毕业设计（论文）开题报告

基于RNN深度神经网络的预测模型研究

一、课题任务及意义

（一）课题任务

1. 调研目前世界上RNN研究现状
2. 分析RNN和其他类型神经网络的区别和特点
3. 总结RNN模型在各个领域的应用方法
4. 熟练掌握几种RNN的结构。
5. 分析RNN几种结构内在的联系和区别
6. 分析各种类型的RNN结构对预测问题性能的影响
7. 根据分析结果，设计一种RNN预测模型
8. 对设计方案进行实现与验证

（二）课题研究意义

人工智能(Artificial Intelligence)，试图理解智能的实质，并制造出能以人类智能相似的方式做出反应的智能机器。如果说机器是人类手的延伸、交通工具是人类腿的延伸，那么人工智能就是人类大脑的延伸，甚至可以帮助人类自我进化，超越自我。

人工智能也是计算机领域最前沿和最具神秘色彩的学科，科学家希望制造出代替人类思考的智能机器，艺术家将这一题材写进小说，搬上银幕，引发人们无限的遐想。近年来，深度学习（Deep Learning）直接尝试解决抽象认知的难题，并取得了突破性的进展。深度学习引爆的这场革命，将人工智能带上了一个新的台阶，不仅学术意义巨大，而且实用性很强，工业界也开始了大规模的投入，一大批产品将从中获益。

2006年，机器学习泰斗Hinton在Science发表文章，提出基于深度信念网络（Deep Belief Networks, DBN）可使用非监督的逐层贪心训练算法，为训练深度神经网络带来了希望。2012年，由人工智能和机器学习顶级学者Andrew Ng和分布式系统顶级专家Jeff Dean领衔的梦幻阵容，开始打造Google Brain项目。2012年，微软首席研究官Rick Rashid在21世纪的计算大会上演示了一套自动同声传译系统。2013年，Google收购了一家叫DNN Research的神经网络初创公司，这家公司只有三个人，Geoffrey Hinton和他的两个学生。这次收购并不涉及任何产品和服务，只是希望Hinton可以将深度学习打造为支持Google未来的核心技术。同年，纽约大学教授，深度学习专家Yann LeCun加盟Facebook，出任人工智能实验室主任，负责深度学习的研发工作，利用深度学习探寻用户图片等信息中蕴含的海量信息。2013年，百度成立了百度研究院及下属的深度学习研究所（IDL），将深度学习应用于语音识别和图像识别、检索，以及广告CTR预估。

如果说Hinton 2006年发表在《Science》杂志上的论文只是在学术界掀起了对深度学习的研究热潮，那么近年来各大巨头公司争相跟进，将顶级人才从学术界争抢到工业界，则标志着深度学习真正进入了实用阶段。

目前，深度学习在几个主要领域都获得了突破性的进展。在自然语言处理（NLP）逐渐引入了深度学习（RNN）的方法，并形成了很多业内最佳的结果。

RNNs已经被在实践中证明对NLP是非常成功的。如词向量表达、语句合法性检查、词性标注等。在RNNs中，目前使用最广泛最成功的模型便是[LSTMs(Long Short-Term Memory，长短时记忆模型)模型](https://en.wikipedia.org/wiki/Long_short_term_memory)，该模型通常比vanilla RNNs能够更好地对长短时依赖进行表达，该模型相对于一般的RNNs，只是在隐藏层做了手脚。

RNNs可以很好的应用在语言模型和文本生成上。给定一个单词序列，我们需要根据之前的单词来预测下次出现的单词。语言模型能够求得一个语句正确的可能性，这是机器翻译的一部分，往往可能性越大，语句越正确。另一种应用是使用生成模型来预测下一个单词的概率，从而生产新的文本。在机器翻译领域，RNNs需要先将源语言语句序列输入，再进行输出。即输出第一个单词时，就需要从完整的输入序列中进行获取。

（三）研究现状分析

人工神经网络（ANN）由多层互相连接的称为人工神经元的单元构成。浅层网络指的是没有循环连接，且只有一个输入层、一个输出层、至多一个隐藏层的ANN。随着网络层数增加，复杂度也增加。层数或循环连接数的增加通常会增加网络的深度，从而使得网络能提供多层次的数据特征，这就是深度学习。一般来说，这些网络由非线性的、简单的单元构成，更深的隐藏层提供数据更抽象的特征，并压制不需要的可变性。由于每个层的非线性结构带来的优化困难，直到2006年出现重大研究进展之前，关于深度网络架构的研究工作都比较少。结合了循环连接的ANNs被称作RNNs，其独特的循环结构特别适用于序列数据的识别和预测。多数RNNs由高维的隐藏层状态和非线性的激活函数组成。隐藏层就像人脑的记忆单元，保存了到当前时间状态整个输入的特征信息。也是有这样的结构，RNNs才得以存储、记忆、处理过去一段时间的复杂的信息。RNNs能够将读一段序列，并且按照某种认为规定的规则，在当前时间状态下，输出另一段序列(Seq2Seq)。

