

目 录

第 1 章 计算机基础知识	1
1.1 计算机概述	1
1.2 计算机系统	5
1.3 计算机信息表示	12
第 2 章 Windows 10 操作系统	19
2.1 Windows 10 基础知识	19
2.2 Windows 资源管理	34
2.3 Windows 设置与控制面板	42
第 3 章 计算机网络与 Internet 应用	47
3.1 计算机网络基础	47
3.2 局域网概述	50
3.3 Internet 基础知识	54
3.4 Internet 的应用	57
3.5 计算机网络安全	65
第 4 章 Word 2016 的应用	67
4.1 Word 2016 文档编辑基础	67
4.2 文档的基本编辑	77
4.3 文档的高级编辑	91
4.4 邮件合并与宏应用	139

第 5 章 Excel 2016 的应用	150
5.1 Excel 2016 编辑操作基础	150
5.2 Excel 2016 基本编辑与格式排版	154
5.3 公式与函数	192
5.4 数据管理	234
5.5 图表制作	252
第 6 章 PowerPoint 2016 的应用	268
6.1 PowerPoint 2016 操作基础	268
6.2 PowerPoint 2016 的基本操作	275
6.3 PowerPoint 2016 的进阶操作	293
6.4 PowerPoint 2016 的综合应用	318
第 7 章 数据库与大数据技术	340
7.1 数据库系统的基本理论	341
7.2 数据模型	343
7.3 关系数据库	346
7.4 大数据技术基础	352
第 8 章 人工智能	357
8.1 人工智能概述	357
8.2 主要的人工智能技术	360
8.3 人工智能应用	368
8.4 人工智能发展面临的趋势与挑战	371
8.5 人工智能工具在办公自动化领域的应用	372
参考文献	386

第 7 章

数据库与大数据技术

数据库和大数据是两个相关但不同的概念，它们在技术体系结构、数据规模、数据结构和管理方式等方面有所区别，但同时数据库技术是大数据处理的基础之一，大数据处理通常需要数据库系统来存储和管理数据。以下是两者区别的详细介绍。

(1) 技术体系结构不同。数据库通常指的是一个集中的数据存储系统，用于存储、管理和检索结构化数据，如关系型数据库（如 MySQL、Oracle、SQL Server），这些数据库使用结构化查询语言（SQL）进行操作；大数据是一个更广泛的技术体系结构，涉及数据的采集、整理、存储、安全、分析和呈现等一系列技术标准，它包括但不限于关系型数据库，还涵盖 NoSQL 数据库（如 MongoDB、Cassandra）和分布式存储系统（如 Hadoop、Spark）。

(2) 数据规模不同。数据库通常用于处理相对较小的数据集；大数据技术用于处理海量数据，适用于如社交网络、电商平台等场景。

(3) 数据结构不同。数据库（尤其是关系型数据库）通常处理结构化数据，如表格形式的数据；大数据系统中的数据结构更加复杂，包括结构化、非结构化和半结构化数据。

(4) 管理方式不同。数据库由集中式数据库管理系统控制，数据存储在同一的服务器或多个服务器上，但访问通过单一的数据库服务器进行；大数据系统采用分布式存储和计算，数据分布在多个服务器上，通过网络实现数据的共享和处理。

尽管存在这些区别，但数据库技术仍然是大数据处理的基础。大数据系统通常依赖于关系型和非关系型数据库来存储和管理数据，并且许多大数据平台提供了与关系型数据库相似的功能，如 NoSQL 数据库的引入使得数据库系统能够更有效地处理非结构化数据。下面分别介绍这两种技术。

7.1 数据库系统的基本理论

7.1.1 数据库系统的组成

数据库系统由数据库、数据库管理系统、数据库管理员、硬件平台和软件平台构成，是一个以数据库为核心的完整的运行实体。

1. 数据

数据就是描述事物的符号记录。数据可以是数字、文字、声音、图形、图像等。计算机中的数据一般分为两部分，其中一部分与程序仅有短时间的交互关系，随着程序的结束而消亡，它们称为临时性数据，这类数据一般存放于计算机内存中；而另一部分数据则对系统起着长期持久的作用，它们称为持久性数据。数据库系统中处理的是持久性数据。软件中的数据是有一定结构的。数据包含数据类型和值两部分。

2. 数据库

数据库是一个长期存储在计算机内的、有组织的、可共享的、统一管理的大量数据的集合，也是一个按数据结构来存储和管理数据的计算机软件系统。数据库中的数据按一定的数据模型组织、描述和储存，具有较小的冗余度、较高的数据独立性和易扩展性，并可对各种用户共享。数据库中的数据具有两大特点：集成和共享。

3. 数据库管理系统

数据库管理系统是数据库的管理机构，它是一种系统软件，负责数据库中的数据组织、操纵、维护、控制、保护和服务等。数据库管理系统是数据库系统的核心。

4. 数据库管理员

由于数据库具有共享性，需要对数据库进行规划、设计、维护、监视等操作，这些需要专人管理，他们就是数据库管理员。数据库管理员的主要工作有数据库设计、数据库维护、改善系统性能和提高系统效率。

7.1.2 数据管理的发展

数据管理的发展至今已经历了3个阶段，即人工管理阶段、文件系统阶段和数据库系统阶段。

1. 人工管理阶段

在20世纪50年代中期以前，计算机主要用于科学计算，硬件无磁盘，直接存取，软件没有操作系统，效率低，不能提供完整、统一的管理和数据共享。

2. 文件系统阶段

20世纪50年代后期到60年代中期，进入文件系统阶段。文件系统是数据库系统发展的初级阶段，它提供了简单的数据共享与数据管理能力，但是它无法提供完整的、统一的管理和数据共享的能力。由于它的功能简单，它附属于操作系统而不成为独立的软件，目

前一般将其看成仅是数据库系统的雏形，而不是真正的数据库系统。

(3) 数据库系统阶段

①层次数据库与网状数据库系统阶段

从20世纪60年代末期起，真正的数据库系统——层次数据库与网状数据库开始发展，它们为统一管理与共享数据提供了有力支撑，这个时期数据库系统蓬勃发展形成了有名的“数据库时代”。但是这两种系统也存在不足，主要是它们脱胎于文件系统，受文件的物理影响较大，给数据库的使用带来了诸多不便，同时此类系统的数据模式构造烦琐不宜于推广使用。

②关系数据库系统阶段

网状数据库和层次数据库已经很好地解决了数据的集中和共享问题，但是在数据独立性和抽象级别上仍有很大欠缺。用户在对这两种数据库进行存取时，仍然需要明确数据的存储结构，指出存取路径。而后来出现的关系数据库较好地解决了这些问题。

关系数据库系统出现于20世纪70年代，在80年代得到蓬勃发展，并逐渐取代前两种系统。关系数据库系统的结构简单，使用方便，逻辑性强、物理性少，因此在80年代以后一直占据数据库领域的主导地位。

7.1.3 数据库系统的基本特点

数据库技术是在文件系统基础上发展产生的，两者都以数据文件的形式组织数据，但由于数据库系统在文件系统之上加入了数据库管理系统对数据进行管理，从而使数据库系统具有以下特点。

1. 数据的集成性

(1) 在数据库系统中采用统一的数据结构方式，如在关系数据库中采用二维表作为统一的结构方式。

(2) 在数据库系统中按照多个应用的需要组织全局的统一的数据库结构（数据模式），数据模式不仅可以建立全局的数据库结构，还可以建立数据间的语义联系从而构成一个内在紧密联系的数据整体。

(3) 数据库系统中的数据模式是多个应用共同的、全局的数据库结构，而每个应用的数据则是全局结构中的一部分，称为局部结构（视图），这种全局与局部的结构模式构成了数据库系统数据集成性的主要特征。

2. 数据的高共享性与低冗余性

由于数据的集成性，数据可为多个应用所共享，特别是在网络发达的今天，数据库与网络的结合扩大了数据关系的应用范围。数据的共享自身又可极大地减少数据冗余性，不仅减少了不必要的存储空间，更重要的是避免了数据的不一致性。

3. 数据独立性

数据独立性是指数据与程序之间的互不依赖性，即数据库中的数据独立于应用程序而不依赖于应用程序。数据的逻辑结构、存储结构与存取方式的改变不会影响应用程序。数

据独立性一般分为物理独立性与逻辑独立性。

(1) 物理独立性即数据的物理结构（包括存储结构、存取方式等）的改变，如存储设备的更换、物理存储的更换、存取方式的改变等都不影响数据库的逻辑结构，从而不致引起应用程序的变化。

(2) 逻辑独立性是指数据库总体逻辑结构的改变，如修改数据模式、增加新的数据类型、改变数据间的联系等，不需要相应地修改应用程序。

4. 数据统一管理与控制

数据库系统为数据提供了高度集成的环境，同时为数据提供了统一管理的手段，主要包含以下三个方面。

- (1) 数据的完整性检查 检查数据库中数据的正确性以保证数据的正确。
- (2) 数据的安全性保护 检查数据库访问者以防止非法访问。
- (3) 并发控制 控制多个应用的并发访问所产生的相互干扰以保证其正确性。

7.2 数据模型

7.2.1 数据模型的基本概念

数据库中的数据模型是将复杂的现实世界要求反映到计算机数据库中的物理世界。

数据是现实世界符号的抽象，而数据模型是数据特征的抽象，它从抽象层次上描述了系统的静态特征、动态行为和约束条件，为数据库系统的信息表示与操作提供一个抽象的框架。数据模型所描述的内容有三个部分，即数据结构、数据操作与数据约束。

1. 数据结构

数据模型中的数据结构主要描述数据的类型、内容、性质及数据间的联系等。数据结构是数据模型的基础，数据操作与数据约束均建立在数据结构上，不同数据结构有不同的数据操作与数据约束。

2. 数据操作

数据模型中的数据操作主要描述在相应数据结构上的操作类型与操作方式。

3. 数据约束

数据模型中的数据约束主要描述数据结构内数据间的语法、语义联系、它们之间的制约与依存关系，以及数据动态变化的规则，以保证数据的正确、有效与相容。

7.2.2 数据模型的分类

数据模型按不同的应用层次，可分为概念数据模型、逻辑数据模型、物理数据模型。

1. 概念数据模型

概念数据模型简称概念模型，是一种面向客观世界、面向用户的模型，与具体的数据库管理系统无关，与具体的计算机平台无关。概念模型着重于描述客观世界复杂事物的结

构及它们之间的内在联系。概念模型是整个数据模型的基础，常见的概念模型有 E-R 模型、扩充的 E-R 模型及谓词模型等。

2. 逻辑数据模型

逻辑数据模型又称数据模型，是一种面向数据库系统的模型，着重于在数据库系统中的实现。概念模型只有在转换为数据模型后才能在数据库中得以表示。常见的逻辑数据模型有层次模型、网状模型、关系模型、面向对象模型等。

3. 物理数据模型

物理数据模型又称物理模型，是一种面向计算机物理表示的模型，是数据模型在计算机上的物理结构表示。

7.2.3 三种数据模型

1. 层次模型

层次模型是最早发展起来的数据库模型。层次模型的基本结构是树形结构，自顶向下、层次分明，如家族结构、行政组织机构等。

层次模型支持的操作主要有查询、插入、删除和更新。层次模型在进行插入、删除、更新操作时，要满足层次模型的完整性约束条件：进行插入操作时，如果没有相应的父节点值就不能插入子节点值；进行删除操作时，如果删除父节点值，则相应的子节点值也被删除；进行更新操作时，应更新所有相应记录，以保证数据的一致性。

层次模型的数据结构比较简单，操作简单；对于实体间的联系来说是固定的，且预先定义好的应用系统，层次模型有较高的性能；层次模型可提供良好的完整性支持。层次模型受文件系统影响大，模型受限制多，物理成分复杂，它不适合于表示非层次性的联系；对于插入和删除操作的限制比较多。

2. 网状模型

网状模型的出现略晚于层次模型，使用图形表示，网状模型是一个不加任何条件限制的无向图。网状模型在结构上较层次模型好，不像层次模型受到多个条件的限制。网状模型中的基本操作是简单二级树中的操作，它包括查询、增加、删除、修改等操作。网状模型相比于层次模型，对数据表示或数据操纵有更高的效率、更为成熟。网状模型数据库系统的不足是，在使用时涉及系统内部的物理因素较多，用户操作和使用并不方便，其数据模式与系统实现比较复杂。

3. 关系模型

(1) 关系的数据结构

关系模型采用二维表（简称表）来表示。一个关系对应一个二维表，二维表就是关系名。在二维表中的一行，称为一个元组；在二维表中的列，称为属性或字段。属性的个数称为关系的元或度。列的值称为属性值，属性值的取值范围称为值域。在一个关系中存在唯一标识一个实体的一个属性或属性集称为实体的键。一个关系中，可能有多个键，这些键可以称为关系的候选码或候选键。在一个关系的若干候选键中指定一个用来

唯一标识该关系的键，该候选键称为主关键字，或简称为主键、关键字、主码。每一个关系都有且只有一主键，通常用较小的属性组合作为主键，如学生表中的“学号”可以作为主键。

(2) 关系操作

关系操作是建立在关系上的数据操作，一般有查询、删除、插入及修改4种操作。

①数据查询 用户可以查询关系数据库中的数据，它包括一个关系内的查询及多个关系间的查询。

一个关系内的查询操作的基本单位是元组分量，其基本过程是先定位后操作。定位包括纵向定位和横向定位。纵向定位是指定关系中的一些属性（称列指定），横向定位是选择满足某些逻辑条件的元组（称行选择）。进行纵向和横向定位后，一个关系中的元组分量即可确定。定位后即可进行查询操作，将定位的数据从关系数据库中取出并放至内存。

对多个关系间的数据查询可分为三步：首先将多个关系合并成一个关系，其次对合并后的一个关系进行定位，最后进行操作。

关系数据库的查询可以分解为一个关系内的属性指定、一个关系内的元组选择、两个关系的合并、三个基本定位操作及一个查询操作。

②数据删除 它的基本单位是一个关系内的元组，是指将关系内指定的元组删除。数据删除可以分解为一个关系内的元组选择与关系中的元组删除两个基本操作。定位只需要横向定位而无须纵向定位，定位后即可执行删除操作。

③数据插入 它仅对一个关系而言，在指定关系中插入一个或多个元组。数据插入不需要定位，直接进行元组插入操作即可。

④数据修改 它是指在一个关系中修改指定的元组和属性。数据修改不是一个基本操作，它可以分解为删除需要修改的元组、插入修改后的元组两个更基本的操作。

(3) 关系中的数据约束

关系模型中有三类数据约束，即实体完整性约束、参照完整性约束及用户定义的完整性约束。实体完整性约束和参照完整性约束由关系数据库系统自动支持。用户定义的完整性约束，则由关系数据库系统提供完整性约束语言，用户利用该语言写出约束条件，运行时由系统自动检查。实体完整性约束和参照完整性约束是关系数据库必须遵守的规则，在任何一个关系数据库管理系统中均由系统自动支持。

①实体完整性约束 该约束要求关系的主键中的属性值不能为空值。

②参照完整性约束 该约束是关系之间相关联的基本约束，不允许关系引用不存在的元组，关系中的外键要么是所关联关系中实际存在的元组，要么为空值。

③用户定义的完整性约束 该约束是针对具体数据环境与应用环境由用户具体设置的约束，它反映了具体应用中数据的语义要求。

7.3 关系数据库

7.3.1 关系代数运算

关系数据库系统的特点之一是它建立在数学理论的基础之上，有很多数学理论可以表示关系模型的数据操作，其中最为著名的是关系代数。关系数据库中使用的 SQL (structure query language, 结构查询语言) 可以支持关系代数的运算和操作。

