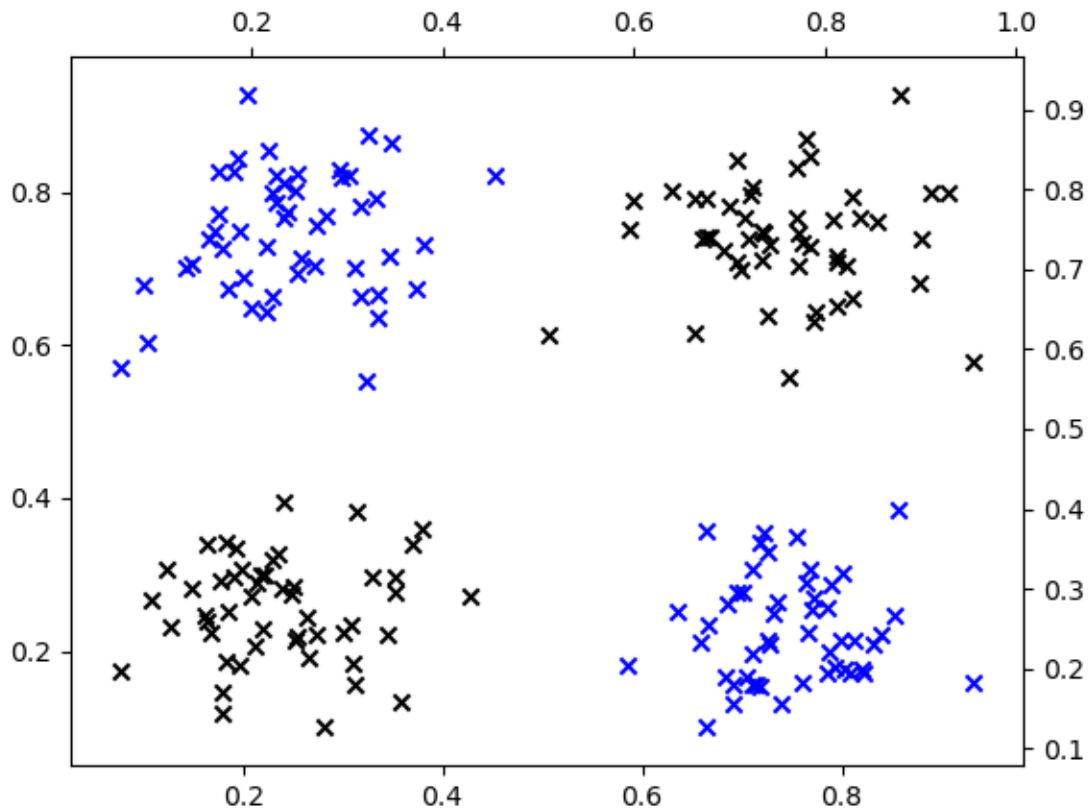


داده‌ها به صورت زیر دسته‌بندی می‌شوند.

0 : سیاه

1 : آبی



در شبکه عصبی اول داریم :

$$A: dcost/dW = dcost/dy * dy/dW = (y-y_0)XS(WX+b)(1-S(WX+b)) = (y-y_0)y(1-y)X$$

$$B: dcost/db = dcost/dy * dy/db = (y-y_0)S(WX+b)(1-S(WX+b)) = (y-y_0)y(1-y)$$

در شبکه عصبی دوم داریم :

$$Y = S((S(XW+b_0), S(XV+b_1))U+b_2)$$

$$d\text{cost}/dW = d\text{cost}/dy * dy/dW = (y-y0)(S(XW+b0)u0)'y(1-y)$$

$$=(y-y0)u0XS(z0)(1-S(z0))y(1-y) = (y-y0)y(1-y)S(z0)(1-S(z0))u0X$$

$$d\text{cost}/dV = d\text{cost}/dy * dy/dV = (y-y0)(S(XV+b1)u1)'y(1-y)$$

$$=(y-y0)u1XS(z1)(1-S(z1))y(1-y) = (y-y0)y(1-y)S(z1)(1-S(z1))u1X$$

$$d\text{cost}/dU = d\text{cost}/dy * dy/dU = (y-y0)Zy(1-y) = (y-y0)y(1-y) (S(XW+b0) , S(XV+b1))$$