近年来，也有大量基于RNNs发表的历史文献，从结构设计到具体应用。

以下表格是近年各专家学者对RNNs的贡献：

|  |  |  |
| --- | --- | --- |
| Year | Author | Contribution |
| 1990 | Elman | 普及了最简单的RNNs |
| 1993 | Doya | 在梯度下降过程用正确标记指导(techer forcing)，预训练RNNs网络 |
| 1994 | Bengio | 发现RNNs无法处理长句依赖的问题 |
| 1997 | Hochreiter | 提出长短期记忆网络（LSTM）用于解决梯度消失问题 |
| 1997 | Schuster | 提出了双向循环网络 |
| 1998 | LeCun | 海森矩阵在解决梯度消失问题的应用 |
| 2000 | Gers | 扩展了LSTM的遗忘门 |
| 2001 | Goodman | 利用单词分类快速训练极大熵 |
| 2005 | Morin | Hierarchical softmax优化输出层 |
| 2005 | Graves | 双向LSTM |
| 2007 | Jaeger | Leaky-integrator神经元的应用 |
| 2007 | Graves | 多维度的RNNs |
| 2009 | Graves | LSTM在手写识别的应用 |
| 2010 | Mikolov | 基于语言模型的RNN |
| 2010 | Neir | ReLU激活函数在解决梯度消失问题的应用 |
| 2011 | Martens | Hessian-free优化RNN学习过程 |
| 2011 | Mikolov | BPTT算法在RNN统计语言模型的应用 |
| 2011 | Duchi | 每个权重的自适应学习率 |
| 2012 | Gutmann | 噪声对比估计（NCE） |
| 2012 | Mnih | NCE在训练概率语言模型神经元的应用 |
| 2012 | Pascanu | 利用梯度消减避免梯度爆炸问题 |
| 2013 | Mikolov | 用负采样代替Hierarchical softmax |
| 2013 | Sutskever | 随机梯度下降法（SGD） |
| 2013 | Graves | 深度LSTM网络 |
| 2014 | Cho | 门循环单元结构的应用 |
| 2015 | Zaremba | 利用Dropout解决过拟合问题 |
| 2015 | Mikolov | 约束结构的递归神经网络（SCRN）在加强长记忆和解决梯度消失问题的应用 |
| 2015 | Visin | ReNet的使用（把CNN中卷积+池化层替换为循环神经网络） |
| 2015 | Gregor | DRAW（Deep recurrent attentive writer）在图片生成的应用 |
| 2015 | Kalchbrenner | Grid LSTM |
| 2015 | Srivastava | Highway networks减轻梯度下降所消耗的时间 |
| 2017 | Jing | 正交的循环单元门 |

循环神经网络（RNN）能够从序列和时序数据中学习特征和长期依赖关系。RNN 具备非线性单元的堆叠，其中单元之间至少有一个连接形成有向循环。训练好的 RNN 可以建模任何动态系统。但是，训练 RNN 主要受到学习长期依赖性问题的影响。为了解决该问题，Sepp Hochreiter 和 Jürgen Schmidhuber 在上世纪九十年代中期提出了基于RNNs的一种长短期记忆网络（LSTM）。

LSTM能解决梯度消失的问题，这是影响RNNs表现的重大障碍。LSTM还可以保留误差，用于沿时间和层进行反向传递。LSTM将误差保持在更为恒定的水平，让RNN能够在许多个时间步学习，这使得RNN学习远距离词组的因果关系变成可能。LSTM 将信息存放在递归网络正常信息流之外的门控单元中。这些单元可以存储、写入或读取信息，就像计算机内存中的数据一样。单元通过门的开关判定存储哪些信息，以及何时允许读取、写入或清除信息。这些门依据接收到的信号而开关。

总的来说，提出LSTM是RNN研究史上的一个重大突破。但是，LSTM的结构比较复杂，收敛速度比较慢。为了更好的解决这个问题，GRU应运而生。GRU和LSTM很像，都通过门将重要特征保留。GRU简化了LSTM的结构，参数更少，训练速度会更快一些。适用于样本数量不多的情况。

二、重点研究内容及技术实现途径

（一）重点研究内容

循环神经网络(Recurrent Neural Networks，RNNs)已经在众多实际问题，例如自然语言处理,图像处理，金融预测等，中取得了巨大成功以及广泛应用。从近年来发表的文献看，RNNs有着许多各种不同的结构设计。其中，最简单的simple RNN包含输入单元（Input units）、输出单元（Output units）和隐藏单元（Hidden units）。通常来说，隐藏单元最为主要。RNN打破了传统ANN的限制，引导信息能从输出单元返回隐藏单元，该行为被称作（Back Projections），同时隐藏层的输入还包括上一隐藏层的状态，即隐藏层内的节点可以自连和互连。

