递归网络结构的实证研究

摘要：递归神经网络(RNN)是一种非常强大的序列模型，通常很难训练。长短时记忆(LSTM)是一个特定的RNN架构，它的设计使得训练更加容易。文章对LSTM最优性框架和各组成部分重要性进行了研究。在一万种RNN结构中选取出来了优于LSTM的框架并引入了封闭复发性单元

方法： 向lstm遗忘门添加偏差1 改动框架结构 超参数搜索 对问题解决评估来获取最优结构

结论：通过选优方法最终选取了三个最优框架分别为MUT1,MUT2,MUT3

①除了语言建模外，GRU在所有任务上都比LSTM做得好

②MUT1在语言建模方面的表现与GRU不相上下，在其他所有任务上的表现都优于GRU

③存在正则化时，LSTM在PTB语言建模方面明显优于所有其他架构

④遗忘偏差较大的LSTM在几乎所有atsk上的表现都优于LSTM和GRU

⑤虽然MUT1在两个音乐数据集上表现最好，但值得注意的是，当使用dropout时，LSTM-i和LSTMo在音乐数据集上获得了最好的结果。

Lstm结构框架：忘记门最重要 输出激活函数次之 遗忘门添加正偏置

LSTM: A Search Space Odyssey NEURAL NETWORKS AND LEARNING SYSTEMS

摘要：在这篇论文中，我们针对语音识别、手写识别和多音音乐建模这三个典型任务，首次对八个LSTM变体进行了大规模分析

方法：超参数每个任务的LSTM变量分别使用随机搜索进行优化，并使用强大的fANOVA框架评估其重要性。

结论：我们得出的结论是，

①最常用的LSTM体系结构(香草LSTM)在各种数据集上表现得相当好。

②耦合输入，忘记门(CIFG)或删除窥视孔连接(NP)在我们的实验中简化了LSTMs，但没有显著降低性能。这有利于减少参数和计算成本。

③遗忘门和输出激活函数是LSTM块最关键的组件。删除其中任何一个都会严重影响性能

④学习率是最关键的超参数，其次是网络规模。然而令人惊讶的是，在我们的在线梯度下降设置中，动量的使用被发现是不重要的。研究发现，输入的高斯噪声对TIMIT有一定的帮助，但对其他数据集有害。

⑤提出了一些新的见解，包括LSTM网络的体系结构选择和超参数调整，它们已经成为解决复杂序列学习问题的选择方法

Named Entity Recognition with Bidirectional LSTM-CNNs

主要内容：对于命名实体的识别使用了一种复杂的双向lstm和CNN网络进行自动检测字和字符级特征，从而消除了大部分特征工程的需要。提出了一种基于神经网络的部分词汇匹配编码方法，并与现有方法进行了比较。系统在CoNLL-2003数据集上是有竞争力的，并超过了之前报告的OntoNotes 5.0数据集的最新性能。超越那些使用大量特性工程、专有词汇和丰富的实体链接信息的系统。解决了Collobert et al. (2011b)文章中的两个问题

①一种抛弃有用的单词间长距离关系的方法

②由于只依赖词嵌入，它无法利用诸如前缀和后缀之类的显式字符级特性，这可能非常有用，特别是在单词嵌入训练不足的罕见单词中。

模型：使用双向长短时记忆(BLSTM)网络代替前馈网络。为了归纳字符级特征，我们使用了卷积神经网络。

步骤：①将提取出的每个词的特征分别输入前向LSTM网络和后向LSTM网络。每个网络在每个时间步长的输出被一个线性层和一个log-softmax层解码为每个标签类别的log- probability。然后将这两个向量简单地相加得到最终的输出

②使用卷积神经网络提取字符特征：我们使用一个卷积和一个max层来从字符特征向量中提取一个新的特征向量

③使用公开词嵌入

④字符嵌入

结论：结合双向LSTM和字符级CNN，并受益于通过dropout的鲁棒训练，在几乎不需要特征工程的情况下实现了命名实体识别的最新成果

Short-term power load forecasting using integrated methods based

on long short-term memory (基于长短时记忆的综合方法进行短期电力负荷预测)

