

清 华 大 学

# 研究生毕业论文开题报告

题目：大型高炉异常炉况早期检测和  
诊断研究

系    别：自动化系  
专    业：控制科学与工程  
姓    名：庞人铭  
指导教师：叶  昊  教授

2015 年 11 月 12 日



## 目录

第一章	选题背景 .....	1
第二章	文献综述 .....	4
2.1	故障诊断方法综述 .....	4
2.2	高炉异常炉况诊断方法 .....	5
第三章	技术难点 .....	6
3.1	热风炉换炉扰动 .....	6
3.2	工作点漂移 .....	6
3.3	故障样本稀少 .....	7
第四章	研究内容与研究方法 .....	8
4.1	深度学习 .....	8
4.1.1	auto-encoder .....	8
4.1.2	softmax .....	9
4.1.3	LSTM .....	9
4.2	基于深度学习的异常炉况诊断研究 .....	11
第五章	研究计划 .....	12
参考文献	.....	13



# 第一章 选题背景

我国每年的钢铁产量世界第一，高炉炼铁是钢铁工业的上游工序，其能耗占了生产链总能耗的 60%，生产成本也占了将近三分之一，因此，改进高炉炼铁的生产工艺、增大自动化技术在高炉炼铁过程中应用的比例，可以带来生产效益的明显提升<sup>[1]</sup>。此外，钢铁生产的规模和复杂度日益增大，而在高炉这样的高温高压情况下，保障安全生产就显得尤为重要。

高炉冶炼工艺如图 1 所示，高炉是一种逆流式圆筒型反应容器。首先将一定配比的铁矿石、焦炭等原料从炉喉装入高炉，热风炉从高炉底部鼓入热风以及一定的煤粉和富氧，形成大量高温煤气，在高温高压下与下降的铁矿原料发生还原反应，生成的铁水会定期从炉缸中排出，剩下的炉渣从渣口排放，煤气从炉顶导出。

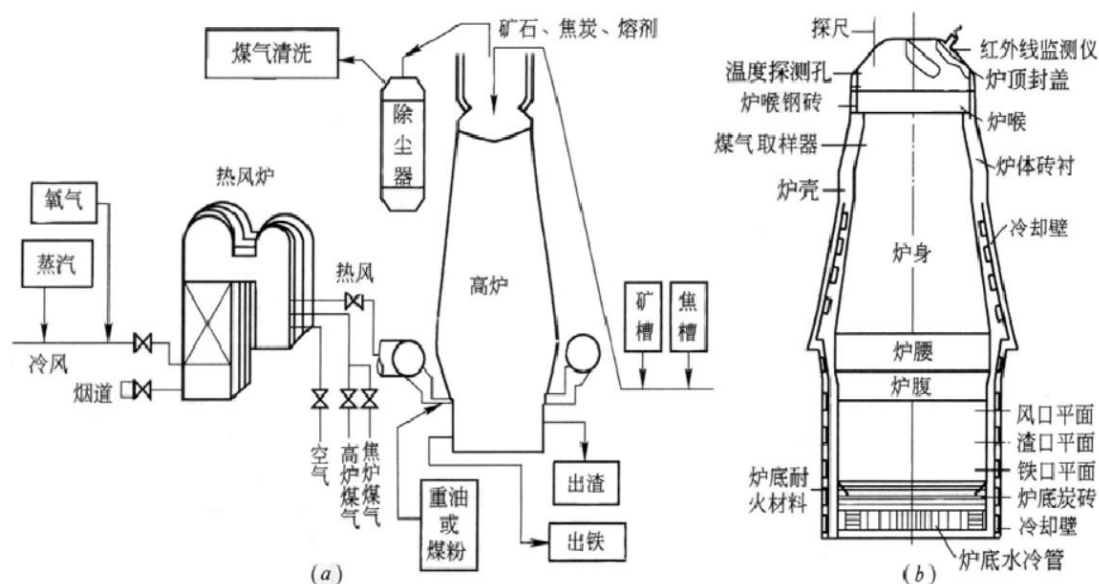


图 1 高炉炼铁工艺<sup>[4]</sup>

在高炉的具体冶炼过程中，物料自顶向下会经历加热、还原、熔融等状态，可以细分为块状带、软融带、滴落带、风口回旋区、渣铁收集区，如图 2 所示。平稳合理的状态区域分布有利于高炉炉况的稳定和生产安全。

高炉的炼铁过程是一个复杂的物料和能量交换过程，可以看作是一个复杂动态系统，具有非线性、大时滞、大噪声等特点<sup>[2]</sup>。高炉炉况的监控和异常炉况的诊断对高炉的高效安全生产有着十分重要的意义，是自动化技术应用于高炉炼铁领域的重要环节。高炉炉况的稳定顺行是使得高炉达到高产、优质、低耗的必要