Simple RNN中，输入层到隐藏层的参数矩阵U、隐藏层到隐藏层的参数矩阵W和隐藏层到输出层的参数V每个时间步参数共享。对于自然语言处理这一类的问题，参数共享有着实际的意义。第一点，参数共享减少了极大的计算开销。其次，参数共享也适合NLP这种有规律可寻的问题。

用捏陶瓷的例子更容易体会共享特性对于数据量的影响。假设不同的角度相当于不同的时间步，若用前馈网络，则网络训练的过程相当于不用转盘，而是徒手将各个角度捏成想要的形状，工作量很大，效果也难以保证。若用循环网络，网络训练的过程相当于在不断旋转的转盘上，以一种收拾捏造所有的角度，工作量大大降低，而且效果也可以保证。这是基于捏出的陶瓷在每个角度的共性所致。

对于RNN的训练和对传统的 ANN 训练一样。同样使用 BP 误差反向传播算法，不过有一点区别。如果将 RNN 进行网络展开，那么参数 W,U,V 是共享的，而传统神经网络却不是的。并且在使用梯度下降算法中，每一步的输出不仅依赖当前步的网络，并且还用前面若干步网络的状态。比如，在 t=4 时，我们还需要向后传递三步，以及后面的三步都需要加上各种的梯度。该学习算法称为 Backpropagation Through Time (BPTT)。需要注意的是，在普通 RNN 训练中，BPTT 无法解决长时依赖问题(即当前的输出与前面很长的一段序列有关，一般超过十步就无能为力了)，因为 BPTT 会带来所谓的梯度消失或梯度爆炸问题(the vanishing/exploding gradient problem)。当然，有很多方法去解决这个问题，如 LSTM 便是专门应对这种问题的。

RNNs也有很多不同的架构，一般的，有1对N、N对1、N对N、N对M（Seq2Seq模型 ）不同的架构使得RNNs应用也极其灵活。

为完成本课题，需要先掌握基本的RNNs原理。在掌握了Simple RNN的原理下，理解LSTM和GRU中各种门的作用，且需要熟练应用Tensorflow实现不同结构的RNNs 。

（二）技术实现途径

RNNs特有的结构设计使得其网络的记忆具有持久性的。若将RNN按照时间展开为链式结构，能清晰关注到RNN与序列和列表的密切关系。因此RNN的这种结构能很自然的使用这类数据。LSTM在Simple RNN的基础上增加了各种门，用来处理各种信息。实践证明LSTM在许多任务中表现的都比Simple RNN出色的多。几乎所有的基于RNN的令人惊叹的结果都是LSTM取得的。

随着输入的上下文信息增多，需要连接起来的信息的距离越来越大，Simple RNN越发无力。这边是长期依赖问题（Long Term Dependencies）。LSTM在设计上明确避免了长期依赖的问题，记住长期信息也不在话下。LSTM也有链状的结构，但是重复模块的结构不同，LSTM重复模块中有4个神经网络层，且它们之间的交互非常特别。

LSTM的关键是细胞状态（Cell State），LSTM的第一步是决定我们将要从细胞状态中扔掉哪些信息，该决定由一个叫做“遗忘门(Forget Gate)”的Sigmoid层控制。下一步是决定把哪些新信息存储到细胞状态。首先，有一个叫做“输入门(Input Gate)”的Sigmoid层决定我们要更新哪些信息。接下来，一个tanh层创造了一个新的候选值，Ct~Ct~，该值可能被加入到细胞状态中。最后，需要决定最终的输出，输出将会基于目前的细胞状态，并且加入一些过滤。过滤需要一个Sigmoid层的输出门(Output Gate)，来决定我们将输出元胞的哪些部分。

除了一开始提出的LSTM，近年来各专家学者也提出了许多观点，改进LSTM。有一种流行的LSTM，它加入了“窥视孔连接(peephole connection)”。也就是说可以让各自门观察到细胞状态。另一种变种是使用对偶的遗忘门和输入门。不再单独的决定遗忘什么信息，加入什么新信息，而是一起做决定：在需要放新信息的某处忘记旧值，在已经忘记旧值的位置放入新值即可。GRU也是LSTM的变种，其将遗忘门和输入们合并为单一的更新门（Update Gate），GRU同时将细胞状态（Cell State）和隐藏状态（Hidden State）合并，同时引入其他一些变化。GRU比LSTM更简化。

