

基于深度置信网络的互联网流量预测方法

孟智慧

(中国移动通信集团设计院有限公司山西分公司, 太原 030000)

摘 要 本文提出了基于深度置信网络的3种不同架构的网络流量预测模型来预测互联网流量。首先, 介绍了深度置信网络的网络结构, 然后构建了3种不同架构的深度置信网络拓扑结构, 最后通过实验对比, 发现隐层的神经元数量对更深层次的网络至关重要, 该模型被证明是一种有效的预测模型。本文所采用的方法在模拟流量数据模式和随机要素的同时, 提供了准确的网络流量预测, 使测试数据集的均方根误差值为0.028。

关键词 深度学习; 深度置信网络; 互联网流量预测

中图分类号 TN915

文献标识码 A

文章编号 1008-5599 (2020) 10-0042-06

DOI:10.13992/j.cnki.tetas.2020.10.007

随着互联网的快速发展和网络资费的下降, 网络流量的规模也直线式上升, 用户对网络资源的需求也逐渐提高, 因此精确规划网络资源使用已成为运营商的首要任务。网络资源的合理分配不仅能充分的利用网络资源, 还能保障用户的使用质量。而网络流量负荷分析在资源优化以及网络服务^[1]的设计、管理和控制中起着至关重要的作用。因此, 网络流量预测对于更熟练的流量工程和异常检测工具的发展是非常重要的。另一方面, 为未来流量预测而开发的模型需要支持检测数据网络中的异常。此外, 通过将实际流量与预测模型的预测结果进行对比, 可以识别出不规律的垃圾信息元素数量和其它安全攻击。

流量建模对网络评估是很重要的, 针对这一问题, 国内外许多专家学者对此进行了研究, 参考文献[2]研究证实了深度神经网络模型在时间序列预测和预测中的

应用; 参考文献[3]采用了叠加式自编码器的方法, 建立了深度学习模型, 对互联网流量进行了预测; 参考文献[4]基于ABC+AFS-LSSVM的模型对网络流量进行了预测; 参考文献[5]把面向自然语言处理提出的全注意力机制引入到时间序列预测问题中, 预测未来分钟级时间窗内秒级的流量; 参考文献[6]利用网络流量历史数据, 引入数据挖掘技术中的状态回声网络对子训练样本集合的网络流量进行预测; 参考文献[7]基于改进的回声状态网络对网络流量进行了预测。这些模型只捕获时间序列中出现的线性相关结构, 而不考虑识别任何非线性模式。

本文重点研究了深度置信网(DBN)的拓扑结构, 使其具有较高的预测精度。所采用的方法在模拟流量数据模式和随机要素的同时, 提供了准确的流量预测, 使测试数据集的均方根误差(RMSE)值为0.028。

收稿日期: 2020-05-11

1 理论背景

由于其对不可预见情况的泛化能力,神经网络被应用在许多领域。利用适当的学习算法对神经网络的输入进行训练,使其达到特定的目标输出,直到神经网络的输出趋向于与给定的目标匹配为止。通过选择网络的参数,可以推导出给定输入和输出之间的非线性关系。为了提高人工神经网络的计算效率,其计算能力需要每层增加更多的神经元或增加隐层的数量,这就形成了神经计算模型的深层结构,其中中间层或隐层作为多层抽象层被执行。就神经网络的深层结构而言,大量的研究正在进行。

与之前浅层的神经网络架构相比,深度神经网络架构能学习到更加复杂的问题,功能也更加强大^[8]。根据隐层的深度,深度神经网络包含了许多非线性层次。深层的层次结构使它们能够有效地表示高度非线性的模式和高度变化的功能抽象。尽管如此,由于网络参数的随机初始化常常陷入较差的解中^[9],如何训练这样的深度网络还不清楚。为了解决这一问题,引入一种贪婪的学习算法^[10],每次只训练一层网络,而深度架构则通过RBM (Restricted Boltzmann Machines, 限制玻尔兹曼机) 网络叠加的深度置信网络来实现。

DBN 是所有深度学习模型中最常见、最有效的方法。DBN 是由 RBM 叠加而成的。RBM 是一个基于能量的概率模型,它依赖于由可见和隐藏节点组成的两层结构,如图 1 所示。