条件<sup>[3]</sup>。而异常炉况频发则会极大地影响生产效率和产品质量，并造成安全隐患。如果能尽早的诊断出异常炉况，就能让操作人员或是专家系统提前决策，做出针对性的调整并进行相应补救，可以降低事故的损失甚至避免事故的发生。

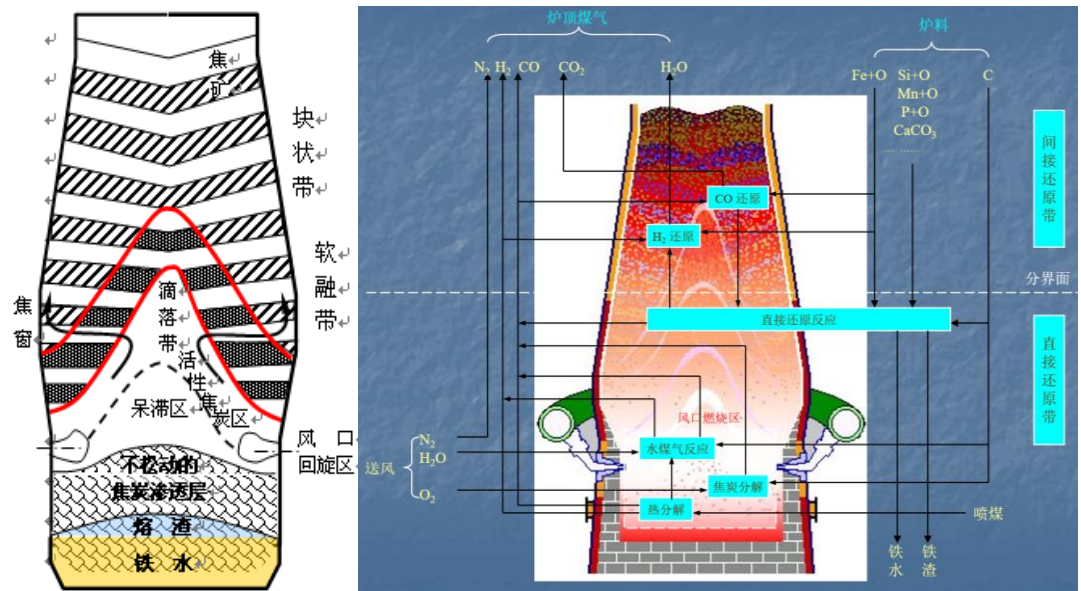


图 2 炉内冶炼还原过程

高炉主要异常炉况有以下几种<sup>[5,6]</sup>:

- (1) 悬料：当高炉内部温度过高时，炉体的透气性变差，处于下降状态的炉料和处于上升状态的边沿气流都在运动过程中发生了停滞，引起炉内风压的急剧上升，造成了一种看似是悬空的状态。
- (2) 管道：管道故障的发生是由于炉内各类化学反应的不平均进行造成的，同时也与物料的不平均混合和分布是密切相关的。因为这种不平均，造成了炉料下降速度的不一致，以至于炉内物料的分布很是无规则。这样的话，就会严重打乱炉内煤气的常态分布，导致资源利用上的浪费。
- (3) 崩料：崩料就是炉料在下降的过程中经过短暂的停歇后，瞬间垮落的现象。这依然是由于物料反应过程中各反应进程不平均进行的后果。由于各反应的速率不一致，会造成物料在炉内的分布不均，进而形成物料高低错落的分布状况，这样就很容易造成物料的塌落。
- (4) 炉缸堆积：炉缸堆积通常分为两类状况。第一类是中心堆积情形，第二类是边沿堆积情形。无论是哪一种状况的炉缸堆积，都是由于炉内煤气总是沿中心或者边沿部分流动造成的。

(5) 炉墙结厚和结瘤：当高炉冷却壁温度不均、各个位置化学反应程度不均时，软融态的炉料在下降过程中遇冷会黏附于炉墙上，当厚度超过正常范围时就被称为炉墙结厚。如果该现象继续持续地进行，那么就会发生炉墙结瘤。

在钢铁厂实际生产制造过程中，积累了丰富的历史数据，对高炉炉况变化过程有着较为全面的刻画，而其中记录的异常炉况和重大事故的数据更是为异常炉况诊断提供了真实样本，并可以基于现场历史数据的离线测试对诊断方法给出客观的效果检验。表 1 是柳钢炼铁厂运行变量列表。

