



Name : Sumit Kumar das

ID: 1731847642

Email : sumit.das@northsouth.edu

CSE 445

Sec : 3

Course Instructor : ITN

Assignment Report

EDA

	Sector	score	LOCATION_ID	PARA_A	Score_A	Risk_A	PARA_B	Score_B	Risk_B	B TOTAL	numbers	Score_B.1	Risk_C	Money_Value	Score_MV	Risk_D	District_Loss	PROB	Risk_E	History_Prob	Score_Inherent	Risk_CONTROL	Risk_Audit	Risk_Risk	Marks_MONEY	Loss_District	Loss_LOSS				
rowspan=10	Sector	score	1.00	-0.05	-0.21	-0.42	-0.13	-0.21	-0.13	-0.15	-0.15	-0.17	-0.16	-0.12	-0.32	-0.11	-0.11	-0.09	-0.13	-0.11	-0.14	-0.33	-0.17	-0.16	-0.09	-0.39	-0.17	-0.32	-0.11	-0.08	-0.09
	LOCATION_ID	-0.05	1.00	-0.01	0.08	-0.00	0.01	0.13	0.01	0.01	-0.02	-0.02	-0.07	0.11	-0.07	-0.11	-0.00	-0.10	-0.08	-0.05	0.09	-0.05	-0.12	-0.09	0.06	-0.02	0.11	-0.11	0.01	-0.00	
	PARA_A	-0.21	-0.01	1.00	0.49	1.00	0.16	0.36	0.16	0.27	0.13	0.14	0.14	0.45	0.28	0.45	0.13	0.04	0.12	0.12	0.17	0.42	0.48	0.15	0.22	0.38	0.14	0.28	0.13	0.04	0.04
	Score_A	-0.42	0.08	0.49	1.00	0.50	0.25	0.57	0.25	0.30	0.24	0.27	0.26	0.20	0.47	0.20	0.09	0.09	0.10	0.18	0.26	0.72	0.32	0.17	0.20	0.62	0.27	0.47	0.09	0.09	0.09
	Risk_A	-0.22	-0.00	1.00	0.50	1.00	0.16	0.36	0.16	0.27	0.13	0.14	0.14	0.45	0.29	0.45	0.13	0.04	0.12	0.12	0.17	0.43	0.48	0.15	0.22	0.38	0.14	0.29	0.13	0.04	0.04
	PARA_B	-0.13	0.01	0.16	0.25	0.16	1.00	0.35	1.00	0.99	0.21	0.23	0.22	0.12	0.31	0.12	0.08	0.04	0.08	0.20	0.32	0.40	0.65	0.19	0.89	0.26	0.23	0.31	0.08	0.04	0.04
	Score_B	-0.21	0.13	0.36	0.57	0.36	0.35	1.00	0.35	0.38	0.28	0.31	0.30	0.20	0.56	0.20	-0.01	0.09	0.01	0.20	0.31	0.90	0.36	0.12	0.21	0.63	0.31	0.56	-0.01	0.10	0.09
	Risk_B	-0.13	0.01	0.16	0.25	0.16	1.00	0.35	1.00	0.99	0.21	0.23	0.22	0.12	0.31	0.12	0.08	0.04	0.08	0.20	0.32	0.40	0.65	0.19	0.89	0.25	0.23	0.31	0.08	0.04	0.04
	TOTAL	-0.15	0.01	0.27	0.30	0.27	0.99	0.38	0.99	0.99	0.22	0.24	0.23	0.17	0.34	0.17	0.09	0.05	0.09	0.21	0.33	0.43	0.69	0.20	0.89	0.29	0.24	0.34	0.09	0.05	0.05
	numbers	-0.15	0.01	0.13	0.24	0.13	0.21	0.28	0.21	0.22	1.00	0.91	0.96	0.19	0.45	0.19	0.12	0.03	0.14	0.20	0.21	0.50	0.27	0.23	0.22	0.31	0.91	0.45	0.12	-0.00	0.03
	Score_B.1	-0.17	-0.02	0.14	0.27	0.14	0.23	0.31	0.23	0.24	0.91	1.00	0.99	0.22	0.51	0.22	0.15	0.04	0.16	0.23	0.25	0.57	0.31	0.26	0.26	0.35	1.00	0.51	0.15	0.00	0.04
	Risk_C	-0.16	-0.02	0.14	0.26	0.14	0.22	0.30	0.22	0.23	0.96	0.99	1.00	0.21	0.49	0.21	0.15	0.04	0.15	0.22	0.24	0.55	0.30	0.25	0.25	0.34	0.99	0.49	0.15	0.00	0.04
	Money_Value	-0.12	-0.07	0.45	0.20	0.45	0.12	0.20	0.12	0.17	0.19	0.22	0.21	1.00	0.39	1.00	0.03	0.03	0.03	0.08	0.11	0.29	0.83	0.07	0.33	0.26	0.22	0.39	0.03	0.02	0.03
	Score_MV	-0.32	0.11	0.28	0.47	0.29	0.31	0.56	0.31	0.34	0.45	0.51	0.49	0.39	1.00	0.39	0.08	0.13	0.10	0.25	0.33	0.76	0.48	0.22	0.29	0.69	0.51	1.00	0.08	0.12	0.13