1. 关系模型的基本操作

关系是由若干个不同的元组组成的，关系可视为元组的集合。 n 元关系是一个 n 元有序组的集合。

关系模型有插入、删除、修改和查询四种基本操作，它们又可以进一步分解为以下六种基本操作。

(1) 关系的属性指定 指定一个关系内的某些属性，用它确定关系这个二维表中的列，它主要用于检索或定位。

(2) 关系的元组选择 使用一个逻辑表达式给出关系中满足此表达式的元组，用它确定关系这个二维表的行，它主要用于检索或定位。

使用上述两种操作即可确定一个二维表中满足一定行、列要求的数据。

(3) 两个关系的合并 将两个关系合并成一个关系。使用此操作可以不断合并从而可以将若干个关系合并成一个关系，以建立多个关系之间的检索与定位。

使用上述三个操作可以进行多个关系的定位。

(4) 关系的查询 在一个关系或多个关系之间进行查询，查询的结果也为关系。

(5) 关系元组的插入 在关系中添加一些元组，用它完成插入与修改。

(6) 关系元组的删除 在关系中删除一些元组，用它完成删除与修改。

2. 关系模型的基本运算

由于操作是对关系的运算，而关系是有序组的集合，因此可以将操作看成集合的运算。

(1) 插入 设有关系 R 需要插入若干元组，要插入的元组组成关系 R' ，则插入操作可使用集合并运算表示为 $R \cup R'$ 。

(2) 删除 设有关系 R 需要删除一些元组，要删除的元组组成关系 R' ，则删除操作可使用集合差运算表示为 $R - R'$ 。

(3) 修改 修改关系 R 中的元组内容时，可使用以下方法实现。

① 设需要修改的元组构成关系 R' ，则先进行删除，得 $R - R'$ 。

② 设修改后的元组构成关系 R'' ，此时将其插入即可得到结果 $R - R' \cup R''$ 。

(4) 查询

用于查询的三个操作无法使用传统的集合运算表示，需要引入一些新的运算。

①投影运算 它是从关系中指定若干个属性组成新的关系的操作。

②选择运算 它是从关系中找出满足给定条件的元组的操作。

③笛卡儿积运算 对于两个关系的合并操作可以使用笛卡儿积表示。设有 n 元关系 R 和 m 元关系 S ，它们分别有 p 、 q 个元组，则关系 R 和 S 的笛卡儿积运算记为 $R \times S$ 。该关系是一个 $n+m$ 元关系，元组个数是 $p \times q$ ，由 R 和 S 的有序组组合而成。

3. 关系代数中的扩充运算

关系代数中除上述几个最基本的运算外，为操纵方便还需增添一些运算，这些运算均可由基本运算导出。常用的扩充运算有交、除、连接及自然连接等。

(1) 交运算 关系 R 与 S 经交运算后所得到的关系是由那些既在 R 中又在 S 中的有序组组成的，记为 $R \cap S$ 。

(2) 除运算 如果将笛卡儿积运算看作乘运算的话，那么除运算就是它的逆运算。当关系 $T=R \times S$ 时，则可将除运算写为 $T \div R=S$ 或 $T/R=S$ 。

(3) 连接与自然连接运算 在数学上，可以使用笛卡儿积建立两个关系之间的连接，但这样得到的关系庞大，而且数据大量冗余。在实际应用中，一般两个相互连接的关系往往须满足一些条件，所得到的结果也较为简单，这样就引入了连接运算与自然连接运算。

连接运算又称为 θ 连接运算，是一种二元运算，即从两个关系的笛卡儿积中选择满足给定属性间的条件的那些元组。

7.3.3 关系的完整性

关系的完整性约束条件主要包括三大类：实体完整性、参照完整性和用户定义完整性。

(1) 实体完整性 确保数据库中的每个元组（行）必须是唯一的，即没有两个完全相同的元组。每个关系（表）都应该有一个或多个属性（字段）作为主码，用来唯一标识每个元组。主码的每个属性都必须有值，这个值在表中是唯一的，不能重复。

(2) 参照完整性 也称为引用完整性，确保了数据库中不同表之间的逻辑关系的一致性和完整性。它定义了建立关系之间联系的主关键字与外部关键字引用的约束条件。关系数据库中通常都包含多个存在相互联系的关系，关系与关系之间的联系是通过公共属性来实现的。

(3) 用户定义完整性 根据应用环境的要求和实际的需要，对某一具体应用所涉及的数据提出约束性条件。这一约束机制一般不应由应用程序提供，而应由关系模型提供定义并检验，用户定义完整性主要包括字段有效性约束和记录有效性。

这三类完整性约束条件是关系模型必须满足的，其中实体完整性和参照完整性是必须满足的完整性约束条件，应该由关系系统自动支持。

7.3.4 关系规范化

规范化理论是数据库逻辑设计的指南和工具，具体步骤如下：

(1) 考察关系模型的函数依赖关系，确定范式等级。逐一分析各关系模式，考察是否

存在部分函数依赖、传递函数依赖等，确定它们分别属于第几范式。

(2) 对关系模式进行合并或分解。根据应用要求，考察这些关系模式是否合乎要求，从而确定是否要对这些模式进行合并或分解，例如对于具有相同主码的关系模式一般可以合并；对于非 BCNF 的关系模式，要考察“异常弊病”是否在实际应用中产生影响，对于那些只是查询，不执行更新操作，则不必对模式进行规范化（分解），实际应用中并不是规范化程度越高越好，有时分解带来的消除更新异常的好处与经常查询需要频繁进行自然连接所带来的效率低相比会得不偿失。对于那些需要分解的关系模式，可以用规范化方法和理论进行模式分解。最后，对产生的各关系模式进行评价、调整，确定出较合适的一组关系模式。

关系规范化理论提供了判断关系逻辑模式优劣的理论标准，帮助预测模式可能出现的问题，是产生各种模式的算法工具，因此是设计人员的有力工具。

7.3.5 关系数据库的设计方法

1. 数据库设计概述

数据库设计是指对一个给定的应用环境，构造最优的数据模式，设计一个能满足用户要求、性能良好的数据库。

数据库设计的基本任务是根据用户对象的信息需求、处理需求和数据库的支持环境（包括硬件、操作系统和数据库管理系统）设计出数据模式。信息需求主要是指用户对象的数据及其结构，它反映了数据库的静态要求；处理需求则表示用户对象的行为和动作，它反映了数据库的动态要求。数据库设计中有一定的制约条件，它们是系统设计平台，包括系统软件、工具软件及设备、网络等硬件。数据库设计是指在一定平台制约下，根据信息需求与处理需求设计出性能良好的数据模式。

数据库设计有两种方法，一种是以信息需求为主，兼顾处理需求，称为面向数据的方法；另一种是以处理需求为主，兼顾信息需求，称为面向过程的方法。早期的应用系统以面向过程的方法较多；近期的大型系统因数据结构复杂、数据量庞大、处理流程简单，以面向数据的方法较多。

数据库设计目前一般采用生命周期法，将整个数据库应用系统的开发分解成目标独立的若干阶段。数据库应用系统的开发过程可分为需求分析阶段、概念设计阶段、逻辑设计阶段、物理设计阶段、编码阶段、测试阶段、运行阶段、进一步修改阶段。

2. 数据库设计的需求分析

需求分析简单地说，就是分析用户的需求，需求分析是设计数据库的起点，需求分析的结果是否准确反映用户的实际要求，将直接影响后面各阶段的设计，并影响设计结果是否合理和实用。

需求分析阶段是数据库设计的第一个阶段，这一阶段收集到的基础数据和一组数据流图是下一步设计概念结构的基础。

需求分析阶段的任务是通过详细调查现实世界要处理的对象（组织、部门、企业等），充分了解原系统的工作概况，明确用户的各种需求，在此基础上确定新系统的功能。新系

统必须充分考虑今后可能的扩充和改变,不能仅按当前应用需求来设计数据库。

调查的重点是“数据”和“处理”,通过调查要从中获得每个用户对数据库的如下要求。

(1) 信息要求 指用户需要从数据库中获得信息的内容与性质。由信息要求可以导出数据要求,即在数据库中需存储哪些数据。

(2) 处理要求 指用户要完成什么处理功能,对处理的响应时间有何要求,处理的方式是批处理还是联机处理。

(3) 安全性和完整性的要求 为了更好地完成调查任务,设计人员必须不断地与用户进行交流,与用户达成共识,以便逐步确定用户的实际需求,然后分析和表达这些需求。需求分析是整个设计活动的基础,也是最困难、最花时间的一步。需求分析人员既要懂得数据库技术,又要对应用环境的业务比较熟悉。

需求分析的方法有结构化分析方法和面向对象的方法。结构化分析方法使用自顶向下、逐层分解的方式分析系统。使用数据流图表达数据和处理过程的关系,数据字典对系统中数据的详尽描述,是各类数据属性的清单。数据字典是进行详细的数据收集和数据分析后所获得的主要结果。

3. 数据库的概念设计

(1) 数据库概念设计概述

数据库概念设计的目的是分析数据间内在的语义关联,在此基础上建立一个数据的抽象模型。

数据库概念设计的方法有以下两种。

①集中式模式设计法 这是一种统一的模式设计方法,它根据需求由一个统一机构或人员设计一个综合的全局模式。其设计简单、方便,强调统一与一致,适用于小型或并不复杂的单位或部门,而对大型的或语义关联复杂的单位则并不适合。

②视图集成设计法 这种方法是将一个单位分解成若干个部分,先对每个部分进行局部模式设计,建立各部分的视图,然后以各视图为基础进行集成。在集成过程中可能会出现一些冲突,这是由于视图设计的分散性形成的不一致所造成的,因此需对视图进行修正,最终形成全局模式。

视图集成设计法是一种由分散到集中的方法,它的设计过程复杂,但它能较好地反映需求,适合大型、复杂的单位。

(2) 数据库概念设计的过程

使用 E-R 模型与视图集成法进行设计时,需要按以下步骤进行。首先选择局部应用,再进行局部视图设计,最后对局部视图进行集成得到概念模式。

①选择局部应用 根据系统的具体情况,在多层的数据流图中选择一个适当层次的数据流图,让这组图中的每一部分对应一个局部应用,以这一层次的数据流图为出发点,设计分 E-R 图。

②视图设计

视图设计一般有以下三种设计次序。

自顶向下 这种方法是先从抽象级别高且普遍性强的对象开始逐步细化、具体化与特殊化。

由底向上 这种设计方法是先从具体的对象开始，逐步抽象、普遍化与一般化，最后形成一个完整的视图设计。

由内向外 这种设计方法是先从最基本与最明显的对象着手逐步扩充至非基本、不明显的其他对象。

上述三种方法为视图设计提供了具体的操作方法，设计者可以根据实际情况灵活掌握，可以单独使用也可以混合使用。有某些共同特性和行为的对象可以抽象为一个实体。对象的组成成分可以抽象为实体的属性。在进行设计时，实体与属性是相对而言的。同一事物，在一种应用环境中作为“属性”，在另一种应用环境中就必须作为“实体”。在给定的应用环境中，属性必须是不可分的数据项，属性不能与其他实体发生联系，联系只发生在实体之间。

4. 数据库的逻辑设计

(1) 从 E-R 图向关系模式转换。数据库的逻辑设计的主要工作是将 E-R 图转换为指定的关系模式。从 E-R 图到关系模式的转换是比较直接的，实体与联系都可以表示成关系，E-R 图中的属性也可以转换为关系的属性。实体集也可以转换为关系。

(2) 逻辑模式规范化及调整、实现。关系数据库设计的关键是关系数据库模式的设计，即确定构造几个关系模式及每一模式各自包含的属性，将相互关联的模式组成合适的关系模型。关系数据库的设计必须在关系数据库规范化理论的指导下进行。

设计不良的关系模式会有数据冗余、插入异常、删除异常及修改异常等问题。

(3) 关系视图设计。逻辑设计的另一个重要内容是关系视图的设计，它又称外模式设计。关系视图是在关系模式基础上所设计的直接面向操作用户的视图，它可以根据用户需求随时创建。

关系视图的作用大致有以下几点：

①提供数据的逻辑独立性，使应用程序不受逻辑模式变化的影响。数据的逻辑模式会随着应用的发展而不断变化，逻辑模式的变化必然会影响应用程序的变化，这就会产生极为麻烦的维护工作。

关系视图则起到逻辑模式与应用程序之间的隔离墙作用，有了关系视图后，建立在其上的应用程序就不会随逻辑模式的修改而产生变化，此时变动的仅是关系视图的定义。

②能适应用户对数据的不同需求。每个数据库有一个非常庞大的结构，每个数据库用户只希望知道他们自己所关心的那部分结构，不必知道数据的全局结构，以减轻用户在此方面的负担。可以使用关系视图屏蔽用户不需要的模式，仅将用户感兴趣的部分呈现出来。

③有一定的数据保密功能。关系视图为每个用户划定了访问数据的范围，从而在应用的各用户间起到一定的保密隔离作用。

5. 数据库的物理设计

数据库的物理设计的主要目标是对数据库内部物理结构进行调整并选择合理的存取路

径，以提高数据库访问速度及有效利用存储空间。在现代关系数据库中，已经大量屏蔽了内部物理结构，因此留给用户参与物理设计的余地并不多，一般留给用户参与物理设计的内容大致有索引设计、集簇设计和分区设计。

7.3.6 常见的关系型数据库

关系模型由关系数据结构、关系操作集合、关系完整性约束三部分组成。简单说，关系型数据库是由多张能互相联接的二维行列表格组成的数据库。当前主流的关系型数据库有 Oracle、DB2、Microsoft SQL Server、Microsoft Access、MySQL 等。

1. Oracle

Oracle 是 1983 年推出的世界上第一个开放式商品化关系型数据库管理系统。它采用标准的 SQL 结构化查询语言，支持多种数据类型，提供面向对象存储的数据支持，具有第四代语言开发工具，支持 Unix、Windows NT、OS/2、Novell 等多种平台。除此之外，它还具有很好的并行处理功能。Oracle 产品主要由 Oracle 服务器产品、Oracle 开发工具、Oracle 应用软件组成，也有基于微机的数据库产品。主要满足对银行、金融、保险等企业、事业开发大型数据库的需求。

2. DB2

DB2 是美国 IBM 公司开发的一套关系型数据库管理系统，它主要的运行环境为 UNIX（包括 IBM 自家的 AIX）、Linux、IBM i（旧称 OS/400）、z/OS，以及 Windows 服务器版本。

DB2 主要应用于大型应用系统，具有较好的可伸缩性，可支持从大型机到单用户环境，应用于所有常见的服务器操作系统平台下。DB2 提供了高层次的数据利用性、完整性、安全性、可恢复性，以及小规模到大规模应用程序的执行能力，具有与平台无关的基本功能和 SQL 命令。

DB2 采用了数据分级技术，能够使大型机数据很方便地下载到 LAN 数据库服务器，使得客户机/服务器用户和基于 LAN 的应用程序可以访问大型机数据，并使数据库本地化及远程连接透明化。

DB2 以拥有一个非常完备的查询优化器而著称，其外部连接改善了查询性能，并支持多任务并行查询。DB2 具有很好的网络支持能力，每个子系统可以连接十几万个分布式用户，可同时激活上千个活动线程，对大型分布式应用系统尤为适用。

