ЧАСТЬ ВТОРАЯ

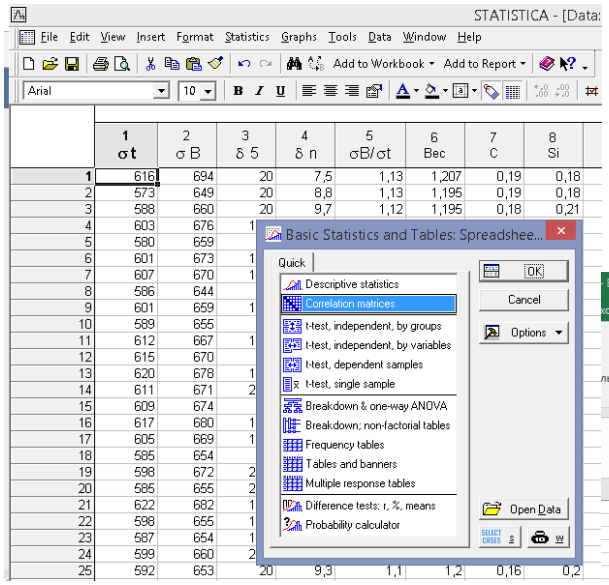
ОБРАБОТКА МАССИВА ДАННЫХ (сортовая прокатка)

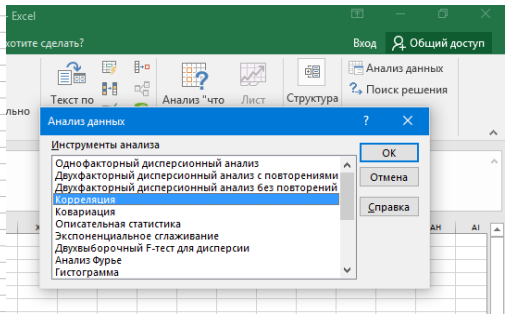
Попробуем рассмотреть вопрос сравнения адекватности двух видов моделей для построения регрессионной зависимости: классическая линейная (или нелинейная) зависимость и математическая модель на основе нейронной сети (можно рассмотреть несколько сетей). Сравнение проведём по методу наименьших квадратов.

8. Постройте регрессионное уравнение типа:

σ_T=f(C,Si,Mn…Вес) или σ_B=f(C,Si,Mn…Вес) (любое на Ваше усмотрение)

Построим матрицу корреляции. Для этого можно использовать среду Microsoft Excel или программу Statistica.



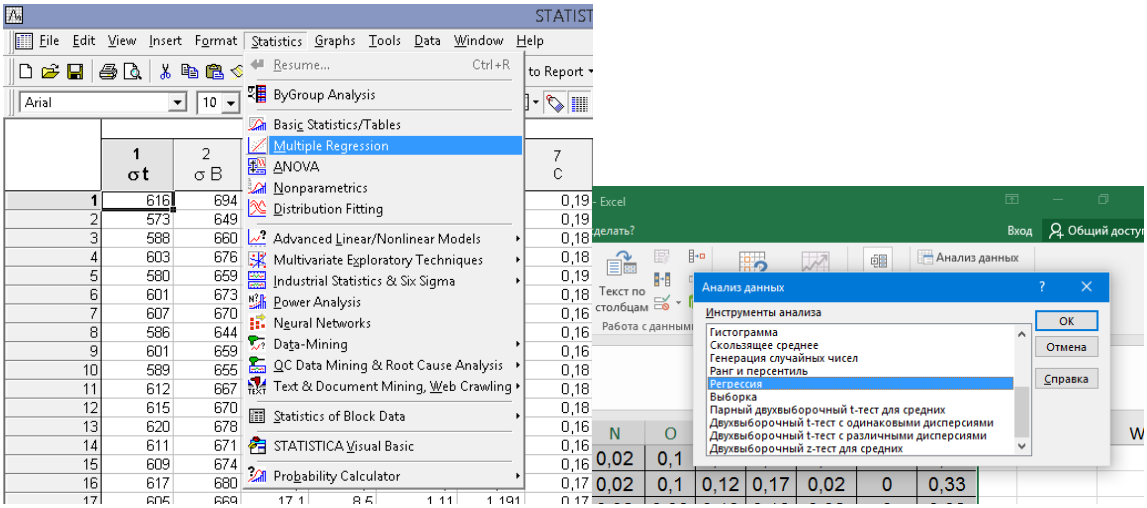


Statistica Excel

Оценив влияние факторов (химия, вес одного метра, расчетный диаметр) на отклик, необходимо определить факторы, которые будут входить в состав регрессионного уравнения.