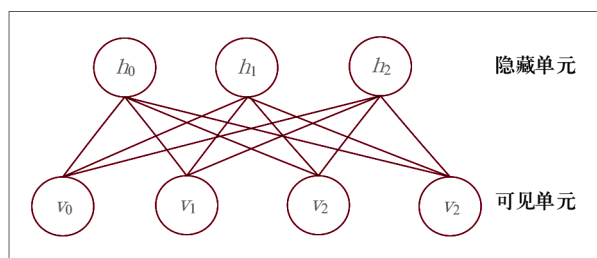


图1 RBM的图形结构

这些节点是二进制单元,即 h 和 $v \in \{0,1\}$, 显层

和隐层内部的神经元都没有互联,只有层间的神经元有对称的连接线,在给定所有显元的值的情况下,每一个隐元取值是互不相关的。

$$p(h_{i=1}|v) = \sigma(\sum_{j=1}^m w_{ij} v_j + c_i) \quad (1)$$

$$p(v_{j=1}|h) = \sigma(\sum_{i=1}^n w_{ij} h_i + b_j) \quad (2)$$

其中, w_{ij} 代表从第 i 个显元到第 j 个隐元的权重, m 代表显元的个数, n 代表隐元的个数,而 b_j 和 c_i 分别是与第 j 个可见变量和第 i 个隐藏变量相关的实值偏置项。RBM 中的采样是使用吉布斯采样,通过运行一个收敛的 Markov 链来获得的,权值参数根据变化率进行更新。

$$\Delta w_{ij} = \varepsilon (\langle v_j h_i \rangle_{data} - \langle v_j h_i \rangle_{model}) \quad (3)$$

2 模型构建

本文提出了 3 种不同架构的 DBN 深度置信网络模型,对未来 1 h 的网络流量数据进行时间序列预测。深度神经网络隐层的输入、输出变量和结构的一般构造如图 2 所示。

在图 2 中可以看出 DBN 模型隐层分为 4 层,每一层的神经元数目分别为 300 个、200 个、100 个、10 个,通过输入变量为 6 个神经元的输入层和 1 个神经元的输出层来预测未来的流量负荷。不能任意选择网络的拓扑结构,即网络的深度和隐层的大小,参考文献 [11] 在考虑隐层的宽度选择时,强调了 3 种不同的拓扑结构来进行模型开发。这表明,模型拓扑的最佳选择是保持隐层的宽度在整个模型中保持不变,或者按大小的递增或递减顺序保持不变。因此,模型架构是基于从下到上减少隐层的大小,通过蒙特卡罗模拟^[12] 计算第一隐层神经元数目范围为 100 ~ 600 个时整个模型的效果最好,因此在本文中第一层隐层神经元数目取中间值 300 个。

DBN 的每一层都被单独训练成一个具有 sigmoid 激活函数的 RBM, DBN 的首选设置见表 1。最初,靠近输入层的第 1 个隐层以无监督的方式进行训练,为神经网络的第 1 层构造一组初始的参数值。然后,从这一

