·先进计算与数据处理· 文章编号：1000-3428(2021)02-0126-07 文献标志码：A 中图分类号：TP391

面向高能物理分级存储的文件访问热度预测

程振京[1,2]，汪 璐[1,2]，程耀东[1,2,3]，陈 刚[1]，胡庆宝[1]，李海波[1,2]
（1.中国科学院高能物理研究所，北京 100049；2.中国科学院大学，北京 100049；
3.中国科学院高能物理研究所天府宇宙线研究中心，成都 610041）

摘 要：高能物理计算是典型的数据密集型计算，其主要采用基于文件的分级存储方案，根据访问热度的不同将数据存储于不同性能的存储设备上，然而当前数据热度预测采用基于人工经验的启发式算法，准确率较低。提出一种借助长短期记忆网络预测文件未来访问热度的方法，包括网络结构设计、训练和预测算法等。该方法通过划分动态时间窗口构造文件访问特征的时序序列，预测不同数据的访问趋势。在LHAASO高能物理实验数据集上的实验结果表明，与SVM、MLP等算法相比，该方法预测准确率提升了30%左右，具有更强的适用性。

关键词：分级存储；文件访问特征；时序数据；长短期记忆网络；文件访问热度

开放科学(资源服务)标志码(OSID)：

中文引用格式：程振京，汪璐，程耀东，等.面向高能物理分级存储的文件访问热度预测[J].计算机工程，2021，47(2)：126-132.
英文引用格式：CHENG Zhenjing，WANG Lu，CHENG Yaodong，et al.File access popularity prediction for hierarchical storage for high-energy physics[J].Computer Engineering，2021，47(2)：126-132.

File Access Popularity Prediction for Hierarchical Storage for High-Energy Physics

CHENG Zhenjing[1,2]，WANG Lu[1,2]，CHENG Yaodong[1,2,3]，CHEN Gang[1]，HU Qingbao[1]，LI Haibo[1,2]
（1.Institute of High Energy Physics，Chinese Academy of Sciences，Beijing 100049，China；
2.University of Chinese Academy of Sciences，Beijing 100049，China；
3.Tianfu Cosmic Ray Research Center，Institute of High Energy Physics，Chinese Academy of Sciences，Chengdu 610041，China）

【Abstract】Computing for high-energy physics is typically data-intensive. It mainly adopts file-based hierarchical storage solutions where data is allocated based on the access popularity to storage devices with different performances. The existing schemes of data popularity prediction generally adopt a heuristic algorithm based on artificial experience，whose prediction accuracy is low.This paper proposes a method of predicting future access popularity using Long Short-Term Memory（LSTM）network，which consists of network structure design，training，and prediction algorithms. The method divides the dynamic time window to construct a time series of file access features，and on this basis predicts the access trends of different data.Experimental results on the data set of LHAASO high-energy physics experiments show that compared with SVM，MLP and other algorithms，the proposed method increases the prediction accuracy by about 30%，and it has stronger applicability.

【Key words】hierarchical storage；file access characteristics；time series data；Long Short-Term Memory（LSTM）network；file access popularity
DOI：10. 19678/j. issn. 1000-3428. 0057273

0 概述

随着高海拔宇宙线观测实验LHAASO[1]建成运行，数据累积规模不断扩大，对数据存储的性能和效率提出了更高的要求。LHAASO实验使用统一命名空间的分级存储系统来存储物理数据，使用的介质包括固态硬盘、机械硬盘和磁带等。Lustre[2]和EOS[3]两个主要的存储管理系统均提供分级存储功能，目前一般采用的是基于系统管理员个人经验的启发式算法，如LRU等本质上是以单一的文件访问

基金项目：国家重点研发计划(2017YFB0203200)；国家自然科学基金(11675201，11805226，11805223)。
作者简介：程振京(1993—)，男，博士研究生，主研方向为分布式存储、机器学习；汪 璐，副研究员、博士；程耀东、陈 刚，研究员、博士；胡庆宝，硕士；李海波，副研究员、博士。
收稿日期：2020-01-20 **修回日期：**2020-02-28 **E-mail：**chengzj@ihep.ac.cn

特征（迁入上级存储时间、访问频率等）设一个阈值。因为这些算法需要在操作系统内核中运行，必须牺牲一些预测精度来提升执行效率，所以存在经验偏差，缺少负载通用性和自适应性。为提升存储系统的资源利用率，按访问热度不同将高能物理数据存储于不同性能、不同容量的存储设备上，根据数据热度改变迁移至合适的存储层级。因此，数据未来的访问热度预测对于设计高效的数据迁移机制十分重要。