	Risk_D	-0.11	-0.07	0.45	0.20	0.45	0.12	0.20	0.12	0.17	0.19	0.22	0.21	1.00	0.39	1.00	0.03	0.03	0.03	0.08	0.11	0.29	0.83	0.07	0.33	0.25	0.22	0.39	0.03	0.02	0.03
	District_Loss	-0.11	-0.11	0.13	0.09	0.13	0.08	-0.01	0.08	0.09	0.12	0.15	0.15	0.03	0.08	0.03	0.00	0.05	0.91	0.07	0.08	0.21	0.08	0.64	0.20	0.40	0.15	0.08	1.00	0.04	0.05
	PROB	-0.09	-0.00	0.04	0.09	0.04	0.04	0.09	0.04	0.05	0.03	0.04	0.04	0.03	0.13	0.03	0.05	1.00	0.41	0.11	0.14	0.17	0.05	0.34	0.07	0.18	0.04	0.13	0.05	0.90	1.00
	Risk_E	-0.13	-0.10	0.12	0.10	0.12	0.08	0.01	0.08	0.09	0.14	0.16	0.15	0.03	0.10	0.03	0.91	0.41	1.00	0.11	0.12	0.23	0.08	0.73	0.20	0.41	0.16	0.10	0.91	0.37	0.41
	History	-0.11	-0.08	0.12	0.18	0.12	0.20	0.20	0.20	0.21	0.20	0.23	0.22	0.08	0.25	0.08	0.07	0.11	0.11	1.00	0.86	0.33	0.18	0.76	0.33	0.24	0.23	0.25	0.07	0.06	0.11
	Prob	-0.14	-0.05	0.17	0.26	0.17	0.32	0.31	0.32	0.33	0.21	0.25	0.24	0.11	0.33	0.11	0.08	0.14	0.12	0.86	1.00	0.44	0.27	0.64	0.43	0.30	0.25	0.33	0.08	0.10	0.14
	Score	-0.33	0.09	0.42	0.72	0.43	0.40	0.90	0.40	0.43	0.50	0.57	0.55	0.29	0.76	0.29	0.21	0.17	0.23	0.33	0.44	1.00	0.46	0.35	0.33	0.78	0.57	0.76	0.21	0.16	0.17
	Inherent_Risk	-0.17	-0.05	0.48	0.32	0.48	0.65	0.36	0.65	0.69	0.27	0.31	0.30	0.83	0.48	0.83	0.08	0.05	0.08	0.18	0.27	0.46	1.00	0.17	0.75	0.36	0.31	0.48	0.08	0.05	0.05
	CONTROL_RISK	-0.16	-0.12	0.15	0.17	0.15	0.19	0.12	0.19	0.20	0.23	0.26	0.25	0.07	0.22	0.07	0.64	0.34	0.73	0.76	0.64	0.35	0.17	1.00	0.36	0.41	0.26	0.22	0.64	0.28	0.34
	Audit_Risk	-0.09	-0.09	0.22	0.20	0.22	0.89	0.21	0.89	0.89	0.22	0.26	0.25	0.33	0.29	0.33	0.20	0.07	0.20	0.33	0.43	0.33	0.75	0.36	1.00	0.22	0.26	0.29	0.20	0.05	0.07
	Risk	-0.39	0.06	0.38	0.62	0.38	0.26	0.63	0.25	0.29	0.31	0.35	0.34	0.26	0.69	0.25	0.40	0.18	0.41	0.24	0.30	0.78	0.36	0.41	0.22	1.00	0.35	0.69	0.40	0.17	0.18
	Marks	-0.17	-0.02	0.14	0.27	0.14	0.23	0.31	0.23	0.24	0.91	1.00	0.99	0.22	0.51	0.22	0.15	0.04	0.16	0.23	0.25	0.57	0.31	0.26	0.26	0.35	1.00	0.51	0.15	0.00	0.04
	MONEY_Marks	-0.32	0.11	0.28	0.47	0.29	0.31	0.56	0.31	0.34	0.45	0.51	0.49	0.39	1.00	0.39	0.08	0.13	0.10	0.25	0.33	0.76	0.48	0.22	0.29	0.69	0.51	1.00	0.08	0.12	0.13
	District	-0.11	-0.11	0.13	0.09	0.13	0.08	-0.01	0.08	0.09	0.12	0.15	0.15	0.03	0.08	0.03	0.00	0.05	0.91	0.07	0.08	0.21	0.08	0.64	0.20	0.40	0.15	0.08	1.00	0.04	0.05
	Loss	-0.08	0.01	0.04	0.09	0.04	0.04	0.10	0.04	0.05	-0.00	0.00	0.00	0.02	0.12	0.02	0.04	0.90	0.37	0.06	0.10	0.16	0.05	0.28	0.05	0.17	0.00	0.12	0.04	1.00	0.99
	LOSS_SCORE	-0.09	-0.00	0.04	0.09	0.04	0.04	0.09	0.04	0.05	0.03	0.04	0.04	0.03	0.13	0.03	0.05	1.00	0.41	0.11	0.14	0.17	0.05	0.34	0.07	0.18	0.04	0.13	0.05	0.96	1.00
	History_score	-0.14	-0.05	0.17	0.26	0.17	0.32	0.31	0.32	0.33	0.21	0.25	0.24	0.11	0.33	0.11	0.08	0.14	0.12	0.86	1.00	0.44	0.27	0.64	0.43	0.30	0.25	0.33	0.08	0.10	0.14