© Левандовский С.А., 2018
ФГБОУ ВО МГТУ им. Г.И. Носова,
Кафедра технологий обработки материалов

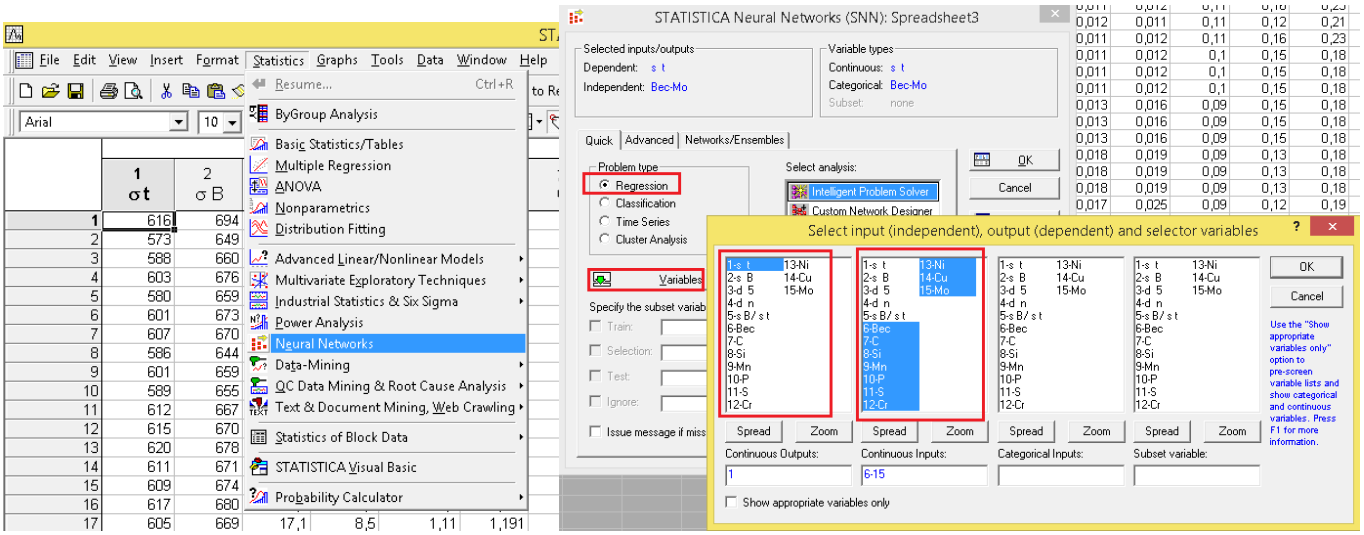
Далее пользуйтесь инструментами «Регрессия» в MS Excel или “Multiple Regression” в Statistica.



Statistica Excel

По полученным уравнениям регрессии посчитайте значения отклика. Например, для σ_T=f(C, Si, Mn … Bec) составьте таблицу из двух столбцов: σ_T (из выборки) и σ_T (исходя из расчётной, созданной формулы).

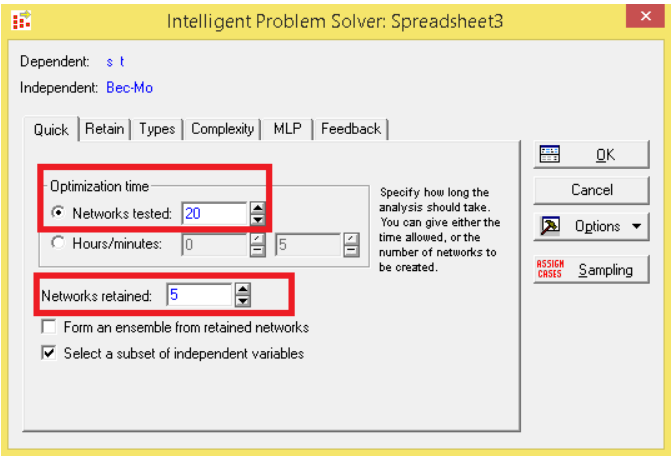
9. Начнём работать с нейронными сетями в программе Statistica (выданный преподавателем теоретический материал необходимо изучить заранее и после чего приступать к работе). Начнём конструировать простейшую нейросеть в программе Statistica:



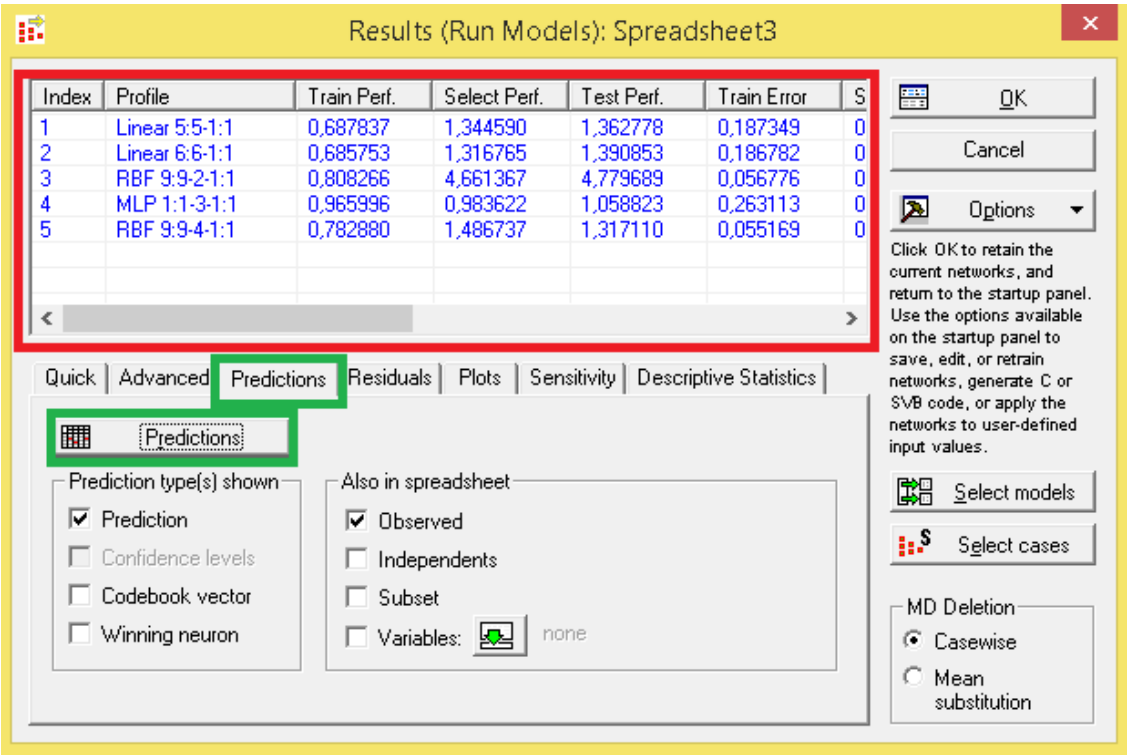
Меню «Statistics-Neural Networks». Далее убедимся, что «Problem Type» у нас установлен как «Regression».

Указываем входящие и выходящие переменные по кнопке «Variables». На примере указаны: Выходящая переменная σ_T, а в качестве входящих «Bec», «C», «Si», «Mn», «P», «S», «Cr», «Ni», «Cu», «Mo».