-----SCIENCE CHINA Technological Sciences

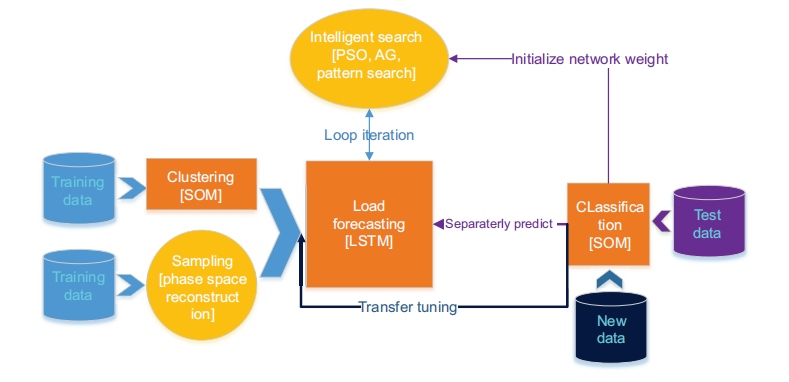
主要内容：提出了一种由自组织映射、混沌时间序列、智能优化算法和长短时记忆(LSTM)组成的集成网络结构，以延长负荷预测时间，减少人工调试，提高短期电力负荷预测精度

与LSTM预测相比，本文算法的预测精度在均方根误差(RMSE)方面提高了61.87%，在某些情况下，40倍预测窗口的预测误差降低了50%。

行文结构： 背景介绍->所涉及到的算法初步知识介绍->现有算法缺陷分析->改进算法去解决现有算法所面临的问题->数值模拟和仿真->结论和未来要做的工作

所做工作：针对LSTM预测模型复杂、预测周期短、人工调试繁琐等问题，提出了三种基于自组织特征图的集成算法--基于自组织特征映射 混沌时间序列 智能优化算法

算法框架



自组织特征映射：通过聚类进行降维度

混沌时间序列：电力负荷的时间序列是混沌时间序列，因此相空间重构方法可以对电力负荷样本进行处理，使其轨道呈现混沌特性，易于LSTM捕获。

智能优化：为了在LSTM中更快地搜索到更好的局部最优权值

结论：试图解决LSTM预测周期短、调试不方便逐步预测，提高LSTM预测的精度

In-process tool condition forecasting based on a deep learning method(基于深度学习的刀具状态预测方法)

------Robotics and Computer Integrated Manufacturing

主要内容：设计了一个基于历史数据的长短时记忆网络来预测刀面磨损量，利用加工过程中获得的原始信号，建立残差卷积神经网络，实现加工过程刀具状态监测，采用中值校正和均值校正来提高精度。

数据来源：IEEE PHM 2010 在加工过程中获取切削力、振动和声发射信号

数据驱动智能已被广泛用于实现过程中刀具状态监测但存在不能实现精确的刀具更换，刀具磨损状况应提前估计并实时更新。提出了一种对加工过程中刀具磨损状况进行预测的方法。通过在时间序列中捕获和使用数据相关性，可以精确地预测未来刀具磨损情况。该方法为数据驱动的智能制造中快速可靠的刀具更换决策提供了支持

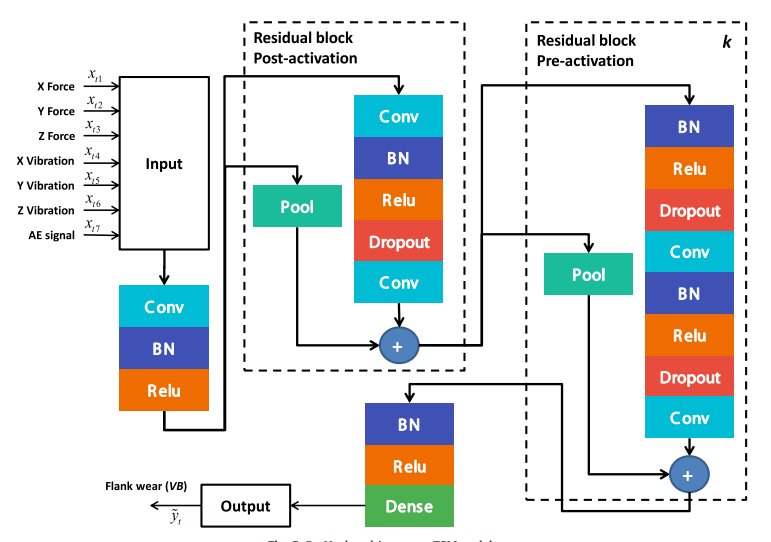
行文结构： 背景介绍(以数据驱动的现代智能制造系统需要准确预测刀具磨损量更换并提前估计并实时跟新)

