

Machine Learning Assignment

Aditya Gavankar (J072)

Exp 9:- Decision Tree and Cross Validation

```
In [16]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline

In [17]: df = pd.read_csv('car_evaluation.csv', header = None)
df.head()

Out[17]:
```

	0	1	2	3	4	5	6
0	vhigh	vhigh	2	2	small	low	unacc
1	vhigh	vhigh	2	2	small	med	unacc
2	vhigh	vhigh	2	2	small	high	unacc
3	vhigh	vhigh	2	2	med	low	unacc
4	vhigh	vhigh	2	2	med	med	unacc

```
In [18]: col_names = ['buying', 'maint', 'doors', 'persons', 'lug_boot', 'safety', 'class']
df.columns = col_names
col_names

Out[18]: ['buying', 'maint', 'doors', 'persons', 'lug_boot', 'safety', 'class']

In [19]: df.head()

Out[19]:
```

	buying	maint	doors	persons	lug_boot	safety	class
0	vhigh	vhigh	2	2	small	low	unacc
1	vhigh	vhigh	2	2	small	med	unacc
2	vhigh	vhigh	2	2	small	high	unacc
3	vhigh	vhigh	2	2	med	low	unacc
4	vhigh	vhigh	2	2	med	med	unacc

```
In [20]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1728 entries, 0 to 1727
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  --
0   buying      1728 non-null   object
1   maint       1728 non-null   object
2   doors       1728 non-null   object
3   persons     1728 non-null   object
4   lug_boot    1728 non-null   object
5   safety      1728 non-null   object
6   class       1728 non-null   object
dtypes: object(7)
memory usage: 94.6+ KB

In [21]: for i in col_names:
print(df[i].value_counts())

med      432
low      432
high     432
vhigh    432
Name: buying, dtype: int64
med      432
low      432
high     432
vhigh    432
Name: maint, dtype: int64
2        432
3        432
5more    432
4        432
Name: doors, dtype: int64
more     576
2        576
4        576
Name: persons, dtype: int64
big       576
med       576
small     576
Name: lug_boot, dtype: int64
med       576
low       576
high      576
Name: safety, dtype: int64
unacc    1210
acc       384
good      69
vgood     65
Name: class, dtype: int64

In [22]: df.shape

Out[22]: (1728, 7)

In [23]: X = df.drop(['class'],axis = 1)
y = df['class']

In [24]: from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.3,random_state=42)

In [25]: from sklearn.preprocessing import OrdinalEncoder
en = OrdinalEncoder()
X_train = en.fit_transform(X_train)
X_test = en.transform((X_test))
```

Gini index as criterion

```
In [26]: from sklearn.tree import DecisionTreeClassifier

In [27]: clf_gini = DecisionTreeClassifier(criterion='gini', max_depth=3, random_state=42)
clf_gini.fit(X_train, y_train)

Out[27]: DecisionTreeClassifier(max_depth=3, random_state=42)

In [28]: y_pred = clf_gini.predict(X_test)
```