高能物理计算主要包括原始实验数据的蒙特卡洛模拟、数据重建以及物理分析等过程，每种计算类型各有其特点，每个用户/应用的计算模式也有可能存在巨大差异。通常情况下在一个较短的连续时间段内，物理学家通常只会专注于分析整体重建数据的一小部分，其他绝大部分重建数据不会被访问。在存储系统中表现为同一文件的访问热度并不是一成不变的，同时访问特征随时间变化。本文研究在高能物理实验LHAASO的真实分级存储中，根据文件访问特征的变化预测访问热度的变化，基于长短期记忆（Long-Short Term Memory，LSTM）神经网络算法，训练一个有监督学习的文件访问热度预测模型，并与传统SVM模型、MLP模型进行对比验证。

1　相关工作

1.1　数据访问预测

数据访问预测和迁移策略一直是存储系统的重要研究领域。文献[4]综述了LRU、CLOCK、2Q、GDSF、LFUDA和FIFO等预测算法驱动的文件迁移策略。这些策略主要用于解决单机环境下数据在易失性主存和外部磁盘间的缓存替换问题，其本质是以某一个单一的文件访问特征（如最后访问的时间等）为阈值，设定启发式的迁入迁出规则。因为需要在操作系统内核中运行，必须在预测精度和执行效率之间做出权衡，不可能利用复杂的文件访问特征或设计非常复杂的算法来辅助预测。文献[5]介绍了LNS模型，在传统Last Successor算法的基础上加入用户信息来提升数据访问预测精度，但同样严重依赖文件访问顺序。文献[6]提出了利用预测文件未来访问热度来制定迁移规则的方法，该方法假设每次用户都会完整、顺序地读完整个文件，然后训练一个基于支持向量机（SVM）算法的监督学习模型来执行预测任务，在特定Web数据集上取得了良好的预测效果，但由于SVM是借助二次规划来支持向量的，模型训练需要耗费大量的CPU运算时间。

近年来，深度学习技术逐渐兴起，训练方法与传统算法相比有很大区别，从而突破了传统神经网络对隐藏层数和每层节点数量的限制，具有很强的自学习和非线性映射能力。在各种深度神经网络模型中，循环神经网络（Recurrent Neural Network，RNN）在网络结构设计上引入了时序的概念，同时结合具有存储能力的网络节点，从而使模型像人一样拥有记忆[7]。循环神经网络能够对输入时序信号逐层抽象并提取特征[8]，目前在语音识别[9]、机器翻译[10]、电力负荷预测[11]、故障预测[12]等领域的时序数据建模中取得了较多的突破，然而在数据访问预测方面应用有限，特别是针对分级存储的数据访问热度预测，目前还未发现类似的研究案例。

1.2　长短期记忆神经网络

由于高能物理数据处理模式的特点，文件访问具有一定的时间特性，和循环神经网络的记忆机制比较契合。因此，循环神经网络在处理访问热度预测问题上具有独特优势。

原始的循环神经网络存在梯度消失和梯度爆炸的问题，长期记忆能力一般，很难学习序列中长期依赖的信息。SCHMIDHUBER等人提出了长短期记忆人工神经网络（LSTM），重新设计了循环神经网络中的计算节点。LSTM使用时间记忆单元用以记录当前时刻的状态，一般称为长短期记忆神经网络的细胞[13]。与每个细胞相连的有遗忘门、输入门和输出门3个信息传递开关门，如图1所示。

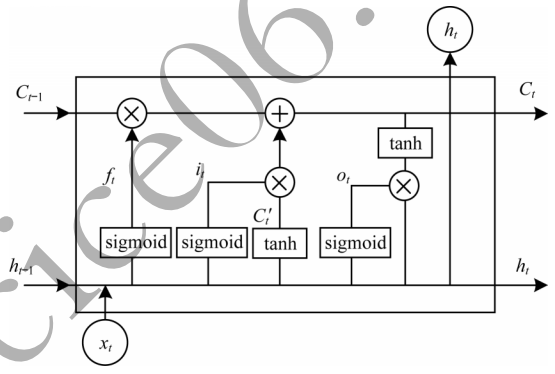

图1　LSTM神经元结构

Fig.1　LSTM neuron structure

信息开关门可以选择性地让信息通过，遗忘门决定某个时刻的序列数据通过时从细胞中丢弃什么信息，输出一个在0到1之间的数值给每个细胞状态C（0代表完全舍弃，1代表完全保留）。其中，h表示LSTM细胞的输出，x表示LSTM细胞的输入。

$$f_t=\mathrm{sigmoid}\left(W_f\cdot\left[h_{t-1},x_t\right]+b_f\right) \tag{1}$$