$$d\text{cost}/db0 = d\text{cost}/dy * dy/db0 = (y-y0)(S(XW+b0)u0)'y(1-y)$$

$$=(y-y0)u0S(z0)(1-S(z0))y(1-y) = (y-y0)y(1-y)S(z0)(1-S(z0))u0$$

$$d\text{cost}/db1 = d\text{cost}/dy * dy/db1 = (y-y0)(S(XV+b1)u1)'y(1-y)$$

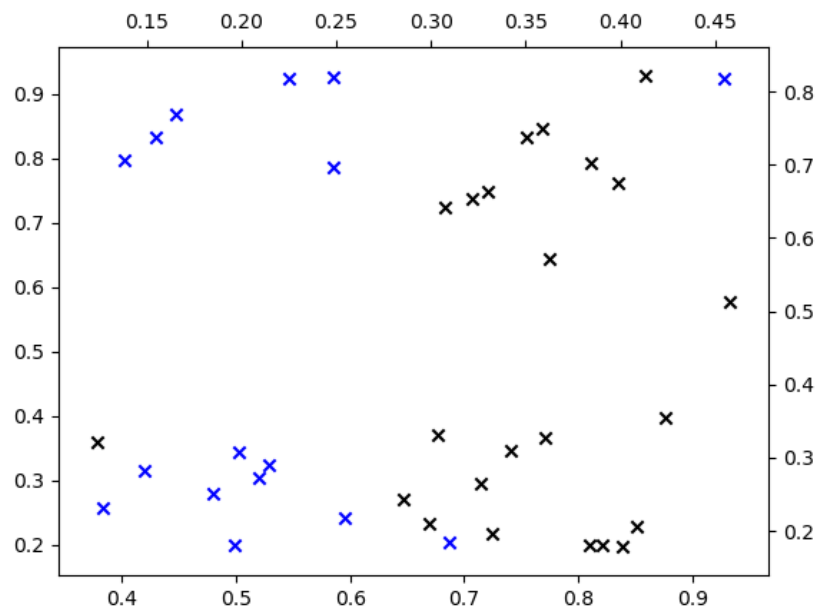
$$=(y-y0)u1S(z1)(1-S(z1))y(1-y) = (y-y0)y(1-y)S(z1)(1-S(z1))u1$$

$$d\text{cost}/db2 = d\text{cost}/dy * dy/db2 = (y-y0)y(1-y)$$

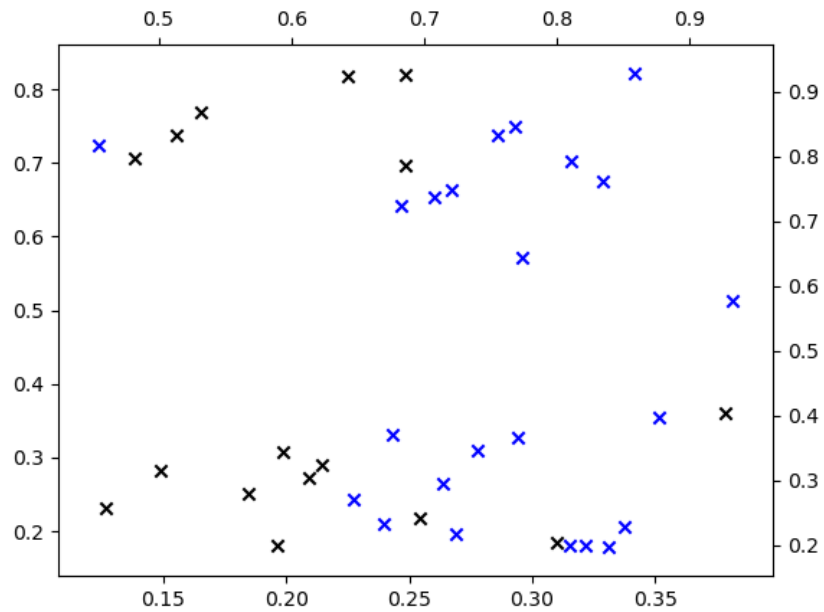
## نتایج مشاهده شده

اگر ¼ داده‌ها آموزشی باشند :

شبکه اول به طور میانگین 46.15384615384615 % داده‌ها را درست جواب میدهد .



شبکه دوم به طور میانگین 58.97435897435897 % داده ها را درست جواب میدهد.



در شبکه عصبی اول وقتی که یک نورون دارد فقط می‌تواند صفحه مختصات را به دو قسمت تقسیم کند اما وقتی در شبکه عصبی دوم تعداد نورون‌ها بیشتر می‌شود، می‌تواند صفحه مختصات را به بخش‌های بیشتری تقسیم کند پس می‌تواند بازده بیشتری داشته باشد.