3. Microsoft SQL Server

SQL Server 是 Microsoft 公司推出的关系型数据库管理系统。具有使用方便可伸缩性好与相关软件集成程度高等优点，可跨越从运行 Microsoft Windows 98 的膝上型电脑到运行 Microsoft Windows 2012 的大型多处理器的服务器等多种平台使用。

Microsoft SQL Server 是一个全面的数据库平台，使用集成的商业智能 BI 工具提供了企业级的数据管理。Microsoft SQL Server 数据库引擎为关系型数据和结构化数据提供了更安全可靠的存储功能，使您可以构建和管理用于业务的高可用和高性能的数据应用程序。

4. Microsoft Access

Access 是由微软发布的关系数据库管理系统。它结合了 MicrosoftJet Database Engine 和图形用户界面两项特点，是 Microsoft Office 的系统程序之一。

Access 是微软把数据库引擎的图形用户界面和软件开发工具结合在一起的一个数据库管理系统。它是微软 OFFICE 的一个成员，在包括专业版和更高版本的 office 版本里面被单独出售。2018 年 9 月 25 日，最新的微软 Office Access 2019 在微软 Office 2019 里发布。

MS ACCESS 以它自己的格式将数据存储存储在基于 Access Jet 的数据库引擎里。它还可以直接导入或者链接数据这些数据存储在其他应用程序和数据库。

5. MySQL

MySQL 是一个关系型数据库管理系统，由瑞典 MySQL AB 公司开发，目前属于 Oracle 旗下产品。MySQL 是最流行的关系型数据库管理系统之一，在 WEB 应用方面，MySQL 是最好的 RDBMS Relational Database Management System，关系数据库管理系统应用软件之一。

MySQL 是一种关系数据库管理系统，关系数据库将数据保存在不同的表中，而不是将所有数据放在一个大仓库内，这样就增加了速度并提高了灵活性。

MySQL 所使用的 SQL 语言是用于访问数据库的最常用标准化语言。MySQL 软件采用了双授权政策，分为社区版和商业版，由于其体积小、速度快、总体拥有成本低，尤其是开放源码这一特点，一般中小型网站的开发都选择 MySQL 作为网站数据库。

7.4 大数据技术基础

7.4.1 大数据的定义

对于“大数据”（Big data）研究机构 Gartner 给出了这样的定义。“大数据”是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力来适应海量、高增长率和多样化的信息资产。麦肯锡全球研究所给出的定义是：一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合，具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低四大特征。

大数据技术的战略意义不在于掌握庞大的数据信息，而在于对这些含有意义的数据进行专业化处理。换言之，如果把大数据比作一种产业，那么这种产业实现盈利的关键，在于提高对数据的“加工能力”，通过“加工”实现数据的“增值”。从技术上看，大数据与云计算的关系就像一枚硬币的正反面一样密不可分。大数据必然无法用单台的计算机进行处理，必须采用分布式架构。它的特色在于对海量数据进行分布式数据挖掘。但它必须依托云计算的分布式处理、分布式数据库和云存储、虚拟化技术。

随着云时代的来临，大数据（Big data）也吸引了越来越多的关注。分析师团队认为，大数据（Big data）通常用来形容一个公司创造的大量非结构化数据和半结构化数据，这些数据在下载至关系型数据库用于分析时会花费过多时间和金钱。大数据分析常和云计算联系在一起，因为实时的大型数据集分析需要像 MapReduce 一样的框架来向数十、数百或

甚至数千的电脑分配工作。大数据需要特殊的技术，以有效地处理大量的容忍经过时间内的数据。

适用于大数据的技术，包括大规模并行处理（MPP）数据库、数据挖掘、分布式文件系统、分布式数据库、云计算平台、互联网和可扩展的存储系统。

7.4.2 大数据处理一般流程

(1) 数据采集 通过传感器、监测设备、物联网等手段，采集来自实际物体或系统的数据，这些数据可以是温度、压力、振动、电流等物理量的测量值，也可以是图像、视频等感知数据。

(2) 数据清洗 对采集到的原始数据进行清洗和处理，去除噪声、异常值和重复数据，确保数据的准确性和一致性。

(3) 数据存储 将清洗后的数据存储到数据库、数据湖或其他存储系统中，选择合适的数据存储技术和架构可以确保数据的可靠性、可扩展性和安全性。

(4) 数据分析和处理 对存储的数据进行分析和处理，提取有价值的信息和模式，数据分析可以包括统计分析、机器学习、深度学习等技术，以实现数据的理解、预测和优化。

(5) 数据可视化 将分析结果以可视化的形式展示，通常使用图表、图像、仪表盘等方式展示数据和分析的结果，数据可视化有助于用户理解和解释数据，支持决策和行动。

(6) 数据安全和隐私保护 在整个大数据处理流程中，数据安全和隐私保护是非常重要的，这包括对数据进行加密、访问控制、身份验证等，以确保数据的机密性和完整性，同时还需要遵守相关的法律法规，保护用户的隐私权益。

综上所述，大数据处理的基本流程包括数据采集，数据清洗，数据存储，数据分析和处理，数据可视化，数据安全和隐私保护等环节，每个环节都需要经过精细的设计和运行，以提高数据处理的效率和准确性。

7.4.3 大数据处理关键技术

大数据处理的关键技术确实包括大数据采集、大数据预处理、大数据存储及管理、大数据分析及挖掘、大数据展现和应用等多个方面。其中大数据采集和预处理是基础步骤，负责收集和准备数据以供后续分析。大数据存储及管理技术则涉及如何高效地存储和管理大量数据。大数据分析及挖掘是对数据进行深入分析，发现数据中的模式和趋势。大数据展现技术则关注如何将分析结果以有效的方式展示给用户。最后，大数据应用技术涉及如何将大数据应用于实际业务场景中，以提升业务价值。

具体到大数据采集和预处理，这两个关键技术涉及到使用多个数据库接收来自客户端（如 Web、App 或传感器等）的各种类型数据，进行简单的查询和处理工作。数据的集成与处理技术则关注将分散的数据库中的数据集成到一个集中的大型分布式数据库或分布式存储集群中，以便进行集中处理。预处理技术主要完成对已接收数据的辨析、抽取、清洗等操作，以确保数据的质量和适用性。

此外，大数据智能感知层和技术基础支撑层也是大数据采集和预处理中重要的组成部分，涉及到数据传感体系、网络通信体系、传感适配体系、智能识别体系及软硬件资源接入系统的设计和实现。这些技术共同构成了大数据处理的基础架构，使得大数据的采集和预处理成为可能。

7.4.4 主流大数据处理平台

大数据平台指的是具体的平台比如某个商业公司用的某个基于大数据开发的平台，大数据平台主要有阿里巴巴、华为云、百度云、浪潮、腾讯等。

大数据平台可以根据应用场景和功能需求，分为多种类型。以下是其中一些类型的大数据平台。

(1) 分布式计算平台 Apache Hadoop、Apache Spark、Apache Flink 等，提供分布式存储和计算能力，支持海量数据处理和分析。

(2) 实时数据处理平台 Apache Kafka、Apache Storm、Apache Ignite 等，专注于实时数据处理和流计算，适用于流媒体、监控和物联网等场景。

(3) 数据仓库平台 Amazon Redshift、Google BigQuery、Snowflake 等，提供基于云的数据仓库解决方案，支持大规模、高速的数据查询和分析。

(4) 海量数据存储平台 Hadoop 分布式文件系统 (HDFS)、Amazon S3、Google Cloud Storage 等，提供海量数据的分布式存储能力，支持多种数据格式和存储方案。

(5) 数据可视化平台 Tableau、QlikView、PowerBI 等，提供数据可视化和报表工具，帮助用户更直观地理解数据和呈现数据分析结果。

(6) 移动端分析平台 Google Firebase、Leanplum、Amplitude 等，专注于移动设备的数据分析和用户行为跟踪，帮助开发人员更好地理解和优化移动应用程序。

综上所述，大数据平台可以按应用场景和功能需求分为多种类型，每种平台都针对不同的数据处理和分析需求提供定制化的解决方案。

7.4.5 NoSQL 数据库

NoSQL 数据库，全称为 Not Only SQL，是一种非关系型数据库管理系统，它相对于传统的关系型数据库（如 MySQL、Oracle 等）具有以下特点：

(1) 非结构化数据存储 NoSQL 数据库通常以键值对、文档、列族或图等非结构化的方式存储数据，与传统的关系型数据库采用表格结构不同，这种非结构化的数据存储模型使得 NoSQL 数据库更加灵活，可以存储各种形式的数据。

(2) 高可扩展性 NoSQL 数据库具有良好的可扩展性，可以方便地在集群中增加新的节点，以满足数据规模的增长需求。它们通过分布式架构和水平扩展来支持大规模数据存储和高并发访问。

(3) 高性能 由于采用了简化的数据模型和灵活的存储结构，NoSQL 数据库在读写操作上通常能够提供较高的性能。一些 NoSQL 数据库还提供了缓存机制和数据分片等功能，进一步提升了性能。

(4) 低一致性 为了获得高可用性和可扩展性，一些 NoSQL 数据库在数据一致性方面进行了一定的弱化。它们通常采用了最终一致性或柔性事务的方式，允许数据在不同节点间存在一定的时间差。

(5) 适用于大数据和实时数据处理 NoSQL 数据库广泛应用于大数据场景和实时数据处理领域。由于其可扩展性和高性能，它们能够有效地处理海量数据和高并发的数据访问。

(6) 灵活的数据模型 NoSQL 数据库采用键/值、列族等非关系模型，允许在一个数据元素里存储不同类型的数据，这种灵活性使得 NoSQL 数据库能够适应不同类型和结构的数据。

(7) 与云计算紧密融合 云计算具有很好的水平扩展能力，NoSQL 数据库可以凭借自身良好的横向扩展能力，充分自由利用云计算基础设施，很好地融入到云计算环境中。

(8) 高可用性 NoSQL 数据库在不太影响性能的情况之下，就能够非常方便的实现高可用的架构，典型的例如 Cassandra、HBase 模型，通过模型的话也可以实现高可用。

根据数据存储方式的不同，NoSQL 数据库可以分为四大类：键值存储、列族存储、文档存储和图形存储。常见的 NoSQL 数据库包括 MongoDB、Cassandra、HBase 等。

7.4.6 MPP 数据库

MPP 是“Massively Parallel Processing”（大规模并行处理）的缩写，MPP 型数据库是一种特定类型的数据库管理系统，它是为处理大规模数据集和复杂查询而设计的。

MPP 型数据库的设计目标是通过水平扩展和并行处理来提供高性能和高吞吐量。它通常在多个计算节点上分布数据，并且在执行查询时，这些节点会并行处理数据的不同部分，从而实现查询的加速。

这些数据库通常用于处理大型数据仓库和分析型应用，特别是在需要进行复杂分析和大规模数据处理时。相较于传统的关系型数据库，MPP 型数据库在处理海量数据和复杂查询时表现更加优异。

MPP 型数据库的特点：

(1) 并行处理 MPP 数据库能够在多个计算节点上同时处理数据，从而加速查询的执行。

(2) 水平扩展 通过添加更多的计算节点，MPP 数据库可以轻松地扩展以适应不断增长的数据量和负载。

(3) 列存储 许多 MPP 数据库采用列存储技术，它可以提供更高的数据压缩比率和查询性能。

(4) 复杂查询优化 MPP 型数据库通常优化复杂查询，以便更有效地处理连接、聚合和其他复杂操作。

(5) 并发控制 由于多个计算节点同时访问数据，MPP 数据库需要有效的并发控制机制来保证数据一致性和可靠性。

总的来说，MPP 型数据库在大规模数据处理和复杂查询场景中表现出色，能够满足

企业对于高性能分析和数据处理的需求。

7.4.7 大数据应用

大数据常见技术应用有农业互联网、金融业互联网、电子商务、医疗器械行业、零售业大数据、生物科技。

1. 农业互联网

农业互联网大数据在农牧业上的运用关键就是指根据将来商业服务的要求来开展农牧业商品生产制造，减少菜贱伤农的几率。

2. 金融业互联网

金融业互联网大数据在金融业运用范畴范围广。

互联网大数据在金融业的运用能够小结为下列两个层面：

(1) 大数据营销 根据顾客消费习惯性、所在位置、消费时间开展强烈推荐。

(2) 风险防控 根据顾客消费和现金流量出示资信评级或股权融资适用，运用顾客社交媒体个人行为纪录透支卡风控。

3. 电子商务

电子商务电商数据比较集中，信息量大，类型较多，未来运用大数据将有大量的空间，包含分析潮流趋势，消费发展趋势、地区消费特性、顾客消费习惯性、各种各样消费者行为的相关性、消费市场、危害消费的关键要素等。

4. 医疗器械行业

医疗器械行业医疗器械行业有着很多的病案，病理报告，痊愈计划方案，药品汇报这些。在将来，凭借数据管理平台人们能够搜集不一样病案和医治计划方案，及其患者的本质特征，能够创建对于病症特性的数据库查询。

5. 零售业大数据

零售业大数据的应用有两个方面，一个方面是零售业能够掌握顾客消费爱好和发展趋势，开展货品的大数据营销，减少营销推广成本费；另一个方面是根据顾客选购商品，为顾客出示将会选购的其他商品，扩张销售总额，也归属于大数据营销层面。此外，零售业能够根据互联网大数据把握将来消费发展趋势，有益于热销产品的拿货管理方法和过季货品的解决。

6. 生物科技

生物科技关键就是指云计算技术在基因分析上的运用，根据数据管理平台人们能够将本身和植物体基因分析的结果开展纪录和储存，运用创建应用场景云计算技术的遗传基因数据库查询。云计算技术将会加快遗传基因技术性的科学研究，迅速协助生物学家开展实体模型的创建和遗传基因组成模拟计算。

第 8 章

人工智能

人工智能（Artificial Intelligence, AI）是一门横跨自然科学、技术科学和社会科学，具有较强综合性的前沿学科。随着人工智能理论和技术在持续迭代中逐步发展成熟，其正被广泛应用于社会生产生活的各个领域，并被认为“将成为第四次工业革命的标准”，推动传统实体经济实现革命性升级与发展，从而催生智能经济新形态，成为发展新质生产力的重要引擎。

8.1 人工智能概述

8.1.1 人工智能的定义

人工智能是一个以计算机科学为基础，由哲学、数学、经济学、计算机、神经科学、心理学、语言学等多学科交叉融合的交叉学科、新兴学科，研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。

现代意义上的“人工智能”概念，最早起源于 1956 年 8 月在美国汉诺斯小镇达特茅斯学院举行的一次头脑风暴讨论会，史称“达特茅斯会议”。这次会议上，约翰·麦卡锡、马文·明斯基、克劳德·香农、艾伦·纽厄尔、赫伯特·西蒙、内森·罗切斯特等科学家共同提出，“人工智能”是指“让机器能像人那样理解、思考和学习，即用计算机模拟人的智能”。具体而言，是通过让机器使用语言，形成抽象和概念，解决在当时认知下只能由人类来求解的各种问题，并改进机器本身。可见，“人工智能”概念被提出时的基本目标是将人的创造性、自我改进和语言应用能力等才能，复制到能在复杂、变化的环境中自主运行的机器上。