输入门决定多少新信息被存储在LSTM细胞中。输入门包含两个处理层次，sigmoid层决定细胞状态中什么值应被更新，tanh层创建一个新的候选值向量$\boldsymbol{C}'$。

$$i_t=\mathrm{sigmoid}\left(W_i\cdot\left[h_{t-1},x_t\right]+b_i\right) \tag{2}$$

$$\boldsymbol{C}'_t=\tanh\left(W_c\cdot\left[h_{t-1},x_t\right]+b_c\right) \tag{3}$$

LSTM细胞状态更新为原始细胞状态丢弃部分信息后，再加上新的候选值向量$\boldsymbol{C}'$的和。

$$\boldsymbol{C}_t = f_t \cdot C_{t-1} + i_t \cdot \boldsymbol{C}'_t \tag{4}$$

输出门基于更新后的LSTM细胞状态，通过一个sigmoid层确定将细胞状态的哪个部分输出。细胞状态通过tanh层后和sigmoid输出相乘。

$$O_t = \text{sigmoid}\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{5}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{6}$$

输入门、输出门和遗忘门三类门共同控制信息流入和流出，以及LSTM细胞状态的更新，因此LSTM模型擅长挖掘时间序列内前后间隔较长的依赖关系，适合预测有时间间隔的延迟事件。传统LSTM模型由于隐藏层需要保留所有的时间序列信息，模型收敛效果容易受序列长度影响，随着序列增长而降低。引入注意力机制能从序列中学习到每一个元素或者事件的重要程度，以后再从相似的场景中学习时，可以把LSTM模型的注意力专注于对预测结果最有意义的部分，从而提高模型效率。

2 访问热度预测

如图2所示，数据访问热度预测系统与现有的Lustre、EOS等高能物理存储系统交互，由特征收集节点、中心数据库、模型训练节点等组成。在每个文件存储服务器FST上部署I/O日志采集组件，过滤掉无关信息后以<时间戳，参数字段，数值>的格式存放在中心key-value数据库中。文件访问特征数据经过计算整合、归一化和批处理，写入模型训练的在线数据队列。模型训练基于Tensorflow[14]、sklearn[15]等深度学习框架，训练后的模型结构存放在本地文件系统进行持久化存储。高能所计算中心的数据迁移系统周期性地在后台扫描磁带、机械硬盘、固态硬盘中的文件列表，根据文件访问预测系统的输出和管理员预先设置的迁移条件，执行迁移动作。

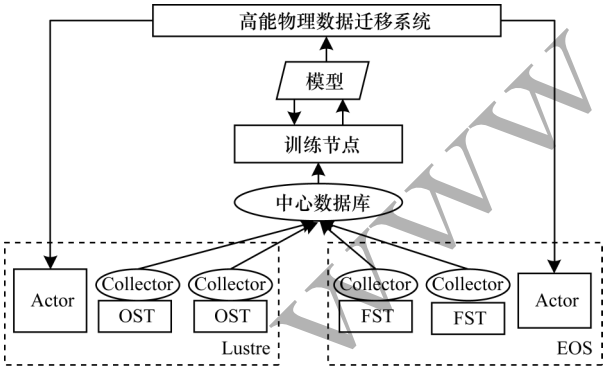

图2 数据迁移管理系统结构

Fig.2 Structure of data migration management system

2.1 模型输入

训练预测模型需要输入大量训练样本。在通常情况下，高能物理存储系统中文件访问与不同用户、不同应用的计算模式都有关系。高能物理Lustre和EOS存储系统提供以文件名为单位的历史访问日志，记录文件创建、打开与关闭操作以及每一次文件指针移动和数据读写等。本文计算各类文件操作的均值和方差，以衡量在时间轴上的离散趋势变化，挖掘文件访问的时间特性，并按照合适的时间窗口整理成文件的时序访问特征序列，如图3所示。在图3中，T为时间戳，C为访问类别，B为读写字节数，N为读写指标移动次数，S为文件大小。