->回顾与此课题相关论文并讲述其主要方法

->将本文方法集成到刀具状态监控（基于残差网络的TCM模型,讲述了模型改进方向）

->实验性研究(公开数据集测试,实验平台验证（三轴力 三轴加速度 声信号)-tensorflow平台）

TIPS:特征的质量决定了特征的准确性、泛化能力和鲁棒性，但敏感特征难以提取。进一步提高估计的准确性、鲁棒性和泛化能力，使估计更加可靠



结论：

①通过捕获时间序列中的数据依赖关系，LSTM网络根据时间序列中的几个实测VB值推断出多个未来的VB值。它弥补了历史和未来刀具磨损曲线之间的差距

②基于resnet的TCM的集成使过程中的TCF成为可能。均值校正和中值校正都能提高精度

③实验研究表明，该方法能够准确预测刀具磨损曲线。比较结果表明，该方法优于现有模型。

通过去除奇异点和离散点，降低了刀具磨损曲线的波动，引出除了刀具的持续磨损外，刀具的破损和削屑都是意外的、离散的、随机的事件，这给未来的研究提出了进一步的挑战。

Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable

Activity Recognition（深卷积和LSTM递归用于多模态可穿戴设备的神经网络活动识别）

------sensor

主要内容：人类活动是由复杂的运动序列构成的，捕捉这种时间动态是HAR（human active recognize）成功的基础,提出了一个基于卷积和LSTM递归单元的活动识别深度框架，该框架适用于多模态可穿戴传感器。该框架不仅适用于同质传感器模式，而且可以融合多模态传感器以提高性能。我们描述了关键架构超参数对性能的影响，以提供关于它们优化的见解

模型特征：(i)适用于多模态可穿戴传感器;(二)能够自然地进行传感器融合;(三)设计特征不需要专业知识;(iv)明确地对特征激活的时间动态进行建模。

①长短时记忆递归(LSTMs)神经网络是一种递归网络，它包含一个记忆来模拟时间序列问题中的时间相关性

②CNN可以作为特征提取器，将几个卷积运算符叠加在一起，形成一个越来越抽象的特征层次结构。这些模型能够自动学习多层特性层次结构(也称为“表示学习”)

文章的贡献：

①提出了一种DeepConvLSTM:它能够自动学习特征表示，并对它们的激活之间的时间依赖关系建模

②适用于人类行为识别与部分手势识别

③证明了该框架可以分别无缝地应用于不同的传感器模式

④证明了该系统可以直接对原始传感器数据进行预处理

⑤所提出的体系结构优于已发表的结果机遇挑战，包括一个深入的CNN。

⑥讨论了结果，包括关键参数对性能的影响，并概述了未来的研究方向，以利用深层建筑的特点

模型构建：卷积层(with four convolutional layers and three dense layers)充当特征提取器，并在特征图中提供输入传感器数据的抽象表示。lstm模拟了特征图激活的时间动态

The dense layer由LSTM代替 不包含池化操作

A Data-driven Model for Milling Tool Remaining Useful Life Prediction with Convolutional and Stacked LSTM Network

(利用卷积和叠加LSTM网络建立铣刀剩余寿命预测的数据驱动模型)

-------------Measurement 2019.10

主要内容：介绍了一种将卷积神经网络(CNN)与堆叠双向和单向LSTM (SBULSTM)网络相结合的混合模型（CNN-SBULSTM）以解决序列数据中刀具剩余使用寿命(RUL)预测任务，介绍模型运行特征(CNN首先用于用于局部特征提取和降维然后利用SBULSTM网络对时间信息进行去噪和编码，最后，在CNN-SBULSTM网络的顶部建立多个全连通层，为输出增加非线性，利用一个回归层生成目标(RUL))，概述实验数据采集条件，模型在公开数据集中验证，并且运用到实际采集数据中。

行文安排思路

①引言：提出刀具磨损问题->引出TCM系统的重要性->各类文献中提到的方法(分析存在的问题)->介绍本文的模型特点->后续文章安排

②methodology：介绍了相关的技术细节，提出了基于CNN-SBULSTM的RUL预测模型的框架。问题建模（介绍一些数学符号和必要附加条件）->模型细节介绍（Temporal Encoder: stacked Bi-directional and Uni-directional LSTM network（引用文献介绍并分析功能效果）、模型介绍（(a)数据预处理;(b)特征提取和模式识别;(c)回归、CNN-SBULSTM网络的层细节）、给出了构建模型的层细节）->模型训练和测试