	Risk_A	Risk_B	Risk_C	Risk_D	Risk_E	Prob	Score	CONTROL_RISK	Audit_Risk	Risk_Risk	MONEY_Marks	Loss
Risk_A	1.00	0.16	0.14	0.45	0.12	0.17	0.43	0.15	0.22	0.38	0.29	0.04
Risk_B	0.16	1.00	0.22	0.12	0.08	0.32	0.40	0.19	0.89	0.25	0.31	0.04
Risk_C	0.14	0.22	1.00	0.21	0.15	0.24	0.55	0.25	0.25	0.34	0.49	0.00
Risk_D	0.45	0.12	0.21	1.00	0.03	0.11	0.29	0.07	0.33	0.25	0.39	0.02
Risk_E	0.12	0.08	0.15	0.03	1.00	0.12	0.23	0.73	0.20	0.41	0.10	0.37
Prob	0.17	0.32	0.24	0.11	0.12	1.00	0.44	0.64	0.43	0.30	0.33	0.10
Score	0.43	0.40	0.55	0.29	0.23	0.44	1.00	0.35	0.33	0.78	0.76	0.16
CONTROL_RISK	0.15	0.19	0.25	0.07	0.73	0.64	0.35	1.00	0.36	0.41	0.22	0.28
Audit_Risk	0.22	0.89	0.25	0.33	0.20	0.43	0.33	0.36	1.00	0.22	0.29	0.05
Risk_Risk	0.38	0.25	0.34	0.25	0.41	0.30	0.78	0.41	0.22	1.00	0.69	0.17
MONEY_Marks	0.29	0.31	0.49	0.39	0.10	0.33	0.76	0.22	0.29	0.69	1.00	0.12
Loss	0.04	0.04	0.00	0.02	0.37	0.10	0.16	0.28	0.05	0.17	0.12	1.00

This is the correlation matrix.

Email : sumit.das@northsouth.edu

Classifiers Used :

Soft Voting Classifier : In **soft voting**, predictions are chosen by summing the prediction probabilities from the individual models and choosing the prediction with the largest sum. In **soft voting**, every individual classifier provides a probability value that a specific data point belongs to a particular target class. The predictions are weighted by the classifier's importance and summed up. Then the target label with the greatest sum of weighted probabilities wins the vote.

Hard Voting Classifier : In **hard voting** (also known as **majority voting**), every individual classifier votes for a class, and the majority wins. In statistical terms, the predicted target label of the ensemble is the mode of the distribution of individually predicted labels.

Tables :

Hard Voting

	Accuracy
KNeighborsClassifier	0.9473684210526315
SVC	0.9526315789473684
LogisticRegression	0.9578947368421052
VotingClassifier	0.9526315789473684

Email : sumit.das@northsouth.edu

Soft Voting

	Accuracy
KNeighborsClassifier	0.9526315789473684
SVC	0.9578947368421052
LogisticRegression	0.9473684210526315
VotingClassifier	0.9578947368421052

Email : sumit.das@northsouth.edu