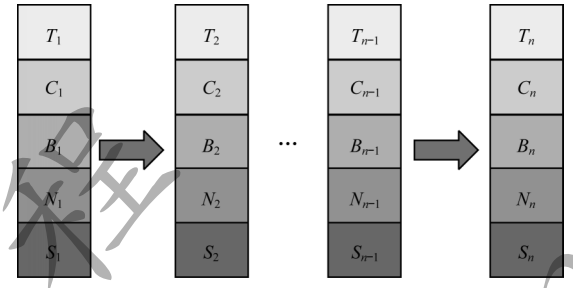

图3 文件访问特征时序序列

Fig.3 File access feature time series sequences

文件访问定义如下：

1）numAccesses：文件的访问次数。一次完整的文件访问包含文件打开、多次读取与写入数据、多次文件指针移动和文件关闭等过程。

2）avgReads：一次访问中读数据次数的均值。

3）stdDevReads：一次访问中读数据次数的均方差。

4）avgBytesRead：一次访问中读数据大小（字节）的均值。

5）stdDevBytesRead：一次访问中读数据大小（字节）的均方差。

6）avgWrites：一次访问中写数据次数的均值。

7）stdDevWrites：一次访问中写数据次数的均方差。

8）avgBytesWritten：一次访问中写数据大小（字节）的均值。

9）stdDevWritten：一次访问中写数据大小（字节）的均方差。

10）avgSeeks：一次访问中文件指针移动次数的均值。

11）stdDevSeeks：一次访问中文件指针移动次数的均方差。

在构建访问特征向量时，需要在海量文件系统日志中筛选出各类文件操作记录并进行持久化存储。高能所计算中心的EOS存储系统具有千万级的文件数目和PB级的数据，每天记录数十万条文件访问日志，需要按照文件名、操作类型和时间窗口3个维度进行整理。本文使用面向列的key-value分布式数据库HBase存储高能物理文件访问特征。Rowkey设计如表1所示。

表1　文件访问I/O数据的Rowkey设计

Table 1　Rowkey design of file access I/O data

Rowkey	字段名称	取值范围
第0字节	文件访问类型	0~15
第1字节~16字节	文件名散列	$0\sim2^{128}$
第17字节~20字节	linux时间戳(ms)	大于0的整数
第21字节~23字节	扩展字段	无

Rowkey第0字节为文件操作类型字段，例如文件打开、关闭、读写等；第1字节~第16字节为文件名散列值字段，散列后的文件名具有统一的长度，从而提高数据均衡分布在每个Region的几率，实现负载均衡以提高查询效率；第17字节~第20字节为文件操作时间字段；第21字节~第23字节为扩展字段，记录文件所属用户名、文件操作权限等。准备模型训练数据时，设定采集文件访问I/O记录起始时间分别为t_s和t_e，时间跨度为：

$$\Delta t = t_e - t_s \tag{7}$$

挖掘文件访问时间特性需要收集前后时间跨度足够长的访问I/O记录，进而做出更准确的访问热度预测。对于高能物理存储系统千万级别的文件数目，假如选取更细粒度的时间窗口，持久化存储文件访问特征向量需要占用更多的物理空间，同时增加预测模型的训练时间。另一方面，假如选取更粗粒度的时间窗口，部分文件访问时间特性可能会丢失。本文设计了动态时间窗口划分机制，设定时刻i的窗口大小为：

$$t_{i-N} = \begin{cases} t_i, N < m \\ t_i \times 2^{N-1}, N \geqslant m \end{cases} \tag{8}$$

2.2　模型输出

文件访问频率预测问题可以看作是一类连续型变量的预测问题。在传统情况下此类问题可以使用回归分析的方法[16]，基于预测结果制定相应的文件迁移决策，以使迁移代价最小化。然而，在实际存储场景中文件数目众多，不同文件时序访问规律差异巨大，不可能为每个文件训练一个回归模型，存在计算复杂、自适应能力差等缺陷。

本文假设高能物理分级存储系统根据存储介质划分为n个存储层次，每个存储层的数据访问性能各不相同，拥有统一名称空间。高能物理分级存储要求尽量减少频繁迁移对正常用户访问的影响。绝大部分情况下访问频率在小范围内波动，并不改变文件应该迁移至哪个存储层级。本文预测文件的访问热度，即访问频率落在哪个区间范围内。预测问题可以被重新表述成一个分类问题，如图4所示。同时，本文使用文件的访问热度级别$(0,1,\cdots,n-1)$，标记训练集中的每一条访问特征序列。n个热度标签分别使用one-hot编码转换为由0和1组成的稀疏向量：

