February 2016

AUG to OCT

AI Bootcamp – 3 months

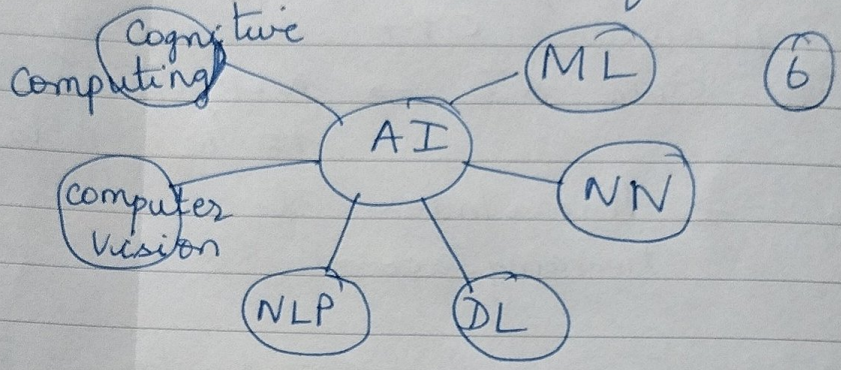
2 033-333 Tuesday

appointments

08/02/2024

1) Beginner Level Understanding of ML Algorithms:

- Core branches of AI:



1) AI : enabling machines to mimic human like behavior

3 034-332 Wednesday

appointments

2) Algorithms : - sequence of steps to perform a specific task

3) Machines → here, refers to computer pgms/ algorithms that do.

4) ML → subset of AI, that learns from data, make decisions with little or no human interference.

Eg: Data about cats/dogs are fed and then learn to make classifications using math based algorithms

4 Thursday 035-331

appointments

Week 5

Types of Machine Learning

- Supervised Learning
  - trained with Label data to classify the data and predict outcome
- Unsupervised Learning
  - Uses unlabelled data to train and identify the hidden data patterns
    - clustering
    - Association
    - Dimensionality Reduction
- Semi-supervised Learning
- Re-inforcement Learning

5 Friday 036-330

appointments

Week 5

Classification

deals with categorical outputs

6

- Logistic Regression
- Naive Bayes
- SVM
- KNN
- Random Forest
- Decision Trees

Regression

deals with continuous values

- Linear Regression
- SVM
- Random Forest
- Decision Trees

6 037-329 Saturday

appointments

Week 5

Unsupervised Learning

- clustering
  - K-Means
  - Hierarchial clustering
- Association
  - Apriori Algorithm
  - Eg :- Market Basket Analysis
- Dimensionality Reduction
  - PCA

7 038-328 Sunday

appointments

Week 5

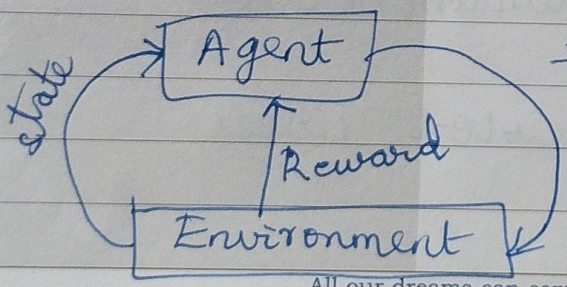
Semi-supervised Learning → partial supervised & unsupervised

- classification
- clustering

↓ useful when abundant untabelled data available & obtaining labelled data is expensive

Re-inforcement Learning → Learning through 'Trial & Error' method

⇒ Learning from mistakes



All our dreams can come true, if we have the courage to pursue them.

039-327
8 Monday

Week 6

appointments

Reinforcement Learning → Q-Learning
→ R-Learning
→ TD Learning

Machine Learning Models:

1) Logistic Regression

↓ counterpart of Linear Regression

$$y = mx + b$$

→ Binary classification → Yes / No

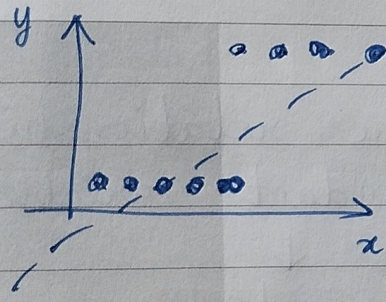
→ Multi-class classification → Positive / Neutral / Negative

040-326
9 Tuesday

Week 6

appointments

For binary classfn.,



↳ Line of Best fit ⇒ this is how Lin. Reg. calculated for Log. Reg pblm

∴ Sigmoid curve / Logistic function / Logit function $f(z)$ used ⇒ to find probability of occurrence and not the value of variable

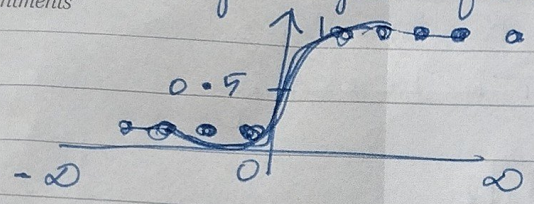
$$z = b_1 x_1 + b_2 x_2 + \dots + b_n x_n + C$$

$$f(z) = \frac{1}{1 + e^{-z}}$$

10 041-325 Wednesday Week 6

appointments

Using Logit function:



Model Parameters:-

Maximum Likelihood Method } Logistic Regression

Least Squares Method } Linear Regression

→ Lin. Reg → independant variable correlates with dependant variable

→ Log. Reg → ind. vari helps in clearly distinguishing dep. variable (Yes / No - 0/1)

11 042-324 Thursday Week 6

appointments

2) Linear Regression:

↳ Aiming to depict the relationship between dependant and independant variables.

