

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2016 年 9 月 22 日 (22.09.2016)



(10) 国际公布号
WO 2016/145676 A1

- (51) 国际专利分类号:
G06N 5/00 (2006.01)
- (21) 国际申请号:
PCT/CN2015/075473
- (22) 国际申请日:
2015 年 3 月 31 日 (31.03.2015)
- (25) 申请语言:
中文
- (26) 公布语言:
中文
- (30) 优先权:
201510112645.9 2015 年 3 月 13 日 (13.03.2015) CN
- (71) 申请人: 中国科学院声学研究所 (INSTITUTE OF ACOUSTICS, CHINESE ACADEMY OF SCIENCES) [CN/CN]; 中国北京市海淀区北四环西路 21 号, Beijing 100190 (CN)。上海尚恩华科网络科技股份有限公司 (SHANGHAI 3NTV NETWORK TECHNOLOGY CO. LTD.) [CN/CN]; 中国上海市闸北区共和新路 1481 号 1 号楼 408 室, Shanghai 200072 (CN)。

- (72) 发明人: 盛益强 (SHENG, Yiqiang); 中国北京市海淀区北四环西路 21 号, Beijing 100190 (CN)。王劲林 (WANG, Jinlin); 中国北京市海淀区北四环西路 21 号, Beijing 100190 (CN)。邓浩江 (DENG, Haojiang); 中国北京市海淀区北四环西路 21 号, Beijing 100190 (CN)。尤佳莉 (YOU, Jiali); 中国北京市海淀区北四环西路 21 号, Beijing 100190 (CN)。
- (74) 代理人: 北京方安思达知识产权代理有限公司 (FASTA INTELLECTUAL PROPERTY LIMITED); 中国北京市海淀区中关村北二条 13 号院 1 号楼 202 室, Beijing 100190 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,

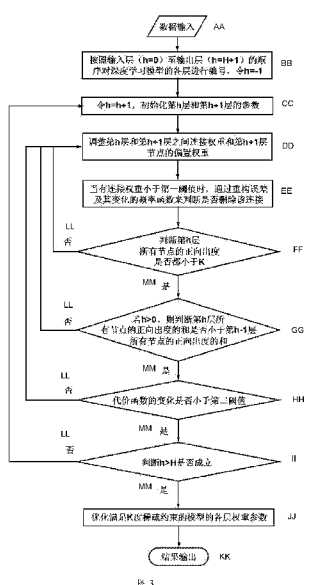
[见续页]

(54) Title: BIG DATA PROCESSING METHOD BASED ON DEEP LEARNING MODEL SATISFYING K-DEGREE SPARSE CONSTRAINT

(54) 发明名称: 基于满足 K 度稀疏约束的深度学习模型的大数据处理方法

(57) Abstract: Proposed is a big data processing method based on a deep learning model satisfying K-degree sparse constraints. The method comprises: step 1), constructing a deep learning model satisfying K-degree sparse constraints using an un-marked training sample via a gradient pruning method, wherein the K-degree sparse constraints comprise a node K-degree sparse constraint and a level K-degree sparse constraint; step 2), inputting an updated training sample into the deep learning model satisfying the K-degree sparse constraints, and optimizing a weight parameter of each layer of the model, so as to obtain an optimized deep learning model satisfying the K-degree sparse constraint; and step 3), inputting big data to be processed into the optimized deep learning model satisfying the K-degree sparse constraints for processing, and finally outputting a processing result. The method in the present invention can reduce the difficulty of big data processing and increase the speed of big data processing.

(57) 摘要: 本发明提出了基于满足 K 度稀疏约束的深度学习模型的大数据处理方法, 所述方法包括: 步骤 1) 通过渐变剪枝法用无标注训练样本构建满足 K 度稀疏约束的深度学习模型; 所述 K 度稀疏约束包括节点 K 度稀疏约束和层次 K 度稀疏约束; 步骤 2) 将更新后的训练样本输入所述满足 K 度稀疏约束的深度学习模型, 优化模型的各层的权重参数; 进而得到优化的满足 K 度稀疏约束的深度学习模型; 步骤 3) 将待处理的大数据输入所述优化的满足 K 度稀疏约束的深度学习模型进行处理, 最后输出处理结果。本发明的方法能够降低大数据处理的难度, 提高大数据处理的速度。