$$\boldsymbol{Y}_i = \{0, 0, 0, \cdots, 1, \cdots, 0\} \tag{9}$$

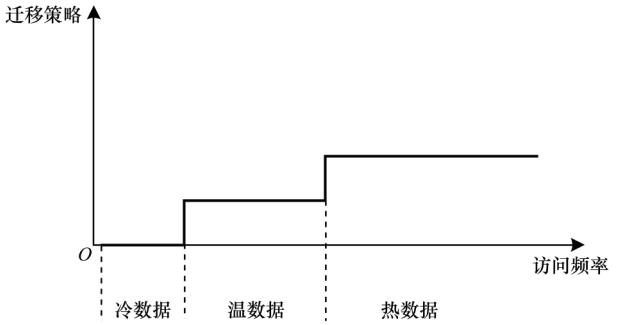

图4　访问频率的多个区间

Fig.4　Multiple intervals of access frequency

2.3　模型训练和文件访问热度预测算法

模型训练主要以LSTM网络的隐藏层作为研究对象。在模型输入层中，定义原始文件访问特征时序序列为：

$$F_o = \{f_1, f_2, \cdots, f_n\} \tag{10}$$

随机划分训练集和测试集，采用标准的z-score标准化方法，标准化后的训练集可以表示为：

$$F'_{\text{train}} = \{f'_1, f'_2, \cdots, f'_n\} \tag{12}$$

$$f'_t = \left(f_t - \sum_{t=1}^{n} f_t/n\right) \sqrt{\sum_{t=1}^{n}\left(f_t - \sum_{t=1}^{n} f_t/n\right)^2 \Big/ n}$$

$$1 \leqslant t \leqslant L, t \in N \tag{13}$$

原始时序序列以小时为分割单位，应用2.1节中的动态时间窗口分割方法对F_o进行二次处理。假设模型展开步长为L，即隐藏层包含L个连接的LSTM神经元。分割后的模型输入为：

$$X = \{X_1, X_2, \cdots, X_L\}, \tag{14}$$

其中，X即为2.1节中挑选出的文件访问I/O特征，对应的理论输出Y为2.2节中定义的文件访问热度标签。

模型输入层将X传递至隐藏层，经过隐藏层后的输出表示为：

$$O = \{O_1, O_2, \cdots, O_L\} \tag{15}$$

$$O_p = \text{LSTM}_{\text{forward}}(X_p, C_{p-1}, H_{p-1}) \tag{16}$$

参考本文2.1节，C_{p-1}和H_{p-1}分别对应前一时刻，即上一LSTM神经元的状态和输出，函数$\text{LSTM}_{\text{forward}}$代表LSTM神经元中的信息前向传递方法。这里假设神经元状态向量大小为S，可知$\boldsymbol{C}_{p-1}$和$\boldsymbol{H}_{p-1}$向量大小也均为S。

在LSTM隐藏层输出后接一个softmax层，以输出各类访问热度的概率。预测时输出最大概率值对应的类标签，即：

$$\tilde{y}_j = \arg\max_j (\log_a \text{soft}\max(ch_i b))_j \tag{17}$$

模型训练选用交叉熵损失函数作为训练过程中的损失函数，定义为：

$$\text{Loss} = -(y_i \log_a(\hat{y}_i) + (1-y_i)\log_a(1-\hat{y}_i)) \tag{18}$$

设定损失函数最小为模型的训练目标，给定随机化种子seed对LSTM网络中的权重和偏差进行随机化，设定LSTM网络的隐藏层数和隐藏节点数分别为layers和nodes，展开长度等于L，设定初始学习率和训练步数为steps。模型训练使用梯度反向传播算法，并使用Adam随机优化算法[17]更新网络中的参数。

模型训练和访问热度预测算法如算法1所示。

算法1 模型训练和访问热度预测算法

输入 F_o，Y，L，layers，nodes，seed，steps

输出 与测试集对应的文件访问热度以及模型准确率

1.get X from F_o by L

2.get X_{train} and X_{test}//划分训练集和测试集

3.X_p=min_max(X_{train})//采用min_max方法对输入数据进行归一化

4.create $LSTM_{cell}$ using layers，nodes and state

5.connect $LSTM_{net}$ by $LSTM_{cell}$ and number L

6.initialize weights and biases by seed

7.for each step in steps：

8.Pred = $LSTM_{forward}$(Xp)

9.Loss = CrossEntryLoss(pred，Y)

10.Update $LSTM_{net}$ by Adam

11.GridSearch(layers，nodes)//使用网格搜索调节网络超参数

12.Optimize end and get $LSTM^*_{net}$

13.for each j in 0：length(X_{test})

14.P = $LSTM_{forward}$(X_j)

15.append P_o to P

16.Error measure ε(P_o，Y_o)

