(ML & Data Science) U-1

DOMS | Page No.
Date / /

Regression Models.

↳ describes the relⁿ b/w one or more independent variables & a respons. dependat or target value.

Linear Model –

- simple models in ML
- (assumes) → data is linearly seperable.
- tries to learn weight of each feature.
- help predict behavior of complex system.

Residuals –

- statistical or machine learning model.
- diff b/w observed & predicted values of data.
- assessing quality of model.
- also known as errors.
- measure of how far away a point

Inference – (Reference).

critical aspect of linear regression analysis involves drawing conclusions & making statistial inference.

Hypothesis Testing
Coefficient Estimation.
Standard Error
confidence Intervals.

pvalue
Goodness of Fit.
Assuption checking
Model Comparison

DOMS | Page No.
Date / /

GLM (Generalized linear models)

- statistical framework that extends traditional linear regression.
- handle broader data range & distributions.

three components → probability distribution, a link func^n & linear predictor
↓ variety of response variable ↓ linear predictors to expected value of response coefficient. ↓ predictors influence response

- Appln → logistic regression for binary outcomes.
- Deviance measures how well GLM fits data & models.
- Interpretability → effected / expected value
- Versatality → epidemiology, finance, etc

Logistic Regression -

- supervised ML.
- classification.
- predict probability that an instance of belonging to class.
- predic categorical dependent variable
- Yes No 0, 1 True & False
- We find S-curve.
- Max. Likelihood estimation method is used for estimation of accuracy

DOMS | Page No.
Date / /

* odds → ratio of some event happening to some event not happening

odds ratio → odds of event occuring in one group to odds in another group

change → predictor variable affects odds of +ve outcome

odds → Ratio → 1 → no effect → on outcome
→ >1 → +ve outcomes
→ <1 → less likely

Helps in feature selection, model evaluation, etc.

* Max. Likelihood Estimation in Logistic Regression in simple :

finds
- best fitting line (mode), predicts binary outcom
- find coefficients for predictor variables
- log likelihood → simplifies calculation.
- optimization → gradient descent
- iterative process
- set of model parameters → max. likelihood fun
- likelihood → 0 to 1.
- max probability of observed data.

DOMS | Page No.
Date / /

Poission Regression.

- regression model → response variables → form of counts & not fractional numbers
- analyze count data.
- answers questions → what factors can predict frequency of event
- rate estimation.
- coefficients represent multiplove effects on rate.
- assumes variance of count is equal to mean.
- to model count & predict events rates.

Interpretting LR -

- coefficient sign & Magnitude:
  - odds & odds Ratios.
  - significance test. (p-values)
  - Model fit. (overall goodness).

+ve → +ve outcome.

Visualizing Fitting LR Curves

- Predicted probabilities.
- Plot the curve
- Interpretation → chag in predictor variable
  - Confidence Intervals. - uncertainty in predicted probabilities
  - Multiple Predictors -