③Experiment set-up and procedure：介绍了实验细节、数据预处理和几种评价机制

④分析RUL结果：将CNN-SBULSTM网络与几种机器学习模型和深度学习模型进行了比较。进一步分析了隐藏特征的学习能力和相对误差。刀具磨损在三个阶段中磨损速度不一样

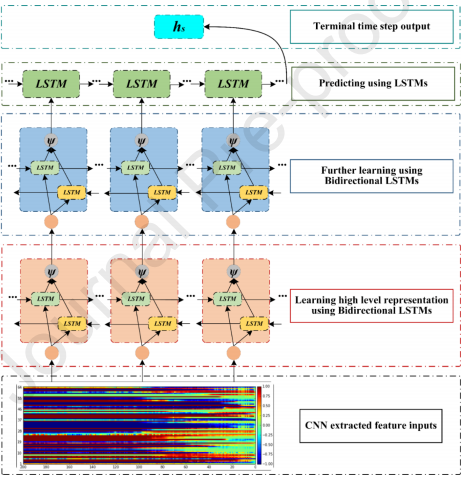
⑤总结：

①由于CNN具有较强的特征学习能力，而LSTM具有较强的长期依赖学习能力，抽象学习和依赖学习都是处理加工过程信号的重要元素。因此，将二者结合可以提高对铣刀性能退化的捕捉能力。统计结果表明，该混合模型可用于不同切削条件下铣削加工的RUL预测。预测的规则误差较小，可以在故障点之前提前更换工具

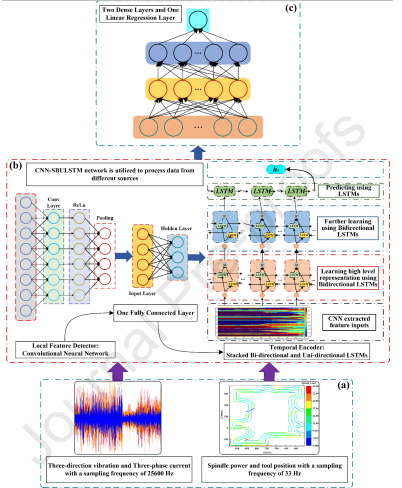
②实验结果验证了该集成模型的有效性和可行性。选择三个评价指(score/RMSE/accuracy)来量化预测性能

③还研究了dropout层和BLSTM层的影响.BLSTM层是一种有效的结构，能够捕捉到每个时间步之前和之后的全部局部关系。

④随着工业物联网的广泛应用，加工过程中产生了大量的数据。深度学习方法具有较强的大机器数据处理能力，能够捕获数据背后的层次/多级表示

SUBLSTMm模型：

CNN-SUBLSTM模型：



Tool remaining useful life prediction method based on LSTM

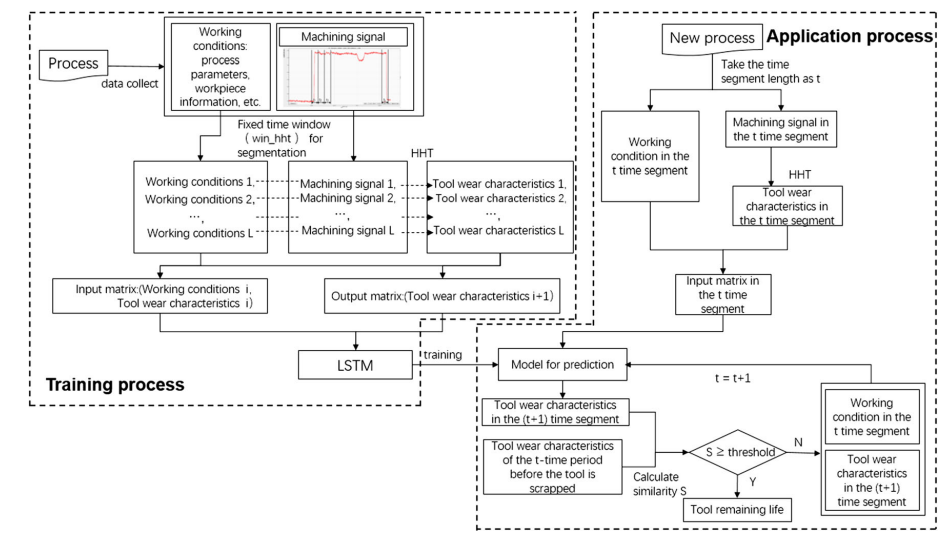
under variable working conditions (基于LSTM的变工况刀具剩余寿命预测方法)