© Левандовский С.А., 2018
ФГБОУ ВО МГТУ им. Г.И. Носова,
Кафедра технологий обработки материалов



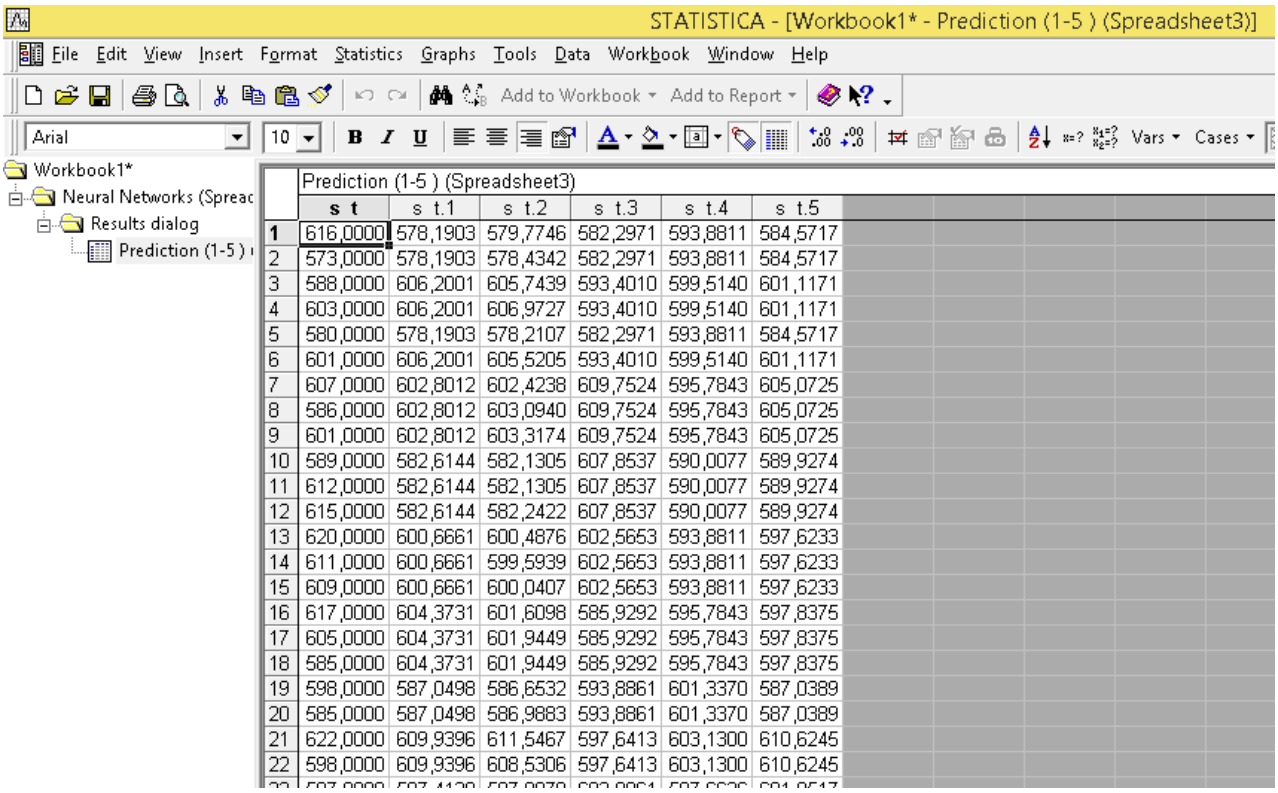
Выбираем число анализируемых сетей (на примере – 20). Выбираем число сетей, которые останутся в качестве результата (на примере - 5). Запускаем расчёт нажав «ОК» и ждём результата.



Результат представлен в красном блоке характеристиками нейросетей. Зелёными блоками обозначен доступ к результатам (предсказанным) значениям по каждой нейросети.

