ML → Supervised → Data → Independent and dependent
unsupervised

$x_1$ $x_2$ $x_3$ $x_4$

y → Numeric (Regression)
→ Categorical → (classification)

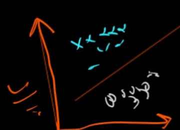
Binary
two outcome
Male | female
Pos | Neg
left | right

(multiclass classification)
Black
blue
white
or
Cum
BJP
AAP
BSP
SP

Classification → ?
1. Ham or spam
2. Pos or neg or neutral
3. Churn Prediction (Bank)
4. buy insurance or not

Machine learning :- Logistic reg ← linear seperable → Data → linear reg
SVM
KNN
DT
Random forest
XGBoost
GBoost
Naive bayes

(Data should be in linear form)



Logistic reg → Linear seperable

1. Probability
2. Geometry (Distance) → SVM

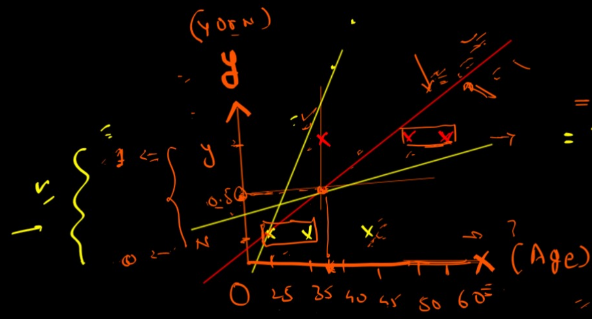
Any machine
1. Calculation
2. Loss
3. learning or optimization

Classification :- Binary and multiclass classification

| Age | insurance buy or not |
|---|---|
| 25 | N → 0 |
| 35 | N → 0 |
| 40 | y → 1 |
| 45 | N → 0 |
| 50 | y → 1 |
| 60 | y → 1 |

Classification (binary)

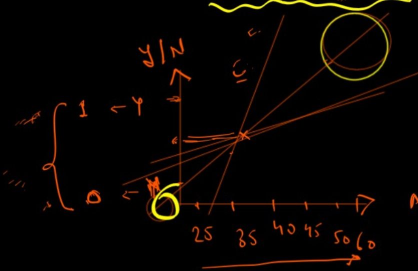
$y = \boxed{m}x + \boxed{C}$ $\quad m = \frac{y}{x}$ $\quad y = mx$




= X

Sigmoid

out of 1

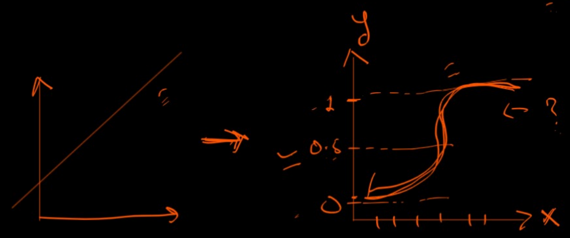



outliers

$y = mx + c$

① Probability (Sigmoid) (logit function)

$P = \frac{1}{1+e^{-x}} \rightarrow (0 \text{ to } 1)$




$y = mx + c$

$P = \frac{1}{1+e^{-(mx+c)}}$

$P = \frac{1}{1 + \frac{1}{e^{(mx+c)}}}$

Probability of happning $\quad \boxed{P(x) = \frac{e^{(mx+c)}}{1+e^{(mx+c)}}}$

Prob of not happ $\quad 1 - P(x) = 1 - \frac{e^{(mx+c)}}{1+e^{(mx+c)}}$

$= \frac{1 + \cancel{e^{(mx+c)}} - \cancel{e^{(mx+c)}}}{1+e^{(mx+c)}}$

$1 - P(x) = \frac{1}{1+e^{(mx+c)}}$

logit function = ? $\quad$ odds = ?

H → $P(x)$

NH → $1 - P(x)$

$= \frac{\frac{e^{(mx+c)}}{\cancel{1+e^{(mx+c)}}}}{\frac{1}{\cancel{1+e^{(mx+c)}}}}$

$\log_{10} 10 = 1$

$\log_e e = 1$

$\log_2 2 = 1$

Probability. 0 - 1 = 1

0.4 - 0.5 - 0.6

0.5

0 $\quad$ 1

$\frac{P(x)}{1-P(x)} = e^{(mx+c)}$

$\log\left(\frac{P(x)}{1-P(x)}\right) = \log_e e^{(mx+c)}$

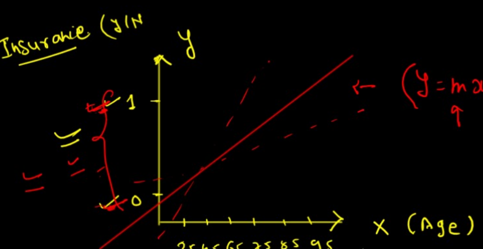
$\boxed{\log\left(\frac{P(x)}{1-P(x)}\right) = mx + c}$

Cost function.

training ?