↓ response variable

↓ predictor variables

↳ Simple Linear Regression → with single ind.var.

↳ Multiple Linear Regression → with multiple independant variables

12 043-323 Friday

Week 6

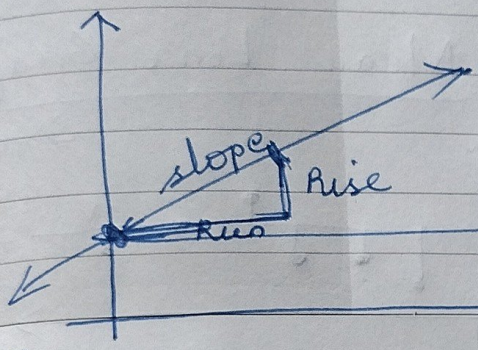
Appointments

↳ Both the types can be explained using a Linear Equation,

$$y = mx + b$$

- y → dependant variable
- x → independant variable
- m → slope
- b → y-intercept.

→ Simple Lin. Reg.



$$m = \frac{\text{Rise}}{\text{Run}}$$

y-intercept

$$y = \beta_0 + \beta_1 x + C$$

13 044-322 Saturday

Week 6

Appointments

Multiple Linear Regression:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_n x_n + \varepsilon$$

↳ Error term

Model Fitting:

↳ Least squares Method:
- minimising the sum of squared (SS) residuals

$$SS(\text{residuals}) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

- y_i → True Value
- $\hat{y}_i$ → Predicted / Estimated value

14 045-321 Sunday

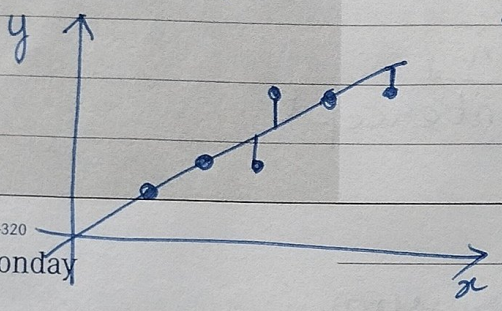
appointments

Model Evaluation:

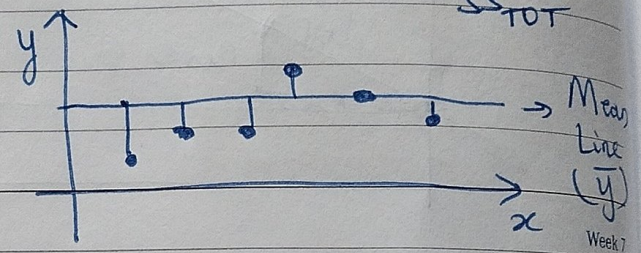
$$R^2 = 1 - \frac{SS_{Res}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

(i.e) $= 1 - \frac{\text{Line of Best Fit}}{\text{Mean Line graph}}$

Eg:- Line of Best Fit: SS_{Res}



Mean Line Graph SS_{TOT}



15 046-320 Monday

appointments

$\therefore$ SS_{Res} will have small term residuals compared to SS_{TOT}

$\therefore$ $R^2 = 1 - \frac{\text{Small Term}}{\text{Large Term}}$

$= 1 - \text{Small Term}$

R^2 = value closes to 1

$\therefore$ Model's best fit lies in smaller residuals, R^2 value closer to 1.

Linear Regression

1) ↳ Predicts continuous dependent variables based on one or more independent variables

Logistic Regression

↳ Predicting binary or categorical dependant variable based on independent variables

2) Linear Eqn :-

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n + \varepsilon$$

→ $\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$

$$f(z) = \frac{1}{1+e^{-z}} ;$$

$$z = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + c$$

17 048-318 Wednesday
appointments

Week 7

3) Model Parameters: Ordinary Least Squares (OLS) method is used to minimise squared residuals

→ Maximum Likelihood Estimation (MLE) is used to find the maximum likelihood of data occurrence

4) Model Evaluation :

↳ R Square
↳ Adjusted R^2
↳ RMSE
↳ MAE

→ Metrics used for evaluation are

↳ Accuracy
↳ Precision
↳ Recall
↳ F1 score
↳ AUC - ROC