
利用SPSS进行神经网络分析过程及结果解读

作者：Doctor Jiang 来源：临床科研与meta分析

本文原地址：<https://www.iikx.com/news/statistics/7819.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

利用SPSS进行神经网络分析

教程。模拟人类实际神经网络的数学方法问世以来，人们已慢慢习惯了把这种人工神经网络直接称为神经网络。

神经网络在系统辨识、模式识别、智能控制

等领域有着广泛而吸引人的前景，特别在智能控制中，人们对神经网络的自学习功能尤其感兴趣，并且把神经网络这一重要特点看作是解决自动控制中控制器适应能力这个难题的关键钥匙之一。

。

本例通过几个自变量预测是否有高血压，2个分类变量，一个性别，一个吸烟;3个连续性变量。



5: MLP_PredictedValue 1.00

	性别	年龄	体重指数	腰臀比	血压水平	吸烟2
1	1.00	72.00	.24	.98	1.00	2.00
2	1.00	64.00	.25	.93	1.00	1.00
3	1.00	66.00	.21	.93	1.00	1.00
4	1.00	53.00	.25	.93	1.00	2.00
5	1.00	65.00	.22	1.05	1.00	1.00
6	1.00	69.00	.24	.92	1.00	1.00
7	1.00	64.00	.21	.91	1.00	1.00
8	1.00	68.00	.21	.90	1.00	1.00
9	1.00	53.00	.26	.85	1.00	1.00
10	1.00	76.00	.21	.84	1.00	2.00
11	1.00	55.00	.20	.98	1.00	1.00
12	1.00	62.00	.21	.79	1.00	1.00
13	1.00	62.00	.35	.96	1.00	2.00
14	1.00	73.00	.20	.85	1.00	1.00
15	1.00	53.00	.24	.89	1.00	1.00
16	1.00	66.00	.28	.92	1.00	2.00
17	1.00	48.00	.25	.90	1.00	2.00
18	1.00	71.00	.22	.86	1.00	1.00
19	1.00	71.00	.29	.97	1.00	2.00

临床科研与meta分析

需要生成一个分组变量，用于区分训练集以及验证集。我们这个样本70%用于训练。通过计算变量，生成分组变量。

计算变量

目标变量(T): = 数字表达式(E):

类型和标签(L)...

- 性别
- 年龄
- 体重指数
- 腰臀比
- 血压水平
- 吸烟2
- 分组
- 血压水平的预测值 [...]
- 血压水平的预测拟概...
- 血压水平的预测拟概...

函数组(G):

- 全部
- 算术
- CDF 与非中心 CDF
- 转换
- 当前日期时间
- 日期运算
- 日期创建

函数和特殊变量(E):

- PdF.Weibull
- Range
- Replace(3)
- Replace(4)
- Rnd(1)
- Rnd(2)
- Rnd(3)
- Rtrim(1)
- Rtrim(2)
- Rv.Bernoulli

RV.BERNOULLI(prob)。数值。返回具有指定概率参数 prob 的 Bernoulli 分布中的随机值。

如果(I)... (可选的个案选择条件)

确定 粘贴(P) 重置(R) 取消 帮助

临床科研与meta分析



5 : MLP_PredictedValue 1.00

	性别	年龄	体重指数	腰臀比	血压水平	吸烟2	分组
1	1.00	72.00	.24	.98	1.00	2.00	1.00
2	1.00	64.00	.25	.93	1.00	1.00	1.00
3	1.00	66.00	.21	.93	1.00	1.00	1.00
4	1.00	53.00	.25	.93	1.00	2.00	1.00
5	1.00	65.00	.22	1.05	1.00	1.00	1.00
6	1.00	69.00	.24	.92	1.00	1.00	1.00
7	1.00	64.00	.21	.91	1.00	1.00	1.00
8	1.00	68.00	.21	.90	1.00	1.00	1.00
9	1.00	53.00	.26	.85	1.00	1.00	1.00
10	1.00	76.00	.21	.84	1.00	2.00	1.00
11	1.00	55.00	.20	.98	1.00	1.00	1.00
12	1.00	62.00	.21	.79	1.00	1.00	.00
13	1.00	62.00	.35	.96	1.00	2.00	.00
14	1.00	73.00	.20	.85	1.00	1.00	1.00
15	1.00	53.00	.24	.89	1.00	1.00	1.00
16	1.00	66.00	.28	.92	1.00	2.00	.00
17	1.00	48.00	.25	.90	1.00	2.00	1.00
18	1.00	71.00	.22	.86	1.00	1.00	1.00
19	1.00	71.00	.29	.97	1.00	2.00	.00
20	1.00	76.00	.26	.94	1.00	1.00	1.00
21	1.00	66.00	.20	.81	1.00	2.00	1.00

