I. Numerical Representation to Value

The Scikit Learn Packages needed all the values to be numerical. For large range of values in Age, Cap-Gain, Cap-Loss, Hours we grouped values into smaller ranges.

For all the categorical columns we changed the values to numerical using catcodes and recorded the corresponding category values to their numbers here.

Age

Range	Numerical
Age <=21	0
21 < Age <= 30	1
30 < Age <= 50	2
50 < Age <= 70	3
Age > 70	4

Work

Category	Numerical
State-gov	5
Self-emp-not-inc	4
Private	2
Federal-gov	0
Local-gov	1
Self-emp-inc	3
Without-pay	6

Edu-Lvl

Category	Numerical
Bachelors	9
HS-grad	11
11th	1
Masters	12
9th	6

Some-college	15
Assoc-acdm	7
Assoc-voc	8
7th-8th	5
Doctorate	10
Prof-school	14
5th-6th	4
10th	0
Preschool	13
12th	2
1st-4th	3

Edu-Years

1-16

Marriage-Status

Category	Numerical
Never-married	4
Married-civ-spouse	2
Divorced	0
Married-spouse-absent	3
Separated	5
Married-AF-spouse	1
Widowed	6

Occupation

Categorical	Numerical
Adm-clerical	0
Exec-managerial	3
Handlers-cleaners	5
Prof-specialty	9
Other-service	7

Sales	11
Craft-repair	2
Transport-moving	13
Farming-fishing	4
Machine-op-inspct	6
Tech-support	12
Protective-serv	10
Armed-Forces	1
Priv-house-serv	8

Relationship

Categorical	Numerical
Not-in-family	1
Husband	0
Wife	5
Own-child	3
Unmarried	4
Other-relative	2

Gender

Category	Numerical
Male	1
Female	0

Cap-Gain

Range	Numerical
<= 2000	0
2000 < Gain <= 4000	1
4000 < Gain <= 6000	2
6000 < Gain < 10000	3
10000 < Gain	4

Cap-Loss

Range	Numerical
<= 1300	0
1300 < Loss <= 1600	1
1600 < Loss <= 1900	2
1900 < Loss < 2200	3
2200 < Loss	4

Hours

Range	Numerical
Hours <= 20	0
20 < Hours <= 40	1
40 < Hours <= 60	2
60 < Hours <= 80	3
Hours > 80	4

Income

Category	Numerical
>50	1
<=50	0

II. Naive Bayes Probability by Attributes

First column is the probability the values would be <=50K and the second column is the probability the value would be >50K, last value is the class the value is categorized in according to the model.

Age :

Age <= 21	0 : [[0.99726059	0.00273941]]	class [0]
21 < Age <= 30	1 : [[0.91769153	0.08230847]]	class [0]
30 < Age <= 50	2 : [[0.65488452	0.34511548]]	class [0]
50 < Age <= 70	3 : [[0.64218234	0.35781766]]	class [0]
Age > 70	4 : [[0.90414855	0.09585145]]	class [0]

Work :

Federal-gov	0 : [[0.66137583	0.33862417]]	class [0]
Local-gov	1 : [[0.74813928	0.25186072]]	class [0]
Private	2 : [[0.77504558	0.22495442]]	class [0]
Self-emp-inc	3 : [[0.75294194	0.24705806]]	class [0]
Self-emp-not-inc	4 : [[0.67276886	0.32723114]]	class [0]
State-gov	5 : [[0.51403004	0.48596996]]	class [0]
Without-pay	6 : [[0.29329474	0.70670526]]	class [1]

Edu-Lvl :

10th	0 : [[0.99310099	0.00689901]]	class [0]
11th	1 : [[0.98608713	0.01391287]]	class [0]
12th	2 : [[0.97383611	0.02616389]]	class [0]
1st-4th	3 : [[0.95422889	0.04577111]]	class [0]
5th-6th	4 : [[0.92567621	0.07432379]]	class [0]
7th-8th	5 : [[0.88809135	0.11190865]]	class [0]
9th	6 : [[0.84358467	0.15641533]]	class [0]
Assoc-acdm	7 : [[0.79630436	0.20369564]]	class [0]
Assoc-voc	8 : [[0.75138619	0.24861381]]	class [0]
Bachelors	9 : [[0.71364204	0.28635796]]	class [0]
Doctorate	10 : [[0.68669521	0.31330479]]	class [0]
HS-grad	11 : [[0.67277098	0.32722902]]	class [0]
Masters	12 : [[0.67287795	0.32712205]]	class [0]
Preschool	13 : [[0.68700879	0.31299121]]	class [0]
Prof-school	14 : [[0.71413833	0.28586167]]	class [0]
Some-college	15 : [[0.75202111	0.24797889]]	class [0]

Edu-Years :

1 : [[0.99108683	0.00891317]]	class [0]