| $x_1$ (Age) | $x_2$ (edu) | y (Y/N) |
|---|---|---|
| 50 | 6.5 | Y |
| 60 | 7.5 | N |
| 70 | 8.6 | N |
| 80 | 9.5 | Y |
| 50 | 10.8 | N |
| 100 | 10.0 | N |

Insurance (Y/N

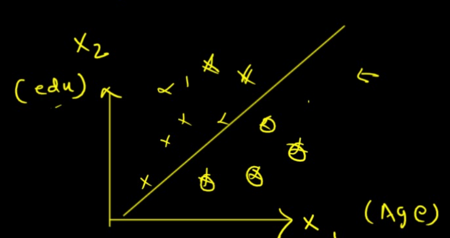


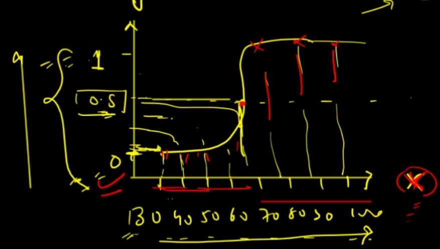

($y = mx + c$)

Classification

Logistic regession

Sigmoid






→ sigmoid (0, 1)

⇒ ?

$P(x) = \frac{1}{1+e^{-x}}$

Probability

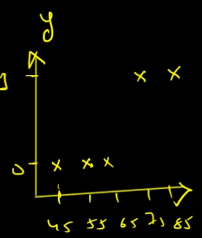
linear

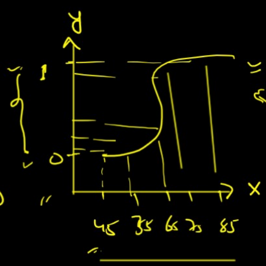
logistic (reg)

Data.

Q.1 why we are calling logistic reg.?

Q.2 why we are not able to solve thing with linear reg.?




$P = \frac{1}{1+e^{-x}}$

$P = \frac{1}{1+e^{-(mx+c)}}$

X  $\hat{y}$ → kinda (logistic regression)

$$\left( \begin{aligned} odd &= \frac{P(x)}{1-P(x)} \\ \log(odd) &= \log\left(\frac{P(x)}{1-P(x)}\right) \end{aligned} \right)$$

(5B + 6R) = 11 Balls

↓

$\frac{1}{11} X \rightarrow \boxed{\frac{5}{11}}$ Black

not getting a black ball

$= 1 - \frac{5}{11}$

$= \frac{11-5}{11} = \frac{6}{11}$

$P(x) = \frac{5}{11}$

$1 - P(x) = \frac{6}{11}$

$odd = \frac{P(x)}{1-P(x)} = \frac{\cancel{5/11}}{\cancel{6/11}}$

$odd = \frac{5}{6}$

$\frac{2}{10}$ (0.2)

2 : 10

1 : 5

$$P(x) = \frac{e^{(mx+c)}}{1+e^{(mx+c)}}$$

$$1 - P(x) = 1 - \frac{e^{(mx+c)}}{1+e^{(mx+c)}}$$

$$(1 - P(x)) = \frac{1}{1+e^{mx+c}}$$

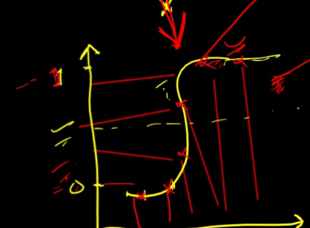
$$odd = \frac{P(x)}{1-P(x)} = \frac{\frac{e^{(mx+c)}}{\cancel{1+e^{(mx+c)}}}}{\frac{1}{\cancel{1+e^{(mx+c)}}}}$$

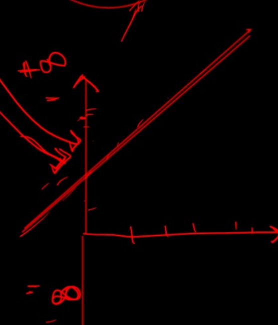
$$odd = \frac{P(x)}{1-P(x)} = e^{(mx+c)}$$

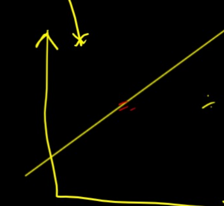
$$\log(odd) = \log\left(\frac{P(x)}{1-P(x)}\right) = \log_e e^{(mx+c)}$$

$$\boxed{P(x) = \frac{e^{(mx+c)}}{1+e^{(mx+c)}}}$$

$$\boxed{\log\left(\frac{P(x)}{1-P(x)}\right) = \boxed{mx+C}}$$



$P(x) = 0.6$

$P(x) = 1$

$P(x) = 0$

$\log\left(\frac{0.6}{0.4}\right)$

$\log\left(\frac{1}{1-1}\right) = \log\left(\frac{1}{0}\right)$

$= \log(1) - \log(0)$

$= 0 - \infty$

$= -\infty$

$\log\left(\frac{0}{1-0}\right)$

$\log\left(\frac{0}{1}\right)$

$\log(0) - \log(1)$

$\infty - 0$

$= \infty$

logistic regression : logit model

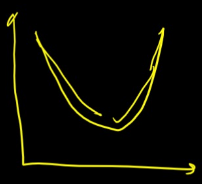
↓

sigmoid

| X (AGE) | y (INS/YORN) |
| --- | --- |
| 23 | Y → 1 |
| 35 | N → 0 |
| 46 | N → 0 |
| 55 | Y → 1 |
| 88 | Y → 1 |

$$\boxed{\text{log\_loss} = -\frac{1}{m}\sum_{i=1}^{m}\left[y\log(P(x)) + (1-y)\log(1-P(x))\right]}$$

