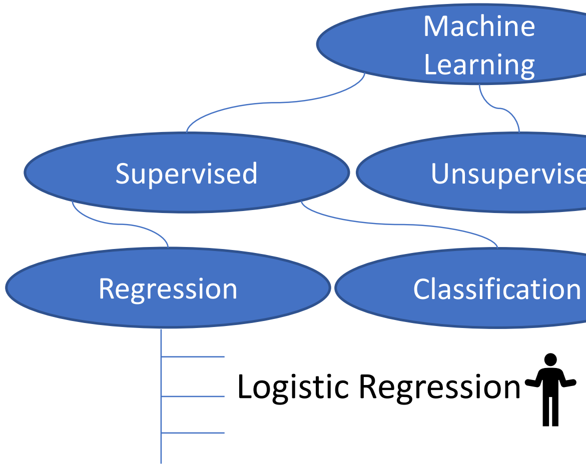
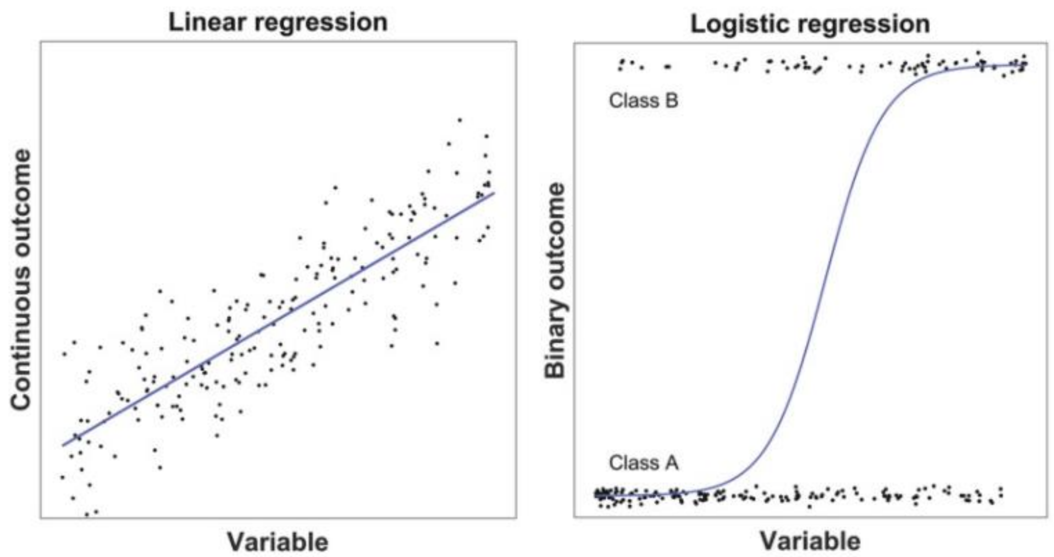
Logistic Regression *developed by statistician David Cox in 1958*



It belongs to class of Generalized Linear Models(GLM) that's spits out probabilities. When threshold is introduced to make decisions it works like classification

- Outcome variable is categorical – Binary values Yes/No
- Want to know probability of outcome
  - Always Positive and is less than or equal to one
- Easily interpretable

<u>Applications</u>: (can be expanded to multi classification)
- Customer buy or pass (Propensity score for an action/behavior)
- Customers who wont renew subscription (Churn Rate)
- Should the Loan be given (What are the odds- credit scoring)
- Medicine field
- Baseline Model