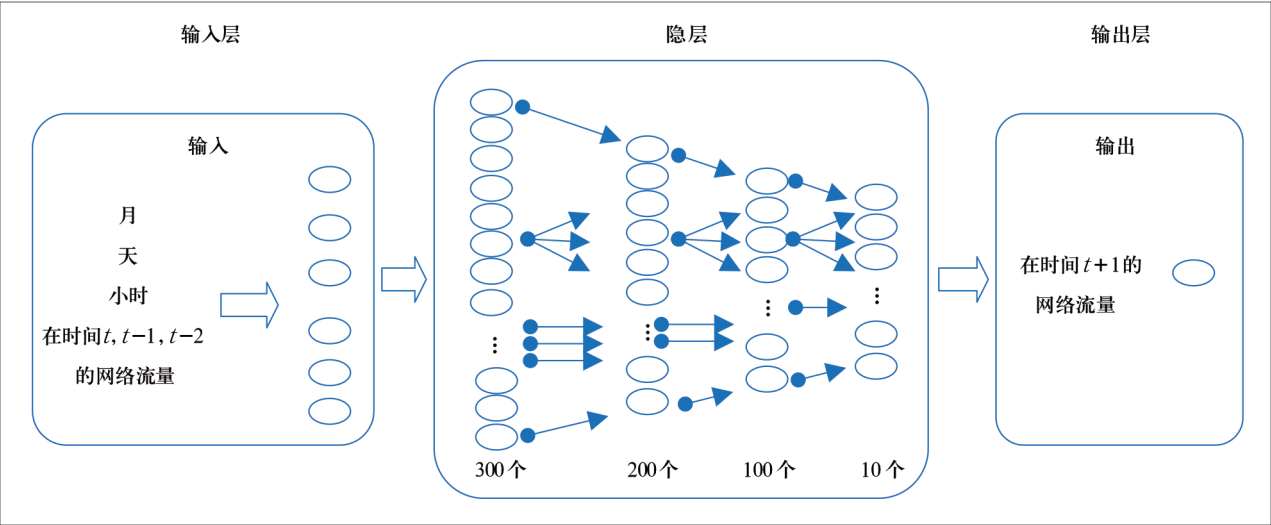


图2 基于DBN的模型1的体系结构

层获得的输出作为 1 个新的特征集，作为下一层的输入。这些层是建设性地添加和独立地训练一个接一个，这个过程被称为贪婪分层预训练。为了获得更多的非线性表示，这个过程可以重复任意次数。这提供了神经网络的参数初始化。DBN 中的 RBM 层一旦经过训练，就会添加一个具有成本函数的输出层。因此，利用监督算法对网络进行全局微调来预测目标，通过增加线性输出层，实现 500 次迭代对网络进行训练。

表1 DBN的参数设置

参数	数值
最大迭代次数（次）	100
初始动量	0.5
最终动量	0.9
学习率	0.1
批量大小	5
辍学率	0.0
转换函数	Sigmoid

为了进行对比分析，本文又研究了两个 DBN 模型。第 2 个模型每个隐层神经元数目都为 300 个节点，第 3 个模型在更高层用更少的隐藏神经元，但这两种模型的深度和参数设置和第 1 个模型是一样的。而模型依赖的区别只在隐层的神经元个数。模型 2 和模型 3 的隐层拓扑如图 3 所示。

3 实验

3.1 数据集

数据集使用英国学术主干网的网络流量数据集，该数据集由每小时从互联网服务提供商记录的 1 600 个样本组成。数据集包含两个属性，一个是记录在后台智能传输服务的流量负载值，另一个是每个记录的特定时间间隔。为了提高预测的准确性，首先将输入集进一步细分为小时、月、日、 t 、 $t-1$ 和 $t-2$ 时刻的互联网流量，通过自相关函数计算作为输入参数的滞后数量。然后通过数据归一化方法对数据进行预处理，将每个输入变量转换成 0 ~ 1 范围内的小数，使算法能够发挥最佳效果。最后选取前 1 200 个样本对模型进行训练，其余数据作为测试集。

3.2 衡量标准

本文采用衡量结果的指标是均方根误差（RMSE）和误差率（ER）。

$$RMSE=\frac{1}{T}\sum_{t=1}^T\sqrt{\frac{1}{m}\sum_{i=1}^m[\log(\hat{X}_{C_{i,t}}+1)-\log(X_{C_{i,t}}+1)]^2}$$
 (4)

$$ER=\frac{1}{T}\sum_{t=1}^T\frac{\sum_{i=1}^m|\hat{X}_{C_{i,t}}-X_{C_{i,t}}|}{\sum_{i=1}^mX_{C_{i,t}}}$$
 (5)

