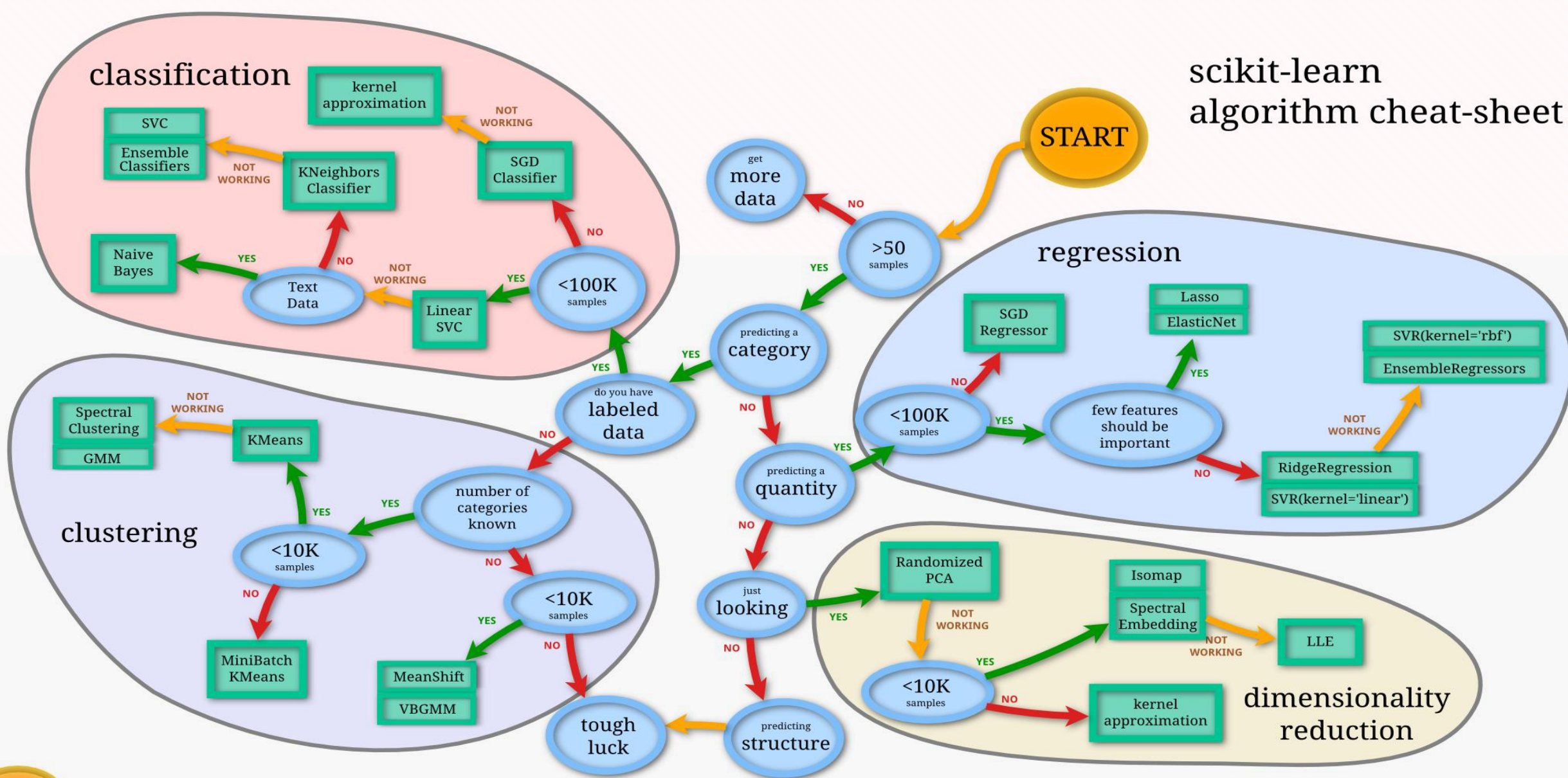
<https://blog.faradars.org/regression-and-gradient-descent/>
<https://blog.faradars.org/implement-gradient-descent-in-python/>
<https://github.com/Vitimir84/Statistics-and-probability/blob/master/Gradient%20and%20programming%20gradient%20descent.ipynb>

<https://www.youtube.com/watch?v=sDv4f4s2SB8>
<https://www.youtube.com/watch?v=vMh0zPT0tLI>

مراحل ساخت یک مدل یادگیری ماشین



scikit-learn algorithm cheat-sheet



Back

scikit
learn

پایان

با تشکر از توجه تان، اوقات خوشی را برایتان آرزومندم.