参数设置

变量(V) 分区 体系结构 训练 输出 保存 导出 选项

变量(V):

- 血压水平的预测值 [MLP_PredictedValue]
- 血压水平的预测拟概率 = 1.00 [MLP_PseudoPr...]
- 血压水平的预测拟概率 = 2.00 [MLP_PseudoPr...]

因变量(D):

- 血压水平

因子(E):

- 性别
- 吸烟2

协变量(C):

- 年龄
- 体重指数
- 腰臀比

协变量重新标度(S):

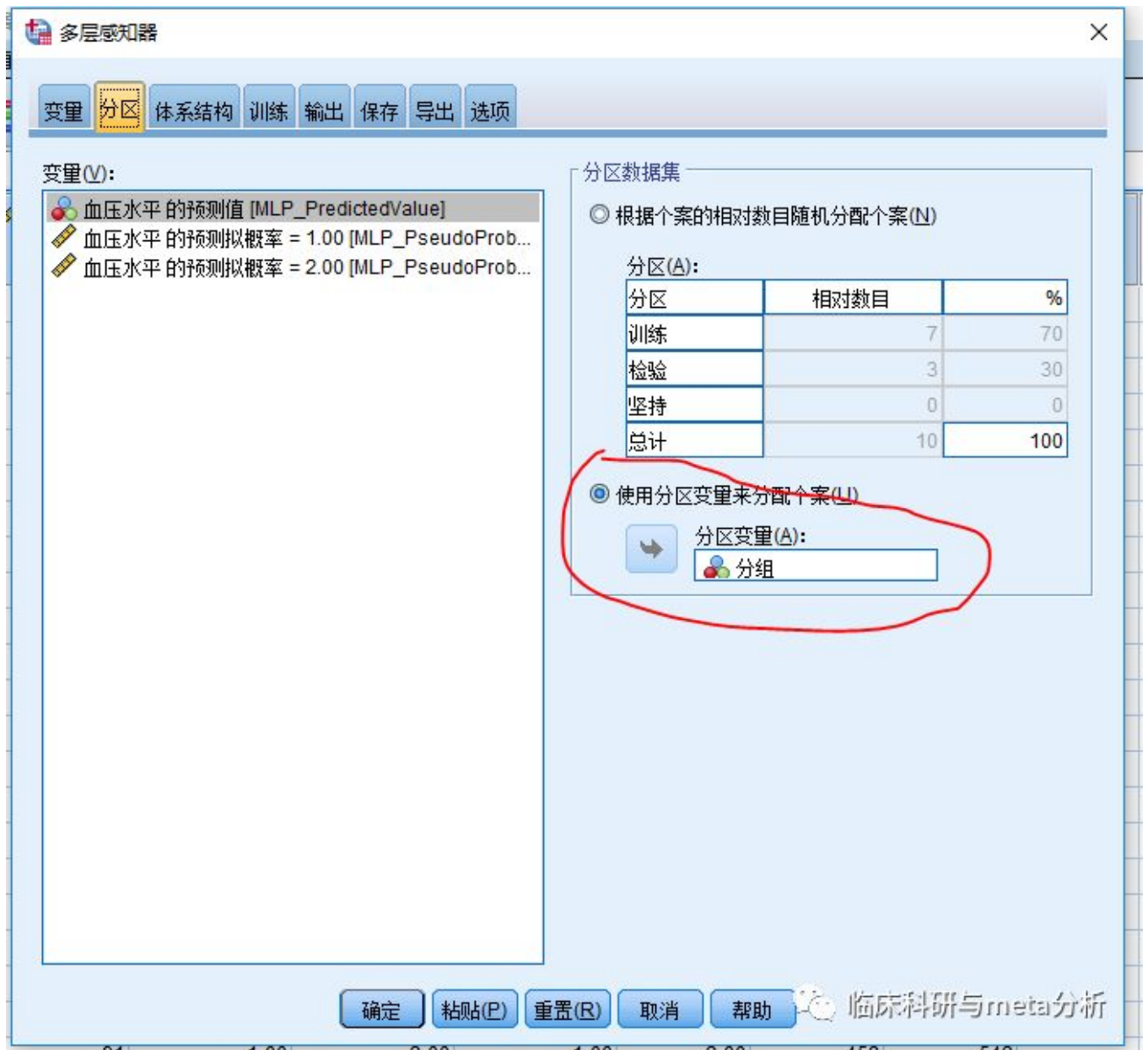
标准化

要更改变量的测量级别，请在“变量”列表中右键单击该变量。

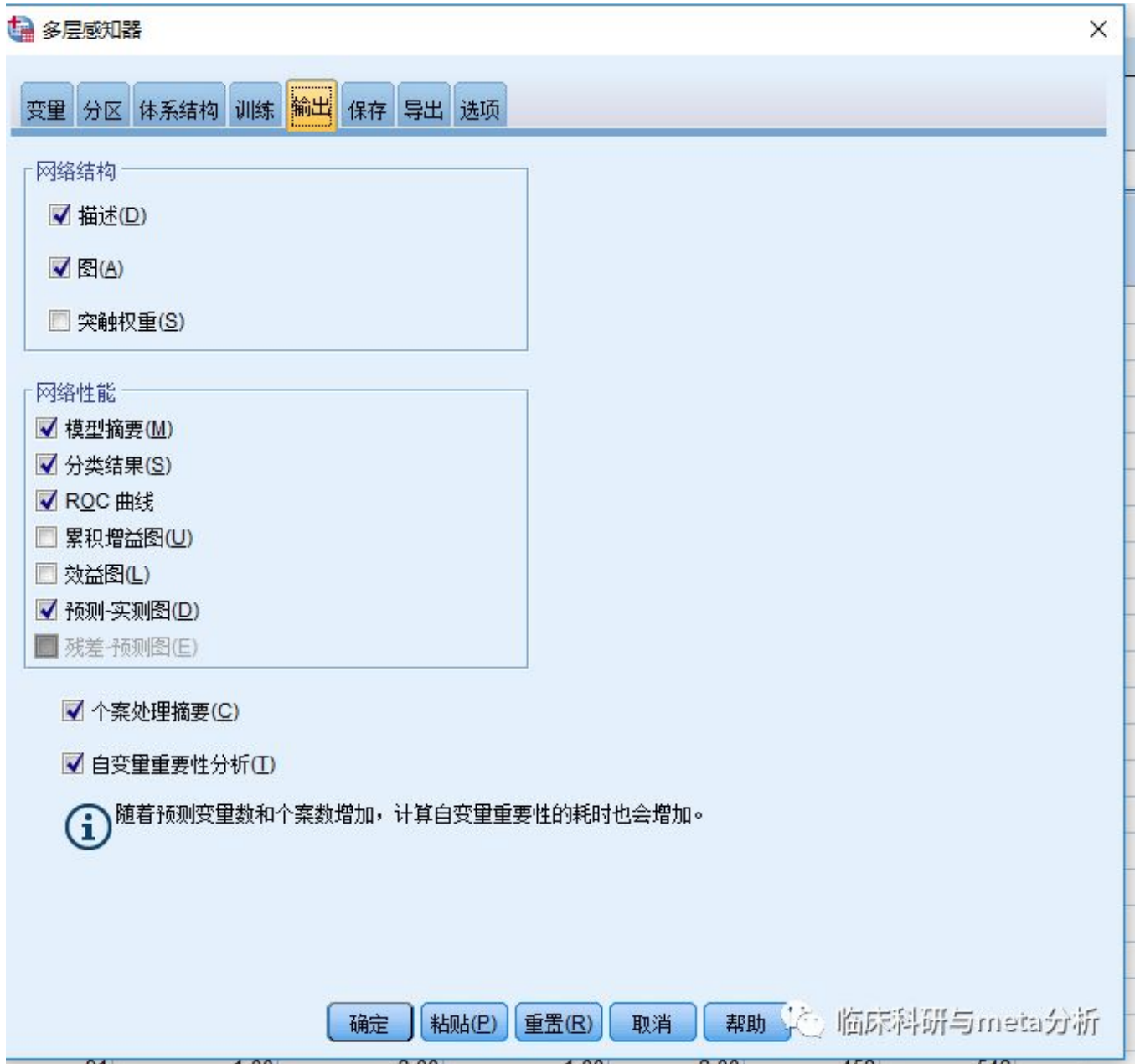