AA: Data input.
BB: Numbering various layers of a deep learning model according to an order from an input layer ($h=0$) to an output layer ($h=H+1$), where H is set that $h=1$.
CC: Setting $h=h+1$, initializing parameters of the h th layer and the $(h+1)$ th layer.
DD: Adjusting a connection weight between the h th layer and the $(h+1)$ th layer and a bias weight of a node of the $(h+1)$ th layer.
EE: When a connection weight is less than a first threshold value, judging whether to delete the connection via a probability function of a reconstruction error and a change thereof.
FF: Judging whether forward out-degrees of all nodes in the h th layer are less than K .
GG: If $h > 0$, judging whether the sum of the forward out-degrees of all the nodes in the h th layer is less than the sum of forward out-degrees of all the nodes in the $(h-1)$ th layer.
HH: Is the change of a cost function less than a second threshold value.
II: Judging whether $h > H$ is valid.
JJ: Optimizing a weight parameter of each layer of a model satisfying K-degree sparse constraints.
KK: Result output.
LL: No.
MM: Yes.

WO 2016/145676 A1



NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

- (84) **指定国** (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT,

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

基于满足 K 度稀疏约束的深度学习模型的大数据处理方法

技术领域

本发明涉及人工智能和大数据领域，特别涉及基于满足 K 度约束的深度学习模型的大数据处理方法。

背景技术

随着网络技术的快速发展，数据的容量和多样性快速增加，而处理数据的算法复杂度却难以改善，如何高效处理大数据已经成为一个紧迫的难题。在现有的依赖个人经验和手工操作来描述数据、标注数据、选择特征、提取特征、处理数据的方法，已经很难满足大数据快速增长的需求。随着人工智能技术的快速发展，特别是深度学习算法的研究突破，为解决大数据处理问题指明了一个值得探索的方向。

Hinton 等人在 2006 年提出了用于深度置信网的逐层初始化训练方法，这是深度学习方法的研究起点，该方法打破了几十年的深度神经网络训练困难且效果不好的局面。此后，深度学习算法替代了传统算法，在图像识别、语音识别、自然语言理解等领域得到广泛的应用。深度学习是通过模拟人脑分层次的抽象，将底层数据逐层映射而获得更抽象的特征，由于它可以从大数据中自动提取特征，并通过海量的样本训练获得很好的处理效果，从而得到了广泛的关注。实际上，大数据的快速增长和深度学习的研究突破是相辅相成的，一方面大数据的快速增长需要一种高效处理海量数据的方法，另一方面深度学习模型的训练需要海量的样本数据。总之，大数据可以使深度学习的性能达到极致。

但是，现有的深度学习模型仍然存在很多严重的问题，例如：模型难以扩展、参数优化困难、训练时间过长、推理效率低下等。在 2013 年 Bengio 的一篇综述论文中，总结了目前深度学习所面临的挑战和难点，包括如何扩展现有的深度学习模型的规模，并应用到更大的数据集；如何减小参数优化困难，如何避免昂贵的推理和采样，以及如何解开变化因素等。

发明内容

本发明的目的在于，克服在大数据应用中现有神经网络深度学习模型存在的上述问题，提出了基于满足 K 度稀疏约束的深度学习模型，该模型通过对各层神经元节点的正向出度进行约束，简化了模型的结构，提高了模型的训练速度和泛化能力，改善模型参数优化困难问题，将该模型应用于大数据处理，能够降低大数据处理的难度，提高大

数据处理的速度。

为了实现上述目的，本发明提出了基于满足 K 度稀疏约束的深度学习模型的大数据处理方法，所述方法包括：

步骤 1) 通过渐变剪枝法用无标注训练样本构建满足 K 度稀疏约束的深度学习模型；所述 K 度稀疏约束包括节点 K 度稀疏约束和层次 K 度稀疏约束；所述节点 K 度稀疏约束是指模型中所有节点的正向出度不超过 K；所述的 K 的取值范围为 $(1, N/H]$ ，其中，N 为所述深度学习模型中所有节点的个数；H 为模型隐层的层数；所述层次 K 度稀疏约束是指第 h 层所有节点的正向出度的和小于第 h-1 层所有节点的正向出度的和；

