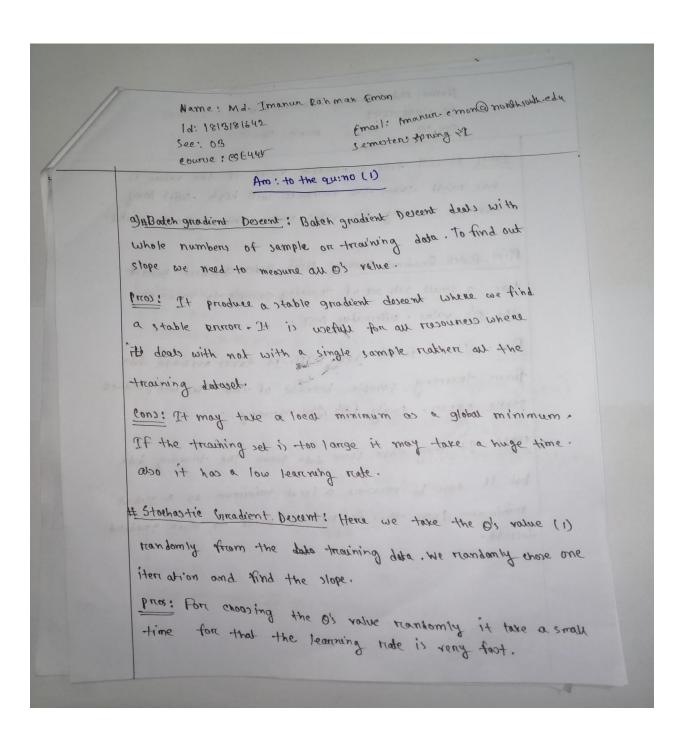
Name: Md Imanur Rahman Emon ID: 1813181642

Course: CSE445 Section: 03



Name: Md. Imanua Rahman Emen

12:1813181642

see 13

wuse: Of445

not right than the variance will high . With May take a wrong o's value which can perform wordst.

Mini-Boten Gradient Descent: Mini-Baten Gradient Descent take a small sub set of treating sample to measure the O's value. Whematly MSE.

Prior: It is faster than BGD. It takes average data from training sample because of that this will produce stable error. Easily fit's the memory.

but it can be measure a local miniman as a global minimum. For the it resects noice to other gradient descent.

Name: Md. [manux fahman &mon]
11. 1813181642

1) B) Momemtum is a term which used in gradual desent algorithms. Anadient desent is a optimization algorithm which works to find the purfect slope on that point Towards that direction in each steps, the value of function to be minimized. The problem in this direction may change gradually in some points of the function where there best path to 90 assumptions not contain a lot of turns. So, it use so called momentum to improve conversione. Which already Soing along for sometime before (t changes it)

Name: Md. Imanur Rahman Emon

12: 1813181642

section: 3

. course : cse 445

semester: spring 2019

NSU Email: Imanua emon@ monthsouth

Am: to the qu: no (4)

Here, the system is failing is A fails & oxiest i of B, C, D fails.

so, the probability is to (A tails) + 3 (13) x (243),

(BD on Be on De fail) = $\frac{1}{3} + \frac{2}{9} = \frac{5}{9} = 0.555$

Probability that only A fail = \frac{1}{3} x (2/3) = 0.0987

Preobability that only A fails given that when system is

Auling = 0.0987

= 0.1778

- a) Probability that each system works is 2/3, and each fulls is 1/3.
- so, the probability that all system work is = 49 x 43 x 43 x 2/3

Name: Md. Imanur Rahman Emost 16:1813181642 course: CSE 445 Sec: 3

= 0.1975 Am: (b) Probability that system does not work proporty 15 = (1-(2/3)4)

$$= 0.803$$

$$= 1-0.101$$

$$= 1-(0.69)$$

Probability that only A system fails is = \frac{1}{3}x (213)

$$=\frac{8}{81}$$

= 0.09876

Probability that only A toil given fulls also system goes not brouk buobellh is = 0.03818

Name: Md Imanun Rahman Emon

14:1813181642

course: CJE445

See! 3

The same is brought out in empire cal probability.

