1. 数据预处理

由于样本的不同特征的单位不同，变异程度不同。，为了消除量纲影响以及变量自身变异大小和数值大小的影响，便于不同单位或数量级的指标能够进行比较和加权，数据的标准化是按照指定区间大小进行缩放。在该学习算法中我们采用 [-1, 1]标准化对原始数据进行线性变换，使得结果落到[-1, 1]区间，以降低各个数据的数量级和差值。转换的公式为y = (ymax-ymin)\*(x-xmin)/(xmax-xmin) + ymin ，其中X是处理的数据，Xmin与Xmax是某行数据最小值和最大值，Ymin和Ymax是期望的某行的最小值与最大值。如果某行的数据全部相同，此时xmax=xmin，除数为0，则此时数据不变。在matlab中，我们通过mapminmax()函数将行最小值和最大值映射到[-1 1]，从而对整个矩阵进行[- 1, 1]标准化。

1. 预测模型的相关理论

2. 1 BP神经网络

神经网络技术在较多领域中得到了广泛的应用，其中BP神经网络的应用范围最广，BP神经网络全称为Back-Propagation Network，即反向传播网络，由Rumehart，McClelland于1985年提出，它是一种单向传播的多层前向网络，其利用误差反向传播方法进行训练。它具有结构简单、可塑性强以及对噪声数据的承受能力高等特点，因此在信息分类、模式识别、以及预测等领域得到了广泛的应用。

BP神经网络是一种具有三层或三层以上的神经网络，包含输入层、隐含层（可包含多层）和输出层，上下层之间实现全互连，而同一层的神经元之间无连接。其利用输出后的值与真实值比较来的误差来估计其直接前导向层的误差，再利用这个误差估计更前一层的误差，如果一层一层的反向传播下去，获得所有其他各层的误差估计，随着误差的逆向传播和修正地不断进行，网络对输入模式响应的正确率不断上升。

2.2 粒子群算法

粒子群优化算法（Particle Swarm Optimization，PSO算法）是一种进化计算技术，由Kennedy与Eberhart在1995年提出，源于对鸟群捕食的行为研究。该算法是一种基于迭代的优化工具，它模拟鸟群飞行觅食的行为，鸟之间通过集体的协作使群体达到最优目的。

在粒子群算法中，每个待优化问题的解抽象为一个微粒，并将其延伸到N维空间。所有粒子都有一个速度来决定搜索的方向和距离，以及一个待优化目标函数决定的适应度值（fitness）。每个粒子知道所经历过的最优位置pbest和当前位置，即粒子有自己的飞行经验，另外，还知道截至目前整个群体中发现的最优位置（gbest），即粒子知道其他同伴的经验，例子根据自己的飞行经验和其他同伴的最优经验来决定下一步的运动位置。

/×××××××××××××××

PSO算法以及基于改进粒子群的BP神经网络算法的具体的公式看论文：

I： <https://ieeexplore.ieee.org/document/5193700/>

II：基于改进粒子群和BP神经网络的网络舆情预测研究\_游丹丹.pdf

×××××××××××/

3基于改进粒子群的BP神经网络预测模型的建立

3.1 标准BP神经网络预测模型

BP神经网络具有自学习的能力，把许多不同的数据以及正常/故障状态输入BP神经网络，网络就会通过自学习，慢慢学会识别类似的数据。自学习能力对于预测而言具有重要意义，并且使其具有高速寻找近似最优解的能力。因此，可利用BP神经网络模型预测电网正常环境下各参数数据的未来数据值，然后将预测的未来数据值与实际值进行简单的比较就可以得出电网的实时状态是否正常。

BP神经网络采用三种拓扑结构，对于电网状态预测而言，本文采用了前三组数据来预测下一组数据，比如1～3组预测第4组，2～4组预测第5组，因此输入层结点数设置为3,输出层结点数设置为1。设置最大迭代次数为10，训练目标误差为1e-5，学习速率为0.05，动量因子为0.9。

关于隐含层结点数，首先根据经验公式（1）计算，



然后再进行2～12依次试错，发现隐含层结点数设为8时，均方误差较小，预测结果拟合程序较好，因此将BP神经网络的隐含层结点数设置为6.