表 1 柳钢炼铁厂运行变量

序号	变量	序号	变量
1	富氧率(%)	2	透气性指数
3	CO 含量(%)	4	H <sub>2</sub> 含量(%)
5	CO <sub>2</sub> 含量(%)	6	标准风速(m/s)
7	富氧流量(m <sup>3</sup> /s)	8	冷风流量(10 <sup>4</sup> m <sup>3</sup> /s)
9	鼓风动能(KJ)	10	炉腹煤气量(m <sup>3</sup> )
11	炉腹煤气指数	12	理论燃烧温度(°C)
13	顶压(KPa)	14	顶压 2(KPa)
15	顶压 3(KPa)	16	顶压 4(KPa)
17	富氧压力(MPa)	18	冷风压力(MPa)
19	冷风压力 2(MPa)	20	全压茶(KPa)
21	热风压力(MPa)	22	热风压力 2(MPa)
23	实际风速(m/s)	24	冷风温度(°C)
25	热风温度(°C)	26	顶温东北(°C)
27	顶温西南(°C)	28	顶温西北(°C)
29	顶温东南(°C)	30	顶温下降管(°C)
31	阻力系数	32	本小时喷煤量(T/h)

## 第二章 文献综述

### 2.1 故障诊断方法综述

“故障”可以理解为一个或者多个变量发生了异常变化，使得整个系统性能恶化的情况或事件<sup>[7]</sup>，广义上，“故障”可以定义为使系统出现所不期望的特征的异常现象<sup>[8]</sup>。故障诊断是通过监控并分析工业过程的运行数据以检测、分离、识别出故障的过程，故障的检测是判断系统是否发生异常，分离是对故障类型与位置的确定，辨识是对故障大小和时间的判定<sup>[8]</sup>。

故障诊断领域已经提出了基于专家系统、解析模型、多元统计分析、机器学习、信号处理等一系列技术的故障诊断方法。

基于专家系统的方法是指计算机利用技术人员在长期生产实践中积累的专家经验建立的专家知识库，模拟专家的决策和判断，对采集的系统信息进行过程监控和故障诊断。由于不需要数学模型，而是基于容易被技术人员所理解的专家经验做出诊断，该方法在电力、机械、化工、冶金等行业获得了广泛的应用。

基于解析模型的故障诊断是通过建立系统的数学解析模型，并基于模型与可测的输入输出量构造出对故障敏感、对噪声鲁棒的残差，并分析和处理残差信号实现故障诊断，该方法可以细分为参数估计法、状态估计法和等价空间法三种类型。由于直接对系统机理建模，该方法往往有很好的仿真效果，但在实际工业生产中，受制于真实过程反应的复杂、非线性、时变、多模态等特性，很难建立精确的数学解析模型。

基于多元统计分析的故障诊断则是利用各个变量间的相关性，将监控数据投影到低维子空间，然后分别计算投影空间的  $T^2$  统计量和残差空间的 SPE 统计量来检测故障，再通过使用贡献图或者重构法实现故障分离。典型的多元统计方法包括主成分分析(PCA)、偏最小二乘(PLS)、独立成分分析(ICA)等等。多元统计分析法不需要对系统结构和运行机理有深入的了解，完全基于历史数据，在缺少故障样本时也能实现过程监控。

基于机器学习的方法则是利用正常工况与各种故障工况下的历史数据样本让机器学习模型进行有监督学习，进而实现故障的检测与辨识。常见的有监督学习的机器学习算法有支持向量机、人工神经网络、决策树、随机森林、逻辑回归



等等。机器学习算法效果的好坏不仅与算法的选取有关，还在很大程度上取决于所提取特征的好坏，提取特征的过程通常被称为特征工程。

## 2.2 高炉异常炉况诊断方法

高炉冶炼过程监控主要普遍采用专家系统，目前已经有一些专家系统能够实现异常炉况诊断，如日本川崎钢铁的 ADS 系统、奥钢联的 SIMETAL BF VAIron、马钢高炉炉况诊断专家系统、宝钢高炉人工智能专家系统、浙大高炉炼铁优化专家系统等。但是在我国高炉厂的实际应用中发现，受到铁矿石来源复杂、质量层次不齐、高炉设定产量的变化、自然环境的变化、设备老化、人工操作误差等主客观因素的影响，不同时段、不同设备的炉况的工作点差异较大，工作点漂移现象十分常见，导致专家系统对异常炉况的命中率并不高。这也就需要我们利用其他故障诊断技术来实现高炉异常炉况的准确诊断。

近年来国内外学者将一系列故障诊断方法应用于异常炉况的检测和诊断。文献<sup>[9]</sup>将主成分分析应用于实验高炉的炉况监控，并对监测数据在主元方向上的变化过程给出物理含义解释。文献<sup>[4]</sup>基于模糊理论对输入特征量进行预处理和模糊化，并建立模糊矩阵，通过稀疏受限非负矩阵分解算法进行高炉故障特征的识别和故障诊断。文献<sup>[10]</sup>利用最小二乘支持向量机算法诊断了高炉向凉、向热、悬料、滑料、低料线等异常炉况。文献<sup>[11]</sup>利用天津钢铁有限公司炼铁厂的历史数据和 matlab 神经网络工具箱，基于 BP 人工神经网络建立了高炉炉况模型，分别建模实现了对炉凉、炉热、悬料、崩料、管道行程等异常炉况的预警。