那么如何才能判断机器是否能模拟人的“智能”呢？1950 年“人工智能之父”阿兰·图灵（Alan Turing）提出著名的“图灵测试”：在测试人与被测试者（机器）不接触的情况下，经过多次问答后，测试人无法根据这些问题判断对方是人还是计算机，那么就可以认为这台机器是具有智能的。简言之，“智能”就是要“让机器会听、会看、会说、

会思考、会行动、会决策，就像人类一样”。图灵测试作为一个虽然主观但可操作的标准，被认为在一定程度上解决了长久以来关于如何定义思考的哲学争论。2014年聊天程序“尤金·古斯特曼”（Eugene Goostman）首次被英国皇家学会认为“通过”图灵测试。

基于这样的认识，一些教材将人工智能定义为“由人制造出来的具有一定智能的系统，能够理解、学习、推理、适应和解决问题”，或者“智能体（intelligent agent）的研究与设计”。人工智能领域经典教材《人工智能：一种现代的方法》按照行为、思维，以及类人与否，将人工智能历史上不同的人工智能定义归为四类：类人行动、类人思考、合理思考、合理行动。人工智能机器主要表现为会听（语音识别、机器翻译等）、会看（图像识别、文字识别等）、会说（语音合成、人机对话等）、会思考（人机对弈、定理证明等）、会学习（机器学习、知识表示等）、会行动（机器人、自动驾驶汽车等）。

需要特别说明的是，随着人工智能理论研究和实践的发展，人们对于“智能”的认识仍在不断变化，并没有完全或严格地使用智能的内涵或外延来定义人工智能，这导致当前流行的“人工智能”概念已与最初诞生时的定义有些不同。尽管如此，人们关于发展人工智能的目标仍是相对一致的，那就是“模拟和扩展人类的智能，使计算机能够执行复杂的任务，提高生产力和生活质量”。

8.1.2 人工智能的分类

人工智能历经半个多世纪的发展，在概念上产生了不同层次的分化，并带来一定程度的概念混乱。学界、产业界根据机器是否应该“思考”，或机器是否能够“思考”、机器是否能够超越人的“思考”等哲学思辨方向，将人工智能划分为弱人工智能、强人工智能和超人工智能三种类型。一些研究者则根据智能的实现程度，将人工智能分为运算智能、感知智能和认知智能三个层次。

1. 弱人工智能、强人工智能和超人工智能

弱人工智能（Artificial Narrow Intelligence, ANI）是专门针对某一特定任务设计的智能系统，例如语音识别系统、图像识别系统、智能助手、推荐系统等。拥有弱人工智能的机器不能实现自我思考、推理和解决问题，只是具有特定功能，能像人类一样或更好完成人类的任务，看起来像拥有智能。例如AlphaGo就是一种典型的弱人工智能产品，其在围棋方面达到很高的专业水平，但其在围棋领域的知识无法适用于象棋等其他领域。

强人工智能（Artificial General Intelligence, AGI）是指具有广泛认知能力和与人类相当的智能水平的系统。拥有强人工智能的机器不仅是一种工具，而且本身拥有意识、知觉和思维能力，能够执行思考、计划、解决问题、抽象思维、理解复杂理念、快速学习和从经验中学习等动作，思考、计划和解决问题，完成任何人类可做到的智力任务。

超人工智能（Artificial Super Intelligence, ASI）最早由牛津哲学教授尼克·波斯托姆所提出，是一种假想的基于软件的人工智能系统，它拥有比人类更先进的尖端认知功能和高度发达的思维能力，在包括科学创造、通识和社交技能等任何领域均比最聪明的人类大脑还要聪明的智能。超人工智能将打破人脑受到的维度限制，其所观察和思考的内容人脑已无法理解，不仅具有高度的创造性和自主性，还能够独立思考和创新。

2. 运算智能、感知智能和认知智能

(1) 运算智能 即“能存会算”，指机器能够像人一样拥有“记忆”“计算”能力。

(2) 感知智能 即“能听会看”，指机器能够通过各种智能感知能力与自然界交互。感知智能的实现主要是通过对大规模数据的采集，以及对图像、视频、声音等类型的数据进行特征抽取，完成结构化处理。例如自动驾驶汽车通过激光雷达等感知设备和人工智能算法主动感知外部环境。

(3) 认知智能 即“能理解会思考”，指机器能主动思考、理解新事物，拥有处理复杂的事实和情形的能力。认知智能需要在数据结构化处理的基础上，理解数据之间的关系和逻辑，并在理解的基础上进行分析和决策。具有认知智能的机器不用人类事先编程就可以实现自我学习，有目的推理并与人类自然交互。例如基于认知智能的客户机器人能根据当前场景以及人和机器的关系，采取不同的语气和情感与客户交流。

上述三种类型的人工智能中，认知智能复杂程度最高，其实现以感知智能和运算智能为基础，被认为是智能时代的第三阶段。许多学者认为通用智能强智能时代才是智能时代发展的最终阶段。

8.1.3 人工智能的发展历程

纵观人工智能近70年的发展史，可谓跌宕起伏、三起三落，学术界通常将其划分为起步发展期、反思发展期、应用发展期、低迷发展期、稳步发展期、蓬勃发展期等六个阶段。

1. 起步发展期（1956年至20世纪60年代初）

自图灵测试面世之后，1956年“人工智能”概念被正式提出，基于逻辑推理的符号主义方法盛行，机器定理证明、跳棋程序等一批令人瞩目的研究成果面世，掀起了人工智能发展的第一个高潮。但由于逻辑主义无法解决现实生活中的一些问题，且当时的计算机性能无法达到很好的运算速度，在一定程度上限制了逻辑主义的发展。

2. 反思发展期（20世纪60年代至70年代初）

进入20世纪60年代，人们对人工智能发展的期望高涨，开始尝试更具挑战性的任务。1966年麻省理工学院（MIT）魏泽鲍姆发布世界上第一个能通过脚本理解简单自然语言，产生类似人类互动的聊天机器人ELIZA。1970年，斯坦福大学计算机教授维诺格拉德开发的人机对话系统SHRDLU能分析指令，通过虚拟方块操作来完成任务。但随后的一些人工智能研发却因目标不切实际而接二连三失败，人们发现人工智能并未取得理想的效果，导致投入资金减少，人工智能研究进入第一个寒冬。1973年，著名数学家莱特希尔向英国政府提交了一份关于人工智能的研究报告，指出简单人工智能在解决多个变量问题时存在困难，甚至可能无法解决，在实验室环境中表现良好的人工智能在实际环境中却收效甚微。这些困难主要系当时计算机有限的内存和处理速度不足以解决任何实际的人工智能问题。

3. 应用发展期（20世纪70年代初至80年代中）

20世纪70年代感知机出现，自然语言处理技术和人机对话技术初步萌芽。人们研发

出首个专家系统 DENDRAL，能模拟人类专家的知识和经验解决特定领域的问题，实现了人工智能从理论研究走向实际应用、从一般推理策略探讨转向运用专门知识的重大突破，打破了人工智能研究领域低迷的局面。专家系统在医疗、化学、地质等领域的成功，推动人工智能走入应用发展的第二次高潮。

4. 低迷发展期（20 世纪 80 年代中至 90 年代中）

然而，随着人工智能应用规模不断扩大，专家系统应用领域狭窄、缺乏常识性知识、知识获取困难、推理方法单一、缺乏分布式功能、难以与现有数据库兼容等问题逐渐暴露出来，许多企业无法继续接受复杂的智能知识库维护与更新。人们认为人工智能应该真正实现智能化，拥有自己的感知系统，并能自主学习，而不是依赖人类知识。

5. 稳步发展期（20 世纪 90 年代中至 2010 年）

20 世纪 90 年代中期，随着互联网技术的快速发展，机器学习算法、知识表示、启发式搜索、计算机视觉、统计机器学习等快速发展，加速了人工智能的创新研究，促使人工智能技术进一步走向实用化。1997 年 IBM “深蓝” 超级计算机战胜了国际象棋世界冠军卡斯帕罗夫，成为首个在标准比赛时限内击败国际象棋世界冠军的计算机系统。2006 年，得益于多层神经网络即深度学习技术的突破，人工智能应用发展迎来第三次浪潮。

6. 蓬勃发展期（2011 年至今）

随着大数据、云计算、互联网、物联网等信息技术的发展，泛在感知数据和图形处理器等计算平台推动以深度神经网络为代表的人工智能技术飞速发展，大幅跨越了科学与应用之间的“技术鸿沟”。诸如图像分类、语音识别、知识问答、人机对弈、无人驾驶等人工智能技术实现了从“不能用、不好用”到“可以用”的技术突破，人工智能迎来爆发式增长的新高潮。其中比较典型的事件包括 2011 年，Watson（沃森）作为 IBM 公司开发的使用自然语言回答问题的人工智能程序参加美国智力问答节目，打败两位人类冠军，赢得了 100 万美元的奖金。2016 年，Google 人工智能 AlphaGo 以 4 : 1 的比分击败世界围棋冠军。

8.2 主要的人工智能技术

人工智能相对漫长而曲折的发展史，导致其形成了庞大复杂的知识体系和技术领域分支。特别是近 10 多年来，由于计算机性能的不不断提升以及互联网技术的发展，人工智能在强化学习、深度学习、机器学习等方向取得巨大进步，形成了智能机器人、模式识别、专家系统、自然语言处理、语音识别、图像识别等诸多研究方向，呈现出多元化发展态势。这些人工智能技术被应用在科学研究和日常生活的不同场景中，催生层出不穷的人工智能产品和服务，导致产业规模高速增长、业态极为丰富。

人工智能技术主要涉及机器学习、深度学习、计算机视觉、自然语言处理技术。

8.2.1 机器学习

机器学习（Machine Learning, ML）是指让计算机通过算法来学习数据中包含的内在

规律和信息，总结出适用于算法的经验和知识，使计算机面对问题时能做出与人类相似的决策。20世纪50年代中期，科学家们就开始利用概率论、统计学、逼近论、凸分析、算法复杂度等理论研究设计一些算法，然后编写程序让计算机执行以获得新知识或技能，不断积累和重构原有知识，从而模拟或实现人类的学习行为，并表现出越来越强的“智能”。因此机器学习与传统的为解决特定任务、硬编码的软件程序截然不同，其本质是设法让计算机通过分析数据、发现规律，自动获得可判断和处理其他类似问题的知识，这样就不必再完全依赖于专家的经验性知识。

1. 机器学习的基本过程

机器学习的基本过程一般可概括为定义模型函数、定义损失函数、训练优化三步。

(1) 定义模型函数 根据业务领域知识，构造一个数学表达式，来描述模型中目标变量（因变量）与各特征变量（自变量）之间的依赖关系，且此函数中的参数未知。

(2) 定义损失函数 利用损失函数来量化模型预测结果与实际训练数据之间的差异。损失函数的值越小，表明模型的预测结果越接近真实情况。在机器学习算法中，常使用随机梯度下降等方法来最小化损失函数。

(3) 训练优化 通过将训练样本输入至模型函数中，利用损失函数指导参数的调整，从而更新模型中的参数，随后进行下一轮迭代。通过多轮迭代，可确保模型函数能更好拟合训练数据，更准确地反映真实的数据生成过程。

为便于理解上述过程，可把机器学习的过程与人类对历史经验归纳的过程做个比对。人类在成长、生活过程中积累了很多的历史与经验。人类定期地对这些经验进行“归纳”，获得了生活的“规律”。当人类遇到未知的问题或者需要对未来进行“推测”的时候，人类使用这些“规律”，对未知问题与未来进行“推测”，从而指导自己的生活和工作的。把这个过程应用到机器上时，就需要为计算机提供大量的样本数据来“训练”，利用各种算法帮助计算机快速归纳、获取数据中隐含的经验性知识，从而对未知样本或新问题自动作出判断或处理。例如当我们浏览网上商城时经常会出现商品推荐的信息，这是网上商城根据你往期的购物记录和冗长的收藏清单，识别出这其中哪些是你真正感兴趣，并且愿意购买的产品（经验性知识），从而帮助商家识别潜在购买者，针对特定类型的客户设计产品营销方案，从而鼓励产品消费。

2. 机器学习算法的三种类型

根据训练样本及反馈方式的不同，机器学习中的算法主要包括监督学习、无监督学习、强化学习等三种类型。

(1) 监督学习算法

监督学习（Supervised Learning）是指利用一组已知类别或数值的样本调整模型的参数，达到对新样本正确分类或回归的目的。“监督”表明所有用于训练模型的样本同时包含特征和代表类别或数值的标签，即存在标准答案帮助修正模型。常用的监督学习算法包括决策树、朴素贝叶斯及支持向量机等。作为机器学习领域最常用且实用的方法，监督学习已在智能安防等多个领域被广泛应用。比如通过对不同天气条件、明暗度、拍摄角度条件下的汽车牌照图像样本进行人工标注，建立汽车牌照样本数据，然后通过卷积神经网络

算法建立牌照自动识别模型，从而自动识别新汽车牌照上的号码。

(2) 无监督学习算法

无监督学习 (Unsupervised Learning) 是指不需要带标签的样本数据来指导学习过程，其本质上是一种统计手段。与监督学习算法相比，无监督学习算法不需要数据的先验标签，亦难以量化评估学习效果。常用的无监督学习算法包括 k 均值聚类、主成分分析、稀疏自编码等。无监督学习算法主要被用于寻找样本数据的内在结构、模式或关联，实现聚类和降维等任务。聚类是通过计算样本之间的相似性，对样本划分类型，帮助寻找典型性和代表性的样本以减少样本规模。降维是将高维度的样本特征数据映射至较低的维度空间，保留样本的主要信息，剔除冗余或无关信息，以降低模型的算法复杂度，因此无监督学习算法在处理未标记的大型数据集方面尤为有价值。

(3) 强化学习算法

强化学习 (Reinforcement Learning) 是指智能体 (Agent) 通过不断“试错”与环境交互获得奖励反馈，并根据反馈进一步改进自身行为和决策的学习方式。强化学习模仿了人类学习的过程。强化学习算法的核心要素包括策略、奖励、价值和环境模型，这些要素共同定义了强化学习算法的工作机制和目标。常用的强化学习算法包括 Q-学习算法、瞬时差分法、自适应启发评价算法等。尽管强化学习算法直观且合理，但由于强化学习面临任务的复杂性，如何行动，如何定义回报，如何根据回报和当前状态决定下一步行动，仍存在较大困难。

除上述三种基础类型的机器学习算法外，还有根据自监督学习 (Self-supervised Learning)、半监督学习 (Semi-supervised Learning)、集成学习 (Ensembler Learning) 等其他学习策略衍生出的大量机器学习算法。在此不作一一阐述，请感兴趣的读者可查阅相关教材。

8.2.2 深度学习

深度学习 (Deep Learning) 是机器学习的子集，它通过使用多层人工神经网络 (Artificial Neural Network) 来学习数据的复杂模式和特征。

1. 人工神经网络

人工神经网络是计算机科学家模仿大脑的神经元结构设计出来的一种网状结构，其中包含若干相互连接的神经元节点。

人工神经网络的出现源于生物神经科学领域。如图 8-1 (a) 所示，生物神经学家研究发现，人和动物的神经系统均由大量神经元通过某种高效的连接方式构成，当生物体接收到外界刺激时，部分神经元就会处于兴奋状态，通过电信号在神经元之间传递信息。当神经元之间的连接程度不同时，传递的电信号强度也不同。接收电信号的神经元，以非简单叠加方式对来自多个神经元的电信号进行综合处理，再将处理结果传送给其他神经元。这种神经信号传导机制使得神经系统表现出强大的处理复杂信息的能力。

