Logistic regression. Logistic regression is used to predict classification problems there are a calegories. D Binary legistic regression (binary classification) O Multipo Multivariable logistic regression (multi class classification) O Logistic Regression is used when the dependent variable (tanget) is categorical. O It is the go-to method for binary classification problems ( problems with two class values). O Logistic regression is a statistial method for omalyzing a dataset in which there are one or more ineleperalent variables. that determine con outcome. The outcome to is measured with a dichotomous variable (in which there are only two possible outcomes) O In logistie regression, the dependent versials is binary or dicho tomous, i.e. ut only contain data coded as I CTRUE, success, pregnant, etc. or O (FALSE, failure, non-pregnant, etc). OThe goal of logistic oregression is to finel the best fitting (yet biologically oreasonable) model to describe the relationship between the dichotomous characteristic of interest (dependent variable = response or outrome raciable) and a set of independent (predictor or explantory) variables.

In simple words, it products the probability occurrence of one event by fitting clater to a logist function. de Hence, et us also teneun as logit regression. since, et predicts the probability, its output since, et predicts the probability, its output sules lies between 0 and I (as expected). A logistic femilion or logistic come is a common 'S' shape (sigmoid curre) with equation f(n) = 1+e-K(n-7p) or 1+e-1 where, es the natural fogarithm base (also known as Euler's number).

To the x-value of the sigmoid's unidpoint. LE the curnels meximum value and K= the logistic growth rate of steepney 6 -4 -2 0 2 4 6 L'near Model 1

Poudiction - Regression
· Prodiction is similar to classification
1) first, construct a model
1) Second, use model to prodict unknown
ralue.  Najor method for prodict unknown  talue.
Olinear and multiple expression.
O Linear and multiple orgression.  O Noor-linear regression.
· Porecliction is different foron classification
O Classification suffers to predict categorical alass label.  O Prediction models continuous valued function
O Prediction models continuous -valued function
Classification problem using linear regression
Have have
TONE SANDO CARROLL SAND
Have have
Have Trosworce? In Alway + + + + + + + + + + + + + + + + + + +
Have The surrounce 9 9 000 Always 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Have Trisurance 9  + + + + + + + + + + + + + + + + + +
Have  The surronce of the Advance of the thirty that the thirt
Have Solution. Logit femetion
Have  The surronce of the Advance of the thirty that the thirt
Have This war conce 9  4  4  4  4  4  4  4  4  4  4  4  4  4

confusion matoria for accuracy OA confusion matrix is a summonly of prodiction results on a classification problem only O The no. of correct and incorrect predictions are summarized with count values and broken down by each class. This is the Key to the confusion matrix othe confusion matoix shows the ways in which your classification model is confused when it makes predictions. O It gives us ansight not only the evenous being made by a classifier but more importantly the types of coross that are being made. Actual | Prediction. on odel. score (x-test) y-test) Accuracy = TP + TN TP + TN + FP+FN impost pondas as polosoldis and import numpy as mp from matplotlib import pyplot as pla from sklear. model-selection import train-testfrom sklear. linear-model jimport Logisticken Data fetch and train - test-split & gelower data = pd. read\_csv ( l'insurance \_data.csv') x = data [['age]] J= data [ 1 bought - insurance 17 x-toain, x-test, y-toain, y-test = toain test-split (x,y, test-size=.2)

Rlot it. Plt. Scatter (xoy, color = lorange), markor = (21) plt. relabel (lage)

pet. relabel (bought insurance) 194 5 4 LU confucion mulois shous the undina a chi os fications model us on 20 30 40 50 60 10 10 00 00 29 A AMONTO It gives its sais by and one of cone on Model it and predict. model = Logistic Regression model. fit (x-train, y-train) model. predict (x\_test) 7 - test model.score (z-test, y-test) Assignment WITTHHIM 1) Now do some exploratory clada analysis to figure out which variables have direct and clear impact on employee vetention (i.e. whether they leave the company or continue to work). 5 com = pd. read\_csv (3 o path) filenave with extension) Date tolch and town - test - split } = qulsusor ( v20. Sale: 170000000000) V20- loose. bg = stob x = data [['apri]] 1support:2) (accounting: 30 meni - Adenogi Joseph = 8 X-team, a-tect, y-sown (25 lasimist -x a ma nagement 1:6,00 E Descorde of Sold of (03)

om ['Department] = com ['Department]. 30/06 ce. . com a (new lap) um. Department. unique () even. salary. vrnique () (O151) = esizeit ) smelt the new 3 al = 7 (low ! : 0, e medium ?: 0: 5, our ton") lodaling com [ 'salony 1] = com [ 'salony 1]. replace now Income sult) x=com[['statisfaction\_level', Nast-evalution', chours, Lime\_spencl\_company, Work\_
accident, promotion\_last\_syears, Department,
(Salary)] y's com [ left ] dention of the land - 2002 from sklear. model 1 - selection import train -from sklear. Ninear model impert Logicula egression. about notest y touin, y test - wouin-test-split (m) lanegel . Ha model = Logisti'c Regression ()
model · fit (xtoain , ytoain)
model · poolict (n test)
ytest plf, advicts (notalion ( ) worke the ortained bland and the variables that model. score (rest, y test) prouped = coon groupsy ("left) ['salary']. meam().

reset indem rename (columns = > !salary': laug-salary';

grouped sound (2) grouped ang-salary

granped = com. grouped ( Looft ) [ sobory]. moon (). colors = [c], (P)] plt. figure (figsize = (12,6)) grouped. plat (Kind = 1 bas', color = colors) plt. xlabel ("Not working")
plt. ylabel ("Average Salary")
plt. show()...[" (3) Plat bors charte showing cosselation between department and employee releation. retention-by-department = df. groupby (Department) ['left]. mean (). recet\_indea() plf. figure (figsize = (10,6)) Sns. bauplat ( data = retention\_by\_ department, m= "Department", y=1 left, plad palette (visidie) pet title ('Employee Retention by Department) plt. xlabel (1 Department) 5 test coins plt. ylabel (1 Retention Rabel) pt. legend () model = Logisti e Repression ( 21 = mollator) exists. Alq plt. show() 4) Now bush logistic suggestion model using steps variables that were narrowed clown in steps By reasone the accuracy of the model = language decheef = rees - dechock model: scroes (my) mendas somenas instantin Red on good of to drawing

Eggistic suggession with mills class classification opis And of poobleon will have multiple classifications. En le se probable de la classifications. its having multiple options. We are going to identify handwritten light impost matplotlib. Pyplot as plo from sklearon. datasets impost load-digits
from sklearon. or odel\_selection impost brain-test-get
- LA from sklearon, metrics import confession - madrix Pouron sklears n. lionear\_model impost Linears Regord digits = load -digits () Sen (digits)
dir (digits) create the boundasie is not possible her digits. Lata [100] digits target [0:5] digits · images (o) di plf-mathow (digits. images[0]) Model creation. xtrain, xtest, ytonin, ytest z train-test\_split(digits.

data, digits. target; test=size = -2) model = Logistic Regeression ( ) in through model. fit (x town, y toain)
model. score (x test, y test) Rosschoon prediction of stood not show II & Romolom prediction means using prodict with randon Plt. matshow (digits. images 8887)
digits, target [888]
model , predict (digits. data [888])