以上文献大都针对仿真数据进行了测试，没有考查算法在噪信比更大、物理化学反应更加复杂的真实现场数据上的性能，而热风炉切换扰动的影响也未被考虑。本文将基于以上问题进行异常炉况的进一步研究。

## 第三章 技术难点

### 3.1 热风炉换炉扰动

高炉的热风炉主要用于加热鼓风，提升炼铁效率，为了保证每座高炉的连续生产，避免断风现象的出现，通常每座高炉配置 3 或 4 座热风炉轮流向高炉供风。在热风炉换炉操作时，需要先为即将给高炉送风的送风炉充压，再将高炉当前的燃烧炉转换为该送风炉，最后将该送风炉转换为燃烧状态。因而在换炉过程中，高炉的风温、风量、风压等数据都会发生短时间的波动，这一波动不仅会对高炉炉况产生影响，也会影响监测高炉炉况的算法的准确率。图 3 是柳钢 3 号高炉一段时间内的冷风流量、热风压力、热风温度的变化情况，其中热风压力尖峰式的波动对应的时间点即为热风炉换炉。

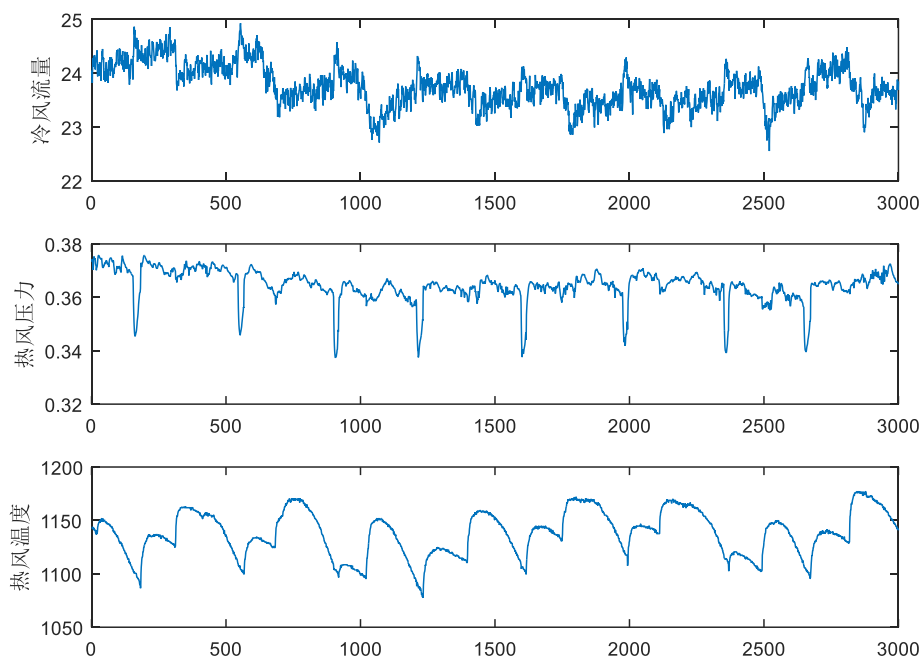


图 3 柳钢 3 号高炉 2012-07-01 00:00:05 至 08:22:37

### 3.2 工作点漂移

由于受到铁矿石和煤炭等物料来源的变化、高炉设定产量的变化、人工操作误差、高炉自身稳态的变化等主客观因素的影响，不同时段、不同设备的炉况的工作点差异较大，工作点漂移现象十分常见。

以柳钢 3 号高炉 2012-07-01 至 2012-07-06 的炉况为例，根据文献<sup>[9]</sup>的方法，对 2012-07-01 约 8 个小时的 4000 个数据点建立 PCA 模型，并作用于作为测试

集的之后 1~5 天的数据，取前两个主成分的结果画出散点图，并给出基于训练集的  $T^2$  统计量的故障检测阈值，如图 4 所示。可以看出，炉况随着时间变化会发生明显漂移，且随着时间增长漂移也越来越严重，基于训练集的  $T^2$  统计量的故障检测阈值会出现明显的误报。