步骤 2) 将更新后的训练样本输入所述满足 K 度稀疏约束的深度学习模型，优化模型的各层的权重参数；进而得到优化的满足 K 度稀疏约束的深度学习模型；

步骤 3) 将待处理的大数据输入所述优化的满足 K 度稀疏约束的深度学习模型进行处理，最后输出处理结果。

上述技术方案中，所述 K 的取值为：

$$K = \left\lfloor \frac{\sqrt{d_{in} d_{out}}}{H} \right\rfloor$$

其中， d_{in} 为模型输入的维度， d_{out} 为模型输出的维度，H 为模型隐层的层数， $\lfloor \cdot \rfloor$ 为取整符号。

上述技术方案中，所述方法中的步骤 1) 进一步包括：

步骤 101) 按照输入层至输出层的顺序对深度学习模型的各层进行编号，令 $h=-1$ ；

设深度学习模型包括输入层、H 个隐层和输出层，从输入层至输出层总共包括 H+2 层；设输入层的编号为 0，第一个隐层的编号为 1，依次类推，输出层的编号为 H+1；

步骤 102) 令 $h=h+1$ ，初始化第 h 层和第 h+1 层的参数；

步骤 103) 将无标注训练样本集 $\mathbf{Y} = \{\mathbf{x}_i^*\}$ 输入第 h 层，在第 h 层和第 h+1 层的代价函数最小化的过程中，调整第 h 层和第 h+1 层之间连接权重和第 h+1 层节点的偏置权重；

步骤 104) 当有连接权重小于第一阈值时, 通过重构误差变化的概率函数来判断是否删除该连接;

如果有连接的权重衰减到小于第一阈值时, 则根据有当前连接和无当前连接的两种情况下重构样本, 得到重构误差变化 ΔE_r , 并以该误差变化的概率函数 $\min[1, \exp(-\Delta E_r/E_r)]$ 来决定是否删除当前连接;

步骤 105) 判断第 h 层所有节点的正向出度是否都小于 K , 如果判断结果是肯定的, 当时, 转入步骤 106); 否则, 转入步骤 103);

步骤 106) 若 $h > 0$, 则判断第 h 层所有节点的正向出度的和是否小于第 $h-1$ 层所有节点的正向出度的和, 如果判断结果是肯定的, 转入步骤 107), 否则, 转入步骤 103);

步骤 107) 判断代价函数变化是否小于第二阈值, 如果判断结果是肯定的, 转入步骤 108), 否则, 转入步骤 103);

步骤 108) 判断 $h > H$ 是否成立, 如果判断结果是肯定的, 步骤 1) 的流程结束; 否则, 转入步骤 102);

上述技术方案中, 所述方法中步骤 2) 的具体过程为:

将更新后的训练样本输入所述满足 K 度稀疏约束的深度学习模型, 当输入的训练样本为无标注样本集 $\mathbf{Y} = \{\mathbf{x}_i^f\}$ 时, 将输入样本得到输出再逆转为输入, 在 K 度约束下自输出层向输入层进行反向重构, 计算重构误差 E_r , 以梯度下降或共轭梯度下降的方式调整各层权重; 直至误差小于临界值; 当输入的训练样本为有标注的样本集 $\mathbf{X} = \{\mathbf{x}_i^f, \mathbf{y}_i^f\}$, 则将输出与 $\mathbf{y}_i^f$ 相比较, 计算出训练误差 E_x , 以梯度下降或共轭梯度下降的方式调整正向权重; 直至误差小于临界值。

本发明的方法可以克服现有神经网络模型的训练速度过长、参数优化困难等缺点, 提高深度前馈神经网络、深度置信网等现有神经网络模型的扩展能力、泛化能力和执行速度, 改善无监督学习难度和参数优化难度, 从而降低深度学习算法进行大数据处理的

难度。

附图说明

图 1 为无层次 K 度稀疏网络及其节点度稀疏约束的示意图；

图 2 为有层次 K 度稀疏网络及其层次度稀疏约束的示意图；

图 3 为本发明的基于满足度稀疏约束的深度学习模型的大数据处理方法的流程图。

具体实施方式

首先对本发明涉及的概念进行解释。

如图 1 所示，无层次 K 度稀疏网络是指所有节点满足节点 K 度稀疏约束，节点 K 度稀疏约束是指：删除节点之间的不必要连接，直到所有节点的正向出度 k_i 都不超过 k 为止，其中 k 为一个设定的参数；正向是指从输入到输出的方向，如果有隐层，那么它是从输入到隐层再到输出的方向。

