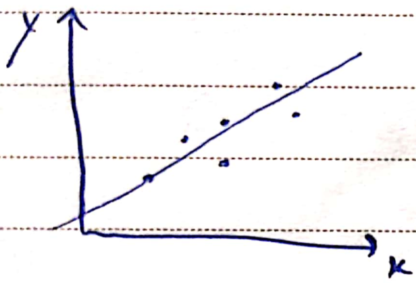
در linear regretion براساس فیچرهایی که به عنوان ورودی داده می‌شود، یک مقدار خروجی می‌دهیم که این مقدار می‌تواند یک range داشته باشد.

مدل Linear Regretion به این صورت است که با داشتن تعدادی نقطه روی صفحه، باید خطی را پیش‌بینی کنیم که نزدیک‌ترین پیش‌بینی را برای مقادیر ورودی انجام دهد.

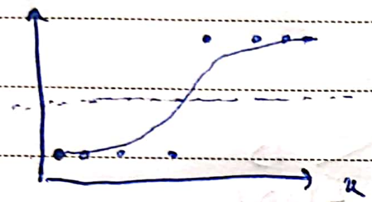


پس از ارائه دادن یک خط، حالا با ورودی دادن x های مختلف می‌توانیم y های متفاوت بگیریم که این y ها دارای یک رنج هستند

* محور می‌تواند چند بعدی باشد

مثال: پیش‌بینی قیمت خانه با استفاده از فیچرهایی مثل متراژ، مکان، طبقه و...

اما در Logistic Regretion براساس فیچر های ورودی، تنها true یا false بودن یک عبارت را پیشبینی می کنیم. یعنی خروجی ما تنها ۰ یا ۱ است نه یک رینج



مثال: پیشبینی معتبر یا نامعتبر بودن یک کارت اعتباری براساس فیچر های مثل اعتبار، سابقه تراکنش ها و...

سایر تفاوت ها:

① در Linear Regretion میزان دقت پیشبینی ما با روش Least square مشخص می شود اما در Logistic Regretion با روش Maximum Likelihood مشخص می شود

② در Linear Regretion، best fit یک خط است. اما در Logistic Regretion، best fit یک منحنی است

③ در Linear Regretion خروجی یک عدد از یک بازه است. اما در Logistic Regretion، خروجی ۰ یا ۱ است

④ از Linear Regretion در پیشبینی های مالی و بیزنسی استفاده می شود اما از Logistic Regretion در مسائل Classification و Image Processing استفاده می شود

$$P_{LAP,k}(x|y) = \frac{c(x,y)+k}{c(y)+k|X|}$$ 2-

$$k=1 \Rightarrow P_{LAP}(x|y) = \frac{c(x,y)+1}{c(y)+|X|}$$

از 4 داده موجود 3 تا c و یکی j است :

$Prior(c) = \frac{3}{4}$, $Prior(j) = \frac{1}{4}$

فرمول بالا برای این مسئله به این شکل در می‌آید $\Rightarrow$ $P(w|c) = \dfrac{count(w,c)+1}{count(c)+|W|}$

$count(c) = 8$ $count(j) = 3$ $|W| = 6$

$count(chinese,c) = 5$ $count(Tokyo,c) = 0$ $count(Japan,c) = 0$

$count(chinese,j) = 1$ $count(Tokyo,j) = 1$ $count(Japan,j) = 1$

حال جدول احتمالات به این صورت در می‌آید $(k=1)$

$P(chinese|c) = \frac{6}{14}$ $P(Tokyo|c) = \frac{1}{14}$ $P(Japan|c) = \frac{1}{14}$

$P(chinese|j) = \frac{2}{9}$ $P(Tokyo|j) = \frac{2}{9}$ $P(Japan|j) = \frac{2}{9}$

$P(class|f_1,f_2,\dots) \propto P(class)\prod_i P(f_i|class)$

$P(c|$ سطر $5) = \frac{3}{4} \times \left(\frac{6}{14}\right)^3 \times \frac{1}{14} \times \frac{1}{14} = 0.000301$ ✓

$P(j|$ سطر $5) = \frac{1}{4} \times \left(\frac{2}{9}\right)^3 \times \frac{2}{9} \times \frac{2}{9} = 0.000135$

احتمال c بیشتر است پس c انتخاب می‌شود.