模式识别作业4

张蔚桐 2015011493 自55 2017 年 4 月 1 日

1 神经网络的训练

如下图所示是隐节点分别为5,10,20,40,100的情况的测试集的confusionTest matrix

从各个图中可以看出,随着隐节点的增多,混淆矩阵指示的准确性提高,训练误差逐渐降低,分析错误率同时降低,收敛速度也得到了提高。但是,当隐节点增多到一定程度后,出现了过拟合的情况,上面的各个评估指数的变化逐渐不明显

Confusion Matrix												
	1	908	0	11	20	1	36	20	0	8	6	39.9%
	1	9.1%	0.0%	0.1%	0.2%	0.0%	0.4%	0.2%	0.0%	0.1%	0.1%	10.1%
	2	0	1095	25	2	11	2	7	19	27	4	91.9%
	2	0.0%	10.9%	0.3%	0.0%	0.1%	0.0%	0.1%	0.2%	0.3%	0.0%	8.1%
	3	4	8	872	54	3	6	7	10	6	0	89.9%
	J	0.0%	0.1%	8.7%	0.5%	0.0%	0.1%	0.1%	0.1%	0.1%	0.0%	10.1%
	4	19	2	59	822	1	74	0	31	44		77.3%
		0.2%	0.0%	0.6%		0.0%				0.4%		22.7%
SS	5	3	1	7	2	876	14	8	18	11		B7.6%
Class	J	0.0%	0.0%	0.1%	A STREET OF THE PARTY OF					0.1%		12.4%
5	6	30	1	1	30	8	655	29	0	79	17	77.1%
Output		0.3%	0.0%	0.0%			6.6%			0.8%		22.9%
Ĕ	7	11	3	22	2	17	21	883	0	24		89.8%
U		0.1%			Name of the last o			10000	International Company	0.2%		10.2%
	8	5	2	6	27	11	6	0	891	4		90.3%
		0.1%			Comment of the Commen		100000	0.0%	450	0.0%	10000000	The second second
	9	0	22	26	39	2	71	4	0	752		B1.0%
		0.0%		The state of the s				0.0%				19.0%
	10	0	1	3	12	52	7	0	59	19		84.9%
		THE PERSON NAMED IN	A PROPERTY AND ADDRESS OF THE PARTY AND ADDRES	STATE OF THE PARTY	The second second	STATE OF THE PARTY	Name and Address of the Owner, where	The Real Property lies, the Parket of the Pa	STATE OF THE PARTY	The second second		15.1%
		Charles III	a Stanton and Advantage	MATERIA CALCULATA	81.4%		A PARTICIPATION OF THE PARTY OF		W 1000000000000000000000000000000000000		- Charles Control	86.2 %
		7.3%	3.5%	15.5%	18.6%	10.8%	26.6%	7.8%	13.3%	22.8%	14.5%	13.8%
		1	2	3	4	5	6	7	8	9	10	
						Targ	get C	lass				