如图 2 所示，训练后的有层次 K 度稀疏网络是指所有层满足层次 K 度稀疏约束，层次 K 度稀疏约束是指：隐层的层次正向出度，即单一隐层节点的正向出度之和从输入到输出单调递减。作为层次 K 度稀疏约束的一种特殊情况，假如每一层中的节点的正向出度相等，那么每一层中的节点数和正向出度的乘积从输入到输出单调递减。

此外，作为上述 K 度稀疏网络的一些简单变化，节点 K 度稀疏网络是指满足 $k_i \leq k$ 的神经网络模型，层次 K 度稀疏网络是指满足 $\sum_i k_i^{(j)} \geq \sum_i k_i^{(j+1)}$ 的神经网络模型，节点上限 K 度稀疏网络是指满足 $k_i = k$ 的神经网络模型，层次上限 K 度稀疏网络是指满足 $\sum_i k_i^{(j)} = \sum_i k_i^{(j+1)}$ 的神经网络模型，理想上限 K 度稀疏网络是指同时满足 $k_i = k$ 和 $\sum_i k_i^{(j)} = \sum_i k_i^{(j+1)}$ 的神经网络模型，它们都可以直接应用本发明所涉及的方法。

下面用数学语言来描述满足 K 度稀疏约束的神经网络模型。

假设神经网络模型共有 N 个节点，先以全连接的方式形成网络，将不连接的权重取值为零，如图 1 所示，那么任何一个节点的输出 x_j 和该节点的输入集合 $X = \{x_i, i=1 \cdots N\}$ 之间都满足以下运算规则：

$$x_j = f(\sum_i w_{ij} x_i + b_j), \text{ 其中, } x_i \in X, x_j \in X$$

这里的 x_j 是任意节点的输出, f 是该节点的激活函数, b_j 是该节点的偏置权重, w_{ij} 是连接到该节点的输入权重, 并允许存在取值为零的权重。

现定义整个神经网络模型的正向为从外部输入到输出的方向, 如图 1 所示, 对于 K 度稀疏网络来说, 任取一个节点的输出都会正向输入到 K_i 个节点上:

$$K_i \leq K$$

这里的 K 是一个超参数, 通常要比全连接时的 N 小一些, 甚至要小很多, 以达到稀疏的效果; K 的取值范围为 $(1, N/H]$, 其中, N 为深度学习模型中所有节点的个数; H 为模型隐层的层数; 优选的, K 的取值为:

$$K = \left\lceil \frac{\sqrt{d_{in} d_{out}}}{H} \right\rceil$$

其中, d_{in} 为模型输入的维度, d_{out} 为模型输出的维度, H 为模型隐层的层数, $\lceil \cdot \rceil$ 为取整符号。

K_i 是第 i 个节点的正向出度, 其中 $i = 1 \cdots N$ 。如果有层次 K 度稀疏网络, 如图 2 所示, 还必须满足所述的层次 K 度稀疏约束:

$$\sum_i K_i^{(j)} \geq \sum_i K_i^{(j+1)}$$

这里的 $K_i^{(j)}$ 是第 j 个隐层中任意节点的正向出度, $K_i^{(j+1)}$ 是第 $j+1$ 个隐层中任意节点的正向出度。

对于第 h 个隐层来说, 任取一个节点的输出都会正向输入到 $K_i^{(h)}$ 个节点上, 有:

$$K_i^{(h)} \leq K^{(h)} \leq K$$

其中, $K^{(h)}$ 是第 h 个隐层的各节点的正向出度的最大值, 根据隐层的不同, $K^{(h)}$ 可以不同, 但 K 值保持不变。

如图 3 所示, 本发明提供了基于满足 K 度稀疏约束的深度学习模型的大数据处理方法, 所述方法包括:

步骤 1) 通过渐变剪枝法用无标注训练样本构建满足 K 度稀疏约束的深度学习模型; 所述步骤 1) 进一步包括:

步骤 101) 按照输入层至输出层的顺序对深度学习模型各层进行编号, 令 $h=-1$;