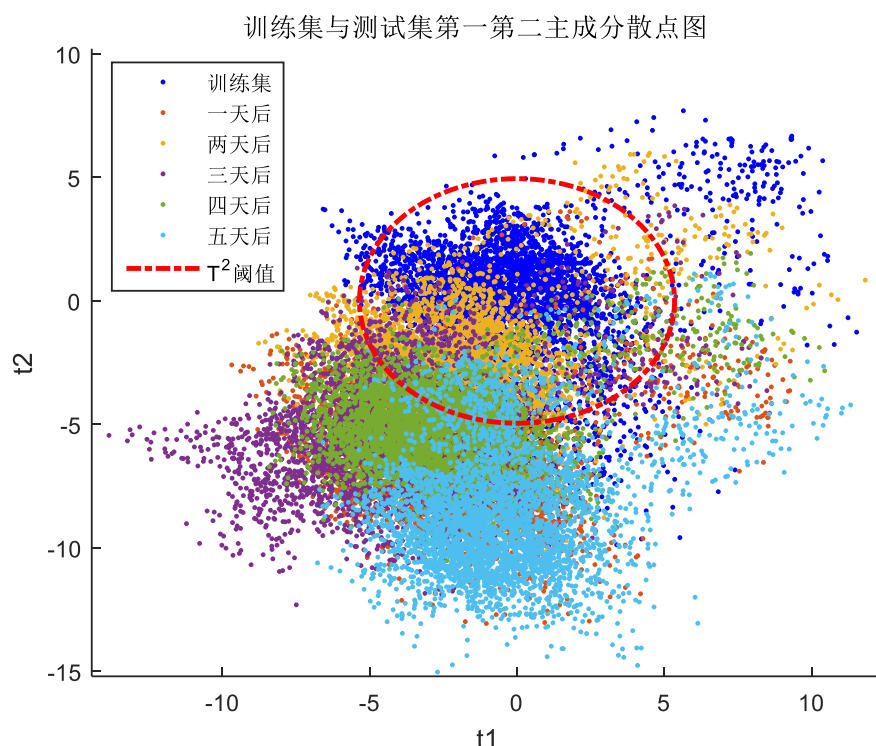


图 4 柳钢 3 号高炉 2012-07-01 至 2012-07-06 主成分散点图

### 3.3 故障样本稀少

高炉的历史数据与其他过程工业的历史数据类似，都是正常工况的数据占绝大多数，导致异常炉况样本稀少、正负样本比例失衡，会对算法的结果造成较大影响，还要警惕模型的过拟合问题。除此之外，判断高炉炉况异常是通过操作工人人工分析历史数据以及根据极少数事故报告得出的结论，由于人力物力有限以及人工判断的不精确，无法通过遍历所有高炉历史数据以给其打上标签，所以高炉的异常炉况诊断也可以认为是大部分无标签数据加上少数有标签数据的半监督学习问题。

## 第四章 研究内容与研究方法

### 4.1 深度学习

深度学习是一种区别于传统浅层机器学习算法，能够自动学习特征的多层架构的复杂网络，把原始数据通过一些简单的非线性模型转变成为更高层次的、更加抽象的表达，通过足够多的转换的组合，可以学习非常复杂的函数<sup>[12]</sup>。深度学习已经被证明能够发现高维数据中的复杂结构，除了在图像识别、语音识别等领域打破了纪录，它还在另外的领域击败了其他机器学习技术，包括预测潜在的药物分子的活性、分析粒子加速器数据、重建大脑回路、预测非编码 DNA 突变对基因表达和疾病的影响，此外，深度学习在主题分类、情感分析、自动问答和语言翻译等自然语言理解的各项任务中也产生了非常可喜的成果<sup>[12]</sup>。