Grid Search Cv

```
In [29]: from sklearn.model_selection import GridSearchCV
option=['gini','entropy']
weight_option=['auto','sqrt','log2']
param_grid = {'criterion': option , 'max_features':[2,3,4,5,6] , 'max_depth':[4,5,6,7] , 'min_samples_split':[2,3,4,5,6,7,8,9,10,15,20,25,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100,110,120,130,140,150,160,170,180,190,200,210,220,230,240,250,260,270,280,290,300,310,320,330,340,350,360,370,380,390,400,410,420,430,440,450,460,470,480,490,500,510,520,530,540,550,560,570,580,590,600,610,620,630,640,650,660,670,680,690,700,710,720,730,740,750,760,770,780,790,800,810,820,830,840,850,860,870,880,890,900,910,920,930,940,950,960,970,980,990,1000,1010,1020,1030,1040,1050,1060,1070,1080,1090,1100,1110,1120,1130,1140,1150,1160,1170,1180,1190,1200,1210,1220,1230,1240,1250,1260,1270,1280,1290,1300,1310,1320,1330,1340,1350,1360,1370,1380,1390,1400,1410,1420,1430,1440,1450,1460,1470,1480,1490,1500,1510,1520,1530,1540,1550,1560,1570,1580,1590,1600,1610,1620,1630,1640,1650,1660,1670,1680,1690,1700,1710,1720,1730,1740,1750,1760,1770,1780,1790,1800,1810,1820,1830,1840,1850,1860,1870,1880,1890,1900,1910,1920,1930,1940,1950,1960,1970,1980,1990,2000,2010,2020,2030,2040,2050,2060,2070,2080,2090,2100,2110,2120,2130,2140,2150,2160,2170,2180,2190,2200,2210,2220,2230,2240,2250,2260,2270,2280,2290,2300,2310,2320,2330,2340,2350,2360,2370,2380,2390,2400,2410,2420,2430,2440,2450,2460,2470,2480,2490,2500,2510,2520,2530,2540,2550,2560,2570,2580,2590,2600,2610,2620,2630,2640,2650,2660,2670,2680,2690,2700,2710,2720,2730,2740,2750,2760,2770,2780,2790,2800,2810,2820,2830,2840,2850,2860,2870,2880,2890,2900,2910,2920,2930,2940,2950,2960,2970,2980,2990,3000,3010,3020,3030,3040,3050,3060,3070,3080,3090,3100,3110,3120,3130,3140,3150,3160,3170,3180,3190,3200,3210,3220,3230,3240,3250,3260,3270,3280,3290,3300,3310,3320,3330,3340,3350,3360,3370,3380,3390,3400,3410,3420,3430,3440,3450,3460,3470,3480,3490,3500,3510,3520,3530,3540,3550,3560,3570,3580,3590,3600,3610,3620,3630,3640,3650,3660,3670,3680,3690,3700,3710,3720,3730,3740,3750,3760,3770,3780,3790,3800,3810,3820,3830,3840,3850,3860,3870,3880,3890,3900,3910,3920,3930,3940,3950,3960,3970,3980,3990,4000,4010,4020,4030,4040,4050,4060,4070,4080,4090,4100,4110,4120,4130,4140,4150,4160,4170,4180,4190,4200,4210,4220,4230,4240,4250,4260,4270,4280,4290,4300,4310,4320,4330,4340,4350,4360,4370,4380,4390,4400,4410,4420,4430,4440,4450,4460,4470,4480,4490,4500,4510,4520,4530,4540,4550,4560,4570,4580,4590,4600,4610,4620,4630,4640,4650,4660,4670,4680,4690,4700,4710,4720,4730,4740,4750,4760,4770,4780,4790,4800,4810,4820,4830,4840,4850,4860,4870,4880,4890,4900,4910,4920,4930,4940,4950,4960,4970,4980,4990,5000,5010,5020,5030,5040,5050,5060,5070,5080,5090,5100,5110,5120,5130,5140,5150,5160,5170,5180,5190,5200,5210,5220,5230,5240,5250,5260,5270,5280,5290,5300,5310,5320,5330,5340,5350,5360,5370,5380,5390,5400,5410,5420,5430,5440,5450,5460,5470,5480,5490,5500,5510,5520,5530,5540,5550,5560,5570,5580,5590,5600,5610,5620,5630,5640,5650,5660,5670,5680,5690,5700,5710,5720,5730,5740,5750,5760,5770,5780,5790,5800,5810,5820,5830,5840,5850,5860,5870,5880,5890,5900,5910,5920,5930,5940,5950,5960,5970,5980,5990,6000,6010,6020,6030,6040,6050,6060,6070,6080,6090,6100,6110,6120,6130,6140,6150,6160,6170,6180,6190,6200,6210,6220,6230,6240,6250,6260,6270,6280,6290,6300,6310,6320,6330,6340,6350,6360,6370,6380,6390,6400,6410,6420,6430,6440,6450,6460,6470,6480,6490,6500,6510,6520,6530,6540,6550,6560,6570,6580,6590,6600,6610,6620,6630,6640,6650,6660,6670,6680,6690,6700,6710,6720,6730,6740,6750,6760,6770,6780,6790,6