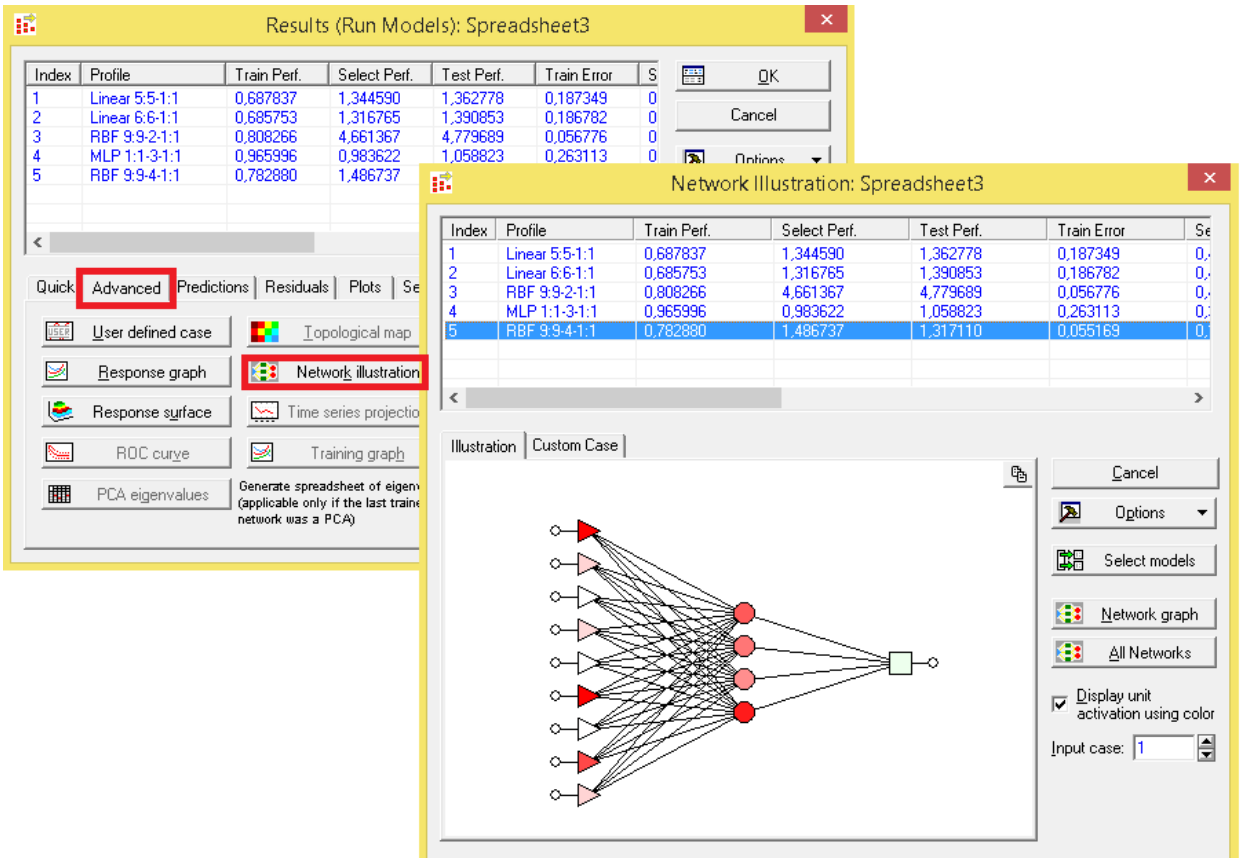
© Левандовский С.А., 2018
ФГБОУ ВО МГТУ им. Г.И. Носова,
Кафедра технологий обработки материалов

Посмотрим теперь на результаты работы нейросетей в виде предсказаний σ_T. В первом столбце приведены значения σ_T из выборки – остальные 5 это результаты прогнозирования на основе нейронных сетей.



	s t	s t.1	s t.2	s t.3	s t.4	s t.5
1	616,0000	578,1903	579,7746	582,2971	593,8811	584,5717
2	573,0000	578,1903	578,4342	582,2971	593,8811	584,5717
3	588,0000	606,2001	605,7439	593,4010	599,5140	601,1171
4	603,0000	606,2001	606,9727	593,4010	599,5140	601,1171
5	580,0000	578,1903	578,2107	582,2971	593,8811	584,5717
6	601,0000	606,2001	605,5205	593,4010	599,5140	601,1171
7	607,0000	602,8012	602,4238	609,7524	595,7843	605,0725
8	586,0000	602,8012	603,0940	609,7524	595,7843	605,0725
9	601,0000	602,8012	603,3174	609,7524	595,7843	605,0725
10	589,0000	582,6144	582,1305	607,8537	590,0077	589,9274
11	612,0000	582,6144	582,1305	607,8537	590,0077	589,9274
12	615,0000	582,6144	582,2422	607,8537	590,0077	589,9274
13	620,0000	600,6661	600,4876	602,5653	593,8811	597,6233
14	611,0000	600,6661	599,5939	602,5653	593,8811	597,6233
15	609,0000	600,6661	600,0407	602,5653	593,8811	597,6233
16	617,0000	604,3731	601,6098	585,9292	595,7843	597,8375
17	605,0000	604,3731	601,9449	585,9292	595,7843	597,8375
18	585,0000	604,3731	601,9449	585,9292	595,7843	597,8375
19	598,0000	587,0498	586,6532	593,8861	601,3370	587,0389
20	585,0000	587,0498	586,9883	593,8861	601,3370	587,0389
21	622,0000	609,9396	611,5467	597,6413	603,1300	610,6245
22	598,0000	609,9396	608,5306	597,6413	603,1300	610,6245

