

本章学习目标

- 掌握决策支持系统的定义及体系结构；
- 了解科学决策与决策支持系统之间的关系；
- 了解人工智能的概念及相关的技术；
- 掌握智能决策支持系统体系结构及系统功能。

全球经济数字化的进程以及信息技术的发展,消除了许多流通壁垒。企业比以往任何时候都面临着更为复杂的生存环境,竞争的压力对企业制定决策的质量、速度都有更高要求,更难以形成科学的“管理决策”并维护其竞争壁垒。

决策支持系统作为一种现代化管理工具,能够为企业提供各种决策信息以及许多商业问题的解决方案,从而减轻了管理者从事低层次信息处理和分析的负担,使得他们专注于最需要决策的工作,因此提高了决策的质量和效率。

前导案例

成立于1992年8月的光大银行,拥有众多客户群,分支机构遍布国内外;同时光大银行以领先的理念为客户提供种类繁多的金融服务。对于一个如此庞大的机构,如此繁多的金融服务,管理的复杂性可想而知。近年来,通过综合柜台业务系统、阳光卡系统、网上银行系统和办公自动化系统等一系列信息化基础建设,光大银行率先实现了业务系统全国联网和总行数据大集中的商务智能系统。

在成功实现业务系统全国联网和总行数据大集中后,经营管理分析方面又出现了一些亟待解决的新问题,如统计数据不够及时准确、

对决策分析缺乏专业化系统化支持、报表处理效率低、数据共享差、难以为以客户为中心的经营管理模式提供充足的信息支持、业绩考核没有理想的 IT 系统作为支撑等。众多新问题的出现是银行管理层始料未及的。

为了尽快突破海量数据的“封锁”，挖掘其中蕴含的知识和信息，光大银行决策层于 2002 年年初开始立项商业智能及数据仓库系统。光大银行根据自身情况，以实际需要为导向，对各家方案的优劣进行仔细分析、反复考查、综合考虑，最终，菲奈特软件公司的高端商务智能产品 BI. Office 以其领先的技术和简便的操作从众多竞争者中脱颖而出，赢得了光大银行决策层的一致青睐。

经过商议，双方在国际结算业务统计分析、对公业务统计分析、信贷风险管理、客户经理业绩考核等方面签订了一系列合作计划。为了降低实施风险，将从国际结算业务统计分析系统开始，各个项目逐步实施。成功的选型是光大银行商业智能决策系统成功实施的开始。国际业务部商业智能的应用证明，光大银行所采取的“以部门为基础实施数据处理”的决定是正确的，也是务实的。

从 2002 年 12 月开始，菲奈特 BI. Office 商业智能决策系统相继应用于光大银行其他几个业务部门，形成相应部门的商业智能系统。这些商业智能系统以数据仓库技术为基础，把分散在各个业务系统的数据进行整合，数据经过清洗、转换，加载到数据仓库；再采用 OLAP(联机分析处理)和数据挖掘(Data Mining)等技术，为管理决策人员提供强大、灵活的日常查询和决策支持。

一个应用实例：有一段时间存款余额持续不断地增长，但是同期的流失客户数也在不断增长，这个问题引起了业务部分析人员的高度重视。该分析人员通过系统进行数据分析后发现，问题的根源在于很多客户经理为了完成揽存目标，费了大量的人力和成本开拓新行业、新客户，而忽略了对老客户的关系管理，才出现了存款余额和流失客户数同时增长的怪现象。于是他马上向主管领导反映，当天就在全行下达了整改通知，及时阻止了不良趋势的蔓延。

目前，光大银行的商业智能系统已经成为管理层进行战略实施、绩效考核不可或缺的工具，各级业务人员的日常经营分析在很大程度上也有赖于该系统的支持。BI. Office 商业智能决策系统在光大银行取得了圆满成功。

3.1 决策支持系统定义及体系结构

3.1.1 决策支持系统的定义

决策支持系统(Decision Supporting System, DSS)是以管理科学、运筹学、控制论和行为科学为基础，以计算机技术、仿真技术和信息技术为手段，解决半结构化的决策问题，支持决策活动的具有智能作用的人机系统。该系统能够为决策者提供决策所需的数据、信息和背景材料，帮助明确决策目标和进行问题的识别，建立或修改决策模型，提供各种备选方案，并且对各种方案进行评价和优选，通过人机交互功能进行分析、比较和判断，为正确决策提供必要的支持。

DSS 的概念是 20 世纪 70 年代提出的，并且在 20 世纪 80 年代获得发展。它的产生基

于以下原因：传统的 MIS 没有给企业带来巨大的效益，人在管理中的积极作用要得到发挥；人们对信息处理规律的认识提高了，面对不断变化的环境需求，要求更高层次的系统来直接支持决策；计算机应用技术的发展为 DSS 提供了技术基础。