------The International Journal of Advanced Manufacturing Technology

摘要：提出刀具剩余使用寿命预测的重要性及期面临的问题->引出本文的解决问题思路（建模->从处理信号中提取磨损特征->将提取出的刀具磨损特征和相应的工作条件组合成预测刀具磨损的输入矩阵(变工况下复杂的时空关系)->通过建立的LSTM模型进行预测刀具磨损量->实验验证）

1. instruction:刀具寿命预测的背景及意义，提出针对对于小批量和多变量的生产制造条件，不像大规模生产，刀具的工作条件经常变化，其寿命往往难以评估问题->现有的解决方案为分为直接法和间接法->变工况下的刀具磨损更具有实用性和挑战性-> 介绍文章的解决问题所用的思路和方法（收集切削过程中的工作条件和加工信号。其次，从加工信号中提取刀具磨损特征。最后，建立了基于lstm的刀具磨损模型，研究了变工况下刀具磨损过程的复杂时空关系。进而得到当前工况下刀具的剩余使用寿命）->文章结构安排（2.介绍有关工具选择和转移学习的文献综述。3.概括了我们的方法并解释了它的原理。4.描述了实验过程。5.给出了实验结果，并讨论了该方法的有效性。6.为结束语。）
2. Related work：刀具磨损预测的研究现状（传统的建模或机器学习的方法解决了在恒定工况下的刀具磨损问题。但由于刀具工况的频繁变化，上述方法不能用于多品种、小批量的生产情况。现有的一些变工况磨损预测大多采用实验和传统方法，不具备灵活性，因此需要更进一部研究）->信号处理研究现状（HHT基于信号的局部特征，可以自适应地进行分解。特别适用于分析非线性非平稳信号，具有较好的应用前景）->LSTM应用状况（LSTM网络内部的更新将根据是否记忆了上一刻的状态和当前时刻的输入来确定。由于它的独特  
   LSTM网络具有“内存”优势，在各个领域得到了越来越广泛的应用）
3. Principle explanation ：介绍了整个模型的实验数据采集和预处理过程以及最后LSTM模型训练过程->对工况、工艺参数和工件材料进行向量化和数学符号统一化处理->HHT->LSTM->刀具预测剩余寿命限定
4. Experimental setup ：介绍了实际实验设备参数
5. Results and discussion ：对实验结果进行了分析，在训练集上得到很好的收敛，但是在测试集上有一定量的偏差
6. Conclusion：提出了一种变工况下刀具剩余寿命的预测方法LSTM网络。首先，根据影响刀具磨损的因素，建立了刀具工作状态的统一表示。从实验数据提取出的磨损特征和其他工作条件构造成一个矩阵，以捕获变工况下复杂的时空关系，利用LSTM特性建立了变工况下刀具剩余寿命预测模型。实验数据的质量可能会对预测的准确性产生一些不利影响。需要进一步研究

刀具预测模型：



Research on a Real-Time Monitoring Method for the Wear State of a Tool Based on a Convolutional Bidirectional LSTM Model

（基于卷积双向LSTM模型的刀具磨损状态实时监测方法研究）

--------------Open Access Journals

摘要：针对CNC在线监控刀具磨损值，提出基于卷积神经网络(CNN)和双向长短时记忆(BiLSTM)网络与注意机制(CABLSTM)融合的实时监控方法，CNN提取时间序列中提取深度特征作为输入，BiLSTM从特征向量中学习到时间序列关系。注意机制是感知到神经网络权值对于刀具磨损分类结果的关系从而分配合适的权值，最后不同权值的信号特征输入到分类器中对刀具磨损进行分类，介绍实验设备。与其他算法比较，提高了精度和一定的泛华能力

1. Introduction ：研究的背景介绍->TCM->direct and indirect method->各类文献总结->ML弊端->DL介绍（各类文献介绍）->现有DL方法的弊端性->提出本文模型想法和先进性以及模型工作大致流程->文章结构介绍：第2节介绍了CABLSTM算法。第三部分介绍了刀