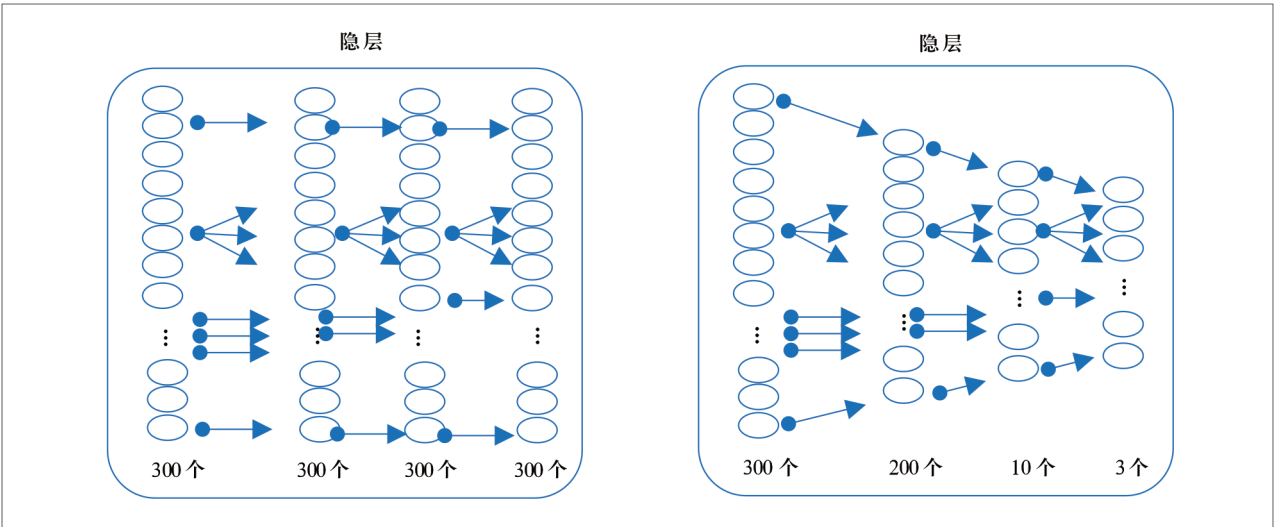


图3 模型2和模型3的隐层拓扑视图

3.3 结果

各模型预测误差值见表 2，模型 3 的效果最差，测试误差大于训练误差。这是因为顶部隐层的尺寸很小。从表 2 的误差值可以明显看出，模型 1 和模型 2 的误差估计值非常接近。这就证明了在更深层次的网络中适当的隐藏神经元，无监督的前训练就像一个数据依赖的正则化矩阵。

表2 模型预测误差

DBN	RMSE		ER	
	训练	测试	训练	测试
模型 1	0.033 4	0.029 6	0.030 0	0.028 6
模型 2	0.032 9	0.032 8	0.031 9	0.031 0
模型 3	0.031 1	0.041 6	0.029 8	0.035 8

在模型 1 中实现的 DBN 的拓扑比其它两个执行得更好，本文重点研究了该模型的结果。互联网流量数据序列的结构和非线性趋势如图 4 所示，从输入层开始的每一层无监督训练期间的 RMSE 值如图 5 所示。

如图 6 所示，以直方图的形式展示了模型 1 对训练数据集估计的误差分布。可以看出，直方图是对称的，误差大多集中在 0 附近。由模型 1 估计的流量负荷预测，以及原始流量值如图 7 所示。该模型对 400 个样品的测