Img. ref: https://www.ncbi.nlm.nih.gov/books/NBK543534/figure/ch8.Fig2/

$Linear\ Regression\ y = w0 + w1x$

a. $Positive \Rightarrow e^y$

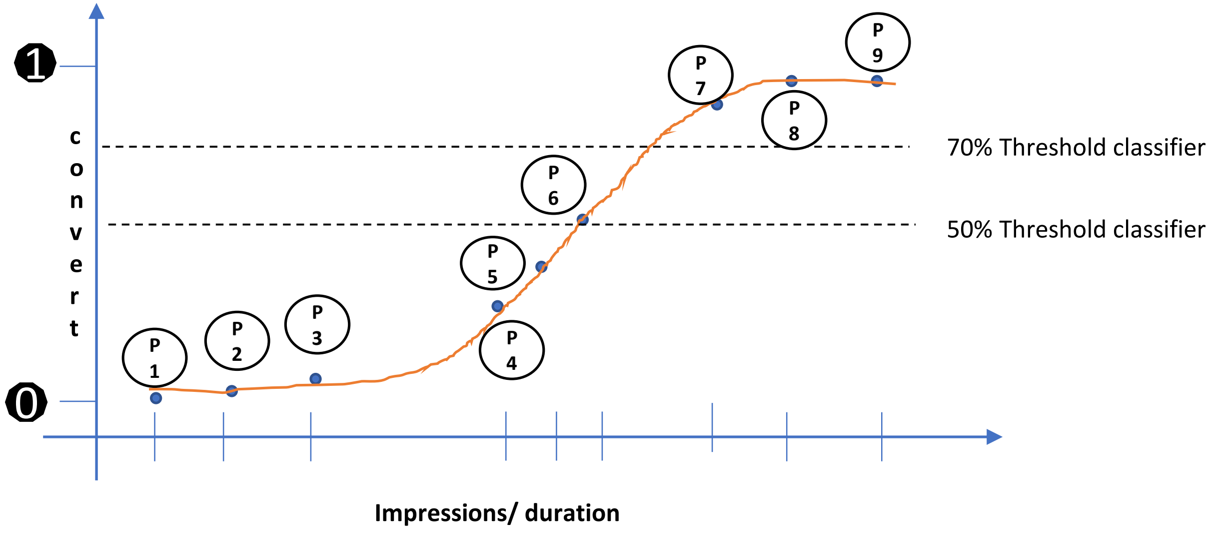
b. $should\ not\ be\ greater\ than\ 1 \Rightarrow \frac{e^y}{e^y+1}$

Probability of "Yes" lets say "P" => $\frac{e^y}{e^y+1}$

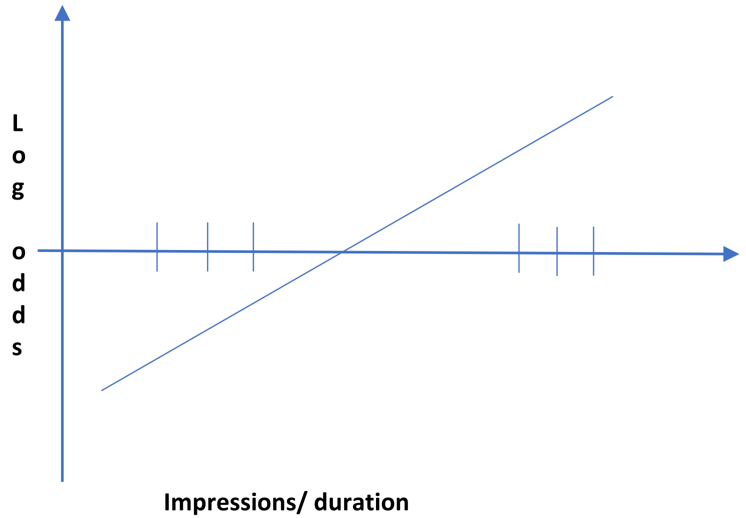
Probability of "No" $(1 - P)$ => $1 - \frac{e^y}{e^y+1}$ => $\frac{1}{e^y+1}$

Odds => $\frac{P}{1-p}$ => e^y

Log of odds => $\log(e^y)$ => $y = w0 + w1x$



Logit =>



Log of odds builds relationship between independent variable x and Probability

Best Curve $_{\text{maximize likelihood}}$ = $P_9 \times P_8 \times P_7 \times P_6 \times (1 - P_5) \times (1 - P_4) \times (1 - P_3) \times (1 - P_2) \times (1 - P_1)$

__Log loss function__ # mostly used in Kaggle competition's submission : when probability is away from actual it penalizes by log as higher value

$$-\frac{1}{n}\sum_{i=1}^{n} y_i \cdot \log p(y_i) + (1 - y_i) \cdot \log(1 - p(y_i))$$