具磨损的监测过程。第四部分给出了刀具磨损状态监测的实验结果。第五部分是本文的结论。

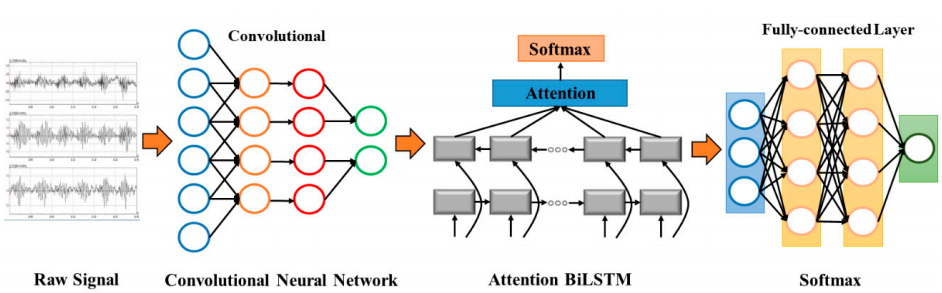
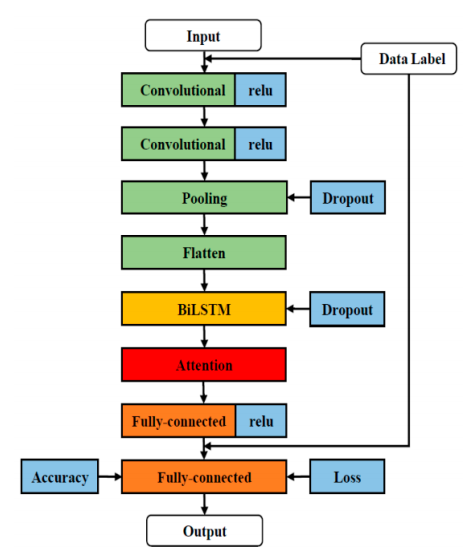
1. CABLSTM Model ：文章模型提出的过程->模型结构介绍（共四部分）->介绍模型各部分原理（第一部分：一维CNN包含卷积层与池化层,卷积层相当于对数据进行了滤波,而池化层为提取信息的主要特征并对来自卷积层高维信息进行降维,第二部分：使用BiLSTM发现长时间数据生成中的时间依赖关系,第三部分监督层进行优化权重,第四部分输出层加入Dropout防止过拟合，对刀具磨损进行阶段分类）

3. Real-Time Monitoring Method of the Tool Wear State ：用加速度传感器进行信号的采集CNC三轴震动信息，并介绍了模型训练参数以及系统流程图信息

1. Experimental ：介绍实验条件参数和数据采样频率，同时也介绍了观测刀具实际磨损值的方法。总共采用四把铣刀进行实验，与其他未经过CNN组合的各类LSTM模型进行比较，分析了结果，CNN可以提取到刀具信号的更加深层次的信息，能够对整个模型的准确度有较大提升。为了验证模型的可靠性，也对机器学习模型进了比较测试。最后也指出模型检测样品速度极快满足在线检测的要求

5. Conclusions ：总结模型的优势所在，未来的工作目标为变工况下的刀具磨损预测以及多源信息整合预测

模型：

流程图：  
·······················

Research on Tool Wear Prediction Based on LSTM and ARIMA

(基于LSTM和ARIMA的刀具磨损预测研究)

----Association for Computing Machinery.

摘要：提出刀具预测寿命的重要性->交代论文主要内容(概括了研究刀具寿命预测发展概括，提出了自己的LSTM and ARIMA模型预测方法)->交代模型的预测结果

CCS concept：计算数学→时间序列分析•应用计算→预测•信息系统→时态数据

1. INTRODUCTION ：现代制造业的发展背景->约束和影响其发展的因素->刀具磨损预测->一系列研究的重要课题和面临的问题是时间序列的分析->时间序列分析所面临的困难->例举了其中研究所用到方法，其主要还是讲解了ARIMA 和 LSTM的研究发展历程->交代后续论文结构安排(第二部分介绍了时间序列模型和LSTM模型的方法。ARIMA (p, d, q)模型应用LSTM模型对刀具磨损时间序列进行了预测第三部分。最后一部分是总结)
2. METHOLOGY ： 分别讲述ARIMA 与 LSTM 原理
3. EVALUATION AND RESULT：数据来源(Fuzzy neural network modelling for tool wear estimation in dry milling operation. In Annual conference of the prognostics and health management society , (Sep. 2009), 1-11.)（根据以往刀具磨损值去预测未来刀具磨损值）交代->数据进行预处理（一阶微分转换）以适应模型（为了避免ARIMA模型的参数过小，在预测计算中将磨损量转换为微米级）->计算ARIMA的参数->介绍综合评价方法MAPE ->介绍使用LSTM进行预测

4. CONCLUSION：全文只是使用ARIMA和LSTM分别预测了刀具未来磨损值，精度也不高不具有参考价值。