图 1: 5隐节点测试集混淆矩阵

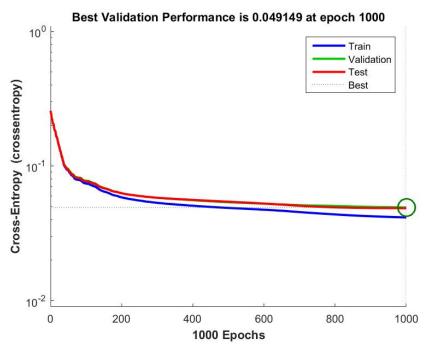


图 2: 5隐节点训练误差曲线

		Confusion Matrix											
	1	951	0	11	3	2	7	14	4	5	16	93.9%	
	ŀ	9.5%	0.0%	0.1%	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.2%	6.1%	
	2	0	1099	6	4	1	5	3	13	9	4	96.1%	
	2	0.0%	11.0%	0.1%	0.0%	0.0%	0.1%	0.0%	0.1%	0.1%	0.0%	3.9%	
	3	1	5	940	18	5	5	8	23	7		92.9%	
	J	0.0%	0.1%	9.4%	0.2%	0.1%	0.1%	0.1%	0.2%	0.1%	0.0%	7.1%	
	4	3	9	14	899	2	45	1	9	35	and the second	B7.9%	
	- 1	0.0%	0.1%	0.1%	9.0%	The same of the sa	0.4%	0.0%	0.1%	0.4%	0.1%	12.1%	
SS	5	3	0	10	2	929	11	10	5	12		91.6%	
Output Class	Ü		0.0%		0.0%	1000		0.1%	The second second		The second second second	8.4%	
5	6	11	3	5	32	0	760	14	1	24		B8.4%	
nd		-	0.0%		0.3%			0.1%				11.6%	
Ħ	7	4	4	10	2	8	18	901	0	16		93.5%	
O			0.0%	100	0.0%				0.0%			6.5%	
	8	1	1	15	20	3	7	1	949	18		90.7%	
	0.55		100			7.00	100000000000000000000000000000000000000	0.0%	9779	The second second	0.3%	Name and State of	
	9	3	14	18	23	4	27	6	2	838	and the same	B8.6%	
	3	-	0.1%					0.1%		S		11.4%	
	10	3	0	3	7	28	7	0	22	10		91.8%	
		DESCRIPTION OF THE PERSON	STREET, SQUARE, SQUARE	0.0%	CANCEL DE COMMUNICO	SELECTION OF THE PERSONS ASSESSED.	ALC: UNKNOWN	THE REAL PROPERTY.	OCCUPATION AND PARTY OF	COLUMN TWO IS NOT THE OWNER.	NAME AND ADDRESS OF THE OWNER, WHEN	STATE OF THE PERSON NAMED IN	
												91.6%	
		3.0%	3.2%	8.9%	11.0%	5.4%	14.8%	5.9%	7.7%	14.0%	11.0%	8.4%	
		1	2	3	4	5	6	7	8	9	10		
Target Class													

图 3: 10隐节点测试集混淆矩阵

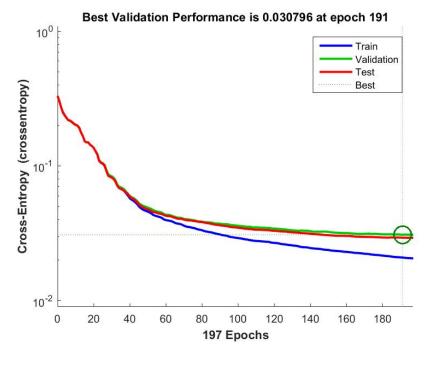


图 4: 10隐节点训练误差曲线

	Confusion Matrix												
	1	947	1	7	2	2	10	7	2	7	6	95.6%	
	1	9.5%	0.0%	0.1%	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.1%	4.4%	
	2	0	1105	4	2	1	1	3	9	3	4	97.6%	
	2	0.0%	11.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	2.4%	
	3	0	4	954	18	5	4	11	17	13	1	92.9%	
	0	0.0%	0.0%	9.5%	0.2%	0.1%	0.0%	0.1%	0.2%	0.1%	0.0%	7.1%	
	4	2	2	10	925	1	35	0	5	19	15	91.2%	
	- 7	0.0%	0.0%	0.1%	9.3%	0.0%	0.4%	0.0%	0.1%	0.2%	0.1%	8.8%	
SS	5	0	0	7	2	916	7	5	7	13		94.0%	
<u>ia</u>	J	0.0%				The second second second		0.1%					
Output Class	6	9	1	4	30	0	783	14	1	18	7	90.3%	
nd	Ĭ	0.1%		0.0%				0.1%					
Ħ	7	13	4	9	1	11	10	908	0	9		93.9%	
U		0.1%	Service Control	0.1%		-	0.1%		The second second second		0.0%		
	8	4	3	9	11	3	8	4	959	8		94.0%	
	Ü	0.0%				7.00		0.0%	70.00	The second second	-	Distance of the last of	
	9	4	15	23	11	6	24	4	4	876	and the same	B9.6%	
		0.0%		Contract Con						C		10.4%	
	10	1	0	5	8	37	10	2	24	8		90.8%	
	. •	District Control of the last	STREET, SQUARE, SQUARE	0.1%	OR RESIDENCE AND ADDRESS.	STATES OF TAXABLE	Name and Address of the Owner, where the Owner, which the Owner, where the Owner, where the Owner, which the	THE PERSON NAMED IN	STREET, SQUARE,	OR OTHER DESIGNATION.	The second named in	THE OWNER OF TAXABLE PARTY.	
		Charles Co. Alex			- Carlotte	A CONTRACTOR OF THE PARTY OF TH	Mark Control	March Control of the	and bed to the country of the	- Contract (1)		93.1%	
		3.4%	2.6%	7.6%	8.4%	6.7%	12.2%	5.2%	6.7%	10.1%	7.3%	6.9%	
		1	2	3	4	5	6	7	8	9	10		
Target Class													