A, B, C, D failures on functioning are simulated wing randombetween function in the range 1 to 9. It I to 3 comes it's failure esse success. Yestern A only failed 98 times out of 810 impropers working of system. The resulting probability in 98/810

= 0.120

Am:

Name! Md. Imanun Rahman Emon

ld: 1813181642 course: CSE 445

Sec:3

Am: to the que no (3)

the Bias: In our model when appearance the pentormance from the known data we trained the model, by doing this with bias. Bias is a source of error in our model it's the reason of anderfitting.

Variance: While we train our model we took

varuance: The varuance is an enror from sensitivity with small function in the training dataset. When we thating our model from unknown data then variance comes.

Bias us variance treadeoff deals with the prediction error Via bias and varience. When testing accuracy is too bad Like we find 80% accuracy while we treating on the other hand when in comes of testing the model with an unknown data and shows yor accuracy that means Name: MJ. Imanun Rahman Emon

ld: 1813181642 Course: CSE 445

sec:3

there's variance is very to high. For measuring the model while we took a data from training data and it shows that the training accuracy is 90% and testing accuracy is 60% it means there state is overfelting. In our machine learning model we need a per-formance where there is low bias and low variance means there is no brentitling on underfitting data. so we need to train our model with a good amount of data, and try to split the Data by 80/20 and the same time try to prepower the dola concerty. Then we will find a good moder where alow bias and a low variance which will helps to prudict the data / system penfectly. It not have to change the model.

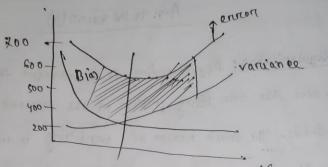
Name: Md. Imanun Rahman Emon

18: 18:13:18:1642

course: ese445

Email: Imanum emon@ northnowth . edu

semesters: spring 2021



Bias - variance tradeoff.

training accuracy is good when the blas. While testing accuracy is too bad then high varience.

Name: Md. Imanur Pahman Ernon
lde 1813181642
see: 03
course: UE445

Am: to the quino (4)

Regularization: Regularization is a technique that is used to solve the overefitting problem of the Machine learning Methods. The main reason of overefitting is making a model more remplex than necessary. If we find away to the duce the complexity, then overefitting issue is solved.

two common methods of regularization are [1 and [2] hegularization.

LI Degularization: LI Regularization is the the penalty for loss regression, it's cost function is defined as, ming (y-x8) (y-x8)+x11 B11

do . Hu) additional cost will cause the resulting weights to do . Hu) additional cost . Unfortunately 11 regularization does not have a closed form solution - because it is not differentiable when a weight B falls to 0. The another

Name: Md. Imanun Rahman Emon

14: 1813181442

see: 03

course: UE445

of the constraint. The constraint have very thank edges which lie on each dimensional ani) at a distance e from the origin.

LI Regulization penalizes (weight)

Lasso regression modifies the overfitted on undertitled by adding the penalty equivalent to the sum of the absolute value of coefficients-

est function = Lon + 8x 21w1

L2 Regularization: Min B (4-48) + 8 11B11

The only difference is the added regularization term & with in the threeoe.

Ridge Regrussion (12) Example:

(cost function = Loss + } x {Ulusis

This is why negularitate are often called a penalty-

Name: Md. Imanua Rahman Emon

14: 1813181642

Course: UE 445

(el: 03 semeters spring 21

Am: to the quinoth) (6)

Have to compute the enticopy of the system. We know,

Enthopy *
$$\mathfrak{All}(x) = \sum_{i=1}^{n} P(x_i) \log_2 \left(\frac{1}{P(x_i)} \right)$$