3 系统部署与验证

文件访问热度预测系统已部署在位于四川稻城的高海拔宇宙线观测实验LHAASO的海量存储系统上，本文首先介绍验证之前如何准备所需的文件访问数据集，训练LSTM模型对文件访问热度进行预测，调整访问热度对应的文件访问频度阈值γ，分别测试在不同阈值下LSTM模型的预测精度，并与目前其他预测模型优劣进行对比。

3.1 数据集准备

本文使用的数据集来自LHAASO EOS存储文件的访问I/O日志，选取在过去30天内曾经活跃过的文件数目为5 842 207个，生成模型训练和测试数据集。文件访问包含EOS存储系统打开文件、读取或写入数据和关闭文件等过程，且每次访问都会在EOS存储系统的FST服务器中生成一条访问日志。高能所计算中心的EOS存储集群系统日志由监控系统定期抓取到ElasticSearch数据库，在数据预处理阶段，从日志中提取文件访问特征并存入HBase数据库。

在验证中设定前27天从文件访问日志中提取的访问特征作为预测模型的输入。应用2.2节中的方法将文件后3天的访问频率Q划分为多个区间，对应多个文件访问热度级别作为预测模型的输出。在一般情况下，在高能物理存储中将Q中为0的文件定义为冷文件。数据迁移系统周期性地将此类文件存储至磁带库。为进一步区分温文件和热文件，本文定义了访问频率阈值γ。频率小于等于阈值γ的文件定义为温文件，迁移系统周期性地将此类文件迁移至机械硬盘HDD层，频率大于γ的文件定义为热文件，迁移系统周期性地将此类文件迁移至固态硬盘SSD层。以γ=3时为例，LHAASO实验训练数据集中冷文件数目约占95.8%，温文件数目约占3.06%，热文件数目约占1.13%。

3.2 现有模型对比

本文将LSTM模型与目前已有的几种预测模型进行对比验证。

3.2.1 SVM模型

SVM是一类按监督学习方式对数据进行二元分类的方法[18]。为解决非线性分类问题，在SVM中引入了核函数，将低维度中线性不可分的数据点映射到更高维度线性可分的新空间中，求解最优的分类面用于分类。本文使用高斯径向基函数(Radial Basis Function，RBF)核作为SVM模型的核函数。针对多分类问题可以通过组合多个二分类器，将某个类别的样本归为一类，其他样本归为另一类，k个类别训练k个SVM模型。由于SVM空间复杂度是样本数据量的二次方，训练数据量很大时训练时间会比较长。但SVM模型训练时的优化目标是结构化风险最小，所以泛化能力较好，在人像识别、文本分类等方面有着广泛应用。

3.2.2 多层感知机模型

多层感知机(MLP)模型又称人工神经网络，是通过模拟大脑神经元处理信息的方式而建立的网络模型[19]。MLP中包含输入层、输出层、隐含层3个部分，每一层包含若干个神经元。每个神经元代表一种特殊的非线性激活函数，MLP层与层之间是全连接的，训练的过程即为确定各层之间的连接权重和偏置等参数的过程。首先随机初始化参数，然后迭代训练不断更新权重和偏置，直到模型输出或模型分类误差满足某个条件为止。MLP模型的优点在于自学习能力和优秀的非线性映射能力，但当层数较多或隐藏节点数较多时会带来参数膨胀的问题，造成训练困难。

3.3 模型评价指标

本文将数据集随机划分为训练集(80%)、验证集(10%)和测试集(10%)3个部分，从耗时、精度和一致性3个方面对预测模型进行评估，耗时为每类模型构建消耗的时间。本文测试并评估了每类模型在训练集和测试集上的预测准确率。根据图4，该问题可以被重新阐述为一个多分类模型的评估问题，与二分类模型的评估方法不同，除预测准确率外，本文还使用基于混淆矩阵的Kappa系数来评估模型的一致性[20]。Kappa系数

是统计学中评估一致性的方法，取值范围为[−1，1]，在实际应用中一般为[0，1]，值越高代表模型分类一致性越好，即在每个类别上的预测置信度都比较高。如果值接近于0，说明模型分类结果接近于随机分类。Kappa系数的计算公式如下：

$$\text{Kappa}=\frac{p_o-p_e}{1-p_e} \tag{19}$$

其中，P_o代表总体分类精度，P_e为：

$$p_e=\frac{a_1\times b_1+a_2\times b_2+\cdots+a_c\times b_c}{n\times n} \tag{20}$$

其中，n为测试集样本数，a_i为每一类真实样本数，b_i为预测出的每一类样本数。

3.4　应用平台与环境

首先介绍模型应用场景，中科院高能物理所计算中心为LHAASO搭建了EOS存储集群，包含2个元数据服务器节点和11个存储服务器节点，总文件数目有6 600万个，占用空间大小为2.9 PB。