2 : [[0.986685	0.013315]]	class [0]
3 : [[0.98028506	0.01971494]]	class [0]
4 : [[0.97109175	0.02890825]]	class [0]
5 : [[0.9580709	0.0419291]]	class [0]
6 : [[0.93993647	0.06006353]]	class [0]
7 : [[0.9151884	0.0848116]]	class [0]
8 : [[0.88223901	0.11776099]]	class [0]
9 : [[0.83966457	0.16033543]]	class [0]
10 : [[0.78658854	0.21341146]]	class [0]
11 : [[0.72313226	0.27686774]]	class [0]
12 : [[0.65077361	0.34922639]]	class [0]
13 : [[0.57240025	0.42759975]]	class [0]
14 : [[0.49192084	0.50807916]]	class [1]
15 : [[0.41351106	0.58648894]]	class [1]
16 : [[0.34078467	0.65921533]]	class [1]

Marriage-Status :

Divorced	0 : [[0.87576716	0.12423284]]	class [0]
Married-AF-spouse	1 : [[0.66206303	0.33793697]]	class [0]
Married-civ-spouse	2 : [[0.59056821	0.40943179]]	class [0]
Married-spouse-absent	3 : [[0.73776478	0.26223522]]	class [0]
Never-married	4 : [[0.93563806	0.06436194]]	class [0]
Separated	5 : [[0.9949999	0.0050001]]	class [0]
Widowed	6 : [[9.99861447e-01	1.38553262e-04]]	class [0]

Occupation :

Adm-clerical	0 : [[0.78538299	0.21461701]]	class [0]
Armed-Forces	1 : [[0.77891458	0.22108542]]	class [0]
Craft-repair	2 : [[0.77258787	0.22741213]]	class [0]
Exec-managerial	3 : [[0.76641992	0.23358008]]	class [0]
Farming-fishing	4 : [[0.76042709	0.23957291]]	class [0]
Handlers-cleaners	5 : [[0.75462502	0.24537498]]	class [0]
Machine-op-inspct	6 : [[0.74902859	0.25097141]]	class [0]
Other-service	7 : [[0.74365188	0.25634812]]	class [0]
Priv-house-serv	8 : [[0.73850814	0.26149186]]	class [0]
Prof-specialty	9 : [[0.73360979	0.26639021]]	class [0]
Protective-serv	10 : [[0.7289684	0.2710316]]	class [0]
Sales	11 : [[0.72459469	0.27540531]]	class [0]
Tech-Support	12 : [[0.72049857	0.27950143]]	class [0]
Transport-moving	13 : [[0.7166891	0.2833109]]	class [0]

Relationship :

Husband	0 : [[0.65743384	0.34256616]]	class [0]
Not-in-family	1 : [[0.7395396	0.2604604]]	class [0]
Other-relative	2 : [[0.80619958	0.19380042]]	class [0]
Own-child	3 : [[0.85786333	0.14213667]]	class [0]
Unmarried	4 : [[0.89660296	0.10339704]]	class [0]
Wife	5 : [[0.92502359	0.07497641]]	class [0]

Gender :

Female	0 : [[0.94645539	0.05354461]]	class [0]
Male	1 : [[0.63791169	0.36208831]]	class [0]

Cap-Gain :

Gain <= 2000	0 : [[0.93938717	0.06061283]]	class [0]
2000 < Gain <= 4000	1 : [[0.08710199	0.91289801]]	class [1]
4000 < Gain <= 6000	2 : [[1.42635987e-08	9.99999986e-01]]	class [1]
6000 < Gain < 10000	3 : [[5.17785633e-20	1.00000000e+00]]	class [1]
10000 < Gain	4 : [[4.56423974e-36	1.00000000e+00]]	class [1]

Cap-Loss :

<= 1300	0 : [[0.87004681	0.12995319]]	class [0]
1300 < Loss <= 1600	1 : [[0.37941338	0.62058662]]	class [1]
1600 < Loss <= 1900	2 : [[4.40642634e-04	9.99559357e-01]]	class [1]
1900 < Loss < 2200	3 : [[2.50991348e-09	9.99999997e-01]]	class [1]
2200 < Loss	4 : [[1.12837309e-16	1.00000000e+00]]	class [1]

Hours :

Hours <= 20	0 : [[0.9048131	0.0951869]]	class [0]
20 < Hours <= 40	1 : [[0.80312386	0.19687614]]	class [0]
40 < Hours <= 60	2 : [[0.63677269	0.36322731]]	class [0]
60 < Hours <= 80	3 : [[0.43002123	0.56997877]]	class [1]
Hours > 80	4 : [[0.24536058	0.75463942]]	class [1]

Combining all these highest probabilities then according to Naives Bayes the profile with the highest probability to have >50K is:

50 < Age <= 70	Relationship = Husband
Work = Without-pay	Gender = Male
Edu-Lvl = Masters	Cap-Gain > 10000
Edu-Years = 16	Cap-Loss > 2200
Marriage-Status = Married-civ-spouse	Hours > 80
Occupation = Transport-moving	