$$= (0.03)(3.43)$$

$$= (0.03)(3.05) + (0.65)(0.689) + (0.56)(1.94)$$

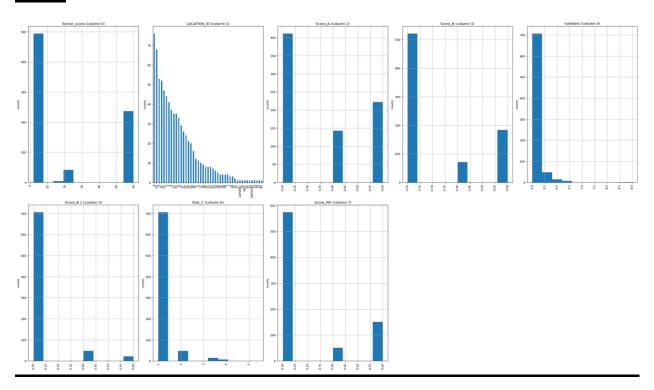


The entrophy of the system is 1:387

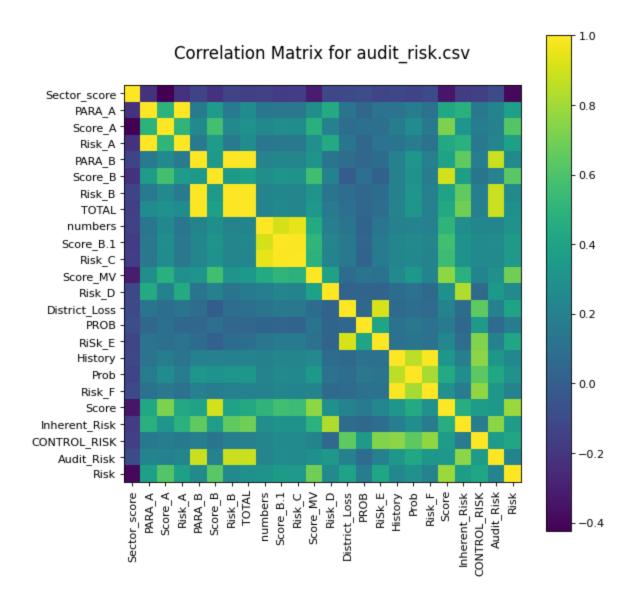
Report:

Problem 1

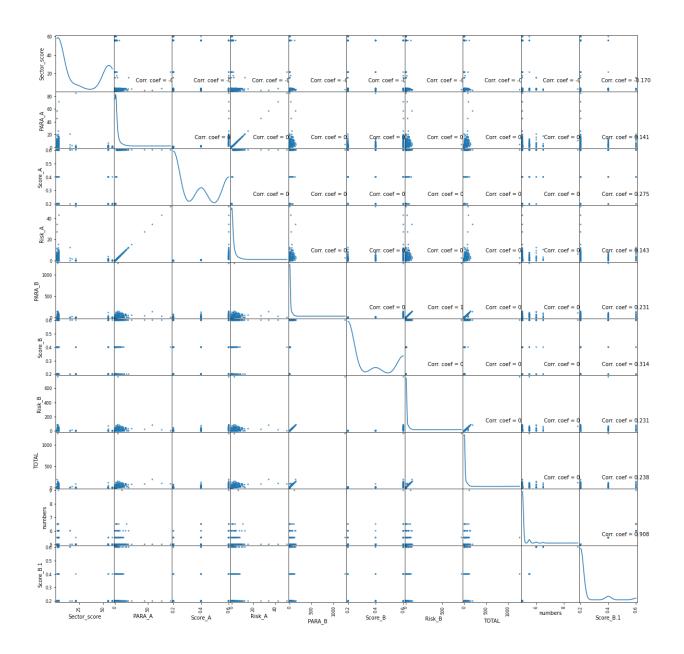
EDA

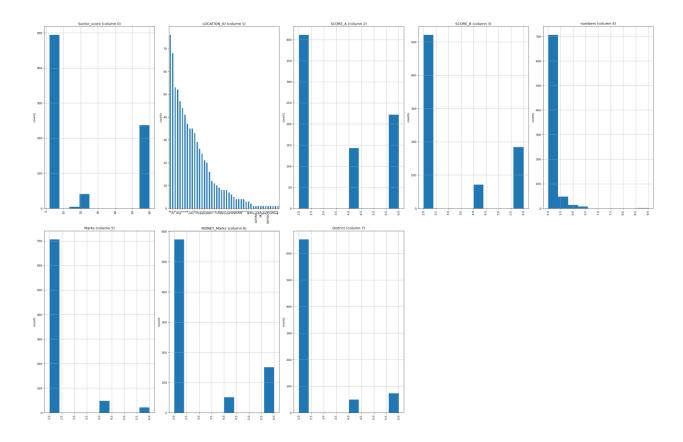


Per column Data



Confusion Matrix





Per Column Data

Sector_score	1	-0.22	-0.43	-0.22	-0.13	-0.22	-0.13	-0.15	-0.15	-0.17	-0.17	-0.12	-0.32	-0.12	-0.11	-0.087	-0.13	-0.11	-0.14	-0.1	-0.34	-0.17	-0.15		-0.092	-0.39	I		1.0
PARA_A	-0.22	1	0.5		0.16	0.36	0.16	0.27	0.13	0.14	0.14	0.45	0.29	0.45	0.13	0.044	0.12	0.12	0.17	0.1	0.43	0.48	0.15		0.22	0.38			
Score_A	-0.43	0.5	1	0.5	0.25	0.57	0.25	0.3	0.24	0.27	0.27	0.21	0.48	0.2	0.089	0.094	0.1	0.18	0.27	0.15	0.72	0.32	0.17		0.2	0.62			
Risk_A	-0.22	1	0.5	1	0.17	0.36	0.17	0.27	0.14	0.14	0.14	0.45	0.29	0.45	0.13	0.044	0.12	0.12	0.18	0.11	0.43	0.48	0.15		0.22	0.39			- 0.8
PARA_B	-0.13	0.16	0.25	0.17	1	0.35	1	0.99	0.21	0.23	0.22	0.13	0.31	0.12	0.083	0.043	0.079	0.2	0.32	0.2	0.4	0.65	0.19		0.89	0.26			
Score_B	-0.22	0.36	0.57	0.36	0.35	1	0.35	0.38	0.28	0.31	0.3	0.21	0.57	0.2	-0.0047	0.093	0.015	0.2	0.31	0.17	0.9	0.37	0.13		0.21	0.64			
Risk_B	-0.13	0.16	0.25	0.17	1	0.35	1	0.99	0.21	0.23	0.22	0.13	0.31	0.12	0.083	0.043	0.08	0.2	0.32	0.2	0.4	0.65	0.19		0.89	0.26			
TOTAL	-0.15	0.27	0.3	0.27	0.99	0.38	0.99		0.22	0.24	0.23	0.17	0.34	0.17	0.093	0.046	0.089	0.21	0.33	0.2	0.43	0.69	0.2		0.89	0.29			- 0.6
numbers	-0.15	0.13	0.24	0.14	0.21	0.28	0.21	0.22	1	0.91	0.96	0.19	0.45	0.19	0.13	0.036	0.14	0.2	0.21	0.2	0.5	0.27	0.23		0.22	0.31			
Score_B.1	-0.17	0.14	0.27	0.14	0.23	0.31	0.23	0.24			0.99	0.22	0.51	0.22	0.15	0.037	0.16	0.23	0.25	0.22	0.57	0.31	0.26		0.26	0.35			
Risk_C	0.17	0.14	0.27	0.14	0.22	0.3	0.22	0.23	0.96	0.99		0.22	0.49	0.22	0.15	0.036	0.15	0.22	0.24	0.22	0.55	0.3	0.25		0.25	0.34			- 0.4
Money_Value	-0.12	0.45	0.21	0.45	0.13	0.21	0.13	0.17	0.19	0.22	0.22	1	0.39	1	0.028	0.032	0.033	0.08	0.11	0.07	0.29	0.83	0.07		0.33	0.26			- 0.4
Score_MV	-0.32	0.29	0.48	0.29	0.31	0.57	0.31	0.34	0.45	0.51	0.49	0.39	1	0.39	0.081	0.13	0.1	0.25	0.33	0.22	0.76	0.48	0.22		0.29	0.69			
Risk_D	-0.12	0.45	0.2	0.45	0.12	0.2	0.12	0.17	0.19	0.22	0.22	1	0.39	1	0.028	0.032	0.033	0.08	0.11	0.07	0.29	0.83	0.07		0.33	0.25			