设深度学习模型包括输入层、 H 个隐层和输出层, 从输入层至输出层总共包括 $H+2$ 层; 设输入层的编号为 0, 第一个隐层的编号为 1, 依次类推, 输出层的编号为 $H+1$;

步骤 102) 令 $h=h+1$, 初始化第 h 层和第 $h+1$ 层的参数;

步骤 103) 将无标注训练样本集 $\mathbf{Y} = \{\mathbf{x}_i^T\}$ 输入第 h 层, 在第 h 层和第 $h+1$ 层的代价函数最小化的过程中, 调整第 h 层和第 $h+1$ 层之间连接权重和第 $h+1$ 层节点的偏置权重;

步骤 104) 当有连接权重小于第一阈值时, 通过重构误差变化的概率函数来判断是否删除该连接;

如果有连接的权重衰减到小于第一阈值时, 则根据有当前连接和无当前连接的两种情况下重构样本, 得到重构误差变化 ΔE_r , 并以该误差变化的概率函数

$$\min[1, \exp(-\Delta E_r / E_r)]$$

来决定是否删除当前连接;

步骤 105) 判断第 h 层所有节点的正向出度是否都小于 K , 如果判断结果是肯定的, 当时, 转入步骤 106); 否则, 转入步骤 103);

步骤 106) 若 $h>0$, 则判断第 h 层所有节点的正向出度的和是否小于第 $h-1$ 层所有节点的正向出度的和, 如果判断结果是肯定的, 转入步骤 107), 否则, 转入步骤 103);

步骤 107) 判断代价函数变化是否小于第二阈值, 如果判断结果是肯定的, 转入步骤 108), 否则, 转入步骤 103);

步骤 108) 判断 $h>H$ 是否成立, 如果判断结果是肯定的, 步骤 1) 的流程结束;

否则，转入步骤 102)。

步骤 2) 将更新后的训练样本输入所述满足 K 度稀疏约束的深度学习模型，优化模型的各层的权重参数；进而得到优化的满足 K 度稀疏约束的深度学习模型；

将更新训练样本输入所述满足 K 度稀疏约束的深度学习模型，当输入的训练样本为无标注样本集 $\mathcal{Y} = \{\mathbf{y}_i^*\}$ 时，将输入样本得到输出再逆转为输入，在 k 度约束下自输出层向输入层进行反向重构，计算重构误差 $\mathcal{E}_s$ ，以梯度下降或共轭梯度下降的方式调整各层权重；直至误差小于临界值；当输入的训练样本为有标注的样本集 $\mathcal{X} = \{\mathbf{x}_i^*, \mathbf{y}_i^*\}$ ，则将输出与 $\mathbf{y}_i^*$ 相比较，计算出训练误差 $\mathcal{E}_t$ ，以梯度下降或共轭梯度下降的方式调整正向权重；直至误差小于临界值。

步骤 2) 将更新后的训练样本输入所述满足 K 度稀疏约束的深度学习模型，优化模型的各层的权重参数；进而得到优化的满足 K 度稀疏约束的深度学习模型；

步骤 3) 将待处理的大数据输入所述优化的满足 K 度稀疏约束的深度学习模型进行处理，最后输出处理结果。

最后所应说明的是，以上实施例仅用以说明本发明的技术方案而非限制。尽管参照实施例对本发明进行了详细说明，本领域的普通技术人员应当理解，对本发明的技术方案进行修改或者等同替换，都不脱离本发明技术方案的精神和范围，其均应涵盖在本发明的权利要求范围当中。

1、基于满足 K 度稀疏约束的深度学习模型的大数据处理方法，所述方法包括：

步骤 1) 通过渐变剪枝法用无标注训练样本构建满足 K 度稀疏约束的深度学习模型；

所述 K 度稀疏约束包括节点 K 度稀疏约束和层次 K 度稀疏约束；所述节点 K 度稀疏约束是指模型中所有节点的正向出度不超过 K；所述的 K 的取值范围为 $(1, \frac{N}{H}]$ ，其中，N