Так же можно ознакомиться с тем как может выглядеть конструкция нейросети.



© Левандовский С.А., 2018
ФГБОУ ВО МГТУ им. Г.И. Носова,
Кафедра технологий обработки материалов

10. Далее возьмём результаты предсказаний нейросетей и сведём их в одну таблицу с результатами расчётов линейной регрессии (пример на рисунке) и туда же внесём данные из выборки.

	A	B	C	D	E	F	G	H
1		Выборка	Н.С. 1	Н.С. 2	Н.С. 3	Н.С. 4	Н.С. 5	Лин Регр
2		616,0000	578,1903	579,7746	582,2971	593,8811	584,5717	593,6902
3		573,0000	578,1903	578,4342	582,2971	593,8811	584,5717	593,6902
4		588,0000	606,2001	605,7439	593,4010	599,5140	601,1171	595,8909
5		603,0000	606,2001	606,9727	593,4010	599,5140	601,1171	595,8909
6		580,0000	578,1903	578,2107	582,2971	593,8811	584,5717	593,6902
7		601,0000	606,2001	605,5205	593,4010	599,5140	601,1171	595,8909
8		607,0000	602,8012	602,4238	609,7524	595,7843	605,0725	601,0052
9		586,0000	602,8012	603,0940	609,7524	595,7843	605,0725	601,0052
10		601,0000	602,8012	603,3174	609,7524	595,7843	605,0725	601,0052
11		589,0000	582,6144	582,1305	607,8537	590,0077	589,9274	583,3211
12		612,0000	582,6144	582,1305	607,8537	590,0077	589,9274	583,3211
13		615,0000	582,6144	582,2422	607,8537	590,0077	589,9274	583,3211

Добавим столбцы для вычисления отклонений прогнозирования каждой из моделей от значений выборки и посчитаем сумму квадратов этих отклонений.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Выборка	Н.С. 1	откл	Н.С. 2	откл	Н.С. 3	откл	Н.С. 4	откл	Н.С. 5	откл	Лин Регр	откл
2		616,0000	578,1903	-37,8097	579,7746	-36,2254	582,2971	-33,7029	593,8811	-22,1189	584,5717	-31,4283	593,6902	-22,3098
3		573,0000	578,1903	5,1903	578,4342	5,4342	582,2971	9,2971	593,8811	20,8811	584,5717	11,5717	593,6902	20,6902
4		588,0000	606,2001	18,2001	605,7439	17,7439	593,4010	5,4010	599,5140	11,5140	601,1171	13,1171	595,8909	7,8909
5		603,0000	606,2001	3,2001	606,9727	3,9727	593,4010	-9,5990	599,5140	-3,4860	601,1171	-1,8829	595,8909	-7,1091
6		580,0000	578,1903	-1,8097	578,2107	-1,7893	582,2971	2,2971	593,8811	13,8811	584,5717	4,5717	593,6902	13,6902
7		601,0000	606,2001	5,2001	605,5205	4,5205	593,4010	-7,5990	599,5140	-1,4860	601,1171	0,1171	595,8909	-5,1091
8		607,0000	602,8012	-4,1988	602,4238	-4,5762	609,7524	2,7524	595,7843	-11,2157	605,0725	-1,9275	601,0052	-5,9948
9		586,0000	602,8012	16,8012	603,0940	17,0940	609,7524	23,7524	595,7843	9,7843	605,0725	19,0725	601,0052	15,0052
10		601,0000	602,8012	1,8012	603,3174	2,3174	609,7524	8,7524	595,7843	-5,2157	605,0725	4,0725	601,0052	0,0052
11		589,0000	582,6144	-6,3856	582,1305	-6,8695	607,8537	18,8537	590,0077	1,0077	589,9274	0,9274	583,3211	-5,6789
12		612,0000	582,6144	-29,3856	582,1305	-29,8695	607,8537	-4,1463	590,0077	-21,9923	589,9274	-22,0726	583,3211	-28,6789
13		615,0000	582,6144	-32,3856	582,2422	-32,7578	607,8537	-7,1463	590,0077	-24,9923	589,9274	-25,0726	583,3211	-31,6789
14		620,0000	600,6661	-19,3339	600,4876	-19,5124	602,5653	-17,4347	593,8811	-26,1189	597,6233	-22,3767	608,9545	-11,0455
15		611,0000	600,6661	-10,3339	599,5939	-11,4061	602,5653	-8,4347	593,8811	-17,1189	597,6233	-13,3767	608,9545	-2,0455
16		609,0000	600,6661	-8,3339	600,0407	-8,9593	602,5653	-6,4347	593,8811	-15,1189	597,6233	-11,3767	608,9545	-0,0455
17		617,0000	604,3731	-12,6269	601,6098	-15,3902	585,9292	-31,0708	595,7843	-21,2157	597,8375	-19,1625	608,1708	-8,8292
18		605,0000	604,3731	-0,6269	601,9449	-3,0551	585,9292	-19,0708	595,7843	-9,2157	597,8375	-7,1625	608,1708	3,1708
19		585,0000	604,3731	19,3731	601,9449	16,9449	585,9292	0,9292	595,7843	10,7843	597,8375	12,8375	608,1708	23,1708