在 DSS 发展过程中，各个研究机构和学者从不同角度对 DSS 进行了定义。下面列举一些主要的观点。

R. H. Sprague 和 E. D. Carlson 对 DSS 的定义：决策支持系统具有交互式计算机系统的特征，帮助决策者利用数据和模型去解决半结构化问题。

在该定义中强调决策支持系统具有如下功能：

- (1) 解决高层管理者常碰到的半结构化和非结构化问题。
- (2) 把模型或分析技术与传统的数据存储和检索功能结合起来。
- (3) 以对话方式使用决策支持系统。
- (4) 能适应环境和用户要求的变化。

P. G. W. Keen 对 DSS 的定义：决策支持系统是决策(Decision)、支持(Support)、系统(System)三者汇集成的一体。即通过不断发展的计算机建立系统的技术，逐渐扩展支持能力，达到更好的辅助决策。传统的支持能力是指提供的工具能适用当前的决策过程，而理想的支持能力是主动地给出备选方案甚至决策备选方案。

S. S. Mittra 对 DSS 的定义：决策支持系统是从数据库找出必要的信息，并利用数学模型的功能，为用户产生所需要的信息。

在该定义中强调决策支持系统具有如下功能：

- (1) 为了做出决策，用户可以试探几种“如果，将如何”(What If...)的方案。
- (2) DSS 必须具备一个数据库管理系统、一组以优化和非优化模型为形式的数学工具和一个能为用户开发 DSS 资源的联机交互系统。
- (3) DSS 结构是由控制模块将数据存取模块、数据变换模块(检索数据，产生报表和图形)、模型建立模块(选择数学模型或采用模拟技术)三个模块连接起来实现决策问题的回答。

综合以上定义，决策支持系统可以定义为：决策支持系统是利用大量数据，有机组合各类模型，在计算机上建立多个方案，通过人机交互，辅助各级决策者实现科学决策的系统。

图 3.1 所示为决策支持系统的结构示意图(两库结构)。

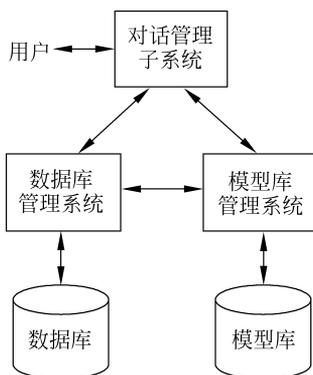


图 3.1 决策支持系统的结构示意图(两库结构)

3.1.2 DSS 与 MIS 的关系

DSS 是从 MIS 的基础上发展起来的,两者都是以数据库系统为基础,都需要进行数据处理,也都能在不同程度上为用户提供辅助决策信息。

DSS 与 MIS 的区别主要体现在以下六个方面:

(1) MIS 是面向中层管理人员,为管理服务的系统。DSS 是面向高层人员,为辅助决策服务的系统。

(2) MIS 是按事务功能(生产、销售、人事)综合多个事务处理的电子数据处理(Electronic Data Processing,EDP)。DSS 是通过多个模型的组合计算辅助决策的。

(3) MIS 是以数据库系统为基础、以数据驱动的系统。DSS 是以模型库系统为基础的、以模型驱动的系统。

(4) MIS 分析着重于系统的总体信息的需求,输出报表模式是固定的。DSS 分析着重于决策者的需求,输出数据的模式是复杂的。

(5) MIS 系统追求的是效率,即快速查询和产生报表。DSS 追求的是有效性,即决策的正确性。

(6) MIS 支持的是结构化决策。这类决策是已知的、可预见的,而且是经常的、重复发生的。DSS 支持的是半结构化决策。这类决策是既复杂又无法准确描述处理,且涉及大量计算,既要应用计算机又要用户干预,才能取得满意结果的决策。

3.1.3 决策支持系统 DSS 的体系结构

对决策支持系统发展影响最大的结构形式有两个:一个是 1980 年 R. H. Sprague 提出的三部件结构;另一个是 1981 年 R. H. Bonczek 等人提出的三系统结构。其中,三部件结构强调模型部件在决策支持系统中的作用;三系统结构强调知识系统在决策支持系统中的作用,它容易与人工智能中的知识系统混淆。在三部件的基础上增加知识部件形成智能决策支持系统,这已成为共识。

1. 三部件结构

1980 年 Sprague 提出著名的决策支持系统的三部件结构。在三部件结构中,决策支持系统由数据库、模型库等子系统与对话管理子系统构成。三库结构示意图如图 3.2 所示。