800,6810,6820,6830,6840,6850,6860,6870,6880,6890,6900,6910,6920,6930,6940,6950,6960,6970,6980,6990,7000,7010,7020,7030,7040,7050,7060,7070,7080,7090,7100,7110,7120,7130,7140,7150,7160,7170,7180,7190,7200,7210,7220,7230,7240,7250,7260,7270,7280,7290,7300,7310,7320,7330,7340,7350,7360,7370,7380,7390,7400,7410,7420,7430,7440,7450,7460,7470,7480,7490,7500,7510,7520,7530,7540,7550,7560,7570,7580,7590,7600,7610,7620,7630,7640,7650,7660,7670,7680,7690,7700,7710,7720,7730,7740,7750,7760,7770,7780,7790,7800,7810,7820,7830,7840,7850,7860,7870,7880,7890,7900,7910,7920,7930,7940,7950,7960,7970,7980,7990,8000,8010,8020,8030,8040,8050,8060,8070,8080,8090,8100,8110,8120,8130,8140,8150,8160,8170,8180,8190,8200,8210,8220,8230,8240,8250,8260,8270,8280,8290,8300,8310,8320,8330,8340,8350,8360,8370,8380,8390,8400,8410,8420,8430,8440,8450,8460,8470,8480,8490,8500,8510,8520,8530,8540,8550,8560,8570,8580,8590,8600,8610,8620,8630,8640,8650,8660,8670,8680,8690,8700,8710,8720,8730,8740,8750,8760,8770,8780,8790,8800,8810,8820,8830,8840,8850,8860,8870,8880,8890,8900,8910,8920,8930,8940,8950,8960,8970,8980,8990,9000,9010,9020,9030,9040,9050,9060,9070,9080,9090,9100,9110,9120,9130,9140,9150,9160,9170,9180,9190,9200,9210,9220,9230,9240,9250,9260,9270,9280,9290,9300,9310,9320,9330,9340,9350,9360,9370,9380,9390,9400,9410,9420,9430,9440,9450,9460,9470,9480,9490,9500,9510,9520,9530,9540,9550,9560,9570,9580,9590,9600,9610,9620,9630,9640,9650,9660,9670,9680,9690,9700,9710,9720,9730,9740,9750,9760,9770,9780,9790,9800,9810,9820,9830,9840,9850,9860,9870,9880,9890,9900,9910,9920,9930,9940,9950,9960,9970,9980,9990,10000,10001,10002,10003,10004,10005,10006,10007,10008,10009,10010,10011,10012,10013,10014,10015,10016,10017,10018,10019,10020,10021,10022,10023,10024,10025,10026,10027,10028,10029,10030,10031,10032,10033,10034,10035,10036,10037,10038,10039,10040,10041,10042,10043,10044,10045,10046,10047,10048,10049,10050,10051,10052,10053,10054,10055,10056,10057,10058,10059,10060,10061,10062,10063,10064,10065,10066,10067,10068,10069,10070,10071,10072,10073,10074,10075,10076,10077,10078,10079,10080,10081,10082,10083,10084,10085,10086,10087,10088,10089,10090,10091,10092,10093,10094,10095,10096,10097,10098,10099,10100,10101,10102,10103,10104,10105,10106,10107,10108,10109,10110,10111,10112,10113,10114,10115,10116,10117,10118,10119,10120,10121,10122,10123,10124,10125,10126,10127,10128,10129,10130,10131,10132,10133,10134,10135,10136,10137,10138,10139,10140,10141,10142,10143,10144,10145,10146,10147,10148,10149,10150,10151,10152,10153,10154,10155,10156,10157,10158,10159,10160,10161,10162,10163,10164,10165,10166,10167,10168,10169,10170,10171,10172,10173,10174,10175,10176,10177,10178,10179,10180,10181,10182,10183,10184,10185,10186,10187,10188,10189,10190,10191,10192,10193,10194,10195,10196,10197,10198,10199,10200,10201,10202,10203,10204,10205,10206,10207,10208,10209,10210,10211,10212,10213,10214,10215,10216,10217,10218,10219,10220,10221,10222,10223,10224,10225,10226,10227,10228,10229,10230,10231,10232,10233,10234,10235,10236,10237,10238,10239,10240,10241,10242,10243,10244,10245,10246,10247,10248,10249,10250,10251,10252,10253,10254,10255,10256,10257,10258,10259,10260,10261,10262,10263,10264,10265,10266,10267,10268,10269,10270,10271,10272,10273,10274,10275,10276,10277,10278,10279,10280,10281,10282,10283,10284,10285,10286,10287,10288,10289,10290,10291,10292,10293,10294,10295,10296,10297,10298,10299,10300,10301,10302,10303,10304,10305,10306,10307,10308,10309,10310,10311,10312,10313,10314,10315,10316,10317,10318,10319,10320,10321,10322,10323,10324,10325,10326,10327,10328,10329,10330,10331,10