在本文中，用于存储文件访问特征的HBase集群配置如下：4台服务器构建的HBase集群，基于Hadoop 2.6.2平台，每台服务器配有2颗AMD Operon 6320 CPU和64 GB内存，1 TB硬盘。

文件访问热度预测模型使用的训练节点配置如下：Intel® Xeon™ CPU E5-2650 v4 @ 2.20 GHz，单CPU24个核，64 GB RAM，配有2个NVIDIA Tesla K80 GPU。MLP、GRU和LSTM模型基于Tensorflow（1.8.0）实现，SVM模型基于Sklearn（0.19.00）实现。

3.5　测试结果

测试过程首先应用3.3节的方法对文件访问热度建立LSTM预测模型，使用棋盘搜索法确定模型的超参数。初始时根据经验确定模型训练时的学习率。不同学习率（η=0.001，0.002，0.005）下LSTM预测模型在训练集上的损失函数变化如图5所示。可以看出，当η=0.002时，损失函数先快速下降后趋于平稳，最终表现优于其他学习率下的模型。

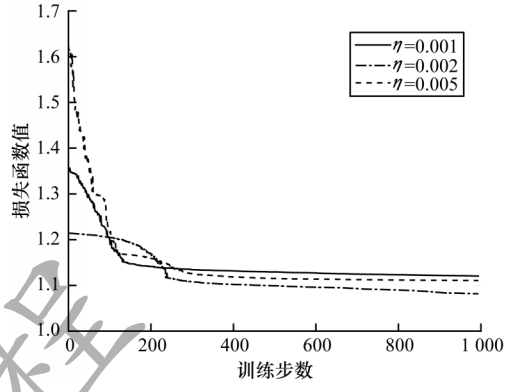

图5　不同学习率对应的损失函数变化

Fig.5　Loss function changes corresponding to different learning rates

根据3.1节中LHAASO高能物理数据集统计，确定访问频率阈值γ的不同取值（γ=1，2，3），如表2和表3所示，其中，表2中最大准确率和最小耗时与表3中最大Kappa值和最小耗时用下划线标记。分别训练MLP、SVM、LSTM、带注意力机制的LSTM模型并在训练集和测试集上进行对比测试，可以看出，LSTM模型在整体预测精度和一致性上表现都较好，在γ=1时预测精度最高，一致性也最好。模型训练耗时和推理耗时高于其他模型，但仍处于可接受的范围内。SVM模型在γ=2，3时测试集一致性上的表现较差。MLP模型训练耗时和推理耗时最短，但在γ=1，2时预测准确率和一致性不如LSTM和带注意力机制的LSTM模型。

表2　不同预测模型预测准确率对比

Table 2　Comparison of prediction accuracy of different predictive models

模型	模型参数	训练集			测试集			训练耗时/h	推理耗时/ms
		γ=1	γ=2	γ=5	γ=1	γ=2	γ=5		
MLP	η=0.1,layers=6,nodes=64,steps=500	0.591	0.658	0.547	0.662	0.693	0.634	0.78	269
SVM	kernel=rbf,c=5.12e+02,gamma=2.50e−01	0.682	0.883	0.481	0.612	0.811	0.352	0.96	513
LSTM	η=0.001,layers=4,nodes=128,steps=500	0.866	0.788	0.772	0.847	0.724	0.763	1.23	1 053
Attention LSTM	η=0.001,layers=4,nodes=128,steps=500	0.873	0.924	0.856	0.866	0.897	0.820	1.37	1 191

表3　不同模型预测一致性对比

Table 3　Comparison of prediction consistency of different models

模 型	模型参数	训练集			测试集			训练耗时/h	推理耗时/ms
		γ=1	γ=2	γ=5	γ=1	γ=2	γ=5		
MLP	η=0.1,layers=6,nodes=64,steps=500	0.564	0.536	0.784	0.408	0.356	0.619	0.78	269
SVM	kernel=rbf,c=5.12e+02,gamma=2.50e−01	0.782	0.978	0.541	0.439	0.221	0.490	0.96	513
LSTM	η=0.001,layers=4,nodes=128,steps=500	0.801	0.682	0.753	0.623	0.512	0.661	1.23	1 053
Attention LSTM	η=0.001,layers=4,nodes=128,steps=500	0.822	0.981	0.795	0.793	0.674	0.727	1.37	1 191