598,0000	587,0498	-10,9502	586,6532	-11,3468	593,8861	-4,1139	601,3370	3,3370	587,0389	-10,9611	588,8571	-9,1429
585,0000	587,0498	2,0498	586,9883	1,9883	593,8861	8,8861	601,3370	16,3370	587,0389	2,0389	588,8571	3,8571
622,0000	609,9396	-12,0604	611,5467	-10,4533	597,6413	-24,3587	603,1300	-18,8700	610,6245	-11,3755	588,6997	-33,3003
598,0000	609,9396	11,9396	608,5306	10,5306	597,6413	-0,3587	603,1300	5,1300	610,6245	12,6245	588,6997	-9,3003
587,0000	597,4120	10,4120	597,0879	10,0879	602,8861	15,8861	597,6626	10,6626	601,9517	14,9517	602,8426	15,8426
599,0000	597,4120	-1,5880	597,7581	-1,2419	602,8861	3,8861	597,6626	-1,3374	601,9517	2,9517	602,8426	3,8426
592,0000	597,4120	5,4120	597,8698	5,8698	602,8861	10,8861	597,6626	5,6626	601,9517	9,9517	602,8426	10,8426
583,0000	598,0697	15,0697	598,6047	15,6047	592,2553	9,2553	591,9549	8,9549	595,4626	12,4626	599,4849	16,4849
568,0000	598,0697	30,0697	598,4930	30,4930	592,2553	24,2553	591,9549	23,9549	595,4626	27,4626	599,4849	31,4849
575,0000	598,0697	23,0697	598,7164	23,7164	592,2553	17,2553	591,9549	16,9549	595,4626	20,4626	599,4849	24,4849
598,0000	604,9473	6,9473	605,2552	7,2552	399,8010	-198,1990	593,8811	-4,1189	562,5854	-35,4146	596,0516	-1,9484
599,0000	604,9473	5,9473	605,8138	6,8138	399,8010	-199,1990	593,8811	-5,1189	562,5854	-36,4146	596,0516	-2,9484
		7469,24		7476,961		85235,32		6270,703		8864,123		7674,854

Из расчётов видно, что наименьшая сумма квадратов отклонений у модели нейронной сети № 4. Следовательно её можно считать наиболее точной и эффективной (в качестве примера была взята выборка объёмом в 30 значений, в Вашем случаем её объём может быть существенно больше).

Проведите аналогичное исследование и сделайте необходимые выводы.

© Левандовский С.А., 2018
ФГБОУ ВО МГТУ им. Г.И. Носова,
Кафедра технологий обработки материалов