受此启发，1943 年美国心理学家麦卡洛克和数学家皮特斯等模拟生物神经元的构造和工作机制，提出了利用神经网络对信息进行处理的数学模型。在这个模型中，神经元接收来自 n 个其它神经元传递过来的输入信号 x_i ，神经元将接收到的输入信号 x_i 按照神经元

之间的连接权重 w_{ij} 叠加起来得到 $\sum x_i w_{ij}$ ，再与将当前神经元阈值 θ 进行比较，然后通过激活函数 f 向外表达输出变换后的信号，从而模拟各种非线性目标 $O_{\text{output}} = f(w^T x + b)$ 。其中，大多数情况下，激活函数 f 均为非线性激活函数，例如阈值函数、Sigmoid 函数、tanh 函数等。理论上，神经元数目足够多的人工神经网络模型可以逼近任意函数，如图 8-1 (b) 所示。

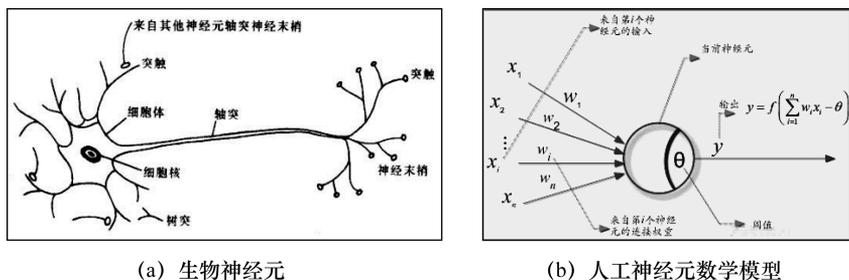


图 8-1 生物神经元与人工神经元数学模型

2. 浅层学习

人工神经网络中，一层神经网络由一组神经元节点共同构成。浅层学习是指层数较少的神经网络模型，例如感知机模型、径向基网络、自编码器以及玻尔兹曼机等。

感知机是只有一层神经元参与数据处理的人工神经网络，也是最基本的神经网络模型。如图 8-2 所示，感知机模型通过输入层接收输入信息 $X = (x_1, x_2, \dots, x_m)^T$ ，输入层神经元个数与输入数据的个数相同。由于感知机的输入层仅负责接收外部信息而不参与数据处理，通常也将输入层称为感知层。感知机模型通过输入层接收到外部信息 $X = (x_1, x_2, \dots, x_m)^T$ 之后，会将这些信息传输至输出层神经元。输出层神经元是感知机的数据处理单元，故也将输出层称为处理层。感知机模型的主要功能是对输入信号进行分类。

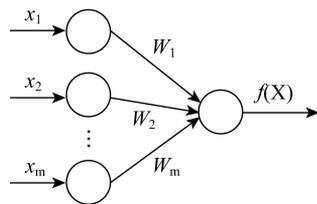


图 8-2 感知机模型

尽管浅层学习模型可逼近任意函数，但其模型容量或灵活性远不及具有较深层次的神经网络模型，难以满足对复杂任务求解的需求。

3. 深度学习

深度学习是与浅层学习相对应的概念。随着计算机硬件的巨大进步和大数据技术的发展，人们开始将多层神经网络叠加组合，构造出能解决比较复杂的实际问题的神经网络模型。这种包含多层神经元的神经网络模型，被称为深度学习模型。

深度学习模型除具有输入层、输出层外，其他各层被统称为隐藏层。1989 年，计算机专家用数学严格证明了通用近似定理，说明前馈神经网络只要隐藏层数足够多，就能以任意精度逼近任意可测函数，深度神经网络的理论基础由此建立。

深度学习模型由于使用多层非线性处理单元进行特征提取和转换，能够建模大量数据间的复杂关系，与浅层学习模型相比具有以下特点：

(1) 多层网络结构 深度学习模型由多层神经网络组成，每一层都学习输入数据的不同特征。这种层次化的结构使得深度学习能够捕捉到数据中更复杂、更抽象的模式。

(2) 自动特征提取 深度学习模型能够自动从原始数据中学习和提取有用的特征，无需人工干预。这是深度学习与传统机器学习方法的一个重要区别，后者通常需要手动选择和提取特征。

(3) 复杂的数据处理能力 由于深度学习模型的多层网络结构和大量的参数，深度学习模型特别适合处理高复杂度的数据，如高分辨率图像、视频和语音等。

(4) 数据需求 深度学习模型需要大量训练数据来优化其数百万个参数，以防止过拟合。而传统的机器学习模型则可以在较少的训练数据上表现良好。

(5) 计算资源 深度学习模型因其复杂性通常需要更高的计算资源，包括高性能GPU和大量内存。相比之下，传统机器学习模型的硬件要求较低。

(6) 模型解释性 深度学习模型因其操作的“黑箱”特性而难以解释。相比之下，许多传统机器学习模型（如决策树）的工作原理更容易理解和解释。

深度学习技术现已被广泛应用于计算机视觉（人脸识别、视频分析、医学图像处理）、自然语言处理（机器翻译、情感分析、语音识别）、游戏和机器人技术（增强现实、自动驾驶、自动化机器人）、医学诊断（临床诊断、疾病自动编码、公共卫生）等领域。

8.2.3 自然语言处理

自然语言处理是一门融合语言学、计算机科学、认知科学和数学于一体的科学。自然语言处理主要包含两个分支，分别是自然语言理解（Natural Language Understanding, NLU）和自然语言生成（Natural Language Generating, NLG）。

1. 自然语言理解

自然语言理解是指计算机能够理解自然语言文本的意义。自然语言理解旨在使机器能像人类一样理解和分析人类语言，实现智能问答、情感分析、信息抽取等任务。例如当一个人说“没有生气”时，这句话可能字面上表示没有生气，但实际语境中可能表达的是极度愤怒的情绪。这种微妙的语言理解能力正是自然语言理解试图在机器上实现的。

自然语言理解主要包括词法分析、句法分析、语义理解、语用理解等四个层次。

(1) 词法分析 指将连续文本切分成一个个单独的词语，并给每个词语标注其语法属性，例如名词、动词、形容词等。

(2) 句法分析 指依据语法规则将句子切分成不同的成分，并确定句子中词语之间的结构关系。句法分析通常会建立一棵句法树来描述主语、谓语、宾语、定语、状语等句子成分之间的关系。

(3) 语义理解 指让计算机能理解句子或文本的真正含义，包括词语和句子所表达的概念、关系和意图等。让计算机结合上下文信息和背景知识来推断和理解文本的真实含义，是自然语言理解的核心。

(4) 语用理解 指让计算机在语义理解的基础上，能更深入地理解特定情境下说话者的意图和态度，以及语句在特定语境中的含义和作用。

自然语言理解要解决的主要问题包括理解句子的领域 (domain)、意图 (intent) 和语义槽 (slot)。这些分别对应着自然语言理解的三个重要子领域：领域分类、意图识别和语义槽填充。举例来说，如果用户说“我想买一张上海飞北京的头等舱机票，下午5点出发，国航的”，系统首先需要确定这句话属于机票预订领域，然后进一步分析其意图是购买机票，接着提取出具体信息如出发地（上海）、目的地（北京）、舱位（头等舱）、时间（下午5点）和航空公司（国航）。这些信息便是语义槽填充的内容。

自然语言理解是实现人机自然语言通信的关键步骤，它不仅具有广泛的实际应用价值，还对深入探索人类语言能力和智能机制具有重要意义。随着深度学习技术的不断进步，自然语言理解有望在未来取得更大的突破和应用。

2. 自然语言生成

自然语言生成是指让计算机能根据一些关键信息及其在机器内部的表达形式，自动生成一段自然语言文本来表达给定的意图，以达到人机交流的目的。其本质是将非语言格式的数据转换为人类可理解的语言格式。

自然语言生成需要综合利用语言学、计算机科学和人工智能技术等多学科技术，经过内容确定、文本结构设计、句子聚合、语法化、参考表达式生成、语言实现等一系列复杂的处理步骤才能实现。

(1) 内容确定 让计算机需要根据关键信息（如提示词等）需要决定哪些信息应该被包含在正在构建的文本中，哪些不应该被包含，并确定被包含信息的内容边界。例如，报道一场篮球比赛时，应优先表达“什么时间”“什么地点”“哪两支球队”等信息，交通、天气、旅游等其他无关信息不应该被包含在内。

(2) 文本结构设计 让计算机根据问题类型或上下文语义关联，设计信息的逻辑结构，合理组织文本的顺序，以帮助读者更好理解和跟随文本内容。例如在报道体育赛事时通常会先介绍赛事基本情况，然后介绍本场比赛概况，接着是本场比赛结局，最后是本场比赛结局对后续比赛的影响或下一场比赛预报等。

(3) 句子聚合 让计算机对每一条表达具体信息的独立句子，按照句子中两个事件的时间关系或因果、实例、归纳等逻辑关系进行合并描述，增强句子的连贯性和信息的密度，使文本更流畅，更易于阅读。

(4) 语法化 让计算机遵照规范的语法形式、拼写方法和标点符号用法，对句子要表达的各种信息之间加上一些连接词，使句子看起来更像是一个完整的自然语言句子。

(5) 文本领域化 也被称为“参考表达式生成”，指让计算机识别内容的领域并使用该领域的特定词汇来构成完整的句子。通过文本领域化，能确保生成的自然语言文本用词更加专业和准确。

(6) 语言实现 让计算机将上述步骤确定的词和短语组合起来，形成一个结构良好的完整句子或文本段落，并通过文本完善确保整个文本的自然流畅和语言上的正确性。

自然语言生成技术现在机器翻译、自动摘要、新闻自动生成、聊天机器人和商业报告生成等领域，已具有十分广泛的应用。

8.2.4 计算机视觉

1. 计算机视觉的概念

计算机视觉是人工智能的一个分支领域，它指让计算机系统能够从图像、视频和其他视觉输入中获取有意义的信息，并根据该信息采取行动或提供建议。换言之，计算机视觉就是要让机器能够理解和解释视觉世界，并具有识别、追踪和决策等能力。如果说人工智能赋予计算机思考的能力，那么计算机视觉就是赋予机器发现、观察和理解的能力。

2. 计算机视觉的基本原理

计算机视觉的基本原理与人类视觉类似。人类视觉系统的优势在于人类依靠视网膜、视神经和视皮质，可在适当的环境下终生训练分辨物体以及辨别物体距离、物体动静与否以及图像是否存在问题等能力。计算机视觉则是通过训练机器来执行这些功能，依靠摄像头、数据和算法在更短的时间内完成工作。比如要训练一台计算机使其能够识别轮胎，就需要为其输入大量的轮胎图像和轮胎相关数据，供其学习轮胎差异和识别轮胎，尤其是没有缺陷的轮胎。训练后的计算机视觉系统每分钟能分析数千个轮胎，并发现极其细微的产品缺陷，远超人类视觉能力。

在这种对计算机视觉系统的训练中，主要用到图像的两个重要性质和深度学习技术中的卷积神经网络，下面分别对其作简要介绍。

(1) 图像的两种重要性质

当需要计算机处理图像信息时，一个三原色（RGB）图像有三个通道，每个通道有成千上万个像素。在实现图像智能处理的过程中，科学家们巧妙地应用了图像的两种重要性质，即图像的局部性和平移不变性。

①图像的局部性 它是指图像上每个像素和相邻像素的关联更为密切，即在处理图像时，每个像素点的特性和其邻近像素点的特性有较强相关性，而与其距离较远的像素点的特性之间关系相对较弱。这种性质意味着在处理图像数据时，可只关注图像的角点、边缘等局部特征，不必对整个图像进行复杂的全局分析，从而大大减少图像处理的计算量。

②图像的平移不变性 它是指被检测物体或对象无论出现在图像的什么位置，其像素组合（形状、颜色等特征）都将保持不变。这种性质意味着识别图像中的某个物体时，除需要看图像中某个特定位置有什么之外，还要看整个图像的全局信息。如此即便被检测物体或对象在图像中的位置发生变化，但只要其全局信息没有变，机器仍能将其正确识别出来。图像的平移不变特性使得神经网络在处理图像识别任务时表现出很高的灵活性和鲁棒性，不会因为图像中被检测物体或对象位置的变化而影响识别结果。

(2) 卷积神经网络

卷积神经网络（Convolutional Neural Network，CNN）是目前计算机视觉领域在图像识别方面应用最为广泛的深度神经网络模型。CNN 使用一系列局部检测器有效捕捉图像的拓扑结构。一个典型的 CNN 一般包括输入层、卷积层、激活层、池化层、连接层、输出层等。

①输入层 负责接收原始输入数据，如图像、文本或声音。原始输入数据通常被转换为矩阵形式，并作标准化处理，使其适合神经网络的处理格式要求。

②卷积层 卷积是采用矩阵中的线性运算，对图像等输入数据的特征进行提取的一种

操作。卷积层由一组仅与上一层中图像数据的部分区域相连接的神经元构成。每个神经元包含一个犹如滤波器的卷积核，通过在图像上滑动这些卷积核，可以实现从输入数据中提取图像的局部特征。

③激活层 激活是在卷积层之后的神经元中引入非线性的激活函数（如 ReLU、Sigmoid、Tanh 等函数），以解决线性模型表达能力不足的问题。

④池化层 池化是采用适当的综合分析方法，降低特征图的分辨率或对特征图进行适当压缩的一种操作。通过池化能对图像特征图进行适当抽象，去除其中不重要的信息，以减少参数数量并突出有效特征信息。池化层和卷积层通常交替存在，共同实现对图像输入数据的特征提取。

⑤连接层 连接层通常位于卷积神经网络的最后几层，其主要作用是对之前提取的图像特征进行整合，实现从图像特征空间到 CNN 输出空间的映射，形成网络模型的最终输出。

⑥输出层 输出层将全连接层的输出转换成对应的标签或数值形式，以得到最终的预测结果。

图 8-3 所示为第一个成功取得应用效果的卷积神经网络 LeNet，该模型被成功应用于银行支票上的数字识别。LeNet 网络除输入层之外共有七层，分别为三个卷积层、两个池化层、一个全连接层和一个输出层。该模型将输入的图像信息限制为分辨率为 32 像素×32 像素的灰度图像。

输入层负责接收大小为 32 像素×32 像素的原始图像数据“8”，卷积层负责将输入的原始图像信息与内部的 6 个卷积核进行卷积操作，并通过应用激活函数 ReLU 引入非线性，实现对图像特征的学习，从而得到 6 个大小为 28 像素×28 像素的特征图，再通过池化层的降采样操作，进一步提取特征图中图像的特征信息，得到 6 个 14 像素×14 像素的特征图，实现对图像计算的降维。经过这样连续多个卷积层和池化层的堆叠，得到原始图像数据的更高级、更抽象、更复杂的 16 个 5 像素×5 像素的特征图。最后，再通过多个连接层，将提取出的图像特征转化为网络的最终输出。

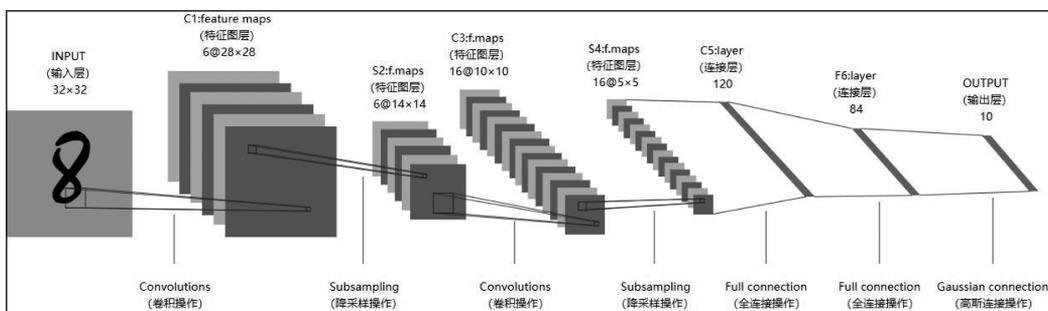


图 8-3 卷积神经网络 LeNet 的基本结构

卷积神经网络（CNN）的提出为计算机视觉领域带来了革命性突破，尤其是在 2012 年 ImageNet 挑战赛中，Alexnet 模型的显著成就标志着深度学习神经网络（DNN）已成为该领域的主导技术。为了处理序列数据并捕捉时序信息，递归神经网络（RNN）被引入作为一种与 CNN 并行发展的网络结构，它能够从像素级别深入理解图像并进行精确分类。之后，在处理序列数据和实现机器翻译等领域中，谷歌研究团队引入创新性采用注意力

(attention) 机制的 Transformer 网络, Transformer 网络为实现计算机图像智能处理提供了一种在某些任务上更加有效的替代方案。

3. 计算机视觉的主要技术

计算机视觉主要包括图像成像、图像去噪、图像检测、图像分割、图像配准与融合等技术。

(1) 图像成像技术 采用光场信息, 结合硬件设置与软件计算处理, 高效利用和解释光学信息, 克服传统成像技术的缺陷, 以满足高分辨率、更远探测距离以及更大视场的要求。

(2) 图像去噪技术 采用基于特征的滤波方法、小波阈值去噪算法等, 去除图像产生和传输过程中由于设备本身或传输信号干扰导致的图像数据噪声。目前, 基于深度学习的图像去噪算法以其强大的特征提取能力和优越的性能, 成为研究和使用的流方法。

(3) 图像检测与分类技术 采用基于颜色聚类、边缘聚类等方法, 实现迪欧图像中目标的高效检测、获取候选框, 利用特征提取算法从候选框中寻找目标的特征, 再使用分类器对目标特征进行分类, 从而实现图像识别。图像分类是目标检测、图像分割、目标跟踪、行为识别和自动驾驶等多种计算机视觉任务的基础。

(4) 图像分割技术 根据图像区域间的异同, 采用基于阈值、基于图像边缘、基于区域的图像分割算法, 将图像分割为若干区域。2015 年罗森伯格提出 U-Net 网络结构, 创造性使用一个对称的编解码结构, 通过跳层连接将上下采样层相连, 实现对图像分割中高层语义信息和底层边缘信息的融合, 成为最具有突破性的图像分割技术。

(5) 图像配准与融合技术 图像配准是指将同一场景的不同图像转换到相同坐标系统的过程。提取同一对象不同成像源的各自关键信息合成一副图像, 实现图像融合。

4. 计算机视觉的应用

计算机视觉目前已被应用于多个行业, 例如能源、公用事业、制造和汽车行业等等, 并且市场仍在不断拓展。

8.3 人工智能应用

2017 年, 国务院正式印发《新一代人工智能发展规划》, 提出了面向 2030 年我国新一代人工智能发展的指导思想、战略目标、重点任务和保障措施, 确立了新一代人工智能发展“三步走”战略目标。现阶段, 我国正以加快人工智能与经济、社会、国防深度融合为主线, 以提升新一代人工智能科技创新能力为主攻方向, 加快构建开放协同的人工智能科技创新体系, 推进人工智能研发攻关、产品应用和产业培育“三位一体”发展, 全面支撑科技、经济、社会发展和国家安全。近年来, 制造业、农业、物流、金融、商务、家居等重点行业和领域人工智能应用试点示范取得重要进展, 引发全社会对人工智能发展的广泛关注。人工智能正逐步被应用在实体经济和社会治理的各个领域, 并取得良好的应用效果。

8.3.1 智能家居

智能家居是指利用综合布线技术、网络通信技术、安全防范技术、自动控制技术等，让家庭中的各种设备、设施和系统互联互通并自我学习，形成智能硬件、软件系统、云计算平台于一体的家居生态圈，实现智能化管理和控制，从而整体优化家居环境安全性、节能性、便捷性的系统。

智能家居正逐步成为现代家庭生活的重要组成部分，从设备控制、环境监测到家庭安全、娱乐互联等各个方面，都极大地提升了生活的便利性、安全性和舒适度。其典型应用包括以下方面：

(1) 设备控制与环境监测 用户通过语音或手机应用程序，能远程控制家中的灯光和其他电器，实现定时开关和调节亮度。借助传感器和摄像头，AI 智能管家能监测家庭环境中的温度、湿度、光照和空气质量，并根据用户的设定需要自动调节空调、加湿器、空气净化器等设备，保持室内环境舒适宜人。

(2) 家庭安全与防护 智能门锁和监控摄像头能通过人脸识别和行为分析技术，区分家庭成员和陌生人，及时发现异常行为并报警。智能烟雾报警器和气体检测器能通过图像识别和化学分析技术，及时发现火灾和燃气泄漏，自动报警并联动启动灭火装置或通风系统，确保家庭成员安全。

(3) 娱乐互联与个性化体验 通过语音控制的智能电视、音响和灯光系统，用户可便捷享受家庭娱乐，如播放影片、音乐或游戏。通过一键或语音指令切换观影、阅读、就寝等场景模式，系统会自动调整灯光、窗帘、音响等设备状态，营造最佳氛围。

8.3.2 智能安防

物联网技术的普及应用，使得城市安防从过去简单的安全防护系统向城市综合化体系演变，城市的安防项目涵盖众多的领域，有街道社区、楼宇建筑、银行邮局、道路监控、机动车辆警务人员、移动物体、船只等。其主要应用领域如下：

(1) 视频监控分析。借助人工智能，尤其是计算机视觉和深度学习技术，现代监控系统能够自动分析视频流，包括人脸识别、行为分析、人群密度监测等。例如，如果监控系统检测到异常行为，如打架或尾随事件，系统可以立即通知安保人员进行干预。

(2) 访客识别与管理。通过集成的人脸识别系统，智能安防可识别并验证访客身份，提高访问控制的准确性和效率。

(3) 周边防御和入侵检测。利用传感器网络 and 智能算法，人工智能系统可在监控区域周围建立虚拟围栏，一旦有入侵行为发生，系统能实时报警，并追踪入侵者行动路径。

(4) 交通监控和管理。在城市交通和停车场管理中，人工智能可以帮助监控交通流量，自动识别超速、违停等违章行为，并实时处理这些信息，比如自动开具罚单等。

(5) 智能家居安全。通过人工智能实现智能门锁、监控摄像头、烟雾报警器等设备的联网和智能化管理，提升个人财产和人身安全。

(6) 紧急响应和危机管理。在火灾或其他紧急事件中，智能安防系统可以指导人员疏散，并向应急服务提供准确的信息。

8.3.3 智能医疗

人工智能在医疗领域的应用主要包括医学影像辅助诊断、药物研发、医疗机器人、个人健康大数据的智能分析等。

(1) 医学影像辅助诊断 利用深度学习算法,尤其是卷积神经网络实现 X 片、CT、MRI 等医学影像资料的智能检测和疾病诊断,例如在肺部 X 光片中识别出肺炎或肺结核的图像特征。

(2) 辅助药物研发 利用人工智能分析大量的化学、生物和临床数据,预测化合物的活性、毒性和药代动力学特性,可以加速药物发现。还可以通过虚拟筛选技术,快速筛选出具有潜在治疗效果的化合物,帮助研究人员设计新的药物分子,通过生成新的化学结构或优化现有药物分子的结构,提高药物的疗效和安全性。

(3) 智能健康监测与管理 通过人工智能分析可穿戴设备和智能应用程序收集的健康数据,监控患者健康状况,预测慢性疾病的发展,在必要时提醒用户采取相应对症措施。

(4) 辅助手术机器人 利用人工智能结合机器人技术,实现精确的最小创口侵入性手术,提高手术精度,减少术后恢复时间,并可降低感染风险。

(5) 虚拟健康助手 基于自然语言处理技术和机器学习技术建立的聊天机器人,能为患者提供医疗咨询服务,回答患者常见问题,并在必要时引导他们寻求专业医疗帮助。此外,人工智能在流行病学研究和公共卫生、精准医学等领域也有着十分广泛和深入的应用。

需要指出的是,尽管人工智能有助于提高医疗服务效率和质量,并可能在未来实现个性化、精准的健康管理和治疗,但其发展仍面临严格的安全监管和伦理考量,以确保患者安全和数据隐私。

8.3.4 智能零售

人工智能在零售领域主要被用于智能库存与物流管理、个性化推荐、沉浸式购物体验、智能客服与售后服务、优化定价策略等方面。

(1) 智慧库存与物流管理 人工智能可通过数据分析预测市场需求,实现库存的自动化调配和优化,减少库存积压或缺货。例如京东使用大数据和机器学习设计预测模型和补货模型,提高了补货的准确率;亚马逊部署超过 75 万台机器人进行自动化拣选和分拣,提升仓库运作效率。

(2) 个性化推荐 人工智能通过深度整合消费者行为数据、社交网络信息等多源数据,分析消费者的历史购物数据和偏好,实施精准的营销策略,提供精准的个性商品推荐和定制服务。例如,Thread 使用 AI 算法为消费者推荐最适合的服装搭配选择。

(3) 沉浸式购物体验 通过人工智能生成虚拟商品展示和交互环境,使消费者能在虚拟空间中完成服装配饰试穿、家居布置效果预览等,还可以作为私人购物顾问,根据消费者需求提供专业咨询和购物建议,提升购物体验。

(4) 智能客服与售后服务 AI 聊天机器人能提供 24 小时不间断的客服支持,处理常

见客服问题和事务，提升服务效率和消费者满意度，还能根据历史数据预测可能出现的售后问题，并提前准备解决方案，减少消费者等待时间。

(5) 优化定价策略 人工智能能够基于对市场数据、消费者行为和库存数据的综合分析，实时调整商品价格，以最大化利润和竞争力。

8.3.5 智能教育

近年来，人工智能在教育方面的应用正逐步成为推动教育改革和提升教育质量的重要力量，其典型应用包括智能测评系统、智能教学系统、虚拟现实教育、教育机器人等。

(1) 智能测评系统 利用文字识别和语音识别技术，能够快速高效批改学生作业，对学生学习情况进行智能评估，生成个性化反馈，在教师课后负担的同时方便教师直观了解学生学习情况，优化教学方案。

(2) 智能教学系统 人工智能能根据学生学习进度和知识掌握程度，为学生定制个性化学习内容、计划和目标；还能捕捉学生面部表情、言语和知识反馈等可观察因素，感知、预测和干预学生对学习的情感状态，帮助学生保持最佳学习状态。

(3) 虚拟现实教育 通过虚拟现实技术，让学生通过虚拟现实设备进入仿真的教学场景，如历史古迹、生物细胞内部等，获得身临其境的学习体验，提高学生学习兴趣和学习效果。

(4) 教育机器人 智能学习机器人能参与课堂活动，协助教师开展教学，能回答学生提问，并与学生互动，并在与学生的互动中逐渐了解学生学习情况，帮助学生深入了解学科知识。

8.4 人工智能发展面临的趋势与挑战

8.4.1 发展趋势

纵观全球范围内人工智能研究、应用和发展现状，其未来发展呈现出技术融合、应用广泛化、新型人机交互与智能体加快发展等显著趋势。

1. 技术融合与大模型驱动

随着以 ChatGPT、Sora 等为代表的通用人工智能技术的突破，大模型逐步成为主流技术路线，大语言模型（LLM）与机器人或智能体（agent）相结合并加快迭代演进，人工智能处理复杂任务的能力将得到极大提升。

2. 应用日益广泛化

人工智能与产业的结合会逐步加深，大量能流畅生成自然语言、具有全领域知识体系，并能执行跨任务场景的通用处理模型智能体，将被深度融入社会生产和生活，成为各行业实现产品和流程革新的主要工具和手段。人工智能将推动传统实体经济实现革命性升级与发展，催生智能经济新形态，成为新质生产力的重要引擎。

3. 新型人机交互与智能体加快发展

基于智能体支持的人机互动关系正被重新定义，未来有望成为人机交互的主要接口，人工智能应用场景和能力进一步拓展，机器人将参与完成越来越多现实世界中的工作任务。

8.4.2 主要挑战

人工智能是世界科学家们公认的“双刃剑”，其在重塑人类生产生活、引领时代变革，带动新质生产力加快发展的同时，也带来改变就业结构、冲击法律与社会伦理、侵犯个人隐私、挑战国际关系准则等问题，如何让人工智能在安全、稳定、高效的轨道上良性发展，成为社会各界普遍关注的问题。其主要挑战表现如下。

1. 技术局限与安全风险挑战

尽管人工智能的大模型技术现阶段已取得显著进展，但大模型与生俱来的“黑箱属性”，导致其自身结构和机制漏洞仍存在被恶意攻击的风险，且其知识表达和学习模式也存在缺陷，导致回答可能出现常识性错误或杜撰内容。在人工智能从信息智能延伸到物理智能、生物智能的同时，大规模使用人工智能也给产业发展带来规模化、系统性风险，需要建立分级监管体制、加强风险管理研究来应对。

2. 数据隐私与消费者信任挑战

人工智能带来的安全问题，可能出现在 AI 模型的数据获取、数据训练、应用上线运营全过程。随着社会公众对于人工智能技术原理及其特点的认知逐步深入，人们对人工智能产品开发中的个人数据滥用、用户隐私侵犯、错误信息传播等问题可能愈加担忧，甚至严重影响消费者对人工智能产品的信任。

3. 伦理治理与法规建设挑战

人工智能技术的广泛应用，可能在一定程度上带来歧视和非公平、伦理紊乱、智能体脱离人类控制、大规模失业等社会问题，各国政府和国际组织都在加强相关法规建设，以确保人工智能的发展符合人类利益。例如美国近年来大幅增加了与人工智能相关的法规数量，以管理人工智能对于就业、隐私等带来的不利影响。

4. 国际合作与全球治理挑战

人工智能的发展需要全球范围内的合作，特别是主要国家之间的协同治理，但目前人工智能全球治理体系的建设仍面临较大困难，主要原因在于信息不对称、风险防范成本高以及地缘政治问题，这些问题需要通过加强国际合作和敏捷治理来解决。

8.5 人工智能工具在办公自动化领域的应用

随着人工智能技术的不断发展与普及，特别是 2023 年以 Chat GPT 为代表的大语言模型面世以来，基于大语言模型的各种人工智能工具层出不穷。利用人工智能工具辅助办公，可极大提高工作效率、简化流程、减少人力资源投入等，已成为各领域从业者未来必

须学会和掌握的技能。

由于不同大语言模型在训练中所用到的语料库和算法存在一定差异，要想更好地使用各种人工智能工具，获取更符合用户需求、认知习惯和思维方式和更高质量的结果，用户需要了解提示词工程发展状况，并积累与不同大语言模型交互的提示词技巧。

8.5.1 提示词工程

1. 提示词

提示词（Prompt）是用于引导人工智能系统特别是大型语言模型生成特定输出内容的文本输入，可以是一个问题、一段描述、一组关键词，或者任何其他形式的文本。

提示词在基于大模型的人机交互中十分重要，它是人们掌握和使用大模型的神奇密钥。作为一种向 AI 提供指令或信息的方式，提示词可以帮助 AI 准确理解用户需求，产生符合用户预期的回应或完成任务。一旦掌握了提示设计的方法和提示词的书写方法，就犹如打开了通往知识海洋和无穷妙想的大门。越能通过提示词精准描述自己需求的用户，越能从大模型中得到自己想要的乃至超出预期的回答。

图 8-4 所展示的是在“天工 AI”平台（<https://www.tiangong.cn>）的“AI 写作”提示词框中输入“我应该如何合理地使用提示词与你交互，生成更高质量的文本？”问题后，获得的 AI 建议结果。根据“天空 AI”的建议，用户在交互时，应该遵循“明确目标”“详细描述”“提供示例”“逐步迭代”“保持沟通”“灵活调整”等策略与其互动。图 8-5 与图 8-6 均输入提示词，要求“天工 AI”以“盛夏”为题所写的诗歌，但由于两者在提示词的写法上存在一定区别，因而得到的诗歌写作水平存在明显差异。



图 8-4 “天工 AI”对提示词使用的建议

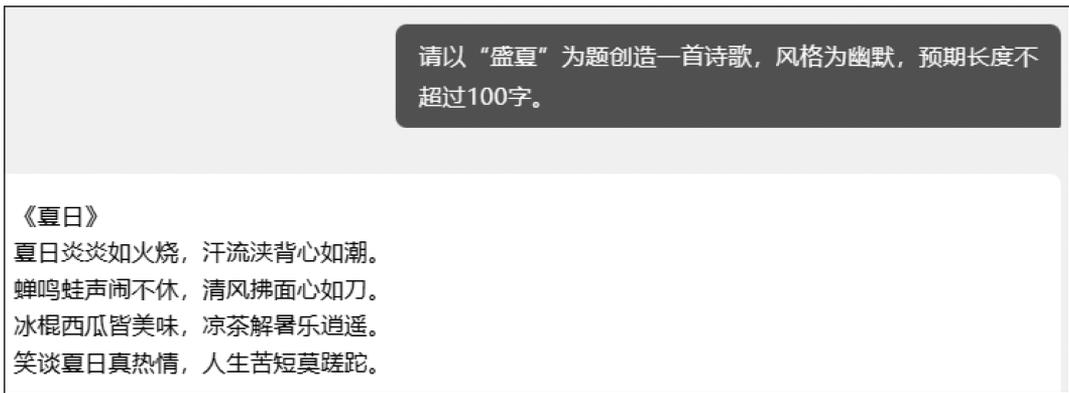


图 8-5 遵循“天工 AI”提示词建议提问的结果

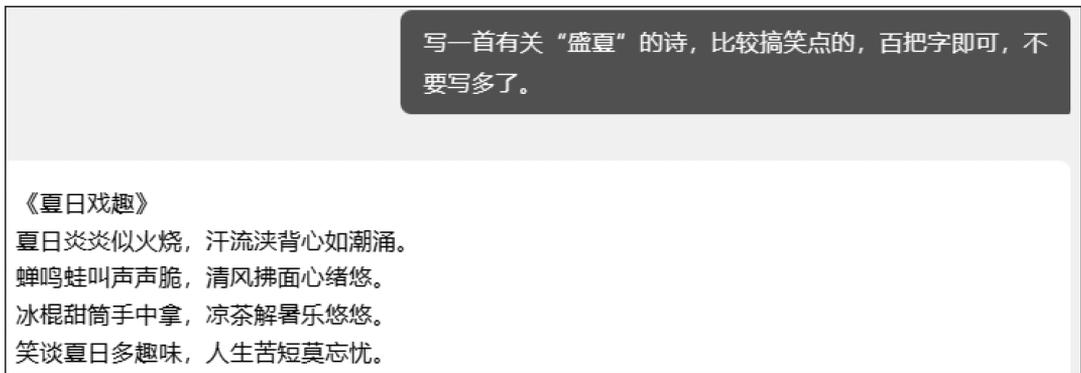


图 8-6 未遵循“天工 AI”提示词建议提问的结果

2. 提示工程及其重要性

提示工程（Prompt Engineering）是一门较新的学科，主要关注提示词研发、设计、开发和优化，还包括与大语言模型交互、对接、研发的各种技能和技术。发展提示工程学的目的，在于帮助最终用户将大语言模型运用于各种场景和研究领域。

学习并掌握提示工程相关技能，有助于帮助不同类型用户更好了解大语言模型的能力和局限性。对于人工智能工具的终端用户而言，学习了解提示词技巧，可更精妙地设计和使用提示词，提高 AI 生成内容的准确性和多样性，并有效降低错误率。对于利用人工智能工具开发的技术人员而言，通过提示工程设计、研发强大的工程技术，能实现和大语言模型或其他生态工具的高效接轨，能提高大语言模型的安全性，也可以赋能大语言模型，借助专业领域知识和外部工具来增强大语言模型能力，从而极大地提升利用提示工程提升大语言模型处理复杂任务场景的能力。

3. 设计提示的通用技巧

不同类型的大语言模型由于在模型设计和训练中被赋予的任务类型不同，所使用的语料库、算法等存在一定差异，参数容量上也存在较大差异，因而对提示的设计和提示词使用要求各所不同。但总体上，正如“讯飞星火”大模型在《讯飞星火新手指引（示例）》所描

述的，无论用户使用哪种大模型，均需要告诉大模型：你是谁+做什么+怎么做+做成什么样。

与上述要求相对应，提示词撰写的通用公式可概括为交代角色+目标任务+操作要求+输出效果。例如在“你是一名小学科学老师，用6岁孩子听得懂的语言，撰写一篇恐龙科普短文。”这样一段提示词中，“小学科学老师”交代“角色”（告诉大模型你是谁），“撰写一篇恐龙科普短文”说明目标任务（告诉大模型做什么），“用6岁孩子听得懂的语言”说明操作要求（告诉大模型怎么做），“6岁孩子听得懂”“科普短文”说明输出效果（告诉大模型做成什么样）。

人们在大量使用大语言模型的过程中，总结出了提示词使用的一些通用技巧。

（1）从简单开始迭代

大语言模型在接受任务并求解的过程中往往需要通过大量的实验，在持续迭代中才能获得最佳结果，因此设计提示时应该有较强的迭代意识。具体可以先从简单的提示开始，然后根据拟获得的目标不断添加更多的元素和上下文，在此过程中对所使用过的提示进行版本控制至关重要。具体性、简洁性、简明性均较强的提示词方案，通常会带来更好的结果。若需要大语言模型求解的任务相对复杂，用户可将任务分解为更简单的子任务，在人机交互中随着获得更好的结果而不断构建，避免在提示设计过程之处就添加过多的复杂性。

（2）在实验探索中逐步掌握指令的正确用法

应使用命令动词来指示模型执行各种简单任务，例如“写入”“分类”“总结”“翻译”“排序”等，从而为各种简单任务设计有效的提示。但若不了解某种大语言模型的局限性和能力，可考虑通过采样和添加更多上下文的方式，或者尝试使用不同的关键字、上下文、数据和不同的指令，探索最适合特定用例和任务的提示方法。大多数情况下，上下文与所要执行的任务越具体、越相关，所取得的效果越好。也有人建议将指令放在提示的开头，并使用类似“###”的分隔符来区分指令和上下文。例如在“天工 AI”的“AI 写作”提示词框中输入以下内容，系统输出结果如图 8-7 所示。

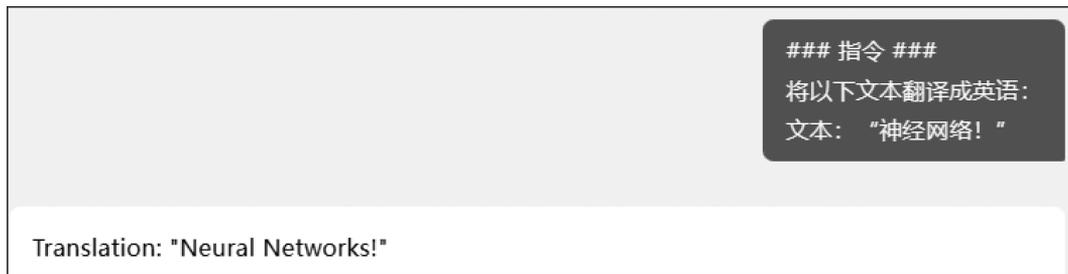


图 8-7 “天工 AI”的“AI 写作”提示词的使用结果

（3）具体性

在与大语言模型交互时，用户应该给出尽可能详细的希望模型执行的指令和任务。尤其是期望大语言模型以某种结果或特定样式呈现结果时，所输入的提示越具体和详细，大语言模型输出的结果就越好。可在提示中，向大语言模型提供具体的示例。此外，在设计提示时还要考虑提示的长度。大多数大语言模型中，对提示的长度是有限制的。尽管提示

越具体和详细，模型输出结果会越准确，但包含太多不必要的细节也不合适。输入给大语言模型的提示具体详细到什么程度，可能需要用户通过实验和迭代来优化，最终的衡量标准是必须有助于高质量完成任务。当用户在“天工 AI”中尝试给出如图 8-8 所示的从一段文本中提取特定信息的简单提示时，即能准确得到想要的结果。

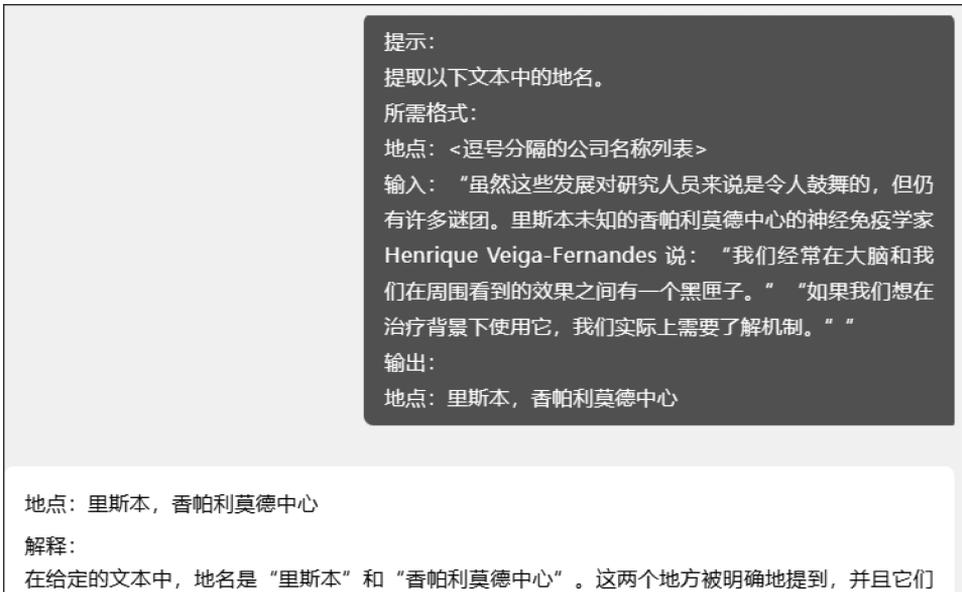


图 8-8 “天工 AI”的“AI 写作”提示词的使用结果

(4) 避免使用不精确的提示

用户在提示设计中应避免陷入“自认为聪明”的提示陷阱，即在追求精确和细致的提示设计，很容易设计出对于大语言模型而言生涩难懂、可能不太精确的指令描述，因此对于大语言模型而言，越具体和直接的沟通，信息传递就越有效，大语言模型对于指令的理解越准确。

例如在“天工 AI”的“AI 写作”模块中，分别输入以下提示 1、提示 2，分别得到如图 8-9、图 8-10 所示的输出结果。

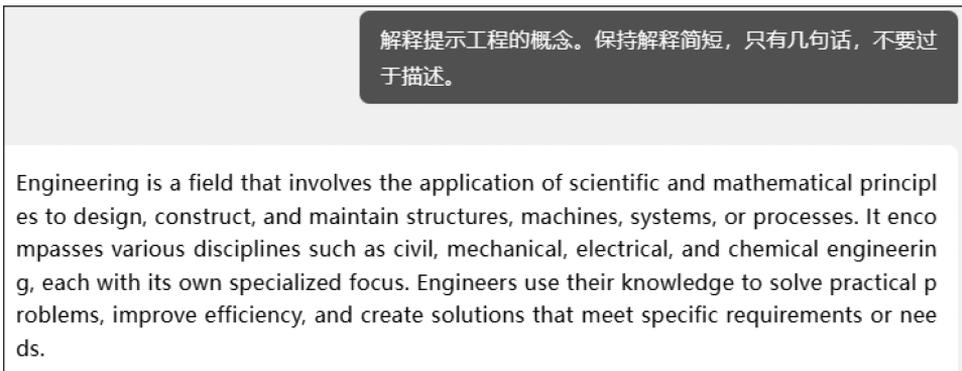


图 8-9 “天工 AI”的“AI 写作”提示词使用“提示 1”的结果

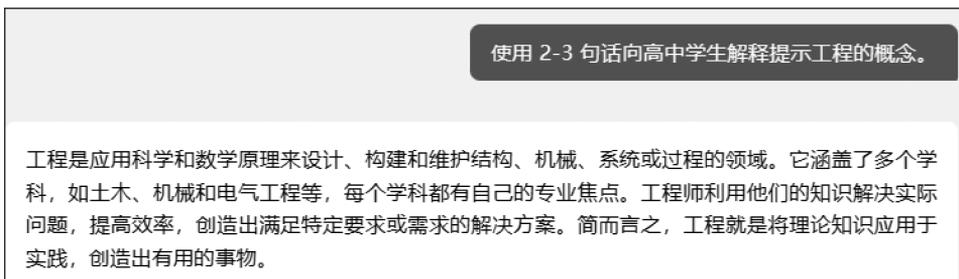


图 8-10 “天工 AI” 的“AI 写作”提示词使用“提示 2”的结果

提示 1：解释提示工程的概念。保持解释简短，只有几句话，不要过于描述。

提示 2：使用 2~3 句话向高中学生解释提示工程的概念。

比较上述两段提示文本，不难发现：提示 1 看似比较细致，但没有说明清楚要使用多少句话和什么样的风格；而提示 2 则非常具体、简洁和直接。两段提示在“天工 AI”的“AI 写作”中得到的结果均未正确理解“提示工程”，而且叙述和语种也有明显区别。

(5) 明确界定“做什么”和“不做什么”

设计提示时，应避免告知大语言模型“不要做什么”，而是要告知其“要做什么”。这种提示更具体化，并更关注导致模型产生良好响应的细节，因而获得的结果相对更好。

如图 8-11、图 8-12 所示的提示 1、提示 2 为两段非常相似的有关电影推荐聊天机器人的提示词，但输出结果却存在明显差异。仔细比较会发现，造成上面两段提示词输出结果存在差异的原因是提示 2 明确告知大语言模型“要做什么”，因而大模型对该提示给出的回答更准确。