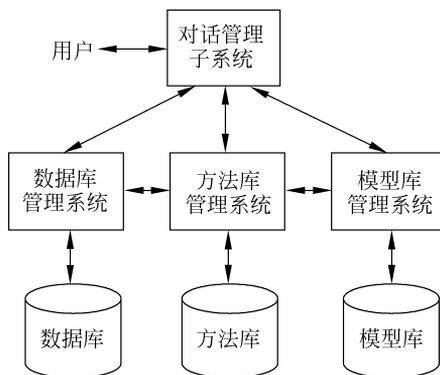


图 3.2 决策支持系统的三库结构示意图

对话管理子系统是 DSS 的人机接口界面,用户通过该子系统提出信息查询的请求或决策支持的请求;对话管理子系统对接收到的请求做检验,形成命令,为信息查询的请求进行数据库操作,提取信息,将所得信息传送给用户。对话部件主要完成以下功能:

- 提供丰富多彩的显示和对话形式;
- 输入输出转换;
- 控制决策支持的有效运行。

数据库子系统由数据库、数据析取模块及数据查询模块等部件组成。数据部件主要完成以下功能:

- 数据库存储的组织形式;
- 数据库管理系统功能;
- 数据库管理语言体系。

对决策支持的请求将识别问题与构建模型,从方法库中选择算法,从数据库读取数据,运行模型库中的模型,运行结果通过对话管理子系统传送给用户或暂存数据库待用。模型部件主要完成以下功能:

- 模型库中模型的表现形式;
- 模型库管理系统;
- 模型库管理系统的语言体系;
- 模型库管理系统的特定功能——模型程序的编辑和编译。

2. 四部件结构

除了图 3.2 所示的三库结构之外,四库结构在模型库、数据库及人机交互系统三个部件基础上,增加了知识库和方法库的管理和应用。四库结构示意图如图 3.3 所示。

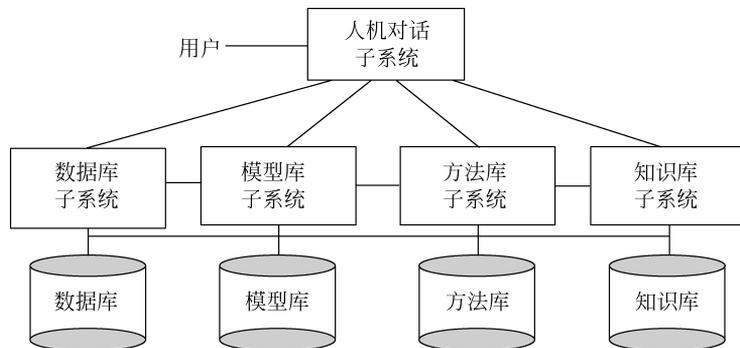


图 3.3 决策支持系统的四库结构示意图

1) 人机对话子系统

人机对话子系统是 DSS 中用户和计算机的接口,在操作者、模型库、数据库和方法库之间起着传送命令和数据的重要作用。

从系统使用角度来看,人机对话接口的设计目标是:

- (1) 能使用户了解系统所能提供的数据、模型及方法的情况;
- (2) 通过运行模型使用户取得或选择某种分析结果或预测结果;
- (3) 在决策过程结束之后,能把反馈结果送入系统,对现有模型提出评价及修正意见;

(4) 当需要的时候,可以按使用者要求的方式,很方便地以图形及表格等丰富的表达方式输出信息、结论及依据等。

从系统维护的检验评价角度,人机对话接口的设计目标是:

- (1) 能帮助维护人员了解系统运行状况,分析存在的问题,找出改进方法;
- (2) 报告模型的使用情况(次数、结果、使用者的评价及改进要求);
- (3) 利用统计分析工具,分析偏差的规律及趋势,为找出症结提供参考;
- (4) 临时性、局部性地修改模型,运行模型,并将结果与实际情况对比,以助于发现问题。

2) 数据库子系统

数据库子系统由数据库、数据析取模块、数据字典、数据库管理系统及数据查询模块等部件组成。

(1) 数据库: DSS 数据库中存放的数据大部来源于 MIS 等信息系统的数据库,这些数据库被称为源数据库。源数据库与 DSS 数据库的区别在于用途与层次的不同,是模型库、方法库和人机对话子系统的基础部分。

(2) 数据析取模块: 负责从源数据库提取能用于决策支持的数据,析取过程也是对源数据进行加工的过程,是选择、浓缩与转换数据的过程。

(3) 数据字典: 用于描述与维护各数据项的属性、来龙去脉及相互关系,也可看作数据库的一部分。

(4) 数据库管理系统: 用于管理、提供与维护数据库中的数据,也是与其他子系统的接口。

(5) 数据查询模块: 用来解释来自人机对话及模型库等子系统的请求,通过查阅数据字典确定如何满足这些请求,并详细阐述向数据库管理系统的数据请求,最后将结果返回人机对话子系统或直接用于模型的构建与计算。

