摘　要

神经网络一直是模式识别领域中研究的热点。由于神经网络具有非常好的学习能力和泛化性能，以及其统一的结构、易于硬件实现和类似生物神经系统的工作实现机制等优点，神经网络已经广泛用于信号处理、信号分类、手写字符识别和图像分类等领域。

近些年，神经网络的研究可以分为两大方向，分别是：以极速学习机为代表的神经网络快速学习方法研究；以深度学习为代表的神经网络性能提升研究。其中，在实际模式识别任务中，关于神经网络快速学习方法的研究更显得重要。这是因为，在实际模式识别任务中，数据比算法更重要。可以进行快速计算的模型可以利用更多的数据，从而性能超过复杂的模型。

本文针对神经网络的快速计算问题，我们使用非迭代策略计算神经网络的隐含层权值，从而达到快速计算的目的。本文主要研究工作包括：

（1）针对传统极速学习机(Extreme Learning Machine, ELM)在回归问题上的病态计算问题，结合支持向量回归，提出鲁棒回归方法，极速支持向量回归模型。该模型在连续回归问题上取得了优越的性能。该方法的隐含层使用极速学习机的随机投影机制。在输出层中，该方法基于ϵ-支持向量回归模型。在目标函数中，对斜率和截距同时求L_2范数最小。同时,在约束项中采用等式近似处理。经过推导求解，我们得到了一个解析解。该方法可以鲁棒性地处理连续回归问题，同时具有非常快的计算速度。

（2）针对极速学习机的完全随机隐层节点使用效率不高的问题，提出基于样本分布构建隐层权值的方法，约束极速学习机(Constrained Extreme Learning Machine, CELM)。隐层节点使用效率不高效，使得在实际中，我们必须使用很多隐节点，才能保证模型的性能达到需求。很多的隐节点意味着更多的训练和测试时间，和更易于过拟合。基于样本分布先验产生隐层权值，可以有效地提升模型的性能，同时保持相同的快速学习速度。

（3）针对当前深度学习模型训练非常耗时，提出基于非监督学习的深度跨层学习模型。该模型使用非监督学习模型学习局部感受野特征，将先前层的特征跨接到最后一层，构建完备表示。同时，使用局部对比标准化和白化预处理各层的输入。我们分别使用极速学习机模型、自动编码机模型和主成分分析模型学习非监督局部感受野特征。实验结果证明，这些模型取得比同类模型优越很多的性能，同时具有比传统深度学习更快的学习速度。

综上所述，本文针对神经网络中学习速度慢的问题，基于隐含层非迭代策略，改进深度学习和极速学习机模型。在设计模型时，我们在考虑模型泛化性能的同时，尽

可能地使模型具有更快的学习速度。大量实验结果表明，本文提出的方法具有相对较好的性能，同时具有非常快速的学习速度。

关键词：极速学习机，病态计算，支持向量回归，极速支持向量回归，隐层节点使用效率，约束极速学习机，深度跨层神经网络，非监督学习局部感受野特征

A Study on Fast Neural Networks Based on Non-iterative Strategy

Zhu Wentao (Computer Application Technology)
Advised by Associate Professor Miao Jun

The research on neural networks is always a hot topic in pattern recognition related areas. Neural networks have been extensively applied to many areas such as signal processing, signal classification, handwritten digital recognition and image classification, because the neural networks have a lot of advantages such as good learning ability, satisfactory generalization performance, easy to be implemented by hardware and its working schema similar to biological neural system.

The research on neural networks can mainly be categorized into two directions during the recent years. They are fast learning speed related research, which is represented by the Extreme Learning Machine (ELM), and better generalization ability related research, which is represented by the deep learning. However, the fast learning speed related research is even more important in practical pattern recognition applications. Because data is always more important than the pattern recognition algorithm in practical applications. A model that learns fast trains with much more data, thus it outperforms a complex model.

In the work, we use non-iterative strategy to calculate the hidden layer weights in the neural networks. Therefore, the enhanced models have fast learning speed. The work can be summarized as follows,

(1) Proposes a robust regression method, Extreme Support Vector Regression (ESVR), combining Support Vector Regression (SVR) with Extreme Learning Machine (ELM) to deal with the ill-conditioned problem in ELM with better performance in regression problems. The hidden layer in ESVR uses random projection schema in ELM. In the output layer, based on ϵ - SVR model, the approach minimizes the L_2 norm of both slop and intercept in the objective function, and uses the equality constraints to approximate the ϵ - SVR model. After the deduction, we obtained an analytical solution. The ESVR robustly deals with the regression problems, whilst maintaining fast learning speed of ELM.

(2) Proposes the Constrained ELMs (CELMs) generating hidden nodes based on sample related distribution prior to deal with the inefficient use of totally randomly generated hidden nodes in ELM. The inefficient use of hidden nodes leads to more hidden nodes to use in practical applications to meet the requirement. More hidden nodes mean more testing time, and easier to over-fitting. Generating hidden nodes based on sample distribution prior boosts the performance effectively, whilst retaining the same extremely fast learning speed as the ELM.

(3) Proposes the deep trans-layer learning model based on unsupervised learning to deal with the high complexity of current deep models. The model uses unsupervised learning to learn the local receptive field features. Then concatenate the features from previous layers into the last layer to construct a more completed representation. We attempt to use Extreme Learning Machine, Auto Encoder and PCA to learn the unsupervised local receptive features respectively. Experimental results show that, these models achieve much better performance than the related models, whilst being much faster learning speed than conventional deep models.

In conclusion, the above works study how to deal with slow learning speed in neural networks based on the non-iterative strategy by enhancing the conventional deep models and Extreme Learning Machine. In the model devising, we consider the generalization of the model and learning speed at the same time. Extensive experiments show, the methods proposed in the paper have better testing accuracy and much faster learning speed at the same time.

Keywords: Extreme Learning Machine (ELM), ill-conditioned computation, Support Vector Regression (SVR), Extreme Support Vector Regression (ESVR), efficiency of hidden nodes, Constrained Extreme Learning Machine (CELM), deep trans-layer neural networks, unsupervised learning local receptive field features