3.2 权重改进的PSO优化BP神经网络

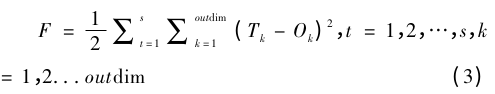
粒子群算法采用常数学习因子c1和c2以及惯性权重w，其中，较大的惯性权重w有利于跳出局部极小点从而提高算法的全局搜索能力，而较小的惯性权重则有利于对当前的搜索区域进行精确局部搜索从而增强算法的局部搜索能力和加快算法的收敛，因此，为了使算法在搜寻的早期具有较强的全局寻优能力，以及在搜索后期可以进行精确局部搜索，采用线性变化的权重来改善PSO算法后期易在全局最优解附近产生震荡和容易早熟的现象。

改进后的粒子群算法优化BP神经网络的关键步骤如下：

步骤1 随机化种群中各粒子的初始位置和速度，设置粒子的速度范围、位置范围、初始惯性权重w、学习因子c1和c2、种群规模、迭代次数以及BP神经网络的拓扑结构等。对于一个具有indim个输入层，hiddennum个隐含层，outdim个输出层的三层BP神经网络来说，粒子群中自变量的个数为：



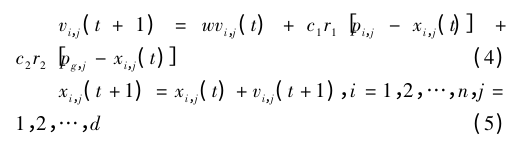
步骤2 根据适应度函数计算每个粒子的适应度值，确定粒子的个体历史最优值和所有粒子的全局最优值。对于每个粒子，将其适应度值与其历史极值作比较，将目前为止各粒子本身的最优位置存储在Pb中，对于每个粒子，将其最优适应度值与全局最优适应度值作比较，将最优Pb存储在Pg中。适应度函数如公式3：



其中s为迭代次数，outdim为输出单元个数，（Tk-Ok）为期望输出和网络预测输出的误差。

步骤3

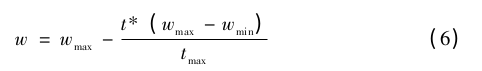
按以下公式更新每个粒子的速度 vi,j(t+1)和位移xi,j(t+1)，其中t为当前的迭代步数。



公式中，r1和r2表示0~1之间的随机数，C1、C2表示学习因子，n表示粒子的数量，d表示解的d维空间，即自变量个数。

步骤4

更新粒子群算法的惯性权重w，其随着算法的迭代次数的线性变化结果为：



其中wmax、wmin分别表示粒子群算法的惯性权重的最大值、最小值，tmax表示最大迭代次数，t表示当前的迭代步数。

步骤5

重复步骤2至步骤4,直至达到预设的运算精度和迭代步数，搜索停止，输出结果，此时得到优化后的链接权值和阈值，将其作为BP神经网络训练的链接权值和阈值

步骤6

进行BP神经网络训练，直至到达预设的最大训练次数或者满足最小误差需求

步骤7

将通过改进粒子群算法优化后的BP神经网络用于预测电网正常环境下各参数数据的未来数据值

改进粒子群算法优化BP神经网络的本质是将BP神经网络的训练误差作为粒子群算法的待优化目标函数，从而评价各粒子的适应度值，通过不断迭代、更新并寻找最优粒子所在的位置，最后得到BP神经网络的较优初始连接权值和阈值

相关参数设置：种群大小n设为20，两个学习因子c1和c2都设置为1.4955,粒子的范围即自变量的个数根据公式（2）设置为59,粒子的初始化位置在[-1, 1]之间，惯性权重范围在[0.90, 0.30]之间，最大迭代次数为10。

改进后的粒子群算法优化BP网络的算法流程图如图所示

（流程图见文件夹）

3.3 粒子群优化的BP神经网络预测模型的实例

我们以预测正常工作环境B的参数7的数据为例论证粒子群优化的BP神经网络预测模型的可靠性。参考留出法（hold-out），取出参数7的前35 个数据，按“3+1”的形式（前3个数据分别样本的3个特征，第四个数据为该样本的输出）设计出训练集、测试集与预测集，如下：

训练集：

测试集：

预测集：

预测结果

图 图 图 ...

神经网络结构图....

我们通过程序运行的预测集数据xxx，发现与实际测量值 xxx很接近 （图）

证明数据满足正常变化趋势，则电网环境正常。如果预测结果与实际测量值偏差过大，则做出电网环境故障的判断。