Machine Learning Homework 3*

Hongzhou Liu
517030910214
deanlhz@sjtu.edu.cn

1 SVM vs. Neural Networks

In this section, I did experiments on the SVM and MLP using following two datasets:

Table 1: Datasets

Dataset	Classes	Size	Features
breast-cancer	2	683	10
dna	3	3186	180

As seen, I tried both binary and multiclass classification tasks on the two classifiers.

1.1 SVM

1.1.1 Experiment on data preprocessing

Table 2: Hyper-parameters

Dataset	training size	C	kernel	dimension
breast-cancer	0.6	1.0	rbf	10
dna	0.6	1.0	rbf	180

Table 3: Experiment on data preprocessing

scale	breast-cancer accuracy	dna accuracy
True	95.26%	95.37%
False	**97.81%**	**95.84%**

Table 4: Hyper-parameters

Dataset	scale	C	kernel	dimension
breast-cancer	False	1.0	rbf	10
dna	False	1.0	rbf	180

Table 5: Experiment on sample size of training set

training size	breast-cancer accuracy	dna accuracy
0.9	95.65%	93.41%
0.8	**99.27%**	**97.65%**
0.7	97.07%	95.86%
0.6	96.35%	95.29%
0.5	97.66%	96.30%

Table 6: Hyper-parameters

Dataset	scale	training size	kernel	dimension
breast-cancer	False	0.6	rbf	10
dna	False	0.6	rbf	180

Table 7: Experiment on penalty factor

C	breast-cancer accuracy	dna accuracy
0.001	64.23%	52.70%
0.005	68.98%	52.70%
0.01	95.26%	49.65%
0.05	**97.08%**	52.55%
0.1	96.72%	68.31%
0.5	96.72%	95.45%
1.0	96.35%	**96.63%**
5.0	96.72%	95.76%
10.0	96.72%	95.69%

Table 8: Hyper-parameters

Dataset	scale	training size	C	dimension
breast-cancer	False	0.6	1.0	10
dna	False	0.6	1.0	180

Table 9: Experiment on kernel of SVM

kernel	breast-cancer accuracy	dna accuracy
linear	96.35%	92.39%
poly	**96.72%**	94.35%
rbf	95.62%	**95.45%**
sigmoid	92.70%	93.41%

Table 10: Hyper-parameters

Dataset	scale	training size	C	kernel
breast-cancer	False	0.6	1.0	rbf
dna	False	0.6	1.0	rbf

Table 11: Experiment on dimension

dimension	breast-cancer accuracy
2	97.08%
3	97.45%
5	97.45%
8	96.72%
10	**97.81%**

Table 12: Experiment on dimension

dimension	dna accuracy
2	76.39%
5	88.94%
10	91.37%
20	93.18%
50	92.39%
100	93.73%
150	95.69%
180	**96.39%**

Table 13: Hyper-parameters

Dataset	training size	hidden layers	activation	α	dimension
breast-cancer	0.6	(100)	relu	0.0001	10
dna	0.6	(100)	relu	0.0001	180

Table 14: Experiment on data preprocessing

scale	breast-cancer accuracy	dna accuracy
True	97.45%	93.80%
False	**97.81%**	**94.20%**

Table 15: Hyper-parameters

Dataset	scale	hidden layers	activation	α	dimension
breast-cancer	False	(100)	relu	0.0001	10
dna	False	(100)	relu	0.0001	180

Table 16: Experiment on sample size of training set

training size	breast-cancer accuracy	dna accuracy
0.9	97.10%	**95.30%**
0.8	**97.81%**	94.83%
0.7	96.59%	94.56%
0.6	96.72%	94.67%
0.5	97.66%	94.98%

Table 17: Hyper-parameters

Dataset	scale	training size	activation	α	dimension
breast-cancer	False	0.6	relu	0.0001	10
dna	False	0.6	relu	0.0001	180

Table 18: Experiment on network structure

hidden layers	breast-cancer accuracy	dna accuracy
(10)	97.08%	93.49%
(100)	97.08%	94.43%
(10, 10)	**97.81%**	93.65%
(100, 100)	95.26%	94.59%
(200, 200)	95.62%	**94.90%**
(100, 200, 100)	97.45%	94.20%

Table 19: Hyper-parameters

Dataset	scale	training size	hidden layers	α	dimension
breast-cancer	False	0.6	(100)	0.0001	10
dna	False	0.6	(100)	0.0001	180