在几乎所有深度网络中，都用到了的一种被称为反向传播的技术，具体的实现则是通过链式求导法则得到。

本文主要介绍在高炉异常炉况诊断中可能使用的三个技术。

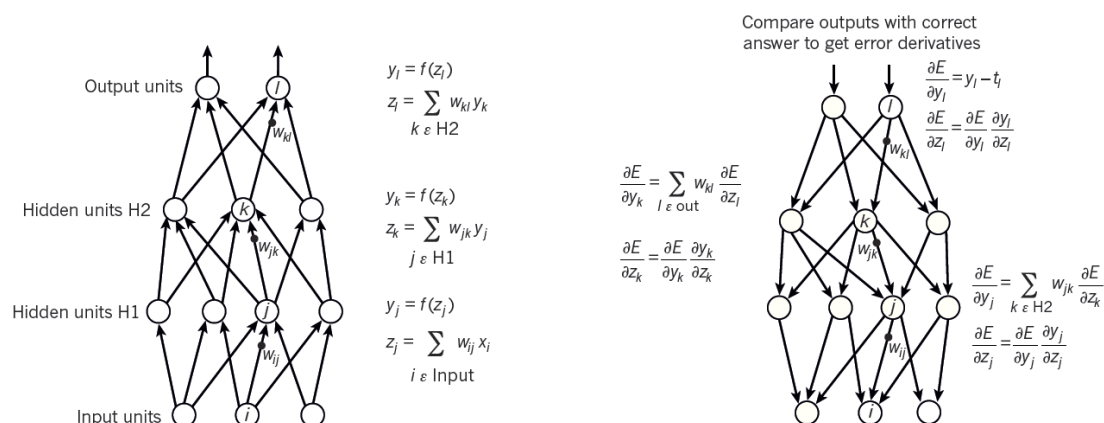


图 5 深度网络的正向传播与反向传播<sup>[12]</sup>

#### 4.1.1 auto-encoder

自动编码器(auto-encoder, AE)在上世纪 80 年代被提出，其基本思路是通过编码和解码来实现降维，实现方法是通过最小化重构误差的损失函数。当编码和解码的激活函数是线性函数时，自动编码器模型训练得到的结果是输入数据的主成分的子空间<sup>[13]</sup>；而当激活函数是非线性函数时，自动编码器能够比主成分分析检测到更多的特征<sup>[14]</sup>。

自动编码器的定义如公式所示，其中  $\mathbf{x} \in \mathbf{R}^n$  是原始输入数据， $\mathbf{y} \in \mathbf{R}^m$  是编码

输出,  $\mathbf{x}'$  是解码后的重构数据,  $J$  是损失函数, 最小化损失函数可以通过随机梯度下降法(Stochastic gradient descent, SGD)实现。

$$\begin{cases} \mathbf{y} = f(\mathbf{x}) = \text{sigmoid}(W\mathbf{x} + b_x) \\ \mathbf{x}' = g(\mathbf{y}) = \text{sigmoid}(W^T\mathbf{y} + b_y) \\ J = \sum L(\mathbf{x}, \mathbf{x}') = \sum_{\mathbf{x} \in D_n} \|\mathbf{x} - g(f(\mathbf{x}))\|^2 \end{cases}$$

在深度学习中, 自动编码器主要用于无监督学习和有监督学习的预训练, 其目标不是降维, 而是将样本非线性投影到高维空间中获得更灵活的特征表达。文献<sup>[15]</sup>通过用自动编码器逐层训练的方式初始化受限玻尔兹曼机(RBM)的神经元参数(pre-training), 再通过反向传播算法进行有监督学习(fine-tuning), 该方法可以避免深度学习网络过早陷入局部最小值的问题。

由于利用自动编码器技术将原始数据投影到高维空间来提取特征容易导致过拟合, 所以衍生出了一系列正则化和约束方法来改进最初的自动编码器方法, 如权重正则化、降噪自动编码器<sup>[16]</sup>(denoising auto-encoders)、饱和自动编码器<sup>[17]</sup>(contractive auto-encoders)等等。

#### 4.1.2 softmax

对于像高炉异常炉况分类这样的多类分类任务, 假设有  $K$  类, 则通常是将深度网络的输出设定为  $K$  个, 通过 softmax 函数<sup>[18]</sup>对输出进行标准化以表示后验概率分布,  $\mathbf{x}$  表示输入,  $\mathbf{y}_k$  代表输出,  $y^i$  代表第  $i$  个样本的类别,  $a_k$  表示前一层网络输出的线性组合:

$$y_k = p(y^i = k | \mathbf{x}) = \frac{e^{a_k}}{\sum_{j=1}^K e^{a_j}}$$

在训练 softmax 时, 输出的目标值是一个逻辑向量, 即  $C_k$  对应的元素为 1, 其他元素都为 0, 对应的损失函数是:

$$J = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{k=1}^K 1\{y^i = k\} \log(p(y^i = k | \mathbf{x}_i)) \right]$$

基于损失函数的随机梯度下降法即可得出多类分类问题的模型参数。

#### 4.1.3 LSTM

长短型记忆(Long-Short Term Memory, LSTM)作为循环神经网络(recurrent

neural network, RNN)的一种，在处理时序数据时能够捕捉变量间长时间的依赖关系<sup>[19]</sup>。LSTM 的关键在于有一个可以储存随时间变化的信息的细胞(cell)，以及输入门、遗忘门、输出门分别控制着该细胞记录哪些输入信息、记录多久以及何时输出细胞储存的信息<sup>[20]</sup>，这样的结构可以避免传统的循环神经网络在反向传播训练时所面临的梯度消失问题<sup>[21]</sup>。图 6 是一个简单的循环神经网络结构及其在时间轴上的展开，将 LSTM 用于该网络结构时，只需直接替代隐藏单元 s 即可。