最近两年，许多专家学者提出，在网络中加入注意力机制（Attention）能够使得该网络拥有更强的效果。所谓注意力机制，即在每一步让RNN从一个更大的信息集合中去选择信息，而不是同时关注所有的信息。就像人在看一幅画的时候，得到聚焦点的信息量最大。

根据以上LSTM及其变种的结构和近年来研究发现的tricks，本课题研究着重关注LSTM网络，在该网络的基础上，针对目前存在的语言翻译、图像处理、金融预测的问题，设计出多种能够适应这些问题的RNN结构，并且利用Google的Tensorflow框架，编程实现其中的结构。因为不同的结构会对问题产生很大的影响，本课题也着重探讨设计的不同结构，对于同一问题的影响，同时也分析讨论在用相同结构应对不同问题时候产生的结果。在此，通过添加或删除具体的门、添加或删除神经网络的层数，来改变其结构。

三、课题预期成果

（一）课题预期成果

本课题将通过设计的RNNs对语言翻译、图形处理和金融预测的实验结果，作图探讨，分析改变RNN结构对问题产生的影响。同时记录各种结构模型在对各问题表现的效果，讨论怎么样才能使得RNN在某个问题上表现的更好。

（二）课题预期特色

本课题设计的RNN结构，综合考虑了近年来各个专家学者提出的问题和优化RNN结构的方法，尽量使得设计出的结构能够很好的适用于本文前面所提及的3个问题。利用Tensorflow平台，降低了编程的难度，让人更专注于本课题研究的本质，即不同结构模型对结果的影响。

|  |  |  |
| --- | --- | --- |
| 四、进度计划 | | |
| 序号 | 起止周次 | 工 作 内 容 |
| 1 | 01周至02周 | 选题和背景基础知识了解，收集资料 |
| 2 | 03周至03周 | 开题报告 |
| 3 | 04周至05周 | 系统概要设计 |
| 4 | 06周至11周 | 系统详细设计 |
| 5 | 12周至14周 | 进行实验，得到实验数据 |
| 6 | 14周至16周 | 总结及撰写论文等工作 |
| 7 | 17周至18周 | 准备答辩 |

参考文献：

主要参考文献：

[1] A. Gupta, Y. S. Ong, and L. Feng, "Multifactorial Evolution: Towards Evolutionary Multitasking", IEEE Transactions on Evolutionary Computation, Accepted 2015..

[2] L. Feng, Y. S. Ong, M.-H. Lim, and I. W. Tsang, "Memetic Search with Inter-Domain Learning: A Realization between CVRP and CARP", IEEE Transactions on Evolutionary Computation, Accepted 2014.

[3] X. S. Chen, Y. S. Ong, M. H. Lim and K. C. Tan, "A Multi-Facet Survey on Memetic Computation", IEEE Transactions on Evolutionary Computation, Vol. 15, No. 5, pp. 591 - 607, Oct 2011.

[4] F. Liang, Y. S. Ong, I. W. Tsang and A.-H. Tan, 'An Evolutionary Search Paradigm that Learns with Past Experiences', IEEE Congress on Evolutionary Computation, June 2012

[5] L. Feng, Y. S. Ong, Q. H. Nguyen and A.-H. Tan, 'Towards Probabilistic Memetic Algorithm: An Initial Study on Capacitated Arc Routing Problem', IEEE World Congress on Computational Intelligence, Congress on Evolutionary Computation 2010.

[6] Y. S. Ong, M. H. Lim and X. S. Chen, "Research Frontier: Memetic Computation - Past, Present & Future", IEEE Computational Intelligence Magazine, Vol. 5, No. 2, pp. 24 -36, 2010

[7] Gers, F. A.; Schmidhuber, E. (November 2001). "LSTM recurrent networks learn simple context-free and context-sensitive languages". IEEE Transactions on Neural Networks. 12 (6): 1333–1340

[8] L. Feng, Y. S. Ong, A. H. Tan and I. W. Tsang, "Memes as Building Blocks: A Case Study on Evolutionary Optimization + Transfer Learning for Routing Problems", Memetic Computing Journal, Accepted, 2015.

[9] Thireou, T.; Reczko, M. (July 2007). "Bidirectional Long Short-Term Memory Networks for Predicting the Subcellular Localization of Eukaryotic Proteins". IEEE/ACM Transactions on Computational Biology and Bioinformatics. 4 (3): 441–446.

[10] Michael Hüsken, Peter Stagge, Recurrent neural networks for time series classification, Neurocomputing, Volume 50, 2003, Pages 223-235,

学生签名： 黄宇霄

2018年 03月 26日

五、指导教师意见

指导教师签名： 冯亮

校外指导教师签名：

2018年 03月 26日

说明：

1．开题报告应根据教师下发的毕业设计（论文）任务书，在教师的指导下由学生独立撰写。

2．本页不够，请加页。