District_Loss	-0.11	0.13	0.089	0.13	0.083	-0.0047	0.083	0.093	0.13	0.15	0.15	0.028	0.081	0.028	1	0.055	0.91	0.069	0.084	0.071	0.21	0.081	0.65		0.2	0.4			- 0.2
PROB	-0.087	0.044	0.094	0.044	0.043	0.093	0.043	0.046	0.036	0.037	0.036	0.032	0.13	0.032	0.055	1	0.41	0.11	0.14	0.11	0.17	0.053	0.34		0.074	0.18			
RiSk_E	-0.13	0.12	0.1	0.12	0.079	0.015	0.08	0.089	0.14	0.16	0.15	0.033	0.1	0.033	0.91	0.41	1	0.11	0.12	0.11	0.24	0.083			0.2	0.41			
History	-0.11	0.12	0.18	0.12	0.2	0.2	0.2	0.21	0.2	0.23	0.22	0.08	0.25	0.08	0.069	0.11	0.11			0.99	0.33	0.19			0.33	0.24			
Prob	-0.14	0.17	0.27	0.18	0.32	0.31	0.32	0.33	0.21	0.25	0.24	0.11	0.33	0.11	0.084	0.14	0.12		1	0.82	0.44	0.27	0.64		0.43	0.3			- 0.0
Risk_F	-0.1	0.1	0.15	0.11	0.2	0.17	0.2	0.2	0.2	0.22	0.22	0.07	0.22	0.07	0.071	0.11	0.11	0.99	0.82	1	0.29	0.17	0.76		0.33	0.21			
Score	-0.34	0.43	0.72	0.43	0.4	0.9	0.4	0.43	0.5	0.57	0.55	0.29	0.76	0.29	0.21	0.17	0.24	0.33	0.44	0.29	1	0.46	0.36		0.33	0.79		ı	
Inherent_Risk	-0.17	0.48	0.32	0.48	0.65	0.37	0.65	0.69	0.27	0.31	0.3	0.83	0.48	0.83	0.081	0.053	0.083	0.19	0.27	0.17	0.46	1	0.17		0.75	0.36			0.2
CONTROL_RISK	-0.15	0.15	0.17	0.15	0.19	0.13	0.19	0.2	0.23	0.26	0.25	0.07	0.22	0.07	0.65	0.34	0.73	0.75	0.64	0.76	0.36	0.17	1		0.36	0.42			
Detection_Risk																													
Audit_Risk	-0.092	0.22	0.2	0.22	0.89	0.21	0.89	0.89	0.22	0.26	0.25	0.33	0.29	0.33	0.2	0.074	0.2	0.33	0.43	0.33	0.33	0.75	0.36		1	0.22			
Risk					0.26	0.64		0.29		0.35		ı.		0.25		٠,	0.41	,	0.3	0.21		0.36		,	0.22	1			0.4
	Sector_score	PARA_A	Score_A	Risk_A	PARA_B	Score_B	Risk_B	TOTAL	numbers	Score_B.1	Risk_C	Money_Value	Score_MV	Risk_D	District_Loss	PROB	RISK_E	History	Prob	Risk F	Score	Inherent_Risk	ONTROL_RISK	Detection_Risk	Audit_Risk	Risk			

Hit Map

Classifiers:

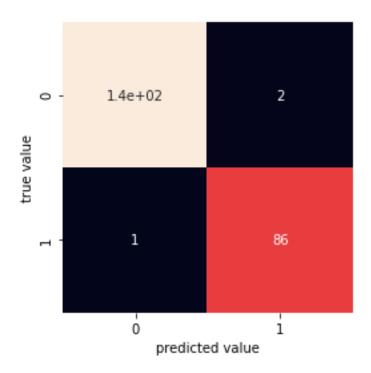
Logistic Regression

Decision Tree

Random Forrest

Accuracy

```
Logistic Regression: 98.71%
K-Nearest Neighbors: 90.13%
Decision Tree: 100.00%
Support Vector Machine (Linear Kernel): 99.14%
Support Vector Machine (RBF Kernel): 97.00%
Neural Network: 99.14%
Random Forest: 100.00%
Gradient Boosting: 100.00%
```



Confusion Matrix

Assignment 2:

