科技导报 SCIENCE & TECHNOLOGY REVIEW

数据挖掘分类算法研究综述

Review of Classification Algorithms in Data Mining

王　刚/WANG Gang[1],黄丽华/HUANG Li-hua[1],张成洪/ZHANG Cheng-hong[1],夏　洁/XIA Jie[2]

1. 复旦大学管理学院,上海　200433
2. 中南大学信息工程学院,长沙　410083

1. School of Management, Fudan University, Shanghai 200433, China
2. School of Information Engineering, Central South University, Changsha 410083, China

[摘要] 随着数据库应用的不断深化,数据库的规模急剧膨胀,数据挖掘已成为当今研究的热点;特别是其中的分类问题,由于其使用的广泛性,现已引起了越来越多的关注。对数据挖掘分类问题的研究现状进行了综述:首先对研究比较多的基于判定树的归纳分类、基于人工神经网络的分类和基于统计的贝叶斯分类作了详细的讨论;然后对目前新提出的几种算法作了简要分析;最后根据数据挖掘的发展现状和研究重点对数据挖掘分类算法的发展趋势作了展望。

[关键词] 数据挖掘;分类;算法;综述

[中图分类号] TP18　　**[文献表识码]** A　　**[文章编号]** 1000-7857(2006)12-0073-04

Abstract: With the application of database deepening and the size of database expanding quickly, Data Mining has recently become the hotspot. Classification, the problem among them especially because of its extensive usage, has acquired more and more concerns presently. On account of this, the article carried on an overview according to the present condition of data mining's classification. Firstly, the article discussed in detail the classification methods that were researched widely, such as Decision Tree, Artificial Neural Network and Bayesian classification. Secondly, the article analyzed the new brought forward algorithms briefly. Lastly, according to the data mining´s developmental conditions and the emphases of research, the article forecasted the trends of the next research of classification.

Key Words: data mining; classification; algorithm; review

CLC Number: TP18　　**Document Code:** A　　**Article ID:** 1000-7857(2006)12-0073-04

1 引言

1989 年 8 月,在第 11 届国际人工智能联合会议的专题研讨会上,首次提出基于数据库的知识发现(KDD,Knowledge Discovery Database)技术。该技术涉及机器学习、模式识别、统计学、智能数据库、知识获取、专家系统、数据可视化和高性能计算等领域,技术难度较大,一时难以应付信息爆炸的实际需求。到了 1995 年,在美国计算机年会(ACM)上,提出了数据挖掘(DM,Data Mining)的概念,由于数据挖掘是 KDD 过程中最为关键的步骤,在实践应用中对数据挖掘和 KDD 这 2 个术语往往不加以区分。

数据挖掘的主要任务有分类分析、聚类分析、关联分析、序列模式分析等,其中的分类分析由于其特殊地位,一直是数据挖掘研究的热点之一。分类作为一类重要的数据挖掘问题,其过程可描述为[1]输入数据,或称训练集(Training Set),是一条条的数据库记录(Record)组成的。每一条记录包含若干个属性(Attribute),组成一个特征向量。训练集的每条记录还有一个特定的类标签(Class Label)与之对应。该类标签是系统的输入,通常是以往的一些经验数据。一个具体样本的形式可为样本向量:$(v_1, v_2, \cdots, v_i, \cdots, v_n; c)$。在这里 v_i 表示字段值,c 表示类别。数据挖掘分类就是分析输入数据,通过在训练集中的数据表现出来的特性,为每一个类找到一种准确的描述或者模型。这种描述常常用谓词表示。由此生成的类描述用来对未来的测试数据进行分类。尽管这些未来的测试数据的类标签是未知的,我们仍可以由此预测这些新数据所属的类。

2 数据挖掘的主要分类算法

2.1 基于判定树的归纳分类

判定树是一个类似流程图的树结构,其中每个内部节点表示在一个属性上的测试,每个分支代表一个测试输出,而每个树叶节点代表类或类分布。树的最顶层节点是根节点。由判定树可以很容易得到 "IF-THEN"形式的分类规则。方法是沿着由根节点到树叶节点的路径,路径上的每个属性-值对形成"IF"部分的一个合取项,树叶节点包含类预测,形成"THEN"部分。一条路径创建一个规则。

判定树归纳的基本算法是贪心算法,它是自顶向下递归的各个击破方式构造判定树。其中一种著名的判定树归纳算法是建立在推理系统和概念学习系统基础上的 ID3 算法[2],如下所示。

1)算法:由给定训练数据产生判定树。

2)输入:训练样本 *samples*,由离散值属性表示;候选属性的集合 *attribute_list*。

收稿日期:2006-8-07

基金项目:国家自然科学基金项目(70571016,70471011)

作者简介:王　刚,男,上海邯郸路 220 号复旦大学管理学院,博士生,主要研究方向为管理信息系统,人工智能等;
E-mail: wg_edison@163.com
张成洪(通讯作者),男,上海邯郸路 220 号复旦大学管理学院,副教授,主要研究方向为信息管理与信息系统;
E-mail: chzhang@fudan.edu.cn

3) 输出:一棵判定树。

4) 方法:

(1) 创建节点 N;

(2) if *samples* 都在同一个类 C then

(3) 返回 N 作为叶节点,以类 C 标记;

(4) if *attribut_list* 为空 then

(5) 返回 N 作为叶节点, 标记 *samples* 中普通的类;//多数表决

(6) 选择 *attribute_list* 中具有最高信息增益的属性 *test_attribute*;

(7) 标记节点 N 为 *test_attribute*

(8) for each *test_attribute* 中的已知值 a_i//划分 *samples*

(9) 由节点 N 长出一个条件为 $test\_attribute=a_i$ 的分枝;

(10) 设 s_i 是 *samples* 中 $test\_attribute=a_i$ 的样本的集合;

(11) if s_i 为空 then

(12) 加上一个树叶,标记为 *samples* 中最普通的类;

(13) else 加上一个由 $Genetate_decision_tree(s_i, attribute_list_test_attribute)$返回的节点。

后来又出现了 ID3 算法的增量版本,包括 ID4 算法和 ID5 算法。但这些算法对于相对小的数据集上是有效的, 对于现实中数据量很大的数据挖掘时, 有效性和伸缩性就成了关注的问题。为此有人提出了一些新的判定树算法,如 SLIQ 算法[3]。其在判定树的构造过程中采用了预排序与广度优先增长策略。使得该算法能够处理更大的训练集, 因此在一定程度上具有良好的随记录个数和属性个数增长的可扩展性。但是它仍然存在着一些不足: ① 由于需要将类别列表存放于内存, 在一定程度上限制了可以处理的数据集的大小; ② 由于采用了预排序技术, 而排序算法的复杂度本身并不是与记录个数成线性关系。因此使得 SLIQ 算法不可能达到随记录数目增长的线性可扩展性。除此之外,为了减少驻留于内存的数据量,还有人提出过 SPRINT 算法[4],其优点是在寻找每个结点的最优分裂标准时变得更简单, 其缺点是对非分裂属性的属性列表进行分裂时变得很困难。

建好判定树并不意味着判定树的完成,还要检查模型是否过适应数据,即模型是否过度贴近训练集的特性。如果模型过适应数据的话, 对其他的数据库数据就不一定奏效了。因为在判定树构造时,有些分枝反映的是训练集中的噪声或孤立点,此时可以用剪枝方法加以处理。树剪枝方法使用统计度量, 检测和减去那些不可靠的分支, 这样可以提高在未知数据记录上分类的准确性。正是由于以上问题,特别是海量数据引起的时间效率问题, 使得基于判定树归纳的分类受到限制。

2.2 基于人工神经网络的分类

人工神经网络的研究至今已有 60 多年的历史。1943 年,心理学家 McCulloch 和数学家 Pitts 合作, 提出了形式神经元的数学模型,即 MP 模型[5]。他们利用逻辑的数学工具研究客观世界的事件在形式神经网络中的表达。1949 年,心理学家 Hebb 提出了改变神经元连接强度的 Hebb 规则。20 世纪 50 年代末,Rosenblatt 设计发展了 MP 模型,提出了多层感知机 Perceptron。60 年代初,Widrow 提出了自适应线性单元模型 Adaline, 以及一种有效的网络学习方法—— Widrow-Hebb 规则[6]。鉴于上述研究,神经网络引起了许多科学家的兴趣。但随着对感知机为代表的神经网络的功能和局限性的深入分析等原因, 使神经网络的研究陷入低潮。但是仍有一些学者坚持研究,并取得了一些成果,出现了 Grossberg 的 ART 模型和 Kohonen 的 SOM 模型。1982 年,通过引入能量函数的概念,Hopfied 研究了网络的动力学性质, 并用电子线路设计出相应的网络, 进而掀起了神经网络新的研究高潮。1986 年,Rumellhart 和 McCllelland 等提出了 PDP 理论, 尤其是发展了多层前向网络的 BP 算法[7],成为迄今应用最普遍的学习算法。其网络拓扑结构如图 1[1],算法如下所示。

1) 算法:BP 算法

2) 输入:训练样本 *samples*,学习率 l,多层前馈网络

3) 输出:一个训练的、对样本分类的神经网络

4) 方法:

(1) 初始化 *network* 的权和阈值

(2) while 终止条件满足{

(3) for *samples* 中的每个训练样本 X {

(4) for 隐含或输出层每个单元 j {

(5) $Ij=\sum_i \omega ij Oi+\theta j$;//相对于前一层 i,计算单元 j 的净输入

(6) $Oj=1/(1+e^{-Ij})$;//计算每个单元 j 的输出

(7) for 输出层每个单元

(8) $Errj=Oj(1-Oj)(Tj-Oj)$;//计算误差

(9) for 由最后一个到第一个隐藏层,对于隐藏层每个单元 j

(10) $Errj=Oj(1-Oj)\sum_k Errk\omega jk$;//计算关于下个较高层 k 的误差

(11) for *network* 中的每一个权 ωij

(12) $\Delta Wij=(l)ErrjOj$;//权增值

(13) $Wij=Wij+\Delta Wij$;//权更新

(14) for *network* 中每个阈值 Θ_j {

(15) $\Delta\Theta_j=(l)Errj$;//偏差增值

(16) $\Theta_j=\Theta_j+\Delta\Theta_j$;//偏差更新 } }

目前, 神经网络用于数据挖掘的最大问题在于对网络的解释, 难以从训练过的网络中抽取规则[8]。主要原因有 2 点。① 网络的连接数可能太多, 以致于无法用 if…then…的形式描述元组及其类别标记之间的关系。如果一个节点有 n 个具有二进制值的输入连接, 那就可能会有多达 2^n 种不同的输入模式,即使 n 很小,规则也会相当长或相当复杂。② 隐节点的激活值可以是区间[-1,1]内依赖于输入元组的任意值。如果有大量的测试数据, 激活值几乎是连续取值的, 那么要在隐节点的激活值与输出层节点的输出值之间导出明确的关系,显然有相当的难度。

基于神经网络的规则抽取算法把隐单元的激活值离散化为易处理的数量的离散值,且不降低网络的分类准确度。离散激活值的较小集合使得要确定输出节点值与隐节点值之间的相关性及隐节点激活值与输入值之间的相关性关系成为可能。为此,已有学者提出了必须在提取神经规则之前对网络进行剪枝[9]。网络剪枝的目的就是删除那些不影响(或影响可以忽略)分类准确性的神经元和链枝,以得到简化的神经网络。

由于人工神经网络用于分类需要较长的训练时间,以及结构的难确定性、可解释性差等问题, 使得在数据挖掘的初期不被看好。但是由于其对噪声数据的高承受能

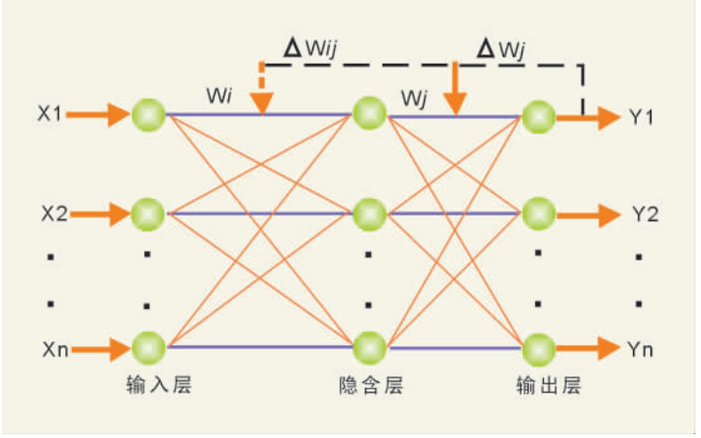


图 1 BP 传播算法的网络结构图

Fig. 1 The network structure of the BP algorithm

力、分类的高精度,以及提出了一些用训练过的神经网络提取规则的算法,使得神经网络用于数据挖掘显示了强大的生命力。但是对海量数据的时间效率仍存在问题,需要同其他方法相结合以期达到理想效果。

2.3 基于统计的贝叶斯分类

贝叶斯分类是统计学的分类方法,基于贝叶斯公式即后验概率公式。朴素贝叶斯分类的分类过程是首先令每个数据样本用一个 N 维特征向量 $X=\{x_1,x_2,\cdots x_n\}$ 表示,其中 x_k 是属性 A_k 的值。所有的样本分为 m 类:$C_1,C_2,\cdots,C_n$。对于一个类别的标记未知的数据记录而言,若 $P(C_i/X)>P(C_j/X)$, $1\leqslant j\leqslant m, j\neq i$,也就是说,如果条件 X 下,数据记录属于 C_i 类的概率大于属于其他类的概率的话,贝叶斯分类将把这条记录归类为 C_i 类。

由贝叶斯公式可知:因为 $P(X)$是常数,那么 $P(C_i/X)\cdot P(C_i)$最大的话,$P(C_i/X)$就最大。一般来讲,计算 $P(X/C_i)$的开销很大。为了降低计算的开销,可以做个假定,假定属性值相互条件独立,也就是说,假定属性之间不存在依赖关系。这样一来概率 $P(X_1/C_i)$,$P(X_2/C_i)$,$\cdots$,$P(X_n/C_i)$由训练样本估值:$P(X_k/C_i)=s_{ik}/s_i(1\leqslant k\leqslant n)$, $P(C_i)=s_i/s$。其中 S_{ik} 是属性 A_k 上具有之 X_k 的类 C_i 的样本数,S_i 是 C_i 中的样本数,S 是总样本数。朴素贝叶斯算法成立的前提是各属性之间互相独立,当数据集满足这种独立性假设时,分类的准确度较高,否则可能较低。

由于贝叶斯定理假设一个属性值对给定类的影响独立于其他属性的值,而此假设在实际情况中经常是不成立的,因此其分类准确率可能会下降。为此,出现了对朴素贝叶斯分类的改进——贝叶斯信念网络[10]。它允许在变量子集间定义类条件独立性,并提供了一种因果关系的图形,可以在其上进行学习。贝叶斯信念网络是描述数据变量之间关系的图形模型,是一个带有概率注释的有向无环图。贝叶斯网 $G=(S,P)$ 由网络的拓扑结构 S 和局部概率分布的集合 P 两部分组成,S 是一个有向无环图,P 代表用于量化网络的一组参数。

建立贝叶斯信念网络可以被分为两个阶段。第一阶段网络拓扑学习,即有向非循环图的学习,利用贝叶斯网络的学习算法,从实例数据建立所有属性变量和类变量构成的贝叶斯网结构。第二个阶段网络中每个变量的局部条件概率分布的学习,采用贝叶斯网的推理算法,计算给定属性变量的值时类变量的最大后验概率。采用这种分类思想的算法有 TAN (tree augmented Bayes network)算法。但是统计上的贝叶斯分类对非线性样本数据,含噪声、孤立点的数据,在分类准确性上仍存在问题。

2.4 其他方法

此外,近些年来又出现了一些新的用于数据挖掘的分类算法。CBA (classification based on association)是基于关联规则的分类算法,该算法分两个步骤构造分类器[11]。① 发现所有形如 $X_{i1}\wedge X_{i2}\Rightarrow C_i$ 的关联规则,即右部为类别属性值的类别关联规则。② 从已发现的 CAR 中选择高优先度的规则来覆盖训练集,也就是说,如果有多条关联规则的左部相同,而右部为不同的类,则选择具有最高置信度的规则作为可能规则。CBA 算法的优点是其分类准确度较高,在许多数据集上比判定树更精确,但是上述两步都具有线性可伸缩性。此外,Liu,Hsu 和 Ma 提出关联分类法和使用显露模式分类的 CREP 算法[12],使用跳跃显露模式的 JEP 算法[13]。

基于数据库的分类算法,虽然数据挖掘的创始人主要是数据库领域的研究人员,然而提出的大多数算法则没有利用数据库的相关技术。在分类算法中,致力于解决此问题的算法有 MIND (mining in database 和 GAC-RDB (grouping and counting- relational database)。其中,MIND 算法是采用数据库中用户定义的函数实现发现分类规则的算法。MIND 采用典型的判定树构造方法构建分类器,具体步骤与 SLIQ 类似。其主要区别在于它采用数据库提供的用户定义的函数和 SQL 语句实现树的构造。该算法的优点是通过采用用户定义的函数实现判定树的构造过程使得分类算法易于与数据库系统集成,其缺点是算法用用户定义的函数完成主要的计算任务,而用户定义的函数一般是用户利用高级语言实现的,无法使用数据库系统提供的查询处理机制,无法利用查询优化方法[11]。GAC-RDB 算法是一种利用 SQL 语句实现的分类算法。该算法采用一种基于分组计数的方法统计训练数据集中各个属性取值组合的类别分布信息,通过最小置信度和最小支持度 2 个阈值,找出有意义的分类规则。该算法的优点是具有与现有的其他分类器相同的分类准确度,执行速度有较大提高,而且具有良好的伸缩性,应用程序易于与数据库系统集成;其缺点是参数的取值需用户完成等[11]。

除此之外,常用的数据挖掘分类算法还有 K-最临近分类,但由于其存储了所有样本,当由非常大的数据集学习时,可能带来困难。最近也出现一些基于人工智能的分类方法,如基于案例推理的分类,基于遗传算法的分类,基于粗糙集的分类,以及基于模糊集的分类,但是这些方法在预测的准确率、速度、强壮性、可伸缩性、可解释性上同前 3 种成熟的方法存在一定的距离。

3 数据挖掘分类算法发展展望

虽然现在关于分类的算法已有很多,并且有的还很成熟,但是面对如今的海量数据,一些算法在时间效率、鲁棒性、精确性上已存在一些问题。Lin、Loh 和 Shih 等在[14]中用实验的方法对各种分类算法进行了比较,但尚未发现一种方法对所有的数据都优于其他方法。并且从数据挖掘在国际上的发展来看,数据挖掘的研究重点现已从提出概念和发现方法转向系统应用和方法创新上,研究注重多种发现策略和技术集成,以及多种学科之间的相互渗透,数据挖掘迫切需要系统科学的理论体系作为其发展的支撑。

基于 Bates 和 Granger 于 1969 年提出的、通过将几个模型相加来提高模型的精度和鲁棒性的方法,现已被大家普遍接受。钱学森教授也提出从定性到定量的综合集成方法[15],将各种方法有机地结合起来,并非杂乱无章的拼凑。通过各种方法解决本领域的问题并相互取长补短,从而形成各种方法的协作,这一思想在各个领域已得到广泛应用,近年来在数据挖掘方面的应用正方兴未艾。

一些学者已开始将数据挖掘三大支柱之一的人工智能领域最新提出的技术——混合智能系统引入,通过将各种方法有机结合起来,扬长避短,以获得各种模型整合功效,从而解决目前数据挖掘分类的新问题。1992 年 Towell 和 Shawlk 提出了 KBAN (Knowledge -Based Artificial Neural Network) 模型,1997 年 Purushothaman 和 Karayiannis 提出了 Fuzzy Feedforward Connetionist Networks 模型,2002 年 Allen 提出了 AFFNN (Aggregate Feedforword Neural Netwoek)模型[16], 2004 年王刚提出了 R-FC-DENN 模型[17]。相比传统的数据挖掘分类算法,这些模型或算法无论在效率还是鲁棒性上都有了大幅提高,但是关于模型的整合还有许多问题,比如模型的整体结构、算法参数的选择等等,特别是与数据挖掘的分类算法的结合仍有许多新的、富有挑战性的问题需要解决。

4 结束语

本文对数据挖掘分类问题的研究现状进行了综述,对基于判定树的归纳分类、基于人工神经网络的分类和基于统计的贝叶斯分类作了详细的讨论,对目前新提出的几种算法:基于关联规则的分类、基于数据库的分类、K-最临近分类、基于案例推理的分类、基于遗传算法的分类、基于粗糙集的分类,以及基于模糊集的分类作了简单分析。并根据目前数据挖掘的研究重点已从提出概念和发现方法转向系统应用和方法创新上,注重研究多种发现策略和技术集成以及多种学科之间的相互渗透的现状,并对数据挖掘分类算法的发展作了展望。

参考文献(References)

[1] HAN J W. **数据挖掘:概念与技术**[M]. **北京:机械工业出版社**,2001.

[2] QUINLAN J R. Induction of decision trees [J]. Machine Learning, 1986(1): 81–106.

[3] METHTA M, AGRAWAL R,RISSANEN J. SLIQ: A fast scalable classifier for data mining [C]. Conf.Extend Database Technology (EDBT´96). Avignon, France, Mar. 1996.

[4] SHAFER J, AGRAWAL R, MTHTA M. SPRINT: A scalable parallel classifier for data mining [C]. Conf.Very large Data Base(VLDB´96), Bombay,India, 1996,544–555.

[5] MCCULLOCH W S, PITTS W H. A logical calculus of the ideas immanent in neuron activity [J]. Bulletin Mathematical Biophysics, 1943(I.5): 115–133.

[6] WIDROW B. An adaptive adaline neuron using chemical memistors [R]. Standford Electronics Lavoratroy Technical Report, 1960.

[7] RUMELLHART D E, MCCLLELAND J L. Parallel distributed processing [M]. Cambridge, MA: MIT Press. 1986.

[8] SAITO K ,NAKANO R. Medical diagnostic expert system based on PDP mode [J]. IEEE international conf.on Neural Networks, 1988(1): 225–262.

[9] LAWRENCE S, GILES C L, TSOI A C. Symbolic conversion, grammatical inference and rule extraction for foreign exchange rate prediction [M]. In Y. Abu–Mostafa, A.S. Weigend and P.N. Refenes, editors, Neural Networks in the capital Markets. Singapore: World Scientific,1997.

[10] HECKERMAN D. Bayesian networks for knowledge discovery [C]/ Fayyad U M, G. Piatesky–shapiro,Smyth P Uthurusamy R., editors. Advances in Knowledge Discovery and Data Mining. Cambridge, MA: MIT Press, 1996: 273–305.

[11] **刘红岩. 数据挖掘中的数据分类算法综述** [J]. **清华大学学报(自然科学版)**, 2002(6): 727–730.

[12] DONG G, LI J. Efficient mining of emerging patterns: Discovering trends and differences [C]. 1999 Int. Conf. Knowledge Discovery and data mining (KDD´99), San Diego, CA, Aug.1999:43–52.

[13] LI J, DONG G. RAMAMOHANRARAO K. Making use of the most expressive jumping emerging patterns for classification[C]. 2000 Pacific –Asia conf. Knowledge discovery and data mining (PAKDD´00).Kyto, Japan, Apr. 2000:220–232.

[14] LIN T S, LOH W Y,SHIH Y S. A comparison of prediction accuracy, complexity and training time of thirty–three old and new classification algorithms [J]. Machine Learning, 2000:39.

[15] **钱学森, 于景元, 戴汝为. 一个科学新领域——开放的复杂巨系统及其方法论** [J]. **自然杂志**,1990,13(1):3–10.

[16] ALLEN S V. An aggregate connectionist approach for discovery association rules [D]. Ph.D. Department of Computer Science and Engineering, Wright State University, 2002.

[17] **王 刚. 基于混合智能系统的数据挖掘分类算法研究**[D]. **长沙: 中南大学**, 2004.

(责任编辑 王士泉)

·科技智力游戏·

轻松一刻——填空补白

“轻松一刻——填空补白”游戏供读者读刊之余自娱自乐,并获取一些小知识。读者可根据下面“横”行、“纵”列的文字提示,分别填上相应的文字。答案见下期。

纵:

1.英国近代伟大的数学家,创建了微积分和经典力学。

2.1998 年, 主导国际空间站阿尔法磁谱仪(AMS)实验的著名华裔物理学家,并在“2006 中国科协年会”开幕式上做过生动有趣的报告。

3.传统道德文化中的五常。

4.我国江苏和浙江两个省的统称。

5.花的一种,凌波仙子,学名:*Narcissus tazeta vai chimensis*,英名:Chinese Narsissus,别名:玉玲珑、金银台、姚女花、女史花、天葱、雅蒜等。

6.丝绸之路南道小国,又名漂沙,位于今新疆叶城县境。汉时其王在呼犍谷,人种与羌人相类。以游牧为主, 地产玉石。北魏时国名改称“悉居半”;唐代称“朱居”,曾吞并帕米尔高原上的蒲犁国、德若国和依耐国。其王族为疏勒人,语言与于阗语大同小异。

7.我国一科学家,发现了并以他的名字命名的一颗小行星。2006 年 9 月 25 日,中科院和中国科协在京举行仪式,正式对外宣布命名行星。

8.“廉则生威,□□生明”。此句源自石成金的《传家宝·绅瑜》, 朱镕基同志将其作为人生信条。请填空白二字。

9.原名《第五交响曲》,它是贝多芬的一部哲理性很强的作品,也是最能代表其艺术风格的作品。

横:

一、一个成语,意思是受到眼前景象的触动而产生某种情感。

二、“□□是第一生产力”,改革开放初期邓小平同志说过的一句名言。请填空白二字。

三、春秋五霸之一,春秋初期的齐国国君,军事统帅,姜姓小白。

四、跪拜礼古代见面时的礼仪。古人席地而坐,臀部紧靠脚后跟。伸腰并使臀部离开脚后跟,用两膝着地则为跪。

五、历史上著名的唐朝女皇帝。

六、导读类性质的报纸简称。

七、抗日战争时期,我敌后根据地军民发明的一种表示十万火急的信件形式。1954 年,著名导演石挥曾拍摄了一部与此同名的经典儿童故事片。

八、早已有之的一种客观唯心主义世界观,源自美索不达米亚、埃及等东方文化中的一派思潮。该理论认为,人的命运是由偶然因素造成,是不可预测的,不可预知而又是注定的、不可改变的一种命观。

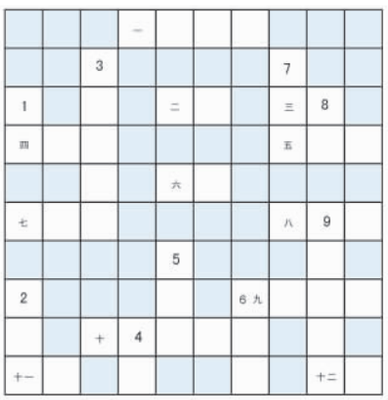
九、一所教育部直属重点大学的简称。其前身是 1896 年创建于上海的南洋公学,1959 年被列为全国重点大学,地处我国中西部地区。

十、唐代张若虚的诗作名, 其中开头若干句“春江潮水连海平,海上明月共潮生。滟滟随波千万里,何处春江无月明”被后人广为传唱。

十一、指竖行书写的长方形书法作品。清朝也代表一种官职的名称。

十二、是我国各种说唱艺术的总称,是以带有表演动作的“口语说唱”来叙述故事、塑造人物、表达思想感情、反映社会生活的表演艺术门类。

(供稿 李建才 责任编辑 赵 佳)



上期答案

		北	约		联		化		
	南	京			合	成	纤	维	
		科	教	兴	国			也	
易		技					雷	纳	
中		大	江	东	去		达		
天	文	学		方				武	直
	字			红	十	字	会		升
地	狱	火			年		计	算	机
中		箭	石		树			盘	
海	岩		英		木	乃	伊		