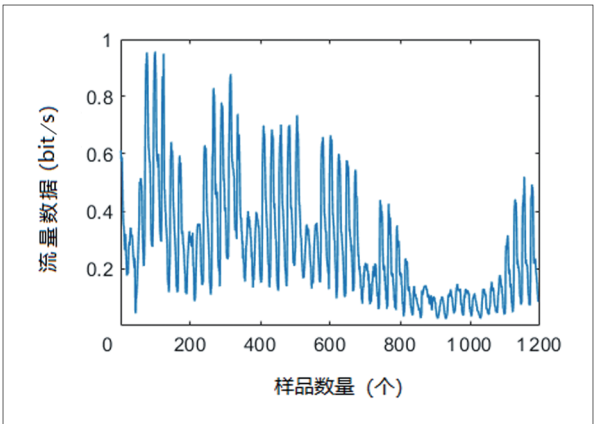


图4 流量数据图

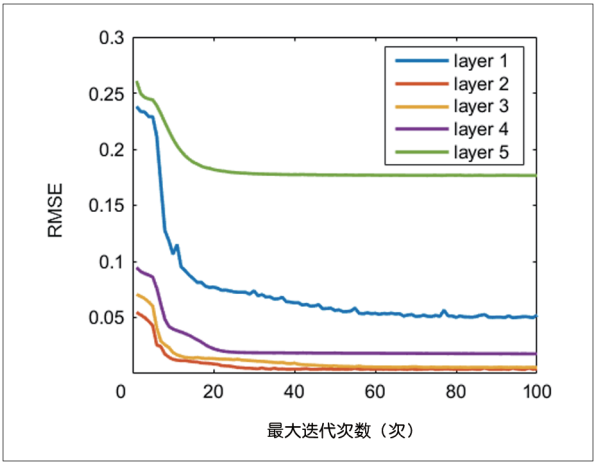


图5 无监督训练RMSE值

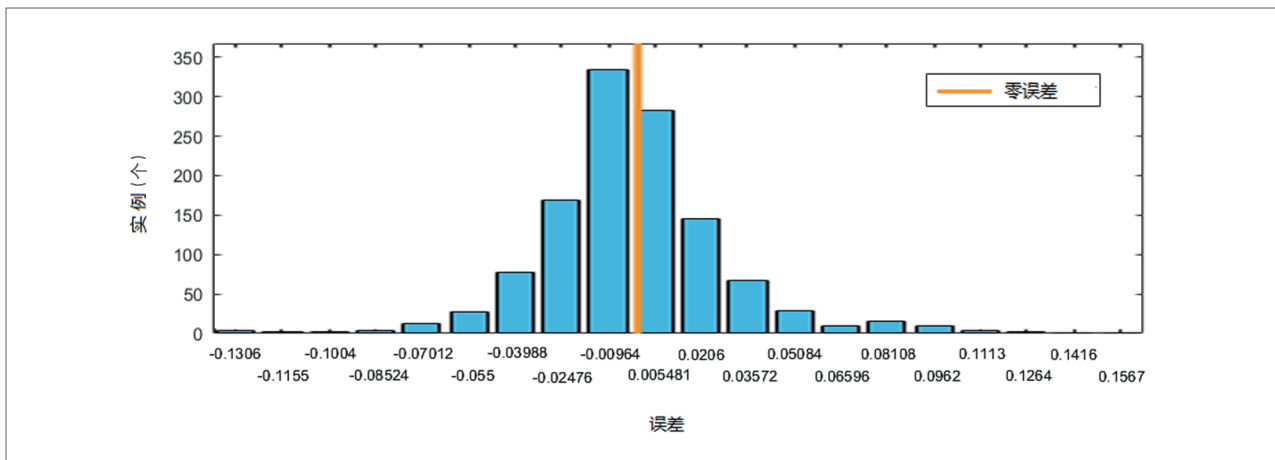


图6 训练数据误差直方图

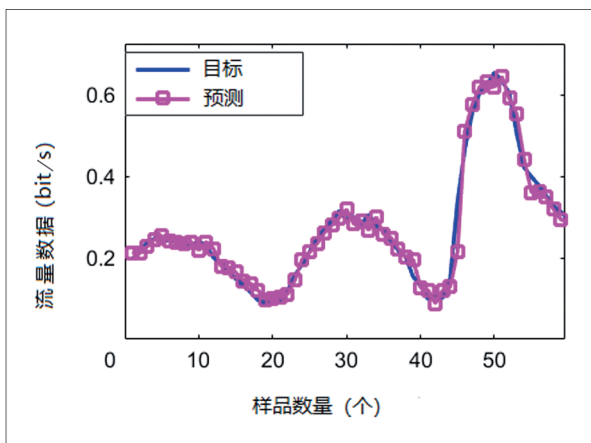


图7 由模型1预估计的预测样本

试数据的 RMSE 为 0.029 6。误差率为 0.0286, 这说明本文的预测方法预测效果较好, 预测结果是可信的。

4 结束语

本文介绍了 DBN 在网络流量预测中的应用, 并研究了深度神经网络中隐层宽度的选择准则。为了确定合适的拓扑结构, 提出了 3 种不同的深层结构预测模型, 根据他们的对比分析, 发现隐层的神经元数量对更深层次的网络至关重要。随着隐层宽度的减小, 该模型被证明是一种有效的预测模型。结果表明, 模型 1 和模型 2