图 5: 20隐节点测试集混淆矩阵

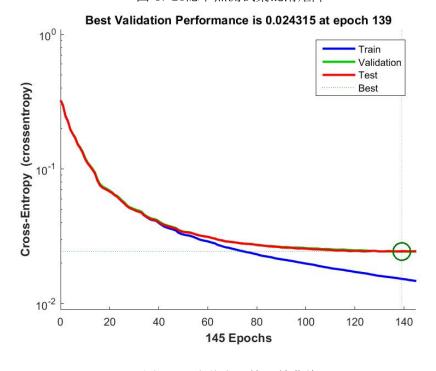


图 6: 20隐节点训练误差曲线

	Confusion Matrix											
	1	958	0	9	3	4	9	12	1	6	2	95.4%
	1	9.6%	0.0%	0.1%	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.0%	4.6%
	2	0	1118	2	0	0	1	4	7	2	7	98.0%
	2	0.0%	11.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%	2.0%
	3	5	3	977	8	1	1	5	20	8	3	94.8%
	3	0.1%	0.0%	9.8%	0.1%	0.0%	0.0%	0.1%	0.2%	0.1%	0.0%	5.2%
	4	1	2	9	958	0	13	0	10	14	8	94.4%
	4	0.0%	0.0%	0.1%	9.6%	0.0%	0.1%	0.0%	0.1%	0.1%	0.1%	5.6%
S	5	0	0	3	3	939	2	9	4	9	23	94.7%
Output Class	J	0.0%	0.0%	0.0%	0.0%	9.4%	0.0%	0.1%	0.0%	0.1%	0.2%	5.3%
5	6	6	2	0	11	0	837	16	0	16	6	93.6%
bū	U	0.1%	0.0%	0.0%	0.1%	0.0%	8.4%	0.2%	0.0%	0.2%	0.1%	6.4%
Ħ	7	5	4	7	1	8	13	905	0	7	2	95.1%
0	- 1	0.1%	0.0%	0.1%	0.0%	0.1%	0.1%	9.0%	0.0%	0.1%	0.0%	4.9%
	8	1	2	7	13	4	3	3	965	7	14	94.7%
	0	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	9.7%	0.1%	0.1%	5.3%
	9	0	4	15	8	2	7	4	6	892	6	94.5%
	9	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.0%	0.1%	8.9%	0.1%	5.5%
	10	4	0	3	5	24	6	0	15	13	938	93.1%
	10	0.0%	0.0%	0.0%	0.1%	0.2%	0.1%	0.0%	0.1%	0.1%	9.4%	6.9%
		97.8%	98.5%	94.7%	94.9%	95.6%	93.8%	94.5%	93.9%	91.6%	93.0%	94.9%
		2.2%	1.5%	5.3%	5.1%	4.4%	6.2%	5.5%	6.1%	8.4%	7.0%	5.1%
		1	2	3	4	5	6	7	8	9	10	
Target Class												