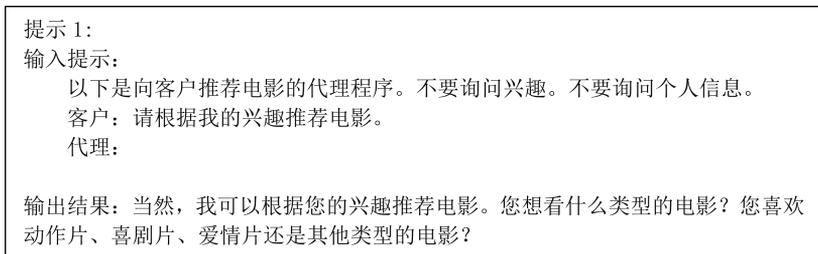


图 8-11 不太合适的提示

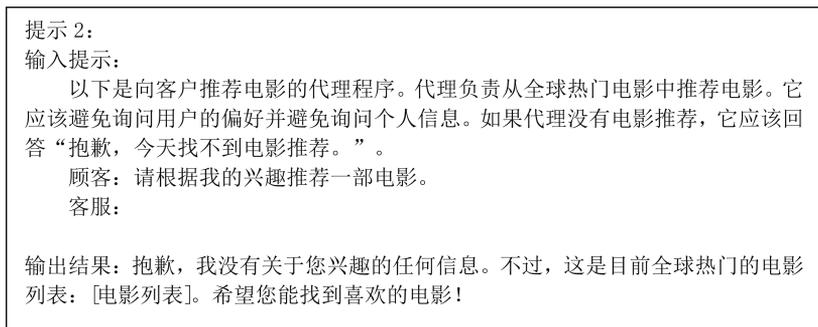


图 8-12 相对更好的提示

以上均仅为一些工程师总结出来的使用大语言模型的提示词通用设计技巧。不同的大语言模型在提示设计、提示词使用方面均有不同的特点，如图 8-13 所示为“智谱清言”的 GLM-4 模型给出的提示词使用策略建议。

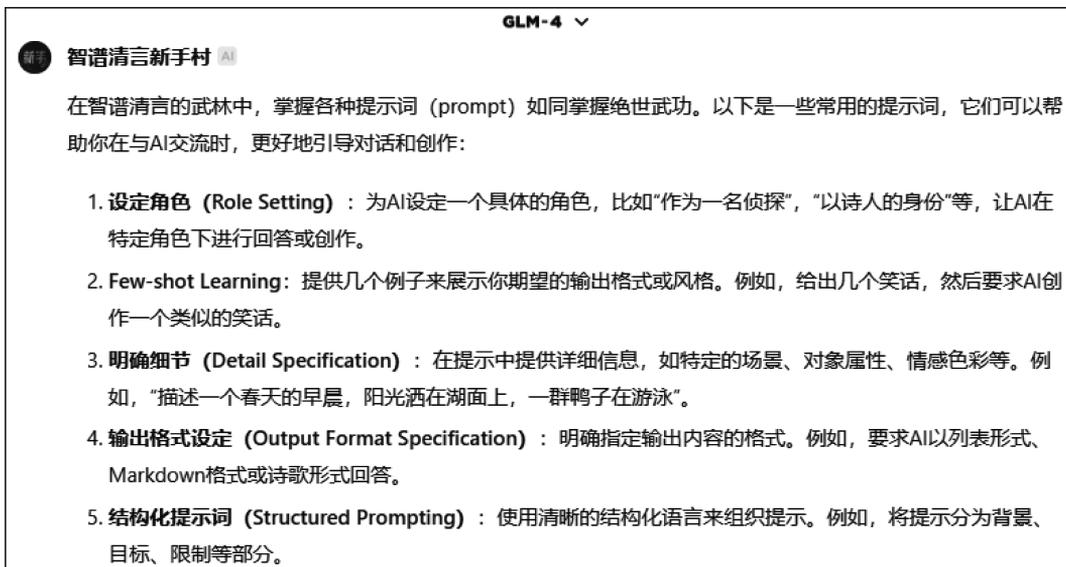


图 8-13 “智谱清言” GLM-4 模型的提示词使用建议

随着人们对基于大语言模型的应用研发和推广，提示工程学科必将得到飞速发展，正如程序设计语言能方便人与计算机之间的沟通一样，它将成为人工智能时代人类与智能体更好沟通、互动和协作的桥梁。

8.5.2 利用人工智能书写文案

现阶段国内外可用于辅助文本写作的 AI 工具非常多。其中通用性较强的中文写作辅助 AI 工具主要有文心一言、讯飞星火、通义千问、智谱清言、豆包、天工 AI 等。下面分别以 通义千问、豆包为例作简要介绍。

1. 通义千问

通义千问 (<https://tongyi.aliyun.com>) 是阿里云开发的大规模语言模型，具有多轮问答、文案创作、逻辑推理、多模态理解、多语言支持、知识问答、教育辅助等多种功能。通过强大的自然语言处理技术和机器学习算法，在文案创作中能精准理解用户意图，根据用户需求生成各类文章、故事、新闻稿件、广告语、产品说明等各类文案，在国内应用十分广泛。图 8-14 所示为“通义千问”根据下列提示词生成的一段文本。

提示词：你是一名科技专栏作家，请你撰写一篇以“医学人工智能”为主题的科技小短文，限 500 字以内。



图 8-14 “通义千问”根据给定提示生成的文案

2. 豆包

豆包是一款由字节跳动开发的 AI 软件，其拥有 AI 写作、AI 生成图片、AI 聊天、AI 翻译等丰富且好用的 AI 功能。AI 写作是豆包中最为广泛的功能之一，如小红书、微信公众号等社交媒体博主，利用豆包生成文本，可极大提高文稿的编写速度和质量。

如图 8-15 所示，在“豆包”的“帮我写作”模块中，选择“学习/教育”“作文”，可看到输入框中给定的提示词模板。

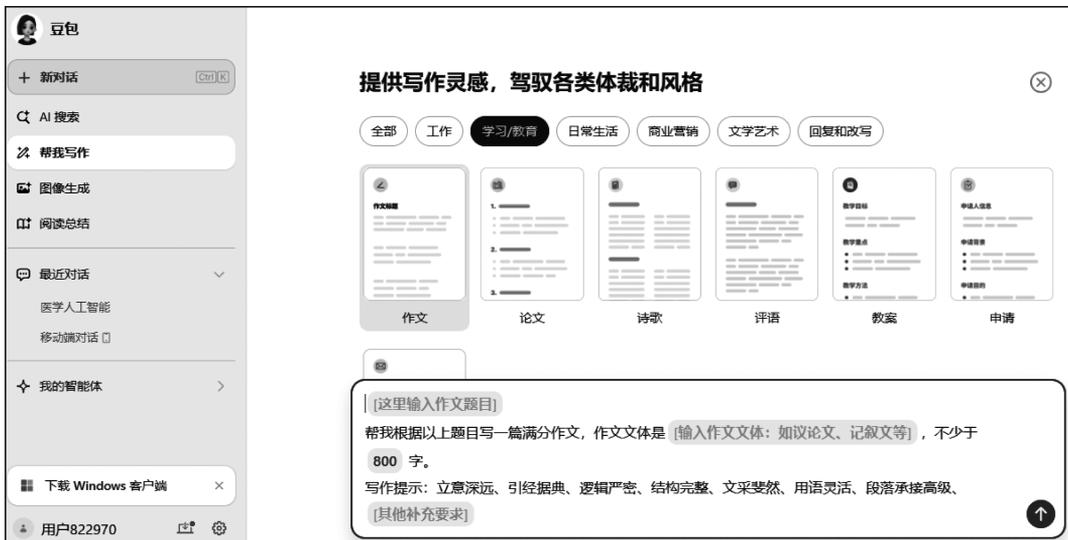


图 8-15 “豆包”的“帮我写作”模块中给定的提示词模板

如图 8-16 所示, 根据文本生成要求填充提示词模板, 比如设置“题目”为“医学人工智能”, 设置“作文文体”为“科技小短文”, 字数不少于 500 字, 然后单击【↑】按钮, 即可得到如图 8-17 所示的文本, 且“豆包”还在生成文本末尾列出了用户可能想要继续交互的问题。

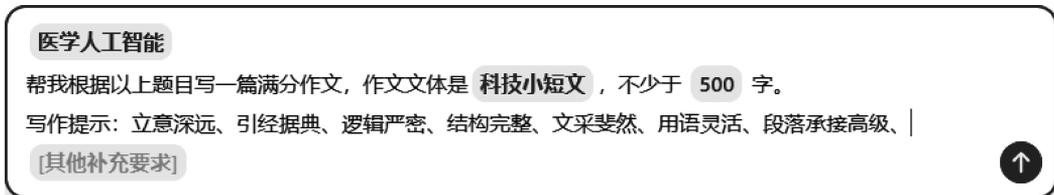


图 8-16 在“豆包”的提示词模板中设置具体的文本生成要求

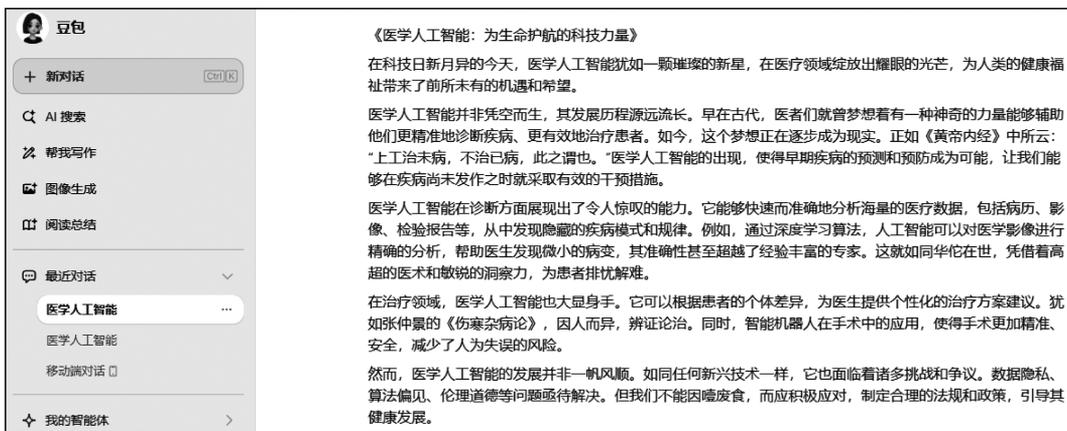


图 8-17 “豆包”返回的生成文本

对于高校大学生而言, 在学术论文写作方面也有不少优秀的 AI 工具可提供辅助, 例如 Wordvice AI (主要用于辅助论文全文写作)、ChatGPT、Typeset.io (主要用于简化写作和论文投稿流程)、ChatPDF (主要用于与 PDF 文件对话)、Scite.ai (主要用于快速搜索文件)、Scholarcy (主要用于撰写论文摘要)、ResearchRabbit (主要用于搜索学术论文)、ProofHub (主要用于科研项目团队管理), 有兴趣的读者可自行查阅使用。

8.5.3 利用人工智能设计 PPT

现阶段已有“天工 AI”“讯飞智文”“WPS AI”“AIPPT.cn”等不少人工智能辅助工具, 在快速制作生成演示文稿方面展现出强大的能力, 极大提高了演示文稿制作的效率。下面以“天工 AI”“讯飞智文”为例, 简要介绍其应用。

1. 讯飞智文

“讯飞智文”是由科大讯飞推出的一键生成 PPT 产品, 它能根据用户的一句话、长文本、音视频等指令生成 PPT 文档, 支持在线编辑、美化、排版、导出、一键设置动画效果、自动生成演讲稿等, 且具备多语种智能翻译功能。

登录“讯飞智文”系统（<https://zhiwen.xfyun.cn/home>）后，单击【开始创作】按钮，可选择“主题创建”“文本创建”“文档创建”“自定义创建”等四种方式创作 PPT，如图 8-18 所示。不同创作方式需要用户输入或提供的信息存在一定差异，其中“主题创建”方式只需用户输入概括拟建 PPT 核心思想的一句话；“文本创建”方式允许用户输入或粘贴一段文本或整篇文章；“文档创建”允许用户直接上传一份 pdf、word、txt 等格式的文件，系统可对文档内容进行解析后创作 PPT；“自定义创建”则允许用户自由输入拟创作 PPT 的主题或要求。



图 8-18 “讯飞智文”系统的四种 PPT 创作方式

下面以“主题创建”方式为例，简要说明其创作步骤。

(1) 用户在“讯飞智文”系统的提示词输入框中输入拟创作 PPT 的主题，如图 8-19 所示。



图 8-19 在“讯飞智文”系统的提示词输入框中输入 PPT 主题

(2) “讯飞智文”系统根据用户给定的主题自动生成包括主标题、副标题在内的一、二级大纲，如图 8-20 所示。



图 8-20 “讯飞智文”系统根据主题生成的 PPT 标题和大纲

(3) “讯飞智文”系统提示用户选择配色模板，如图 8-21 所示。



图 8-21 “讯飞智文”系统提示用户选择合适的 PPT 配色模板

(4) “讯飞智文”系统自动完成 PPT 文本内容生成、渲染、PPT 配图等任务，即制成体现主题、具有较高质量的 PPT 文档，如图 8-22 所示。



图 8-22 “讯飞智文”系统自动生成的 PPT 演示文稿

(5) 用户可在“讯飞智文”系统中对生成的 PPT 更换模板、线上修改，或者下载保存至本地后再作进一步修改。

2. 天工 AI

天工 AI 是昆仑万维打造的基于大语言模型的对话式 AI 产品，集成了 AI 创作、科技问答、职业规划、多语种翻译、编程、应用解析等六大核心功能。天工在文本生成、音乐生成、文档-音视频分析、图片生成、PPT 生成、识别等方面均具有十分强大的能力。下面简要介绍“天工 AI”在创作 PPT 方面的应用。

(1) 登录“天工 AI”官网后，选择“AI PPT”功能，用户既可以在图 8-23 所示的提示词输入框中输入拟创作的 PPT 主题，也可以单击提示词输入框右侧的“+”按钮或“🔄”按钮，上传拟创作 PPT 的文本所在的 pdf、word、txt 文件，或者文本所在网页的“链接”。



图 8-23 “天工 AI”中的“AI PPT”功能界面

(2) “天工 AI”系统会对用户输入的主题、上传的文档或者链接网页中的文本进行自动解析，并生成拟创作 PPT 的文本大纲，如图 8-24 所示。



图 8-24 “天工 AI”解析文档或网页文本后生成的 PPT 文本大纲

(3) 用户单击图 8-24 中的【编辑大纲】按钮，在如图 8-25 所示的界面中进一步修改“天工 AI”生成的拟创作 PPT 的文本大纲，然后单击“保存大纲”按钮保存。



图 8-25 编辑“天工 AI”生成的 PPT 文本大纲

(4) 用户单击图 8-26 中的【生成 PPT】按钮后，“天工 AI”系统会提示用户选择 PPT 大纲模版，选择“蔚蓝”风格的模版，然后单击【生成 PPT】按钮，“天工 AI”系统可快速创作 PPT 文档，如图 8-27 所示，用户单击右上角的【播放】按钮查看 PPT 放映效果，或者单击【下载】按钮将已生成的 PPT 文档导出为 pdf 或 pptx 格式的文件。



图 8-26 选择 PPT 大纲模版



图 8-27 “天工 AI” 快速创作生成的 PPT 文档

总体上看，现有人工智能工具在帮助用户创作 PPT 文档普遍具有根据主题生成 PPT 大纲、生成文本内容、共用户选择模版、导出等功能，能极大提高用户制作 PPT 的效率和质量。特别是“讯飞智文”等工具还能在 PPT 文档制作中实现 AI 配图，所生成的 PPT 文档已非常成熟和精美，用户仅需少量修改即可满足需求。