确定 粘贴(P) 重置(R) 取消 帮助

临床科研与meta分析

点击分区，选择生成的分组变量：



点击输出，选择如下参数：



点击保存，勾选预测值和预测概率



点击导出，可以保存相应模型，用于新数据的预测。

变量 分区 体系结构 训练 输出 **保存** 导出 选项

- 保存每个因变量的预测值或类别(S)
- 保存每个因变量的预测拟概率(E)

变量(V):

	预测值或类别	预测拟概率	
因变量	保存的变量的名称	保存的变量的根名称	要保存的类别
血压水平	MLP_PredictedValue	MLP_PseudoProbability	25

保存的变量的名称

自动生成唯一名称(A)

如果要在每次运行模型时都将一组新的保存变量添加到数据集，请选择此选项。

定制名称(C)

请为变量指定名称。如果选择此选项，那么每次运行模型时，都将替换所有具有相同名称或根名称的现有变量。

临床科研与meta分析

结果浏览：

个案处理摘要

	个案数	百分比	
样本	训练	1417	71.0%
	检验	578	29.0%
有效		1995	100.0%
排除		5	
总计		2000	

网络信息

输入层	因子	1	性别
		2	吸烟2
	协变量	1	年龄
		2	体重指数
		3	腰臀比
	单元数 ^a		7
	协变量的重新标度方法		标准化
隐藏层	隐藏层数		1
	隐藏层 1 中的单元数 ^a		4
	激活函数		双曲正切
输出层	因变量	1	血压水平
	单元数		2
	激活函数		Softmax
	误差函数		交叉熵