图 7: 40隐节点测试集混淆矩阵

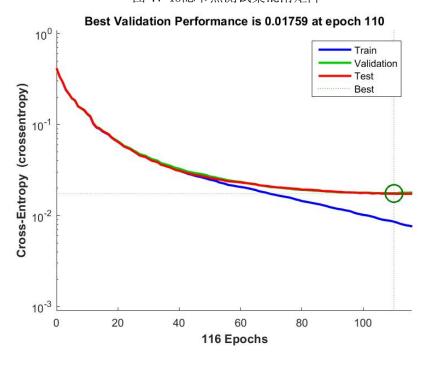


图 8: 40隐节点训练误差曲线

		Confusion Matrix											
	1	965	0	6	0	3	3	5	0	2	3	97.8%	
	į	9.7%	0.0%	0.1%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	2.2%	
	2	0	1122	2	0	1	1	3	3	2	6	98.4%	
	2	0.0%	11.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	1.6%	
	3	1	3	1000	9	4	0	1	11	3	2	96.7%	
	3	0.0%	0.0%	10.0%	0.1%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	3.3%	
	4	0	2	3	979	0	15	0	4	8	9	96.0%	
	4	0.0%	0.0%	0.0%	9.8%	0.0%	0.1%	0.0%	0.0%	0.1%	0.1%	4.0%	
SS	5	1	0	3	0	954	3	4	4	5	13	96.7%	
Output Class	5	0.0%	0.0%	0.0%	0.0%	9.5%	0.0%	0.0%	0.0%	0.1%	0.1%	3.3%	
Ç	6	3	0	2	7	0	849	5	0	11	6	96.1%	
nd	U	0.0%	0.0%	0.0%	0.1%	0.0%	8.5%	0.1%	0.0%	0.1%	0.1%	3.9%	
Ħ	7	5	0	5	0	3	7	935	0	1	3	97.5%	
0	-	0.1%	0.0%	0.1%	0.0%	0.0%	0.1%	9.3%	0.0%	0.0%	0.0%	2.5%	
	8	1	2	6	7	1	2	3	989	4	12	96.3%	
	U	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	9.9%	0.0%	0.1%	3.7%	
	9	4	6	5	2	2	9	2	1	933	- and the land	96.3%	
	3	0.0%	0.1%	0.1%	0.0%	0.0%	0.1%	0.0%	0.0%	9.3%	0.1%	3.7%	
	10	0	0	0	6	14	3	0	16	5	7.00	95.6%	
	10	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.2%	0.1%	9.5%	4.4%	
		98.5%	98.9%	96.9%	96.9%	97.1%	95.2%	97.6%	96.2%	95.8%	94.2%	96.8%	
		1.5%	1.1%	3.1%	3.1%	2.9%	4.8%	2.4%	3.8%	4.2%	5.8%	3.2%	
		1	2	3	4	5	6	7	8	9	10		
Target Class													

图 9: 100隐节点测试集混淆矩阵

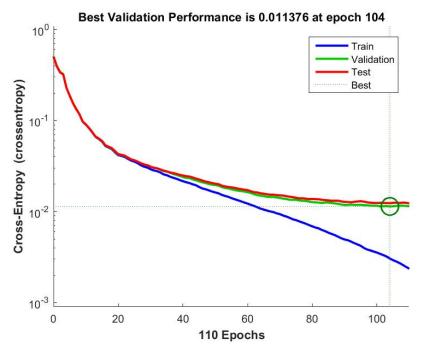


图 10: 100隐节点训练误差曲线

0128456728

图 11: 训练集中的图像

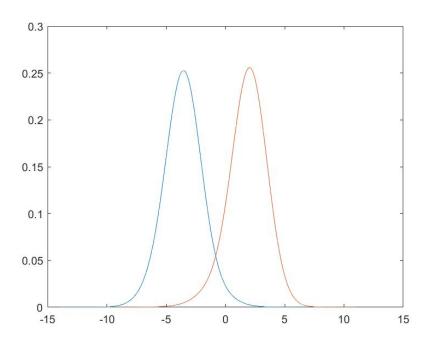


图 12: 概率密度图像

2 神经网络准确性和线性方法的比较

从图9可得神经网络中出现误差较大的识别组为'7'和'9'两组,正确率达到98.57

2.1 Fisher方法

如图12所示是Fisher得到的概率密度图像,经过学习可以得到正确率为95.73%

2.2 Logistic方法

Logistic方法学习迭代1000次可以得到正确率96.02%

2.3 总结

可以看出神经网络在准确度上明显优于Logistic方法,这可以理解为神经网络是一种复杂化了的Logistic方法。而Fisher方法相对于Logistic方法的准确度相对较差,但并不明显