Table 20: Experiment on activation function

activation	breast-cancer accuracy	dna accuracy
identity	**97.45%**	91.45%
logistic	96.35%	93.96%
tanh	94.16%	**94.59%**
relu	97.08%	94.20%

Table 21: Hyper-parameters

Dataset	scale	training size	hidden layers	activation	dimension
breast-cancer	False	0.6	(100)	relu	10
dna	False	0.6	(100)	relu	180

Table 22: Experiment on regularization factor

α	breast-cancer accuracy	dna accuracy
0.001	**96.35%**	93.41%
0.0001	95.99%	**94.04%**

Table 23: Hyper-parameters

Dataset	scale	training size	hidden layers	activation	α
breast-cancer	False	0.6	(100)	relu	0.0001
dna	False	0.6	(100)	relu	0.0001

Table 24: Experiment on dimension

dimension	breast-cancer accuracy
2	96.35%
3	**98.18%**
5	96.72%
8	98.18%
10	97.44%

Table 25: Experiment on dimension

dimension	dna accuracy
2	77.10%
5	89.02%
10	90.27%
20	89.18%
50	91.76%
100	92.31%
150	**96.31%**
180	94.51%

Table 26: Hyper-parameters

No.	Dataset	scale	C	kernel	dimension
0	CIFAR-10	False	5.0	rbf	3072
1		True			3072
2		False			50
3		False			500
4		False			1500

Table 27: SVM on CIFAR-10

No.	accuracy	time
0	57.21%	17106.87s
1	**57.29%**	19272.48s
2	53.93%	**452.99s**
3	54.78%	9371.59s
4	44.46%	30315.99s

1.1.2 Experiment on sample size of training set

1.1.3 Experiment on penalty factor C

1.1.4 Experiment on kernel of SVM

1.1.5 Experiment on dimension

1.2 MLP

1.2.1 Experiment on data preprocessing

1.2.2 Experiment on sample size of training set

1.2.3 Experiment on network structure

1.2.4 Experiment on activation function

1.2.5 Experiment on regularization factor α

1.2.6 Experiment on dimension

1.3 SVM on Big Datasets

2 Causal discovery algorithms

2.1 LiNGAM

2.2 Experiment

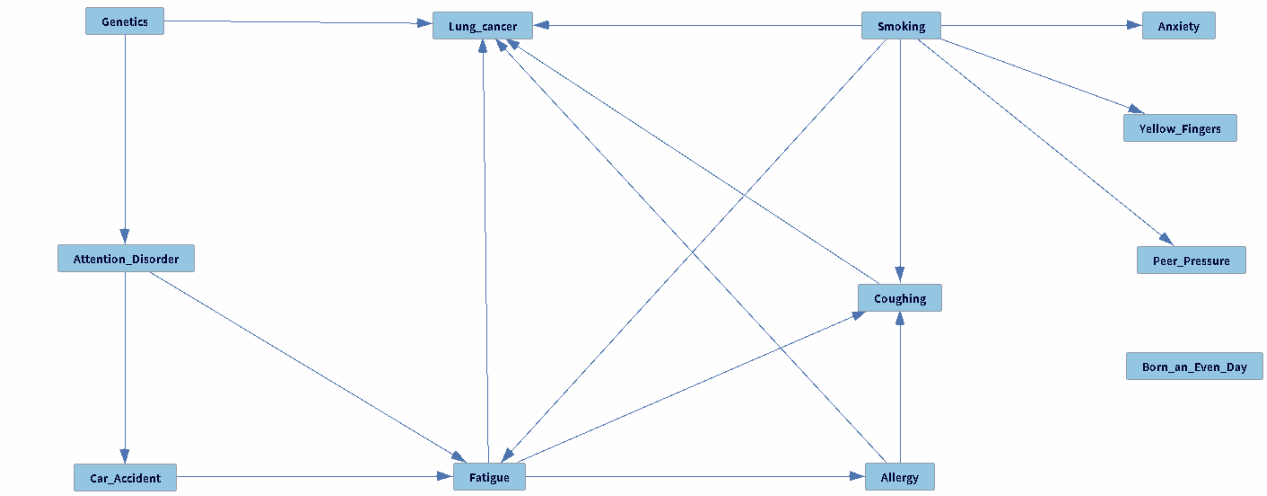

Figure 1: LiNGAM Result

*GitHub repo: https://github.com/DeanAlkene/CS420-MachineLearning/tree/master/A3