4 结束语

本文提出一种分级存储中基于LSTM深度学习模型预测文件访问热度的方法，主要包括数据集准备、文件访问特征构建、训练和预测等。相对于基于管理员经验和基于统计的迁移方法，LSTM模型能更准确地预测文件访问热度变化，从而提供更有效的文件迁移依据。本文提出的预测模型输入为高能物理分布式存储EOS中的文件访问特征，可推广到其他分布式存储系统中，该预测模型是通过批量学习来训练的，需要训练数据集在学习任务开始前准备，模型需要定期迭代更新，下一步将尝试构造训练数据流，引入预测模型在线学习，利用存储系统线上数据实时更新模型。

参考文献

[1] CAO Zhen, CHEN Mingjun, CHEN Songzhan, et al. Introduction to large high altitude air shower observatory[J]. Chinese Astronomy and Astrophysics, 2019, 43(4): 457-478.

[2] SCHWAN P. Lustre: building a file system for 1 000-node clusters [EB/OL]. [2019-12-10]. http://www.clusterfs.com.

[3] PETERS A J, SINDRILARU E A, ADDE G. EOS as the present and future solution for data storage at CERN[J]. Journal of Physics, 2015, 664(4): 42-62.

[4] MEDDEB M, DHRAIEF A, BELGHITH A, et al. Least fresh first cache replacement policy for NDN-based IoT networks[J]. Pervasive and Mobile Computing, 2019, 52: 60-70.

[5] LIU Aigui, CHEN Gang. A user-based LNS file prediction model[J]. Computer Engineering and Applications, 2007, 43(29): 14-16. (in Chinese)
刘爱贵，陈刚. 一种基于用户的LNS文件预测模型[J]. 计算机工程与应用，2007，43(29)：14-16.

[6] ZHANG G, CHIU L, DICKEY C, et al. Automated lookahead data migration in SSD-enabled multi-tiered storage systems[C]//Proceedings of IEEE Symposium on Mass Storage Systems & Technologies. Washington D. C., USA: IEEE Press, 2010: 135-148.

[7] SUN Ziheng, DI Liping, FANG Hui. Using long short-term memory recurrent neural network in land cover classification on landsat and cropland data layer time series [J]. International Journal of Remote Sensing, 2019, 40(2): 593-614.

[8] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning[M]. [S. 1.]: MIT Press, 2016.

[9] CHENG Gaofeng, LI Xin, YAN Yonghong. Using highway connections to enable deep small-footprint LSTM-RNNs for speech recognition[J]. Chinese Journal of Electronics, 2019, 28(1): 107-112.

[10] FARHAN W, TALAFHA B, ABUAMMAR A, et al. Unsupervised dialectal neural machine translation [J]. Information Processing and Management, 2020, 57(3): 334-348.

[11] KONG Weicong, DONG Zhaoyan, JIA Youwei, et al. Short-term residential load forecasting based on LSTM recurrent neural network [J]. IEEE Transactions on Smart Grid, 2017, 10(1): 841-851.

[12] LIMA F D, AMARAL G M R, DE MOURA L G, et al. Predicting failures in hard drives with lSTM networks[C]// Proceedings of 2017 Brazilian Conference on Intelligent Systems. Washington D. C., USA: IEEE Press, 2017: 222-227.

[13] GREFF K, SRIVASTAVA R K, KOUTNIK J, et al. LSTM: a search space odyssey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 28(10): 2222-2232.

[14] ABADI M, BARHAM P, CHEN J, et al. TensorFlow: a system for large-scale machine learning[C]//Proceedings of the 12th IEEE Symposium on Operating Systems Design and Implementation. Washington D. C., USA: IEEE Press, 2016: 265-283.

[15] GERON A. Hands-on machine learning with Scikit-learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems[M]. [S. 1.]: O'Reilly Media, 2019.

[16] RUBIO G, POMARES H, ROJAS I, et al. A heuristic method for parameter selection in LS-SVM: application to time series prediction [J]. International Journal of Forecasting, 2011, 27(3): 725-739.

[17] KINGMA D P, BA J. Adam: a method for stochastic optimization[EB/OL]. [2019-12-10]. https://arxiv.org/abs/1412.6980.

[18] JOACHIMS T. Making large-scale SVM learning practical[M]. Cambridge, USA: MIT Press, 1998.

[19] TANG Jiexiong, DENG Chenwen, HUANG Guangbin. Extreme learning machine for multilayer perceptron[J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 27(4): 809-821.

[20] TSOU T S. A robust likelihood approach to inference about the kappa coefficient for correlated binary data [J]. Statistical Methods in Medical Research, 2019, 28(4): 1188-1202.

编辑 索书志