为所述深度学习模型中所有节点的个数；H 为模型隐层的层数；所述层次 K 度稀疏约束是指第 h 层所有节点的正向出度的和小于第 h-1 层所有节点的正向出度的和；

步骤 2) 将更新后的训练样本输入所述满足 K 度稀疏约束的深度学习模型，优化模型的各层的权重参数；进而得到优化的满足 K 度稀疏约束的深度学习模型；

步骤 3) 将待处理的大数据输入所述优化的满足 K 度稀疏约束的深度学习模型进行处理，最后输出处理结果。

2、根据权利要求 1 所述的基于满足 K 度稀疏约束的深度学习模型的大数据处理方法，其特征在于，所述 K 的取值为：

$$K = \left\lceil \frac{\sqrt{d_{in} d_{out}}}{H} \right\rceil$$

其中， d_{in} 为模型输入的维度， d_{out} 为模型输出的维度，H 为模型隐层的层数， $\lceil \cdot \rceil$ 为取整符号。

3、根据权利要求 1 所述的基于满足 K 度稀疏约束的深度学习模型的大数据处理方法，所述方法中的步骤 1) 进一步包括：

步骤 101) 按照输入层至输出层的顺序对深度学习模型的各层进行编号，令 $h=-1$ ；

设深度学习模型包括输入层、H 个隐层和输出层，从输入层至输出层总共包括 H+2 层；设输入层的编号为 0，第一个隐层的编号为 1，依次类推，输出层的编号为 H+1；

步骤 102) 令 $h=h+1$ ，初始化第 h 层和第 h+1 层的参数；

步骤 103) 将无标注训练样本集 $\mathbf{Y} = \{\mathbf{x}_i^*\}$ 输入第 h 层, 在第 h 层和第 $h+1$ 层的代价函数最小化的过程中, 调整第 h 层和第 $h+1$ 层之间连接权重和第 $h+1$ 层节点的偏置权重;

步骤 104) 当有连接权重小于第一阈值时, 通过重构误差变化的概率函数来判断是否删除该连接;

如果有连接的权重衰减到小于第一阈值时, 则根据有当前连接和无当前连接的两种情况下重构样本, 得到重构误差变化 ΔE_r , 并以该误差变化的概率函数 $\min[1, \exp(-\Delta E_r / E_r)]$ 来决定是否删除当前连接;

步骤 105) 判断第 h 层所有节点的正向出度是否都小于 K , 如果判断结果是肯定的, 当时, 转入步骤 106); 否则, 转入步骤 103);

步骤 106) 若 $h > 0$, 则判断第 h 层所有节点的正向出度的和是否小于第 $h-1$ 层所有节点的正向出度的和, 如果判断结果是肯定的, 转入步骤 107), 否则, 转入步骤 103);

步骤 107) 判断代价函数变化是否小于第二阈值, 如果判断结果是肯定的, 转入步骤 108), 否则, 转入步骤 103);

步骤 108) 判断 $h > H$ 是否成立, 如果判断结果是肯定的, 步骤 1) 的流程结束; 否则, 转入步骤 102)。

4、根据权利要求 3 所述的基于满足 K 度稀疏约束的深度学习模型的大数据处理方法, 所述方法中步骤 2) 的具体过程为:

将更新后的训练样本输入所述满足 K 度稀疏约束的深度学习模型, 当输入的训练样本为无标注样本集 $\mathbf{Y} = \{\mathbf{x}_i^*\}$ 时, 将输入样本得到输出再逆转为输入, 在 K 度约束下自输出层向输入层进行反向重构, 计算重构误差 E_r , 以梯度下降或共轭梯度下降的方式调整各层权重; 直至误差小于临界值; 当输入的训练样本为有标注的样本集 $\mathbf{X} = \{\mathbf{x}_i^*, \mathbf{y}_i^*\}$, 则将输出与 $\mathbf{y}_i^*$ 相比较, 计算出训练误差 E_t , 以梯度下降或共轭梯度下降的方式调整正向