(Avg)



$y=1$

$= 1 \times \log(P(x)) + (1-1)\log(1-P(x))$

$= \log(P(x)) + 0 \times \log(1-P(x))$

$$\boxed{\text{log\_loss} = -\log(P(x))} \quad y=1 \text{ (Actual)}$$

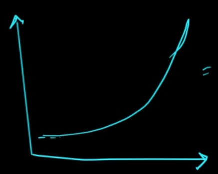
$= -\left[0 \times \log(P(x)) + (1-0)\log(1-P(x))\right]$

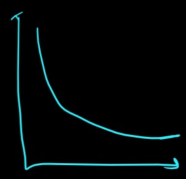
$= -\left[0 + 1 \times \log(1-P(x))\right]$

$$\boxed{\text{log\_loss} = -\log(1-P(x))} \quad y=0 \text{ (Actual)}$$

Predict : $\boxed{\hat{y}=0}$ $\quad -\log(0)$

$-\infty$





$\boxed{\hat{y}=1}$

$-\log(1-1) =$

$-\log(0) = -\infty$

$$P(x) = \frac{e^{(mx+c)}}{1+e^{(mx+c)}}$$

③ Optimize :
- MLE (Probability)
- Iterative (Gradient decent)

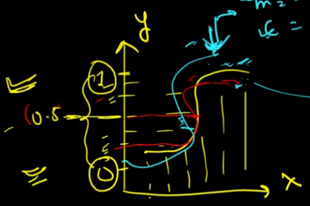
} → m, c ?

① cal
② Loss =
③ optimize

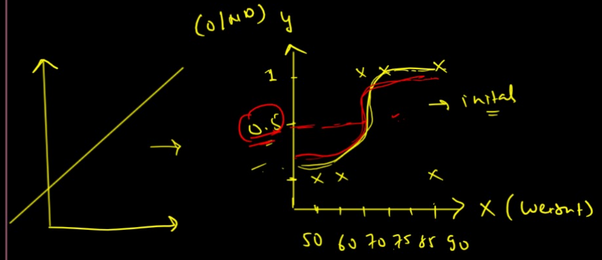
$$\boxed{m_{new} = m_{old} - \eta \frac{\partial L}{\partial m}} \quad \text{(log\_loss)}$$

① cal
②
log-loss
MLE
↓
Goal :- to set a best fit Probability curve

m = ✓
c = ✓




Accuracy = 0.75
Accuracy = 0.80
Accuracy = 0.85




| Weight_value | Obese/Not_obese | | Predict Prob | Prob |
|---|---|---|---|---|
| 50kg | N.O. → 0 | | 0.4 | 0.3 |
| 60kg | N.O → 0 | → Actual | 0.5 | 0.4 |
| 70kg | Ob → 1 | | 0.6 | 0.6 |
| 75kg | Ob → 1 | | 0.7 | 0.7 |
| 85kg | NO → 0 | | 0.3 | 0.2 |
| 90kg | Ob → 1 | | 0.8 | 0.9 |

→ logistic reg

main goal

(best Possibility)

Accuracy
error
TPR
FPR
TNR
Recall - Precision

$\hat{y}$ (output)

0
0
1
1
0
1

0.5

MLE

$$P(x) = \frac{e^{(mx+c)}}{1+e^{(mx+c)}}$$

→ threshold → $\boxed{0.5}$

$$\text{loss\_value} = -\frac{1}{m}\sum_{i=1}^{m}\left[ y\,(P(x)) + (1-y)\,(1-P(x)) \right] = \boxed{\text{Some value}}$$

Gradient decend for optimization ⇒

$$m_{new} = m_{old} - \eta \frac{\partial L}{\partial m} \leftarrow \text{logloss}$$

$$c_{new} = c_{old} - \eta \frac{\partial L}{\partial c} \leftarrow \text{logloss}$$

$\boxed{\text{Regularization}}$

Doubt → 10:05 pm.

Overfitting

Accuracy

Linear Reg → $R^2$

$\boxed{\text{metrics MSE} \rightarrow}$

Goodness of fit

0.85 0.85 $R^2$

$\boxed{R^2}$ → $\boxed{75\%}$ (80%)

Data → $\boxed{\text{model}}$ → training + $\boxed{\text{validation}}$

$R^2 = 0.8$

(80%) $R^2 = 0.5$ (50)

Overfit