

NAME : MUHAMMAD RAEES (19122041)

SECTION : 6TH A

NAME : MUHAMMAD HUSSAIN (19122010)

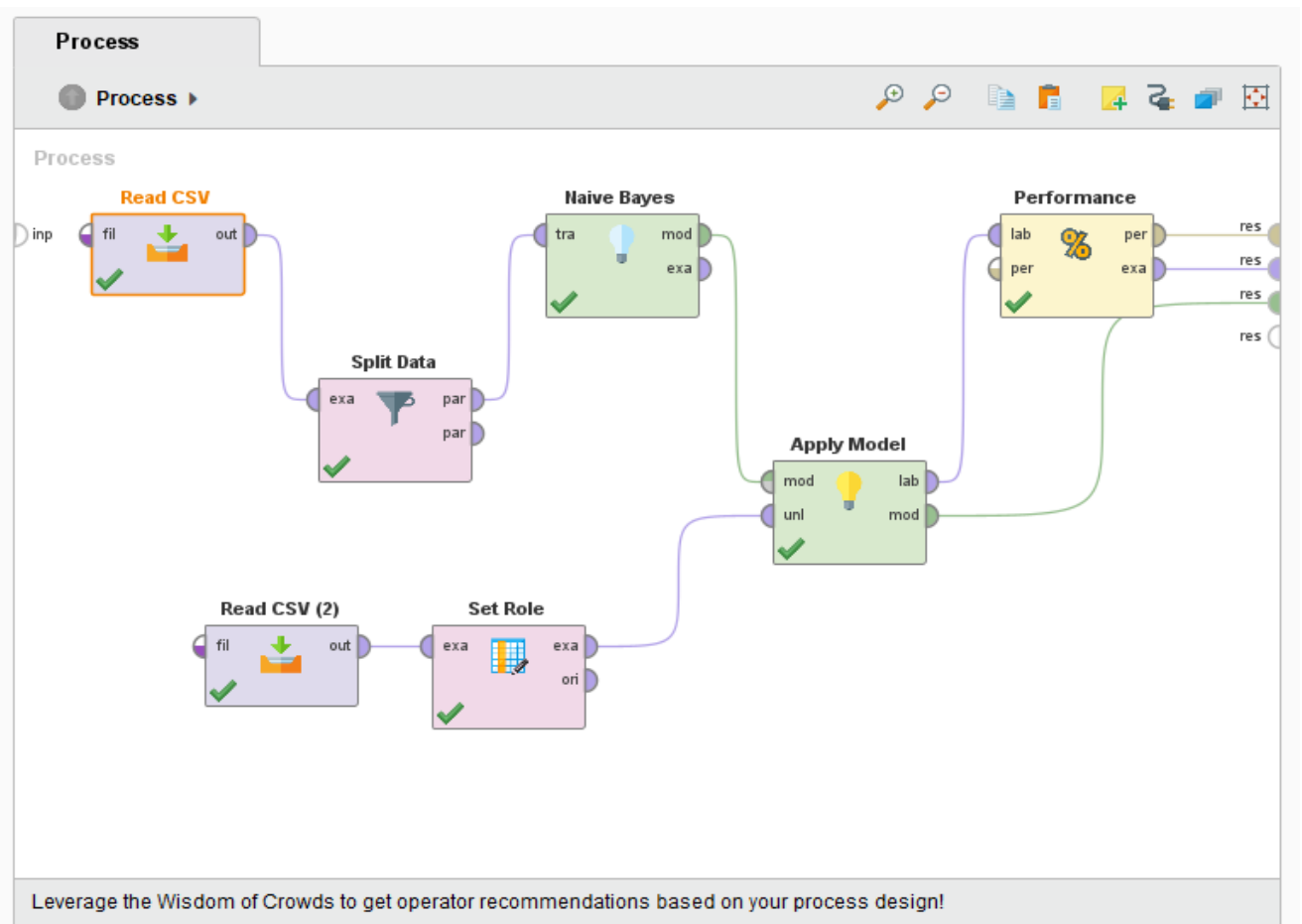
SUBJECT : DATA MINING

Term Project Final Design and Analysis

INSTALLATION OF TRAFFIC LIGHTS TO PREVENT ACCIDENT

TRAINING AND TESTING DESIGN:

IMAGE:1



SIMPLE DISTRIBUTION NAIVE BAYES:

IMAGE:2

SimpleDistribution (Naive Bayes)

ExampleSet (Apply Model)

SimpleDistribution

Distribution model for label attribute ACCIDENT

Class Not Possible (0.495)
8 distributions

Class Possible (0.505)
8 distributions

APPLY MODEL PREDICTION OF ALL DATA SET:

IMAGE:3

SimpleDistribution (Naive Bayes)

ExampleSet (Apply Model)

PerformanceVector (Performance)

Open in

Turbo Prep

Auto Model

Filter (159 / 159 examples): all

Row No.	ACCIDENT	prediction(ACCIDENT)	confidence(Not Possible)	confidence(Possible)	LOCATION D...	ROAD COND...	WEATHER C...	TIME OF DAY	VEH
1	Possible	Possible	0.004	0.996	Gulshan Iqbal	Dry	Clear Sky	Morning	Slov
2	Not Possible	Not Possible	0.959	0.041	Tower	Wet	Rainy	Evening	Nor
3	Possible	Possible	0.068	0.932	Islamabad	Icy	Snowy	Afternoon	Fas
4	Not Possible	Not Possible	0.999	0.001	Kharadar	Wet	Rainy	Night	Fas
5	Not Possible	Not Possible	1.000	0.000	Multan	Icy	Snowy	Morning	Slov
6	Possible	Possible	0.005	0.995	Agratag	Dry	Clear Sky	Morning	Slov
7	Possible	Possible	0.000	1.000	Korangi	Wet	Rainy	Night	Fas
8	Not Possible	Not Possible	1.000	0.000	Bahawalpur	Sandy	Clear Sky	Afternoon	Fas
9	Not Possible	Not Possible	0.875	0.125	Quetta	Icy	Sleet	Evening	Nor
10	Not Possible	Not Possible	1.000	0.000	Malir	Dry	Windy	Evening	Nor
11	Not Possible	Not Possible	0.997	0.003	Abbottabad	Muddy	Rainy	Morning	Slov
12	Possible	Possible	0.072	0.928	Dadu Johi mor	Dry	Windy	Afternoon	Fas
13	Possible	Not Possible	0.620	0.380	Quetta	Icy	Sleet	Night	Fas
14	Not Possible	Not Possible	0.876	0.124	Sehwan	Flooded	Rainy	Morning	Slov

ExampleSet (159 examples, 4 special attributes, 8 regular attributes)

CORRECT PREDICTION OF 124 ROWS:

IMAGE:4

SimpleDistribution (Naive Bayes) ExampleSet (Apply Model) PerformanceVector (Performance)										
Open in Turbo Prep Auto Model		Filter (124 / 159 examples): correct_predictions								
Row No.	ACCIDENT	prediction(ACCIDENT)	confidence(Not Possible)	confidence(Possible)	LOCATL...	ROAD C...	WEATH...	TIME OF...	VEHICA...	DRIVER ...
1	Possible	Possible	0.004	0.996	Gulshan ...	Dry	Clear Sky	Morning	Slow	Speeding
2	Not Possible	Not Possible	0.959	0.041	Tower	Wet	Rainy	Evening	Normal	Reckles...
3	Possible	Possible	0.068	0.932	Islamabad	Icy	Snowy	Afternoon	Fast	Impaired...
4	Not Possible	Not Possible	0.999	0.001	Kharadar	Wet	Rainy	Night	Fast	Fatigue
5	Not Possible	Not Possible	1.000	0.000	Multan	Icy	Snowy	Morning	Slow	Tailgating
6	Possible	Possible	0.005	0.995	Agratag	Dry	Clear Sky	Morning	Slow	Distracted
7	Possible	Possible	0.000	1.000	Korangi	Wet	Rainy	Night	Fast	Speeding
8	Not Possible	Not Possible	1.000	0.000	Bahawal...	Sandy	Clear Sky	Afternoon	Fast	Fatigue
9	Not Possible	Not Possible	0.875	0.125	Quetta	Icy	Sleet	Evening	Normal	Impaired...
10	Not Possible	Not Possible	1.000	0.000	Malir	Dry	Windy	Evening	Normal	Fatigue
11	Not Possible	Not Possible	0.997	0.003	Abbottab...	Muddy	Rainy	Morning	Slow	Tailgating
12	Possible	Possible	0.072	0.928	Dadu Jo...	Dry	Windy	Afternoon	Fast	Reckles...
13	Not Possible	Not Possible	0.876	0.124	Sehwan	Flooded	Rainy	Morning	Slow	Impaired...
14	Possible	Possible	0.000	1.000	Sekerand	Wet	Rainy	Afternoon	Fast	Impaired...