权重；直至误差小于临界值。

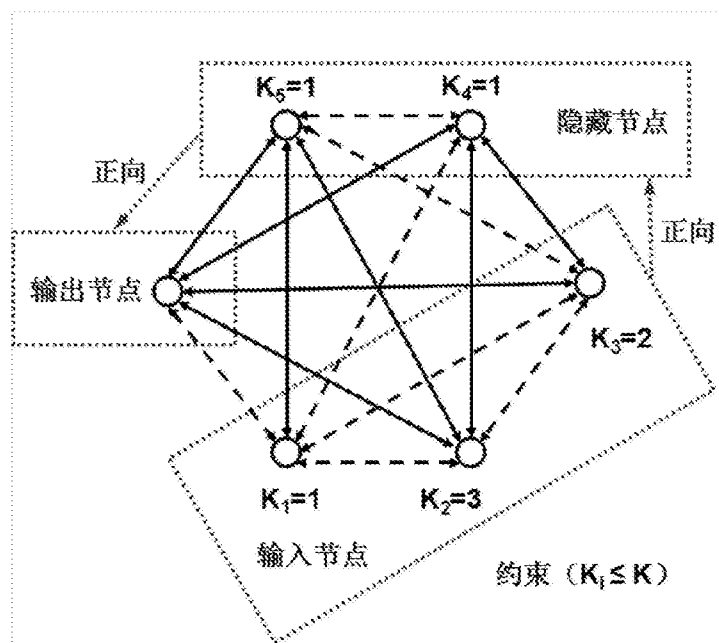


图 1

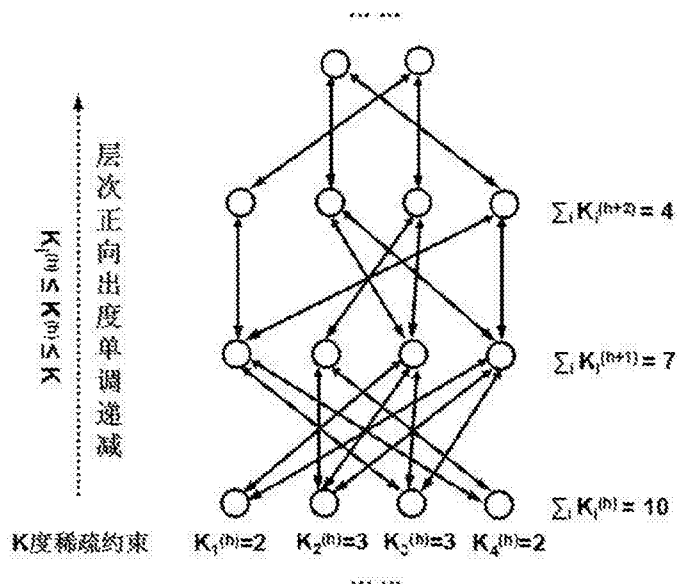


图 2

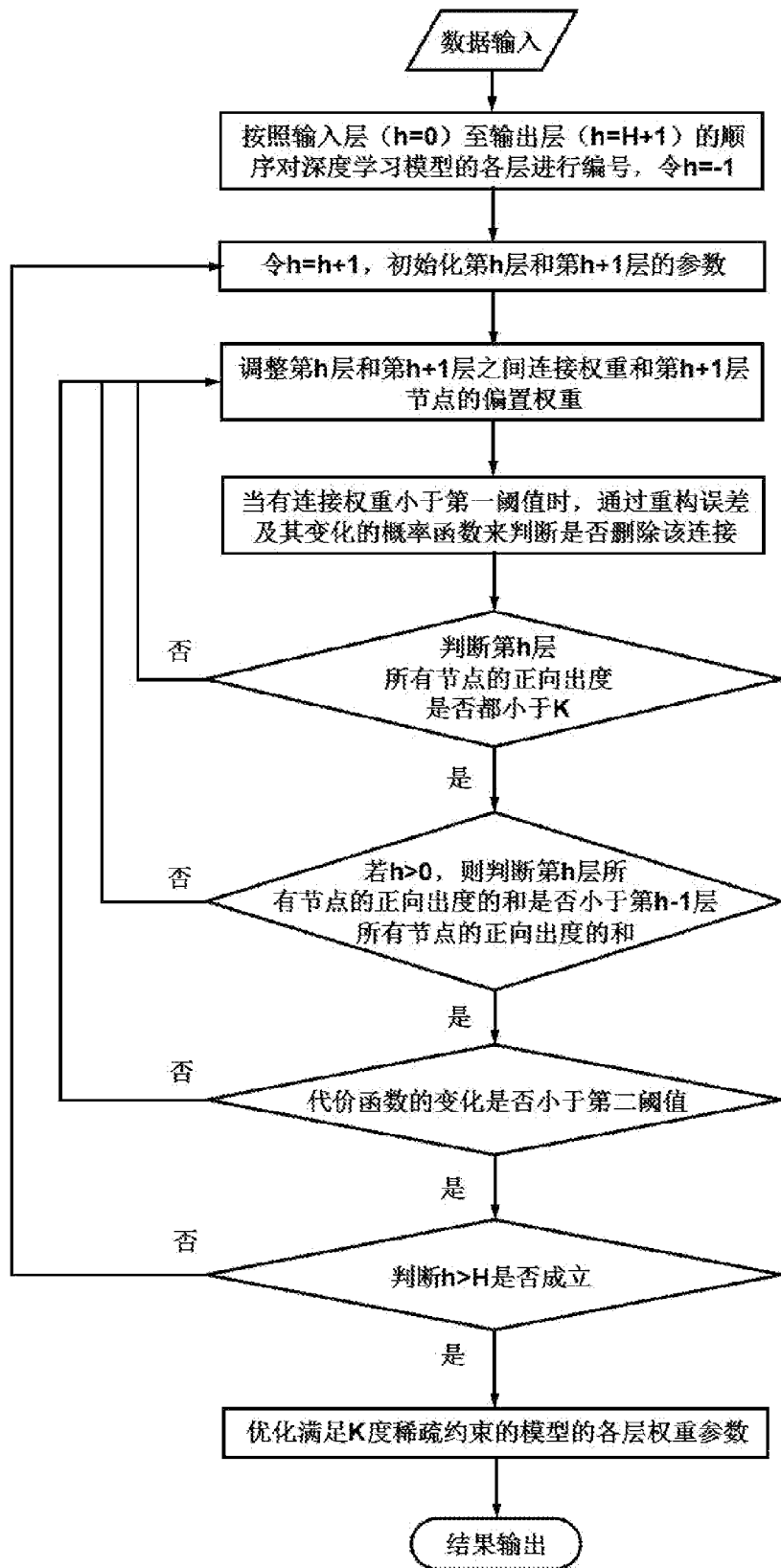


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2015/075473**A. CLASSIFICATION OF SUBJECT MATTER**

G06N 5/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06N, G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, DWPI, SIPOABS: train, depth, belief network, sparsity constraint, big data, process, learn, model, deep, belief, confidence, net, sparsity constraint, sample, example, optimize, parameter, node, layer, level, hierarchical

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 103530689 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES), 22 January 2014 (22.01.2014), the whole document	1-4
A	CN 103838836 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES), 04 June 2014 (04.06.2014), the whole document	1-4
A	CN 103489033 A (NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY), 01 January 2014 (01.01.2014), the whole document	1-4

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
07 December 2015 (07.12.2015)Date of mailing of the international search report
23 December 2015 (23.12.2015)Name and mailing address of the ISA/CN:
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No.: (86-10) 62019451Authorized officer