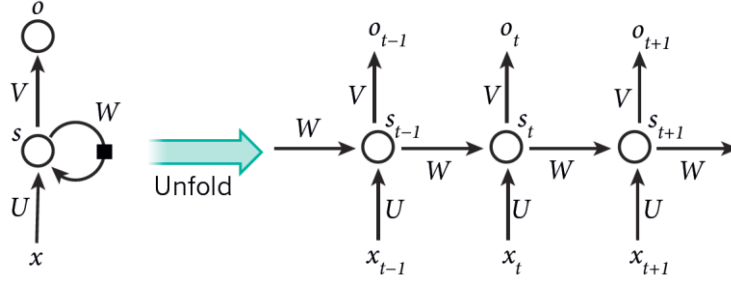


图 6 一个循环神经网络结构及其在时间轴上的展开<sup>[12]</sup>

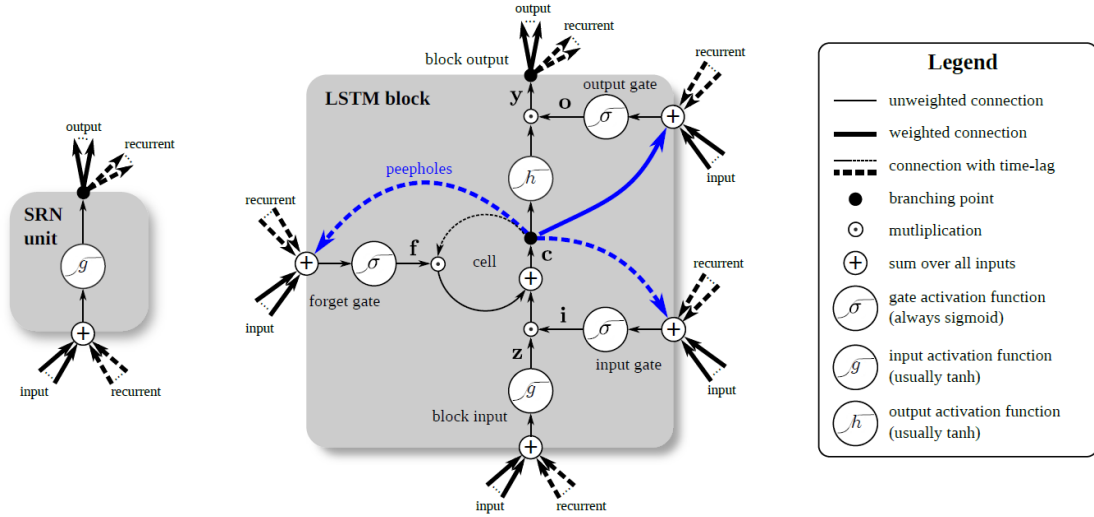


图 7 简单 RNN 单元与 LSTM 结构的对比<sup>[22]</sup>

由于长短型记忆神经元善于利用上下文信息，并克服了长时间范围内梯度消失的问题，因此在英文手写字体识别和语音识别领域取得了最好效果<sup>[23, 24]</sup>。

一个典型的 LSTM 结构如图 7 所示，输入序列  $x = (x_1, x_2, \dots, x_T)$  经过 LSTM 神经元的处理得到输出序列  $y = (y_1, y_2, \dots, y_T)$ ，被称为前馈过程，具体处理如下：

输入(input)

$$\begin{cases} z_1^t = W_z x^t + R_z y^{t-1} + b_z \\ z_2^t = g(z_1^t) \end{cases}$$

输入门(input gate)



$$\begin{cases} i_1^t = W_i x^t + R_i y^{t-1} + p_i \odot c^{t-1} + b_i \\ i_2^t = \sigma(i_1^t) \end{cases}$$

遗忘门(forget gate)

$$\begin{cases} f_1^t = W_f x^t + R_f y^{t-1} + p_f \odot c^{t-1} + b_f \\ f_2^t = \sigma(f_1^t) \end{cases}$$

细胞(cell)

$$c^t = i_2^t \odot z_2^t + f_2^t \odot c^{t-1}$$

输出门(output gate)

$$\begin{cases} o_1^t = W_o x^t + R_o y^{t-1} + p_o \odot c^{t-1} + b_o \\ o_2^t = \sigma(o_1^t) \end{cases}$$

输出(output)

$$y^t = o_2^t \odot h(c^t)$$

该模型的训练一般通过基于随机梯度下降法的反向传播实现。

## 4.2 基于深度学习的异常炉况诊断研究

高炉的运行情况是通过一系列运行变量记录的，可以认为是一组多维时序数据，通过操作工的临场判断和对历史数据的分析，可以判断出高炉是否发生了异常炉况以及异常炉况的类别，可以认为是一组时序标签，这样就将高炉异常炉况的诊断问题转换为了对多维时序输入数据和目标时序标签建模的机器学习问题。