332,10333,10334,10335,10336,10337,10338,10339,10340,10341,10342,10343,10344,10345,10346,10347,10348,10349,10350,10351,10352,10353,10354,10355,10356,10357,10358,10359,10360,10361,10362,10363,10364,10365,10366,10367,10368,10369,10370,10371,10372,10373,10374,10375,10376,10377,10378,10379,10380,10381,10382,10383,10384,10385,10386,10387,10388,10389,10390,10391,10392,10393,10394,10395,10396,10397,10398,10399,10400,10401,10402,10403,10404,10405,10406,10407,10408,10409,10410,10411,10412,10413,10414,10415,10416,10417,10418,10419,10420,10421,10422,10423,10424,10425,10426,10427,10428,10429,10430,10431,10432,10433,10434,10435,10436,10437,10438,10439,10440,10441,10442,10443,10444,10445,10446,10447,10448,10449,10450,10451,10452,10453,10454,10455,10456,10457,10458,10459,10460,10461,10462,10463,10464,10465,10466,10467,10468,10469,10470,10471,10472,10473,10474,10475,10476,10477,10478,10479,10480,10481,10482,10483,10484,10485,10486,10487,10488,10489,10490,10491,10492,10493,10494,10495,10496,10497,10498,10499,10500,10501,10502,10503,10504,10505,10506,10507,10508,10509,10510,10511,10512,10513,10514,10515,10516,10517,10518,10519,10520,10521,10522,10523,10524,10525,10526,10527,10528,10529,10530,10531,10532,10533,10534,10535,10536,10537,10538,10539,10540,10541,10542,10543,10544,10545,10546,10547,10548,10549,10550,10551,10552,10553,10554,10555,10556,10557,10558,10559,10560,10561,10562,10563,10564,10565,10566,10567,10568,10569,10570,10571,10572,10573,10574,10575,10576,10577,10578,10579,10580,10581,10582,10583,10584,10585,10586,10587,10588,10589,10590,10591,10592,10593,10594,10595,10596,10597,10598,10599,10600,10601,10602,10603,10604,10605,10606,10607,10608,10609,10610,10611,10612,10613,10614,10615,10616,10617,10618,10619,10620,10621,10622,10623,10624,10625,10626,10627,10628,10629,10630,10631,10632,10633,10634,10635,10636,10637,10638,10639,10640,10641,10642,10643,10644,10645,10646,10647,10648,10649,10650,10651,10652,10653,10654,10655,10656,10657,10658,10659,10660,10661,10662,10663,10664,10665,10666,10667,10668,10669,10670,10671,10672,10673,10674,10675,10676,10677,10678,10679,10680,10681,10682,10683,10684,10685,10686,10687,10688,10689,10690,10691,10692,10693,10694,10695,10696,10697,10698,10699,10700,10701,10702,10703,10704,10705,10706,10707,10708,10709,10710,10711,10712,10713,10714,10715,10716,10717,10718,10719,10720,10721,10722,10723,10724,10725,10726,10727,10728,10729,10730,10731,10732,10733,10734,10735,10736,10737,10738,10739,10740,10741,10742,10743,10744,10745,10746,10747,10748,10749,10750,10751,10752,10753,10754,10755,10756,10757,10758,10759,10760,10761,10762,10763,10764,10765,10766,10767,10768,10769,10770,10771,10772,10773,10774,10775,10776,10777,10778,10779,10780,10781,10782,10783,10784,10785,10786,10787,10788,10789,10790,10791,10792,10793,10794,10795,10796,10797,10798,10799,10800,10801,10802,10803,10804,10805,10806,10807,10808,10809,10810,10811,10812,10813,10814,10815,10816,10817,10818,10819,10820,10821,10822,10823,10824,10825,10826,10827,10828,10829,10830,10831,10832,10833,10834,10835,10836,10837,10838,10839,10840,10841,10842,10843,10844,10845,10846,10847,10848,10849,10850,10851,10852,10853,10854,10855,10856,10857,10858,10859,10860,10861,10862,10863,10864,10865,10866,10867,10868,10869,10870,10871,10872,10873,10874,10875,10876,10877,10878,10879,10880,10881,10882,10883,10884,10885,10886,10887,10888,10889,10890,10891,10892,10893,10894,10895,10896,10897,10898,10899,10900,10901,10902,10903,10904,10905,10906,10907,10908,10909,10910,10911,10912,10913,10914,10915,10916,10917,10918,10919,10920,10921,10922,10923,10924,10925,10926,10927,10928,10929,109
```