能够准确地捕获网络流量的显著特征。未来, 将进一步扩展所采用的模型, 以产生更好的预测结果。

参考文献

- [1] FENG H, SHU Y. Study on network traffic prediction techniques[C]. International Conference on Wireless Communications (IEEE), 2005.
- [2] DALTO M. Deep neural networks for time series prediction with applications in ultra-short-term wind forecasting[C]. International Conference on Industrial Technology (ICIT), 2015.
- [3] OLIVERIRA T P, BARBAR J S, SOARES A S. Multilayer perceptron and stacked autoencoder for internet traffic prediction[J]. In IFIP International Conference on Network and Parallel Computing, 2014(1).
- [4] 孙群, 袁宏俊. 基于ABC+AFS-LSSVM的网络流量预测模型[J]. 淮阴师范学院学报(自然科学版), 2019(2).
- [5] 郭佳, 余永斌, 杨晨阳. 基于全注意力机制的多步网络流量预测[J]. 信号处理, 2019(5).
- [6] 郭海蓉. 大数据环境下的网络流量非线性预测建模[J]. 微型电脑应用, 2019(8).
- [7] 杨欣颜. 基于改进ESN的网络流量预测方法研究[D]. 南京: 南京邮电大学, 2019.
- [8] ROMEU P, PARDO J. Time-series forecasting of indoor temperature using pre-trained deep neural networks[C]. International Conference on Artificial Neural Networks, 2013.
- [9] BENGIO Y, LAMBLIN P, POPOVIVI D, et al. Greedy layer-wise training of deep networks[C]. Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems, 2007.
- [10] HIOTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural computation, 2006(7).
- [11] BENGIO Y. Learning deep architectures for AI[J]. Foundations & Trends in Machine Learning, 2009(1).
- [12] 宾德, 赫尔曼. 统计物理学中的蒙特卡罗模拟方法[M]. 北京: 北京大学出版社, 1994.

Internet traffic prediction method based on deep belief network

MENG Zhi-hui

(China Mobile Group Design Institute Co., Ltd. Shanxi Branch, Taiyuan 030000, China)

Abstract This paper proposes three different network traffic prediction models based on deep belief network to predict the Internet traffic in the next 1 hour. First, the network structure of deep belief network is introduced. Then, the deep belief network topology of three different architectures was constructed. Finally, through experimental comparison, it was found that the number of neurons in the hidden layer was crucial to the deeper level of the network. This model proved to be an effective prediction model. The method adopted in this paper provides accurate network traffic prediction while simulating traffic data patterns and random elements, so that the root mean square error value of the test data set is 0.028.

Keywords deep learning; deep belief network; Internet traffic prediction

News

云智融合 揽胜九天 中国移动云智融合峰会在苏州召开

近日,由中国移动政企事业部、技术部主办,中国移动云能力中心、研究院承办,中国移动科协协办的中国移动云智融合峰会在苏州召开,大会通过移动云官网平台直播,同时在线上线下举行。本次峰会以“云智融合,揽胜九天”为主题,来自全国各大头部厂商、高等院校的云计算、人工智能顶尖人才齐聚苏州,聚焦人工智能前沿科技,探讨“智慧云”发展方向,共创云智融合行业价值。中国工程院院士潘云鹤,中国移动通信集团有限公司副总经理高同庆,江苏省工信厅副厅长胡学同,苏州市副市长陆春云,苏州市高新区党工委书记、虎丘区委书记方文浜等领导嘉宾出席峰会。中国移动通信集团有限公司党组书记、董事长杨杰致信祝贺峰会顺利召开。

在当天召开的中国移动云智融合峰会上,中国移动正式发布“融智”战略。新的时代,经济社会正在加速向数字化、网络化、智能化的转型,并呈现出“五纵三横”的新特征。智能化作为重要的赋能方式,将深度融合到基础设施、社会治理、生产方式、工作方式、生活方式中,融智已经成为战略性创新要素和驱动要素,将对经济发展、社会进步、全球治理等方面产生重大而深远的影响。中国移动通过打造融智引擎、网智融合、云智融合、企智融合、产智融合深入推进“融智”创新发展。通过融智加速业务创新,实现多元丰富的智慧生活,让人民更幸福;通过融智提升城市治理水平,实现社会智慧治理,让城市更美好;通过融智促进行业转型,赋能千行百业,让产业更繁荣。

会上,移动云品牌口号由“5G时代,你身边的云”升级为“5G时代,你身边的智慧云”,与会专家对移动云“智慧云”的品牌形象作了进一步诠释。新添加的“智慧”二字虽然简短,却已明确传达出移动云未来将在云智融合方面集中发力,以期赋能百业的展望与决心。

移动云的“智慧化”,离不开强有力的人工智能平台。中移九天人工智能平台是中国移动积极响应国家科技和网络强国战略自主研发的人工智能领域创新成果,是中国移动“5G+AI CDE”战略落地的重要引擎。

当天,中国移动继去年在全球合作伙伴大会上发布万象计划,提出构建开放云生态以来,首次对外公开展示移动云AI生态布局的最新成果。

(来源:中国移动通信集团有限公司官网)