WRONG PREDICTION OF 35 ROWS:

IMAGE:5

SimpleDistribution (Naive Bayes)

ExampleSet (Apply Model)

PerformanceVector (Performance)

Open in Turbo PrepAuto Model

Filter (35 / 159 examples): wrong_predictions

Row No.	ACCIDENT	prediction(ACCIDENT)	confidence(Not Possible)	confidence(Possible)	LOCATL...	ROAD C...	WEATH...	TIME OF...	VEHICA...	DRIVER ...	SURRO...
1	Possible	Not Possible	0.620	0.380	Quetta	Icy	Sleet	Night	Fast	Distracted	Commer
2	Possible	Not Possible	0.576	0.424	Larkana	Wet	Rainy	Night	Fast	Distracted	Roadsid
3	Possible	Not Possible	0.649	0.351	Quetta	Slushy	Windy	Night	Fast	Inattention	Pedstria
4	Not Possible	Possible	0.022	0.978	Gujrat	Muddy	Rainy	Morning	Slow	Impaired...	Bike lan
5	Possible	Not Possible	0.995	0.005	Gilgit Bal...	Slushy	Snowy	Night	Fast	Ignor Trf...	Houses
6	Possible	Not Possible	0.996	0.004	Malir	Muddy	Clear Sky	Night	Fast	Distracted	Public Ar
7	Possible	Not Possible	0.740	0.260	Ziyarat	Slushy	Snowy	Night	Fast	Overconf...	Bank
8	Possible	Not Possible	0.986	0.014	Agrataj B...	Slushy	Foggy	Afternoon	Fast	Fatigue	Construc
9	Not Possible	Possible	0.174	0.826	Chaman	Slushy	Foggy	Morning	Slow	Evasive ...	Pedstria
10	Possible	Not Possible	0.745	0.255	Malir	Wet	Rainy	Afternoon	Fast	Inattention	Crosswa
11	Not Possible	Possible	0.023	0.977	Gulshan ...	Muddy	Snowy	Evening	Normal	Inattention	Commer
12	Not Possible	Possible	0.000	1.000	Liyari Ex...	Wet	Clear Sky	Night	Fast	Overuse ...	Public Ar
13	Possible	Not Possible	0.862	0.138	Quetta	Slushy	Sleet	Morning	Slow	Speeding	Construc
14	Not Possible	Possible	0.017	0.983	Sukkur	Muddy	Clear Sky	Afternoon	Fast	Evasive ...	Houses

NAIVE BAYES PERFORMANCE ACURACY:

IMAGE:6

	true Possible	true Not Possible	class precision
pred. Possible	59	13	81.94%
pred. Not Possible	22	65	74.71%
class recall	72.84%	83.33%	

RESULT:

WE CREATE ANOTHER COLUMN OF LIGHTS INSTALLATION IN OUR **159 ROWS OF DATA** SET AS PER NAIVE BAYES PREDICTION WHERE WE PREDICT “POSSIBLE” AND “NOT POSSIBLE” CONDITION. NAIVE BAYES GIVE US **124 ROWS** OF DATA CORRECT PREDICTION ABOVE MENTION IN **IMAGE 4** AND **35 ROWS** OF DATA WRONG PREDICTION ABOVE METION IN **IMAGE 5**. PERFORMANCE **ACURACY IS 77.99%** WHICH IS BEST AS PER PREDICTION WE ANALYSIS AND DECIDE WHERE WE INSTALL THE LIGHTS OR WHER NOT THEN WE UPDATE LIGHTS INSTALLATION COLUMN “YES” AND “ NO”, YES PRESENT INSTALL THE LIGHT NO PRESENT DON’T INSTALL.