ZHAN, Qianqian
Telephone No.: (86-10) **62089525**

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2015/075473

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103530689 A	22 January 2014	None	
CN 103838836 A	04 June 2014	None	
CN 103489033 A	01 January 2014	None	

国际检索报告

国际申请号

PCT/CN2015/075473

A. 主题的分类

G06N 5/00(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06N, G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, DWPI, SIPOABS :大数据, 处理, 学习, 训练, 模型, 深度, 置信网, 稀疏约束, 样本, 优化, 参数, 节点, 层; big data, process, learn, model, deep, belief, confidence, net, sparsity constraint, sample, example, optimize, parameter, node, layer, level, hierarchical

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 103530689 A (中国科学院自动化研究所) 2014年 1月 22日 (2014 - 01 - 22) 全文	1-4
A	CN 103838836 A (中国科学院自动化研究所) 2014年 6月 4日 (2014 - 06 - 04) 全文	1-4
A	CN 103489033 A (南京理工大学) 2014年 1月 1日 (2014 - 01 - 01) 全文	1-4

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2015年 12月 7日

国际检索报告邮寄日期

2015年 12月 23日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

詹芊芊

电话号码 (86-10)62089525

国际申请号	PCT/CN2015/075473
-------	-------------------

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	103530689	A	2014年 1月 22日	无	
CN	103838836	A	2014年 6月 4日	无	
CN	103489033	A	2014年 1月 1日	无	

表 PCT/ISA/210 (同族专利附件) (2009年7月)