将深度学习用于异常炉况的诊断预计有以下 4 种方案：

- 使用类似文献<sup>[24]</sup>解决语音识别问题那样 end-to-end 的训练方式，即直接将原始数据序列作为循环神经网络的输入，而使输出拟合标签序列。
- 将时序数据序列的下一时刻的输入值作为当前时刻的预测目标值，通过构建深度网络拟合该目标值，再将隐层的输出作为特征来训练分类器。
- 将时序数据的每个时间切片作为单个样本，采用堆叠自动编码器(stacked auto-encoder)预训练所有数据，再在最后一个隐层上加入 softmax 结构拟合样本标签。
- 根据传统的机器学习框架，先利用工程技术和专业知识从原始数据中提取对高炉工作点变化鲁棒、而对炉况变化敏感的特征，比如将 PCA 的主元、重构误差以及近段时间的变化趋势作为特征，再用深度网络或者其他分类器来拟合特征输入与标签序列。

## 第五章 研究计划

- 2015.6-2015.9: 前期调研、文献综述
- 2015.10-2015.12: 高炉数据预处理、深度学习算法、多元统计分析算法的实现
- 2016.1-2016.7: 实现异常炉况的分类
- 2016.7-2016.12: 结合深度学习与传统故障诊断技术, 研究实现异常炉况检测与分类
- 2017.1-2017.7: 毕业论文的撰写



## 参考文献

- [1] 曲飞, 吴敏, 曹卫华, 等. 基于支持向量机的高炉炉况诊断方法[J]. 钢铁, 2007(10): 17-19.
- [2] 李启会. 高炉冶炼过程的模糊辨识、预测及控制[D]. 浙江大学, 2005.
- [3] 刘振. 基于贝叶斯网络 (Bayesian Networks) 方法的高炉故障诊断研究[D]. 武汉科技大学, 2015.
- [4] 彭鑫. 基于稀疏矩阵的高炉故障识别研究[D]. 武汉科技大学, 2014.
- [5] 张寿荣, 于仲洁. 高炉失常与事故处理[M]. 北京: 冶金工业出版社, 2012.
- [6] 李芳. 高炉异常炉况预报专家系统研究[D]. 重庆大学, 2007.
- [7] 葛志强. 复杂工况过程统计监测方法研究[D]. 浙江大学博士学位论文, 2009.
- [8] 张萍, 王桂增, 周东华. 动态系统的故障诊断方法[J]. 控制理论与应用, 2000, 17(2): 153-158.
- [9] Vanhatalo E. Multivariate process monitoring of an experimental blast furnace[J]. Quality and Reliability Engineering International, 2010, 26(5): 495-508.
- [10] Liu L, Wang A, Sha M, et al. Multi-Class Classification Methods of Cost-Conscious LS-SVM for Fault Diagnosis of Blast Furnace[J]. Journal of Iron and Steel Research, International, 2011, 18(10): 17-33.
- [11] 赵明. 基于神经网络的高炉炉况诊断与预报研究[D]. 东北大学, 2010.
- [12] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [13] Baldi P, Hornik K. Neural networks and principal component analysis: Learning from examples without local minima[J]. Neural networks, 1989, 2(1): 53-58.
- [14] Japkowicz N, Hanson S J, Gluck M. Nonlinear autoassociation is not equivalent to PCA[J]. Neural computation, 2000, 12(3): 531-545.
- [15] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. SCIENCE, 2006, 313(5786): 504-507.
- [16] Vincent P, Larochelle H, Lajoie I, et al. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion[J]. The Journal of Machine Learning Research, 2010, 11: 3371-3408.
- [17] Rifai S, Vincent P, Muller X, et al. Contractive auto-encoders: Explicit invariance during feature extraction[C]. 2011.
- [18] Bridle J S. Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition[M]. Neurocomputing, Springer, 1990, 227-236.
- [19] Graves A. Supervised sequence labelling with recurrent neural networks[M]. Springer, 2012.
- [20] Gers F A, Schraudolph N N, Schmidhuber J. Learning precise timing with LSTM recurrent networks[J]. JOURNAL OF MACHINE LEARNING RESEARCH, 2003, 3(1): 115-143.
- [21] Hochreiter S, Bengio Y, Frasconi P, et al. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies[Z]. A field guide to dynamical recurrent neural networks. IEEE Press, 2001.
- [22] Greff K, Srivastava R K, Koutn k J. LSTM: A Search Space Odyssey[J]. arXiv preprint arXiv, 2015.
- [23] Graves A, Liwicki M, Fernandez S, et al. A Novel Connectionist System for Unconstrained Handwriting Recognition[J]. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2009, 31(5): 855-868.
- [24] Graves A, Mohamed A R, Hinton G. SPEECH RECOGNITION WITH DEEP RECURRENT NEURAL NETWORKS[M]. International Conference on Acoustics Speech and Signal Processing ICASSP, NEW YORK: IEEE, 2013, 6645-6649.