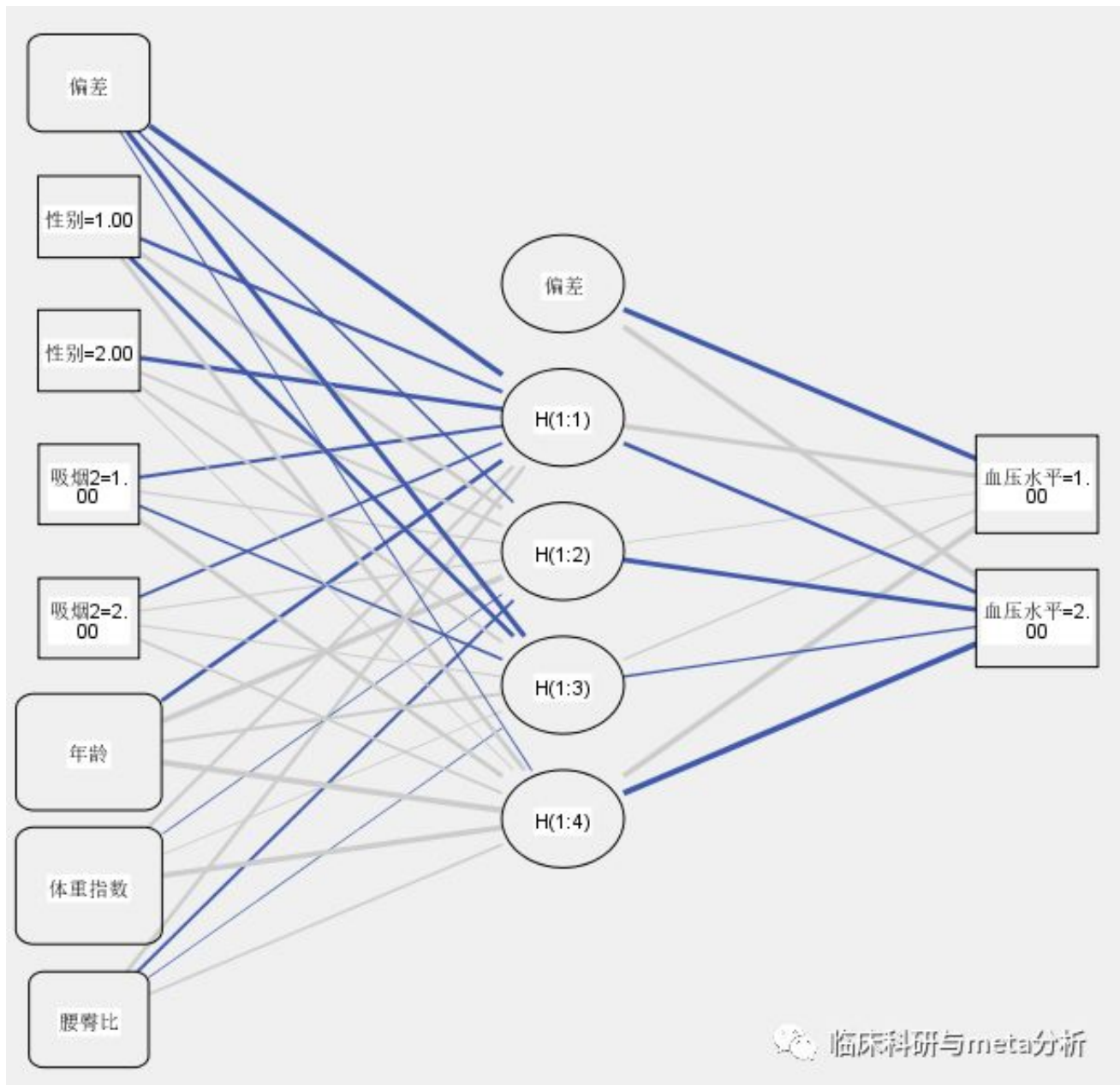
a. 排除偏差单元

临床科研与meta分析

首先是对训练集合检验集的描述

网络信息对神经网络的输入层，隐藏层以及输出层进行描述。

下图为程序运行后的神经网络图，线条的粗细代表了权重的大小。



模型摘要以及分类对具体的分类结果以及预测模型分类结果进行了比较

模型摘要

训练	交叉熵误差	607.856
	不正确预测百分比	22.3%
	使用的中止规则	误差在 1 个连续步骤中没有减小 ^a
	训练时间	0:00:00.20
检验	交叉熵误差	251.639
	不正确预测百分比	22.8%

因变量：血压水平

a. 误差计算基于检验样本。

分类

样本	实测	预测		正确百分比
		高血压	非高血压	
训练	高血压	79	246	24.3%
	非高血压	70	1022	93.6%
	总体百分比	10.5%	89.5%	77.7%
检验	高血压	40	97	29.2%
	非高血压	35	406	92.1%
	总体百分比	13.0%	87.0%	77.2%

因变量：血压水平

临床科研与meta分析

校准箱型图

ROC曲线下面积评估模型好坏

自变量对模型的重要性排行

更多 统计方法 请访问 <https://www.iikx.com/news/statistics/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](http://www.iikx.com)转发