3) 模型库子系统

模型库子系统由模型库和模型库管理系统两部分组成,它是 DSS 中最复杂与最难实现的部分。DSS 用户是依靠模型库中的模型进行决策的,因此认为 DSS 是由模型驱动的。

(1) 模型库主要存储的是能让各种决策问题共享或专门用于某特定决策问题的模型基本模块或单元模型,是模型库子系统的核心部件;

(2) 使用 DSS 支持决策时,根据具体问题构造或生成决策支持模型,这些决策支持模型若有再用的可能性则也可存储于模型库;

(3) 从理论上讲,利用模型库中的“元件”可以构造出任意形式且无穷多的模型,以解决任何所能表述的问题。

4) 方法库子系统

方法库子系统是存储、管理、调用及维护 DSS 各部件要用到的通用算法、标准函数等方法的部件,方法库中的方法一般用程序方式存储。方法库子系统由方法库与方法库管理系统组成。

(1) 方法库是存储方法和程序模块的工具,具体有排序算法、分类算法、最小生成树算法、最短路径算法、计划评审技术(PERT)、线性规划、整数规划、动态规划、各种统计算法、各种组合算法等。决策支持系统的方法库如图 3.4 所示。

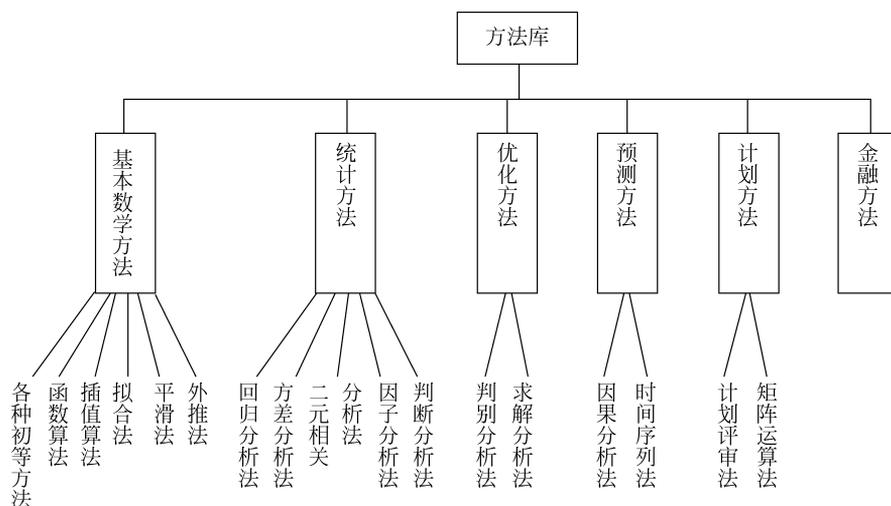


图 3.4 决策支持系统的方法库

按方法的存储方式,方法库可以被分为层次结构型方法库、关系型方法库、语义网络模型结构方法库和含有人工智能技术的方法库等。

(2) 方法库管理系统是方法库系统的核心部分,是方法库的控制机构。

3.2 科学决策与决策支持系统

3.2.1 科学决策

决策自古有之。从宏观上讲,决策就是制定政策;从微观上讲,决策就是做出决定。决策是指个人或集体为了达到或实现某一目标,借助一定的科学手段和方法,从若干备选方案中选择或综合成一个满意合理的方案,并付诸实施的过程。

科学决策是决策者依据科学方法、科学程序、科学手段所进行的决策工作。科学决策的主要特点是:

(1) 有科学的决策体系和运作机制。决策体系包括决策系统、参谋系统、信息系统、执行系统和监督系统。

(2) 遵循科学的决策过程。决策过程包括:提出问题和确定目标;拟定决策方案;决策方案的评估和优选;决策的实施和反馈。

(3) 重视“智囊团”在决策中的参谋咨询作用。

(4) 运用现代科学技术和科学方法。

决策按其性质可分为如下三类:

(1) 结构化决策,是指对某一决策过程的环境及规则,能用确定的模型或语言描述,以适当的算法产生决策方案,并能从多种方案中选择最优解的决策。

(2) 非结构化决策,是指决策过程复杂,不可能用确定的模型和语言来描述其决策过程,更无所谓最优解的决策。

(3) 半结构化决策,是介于以上二者之间的决策,这类决策可以建立适当的算法产生决策方案,从决策方案中得到较优的解。

非结构化决策和半结构化决策一般用于一个组织的中、高层管理,其决策者一方面需要根据经验进行分析判断,另一方面也需要借助计算机,通过决策支持系统为决策提供各种辅助信息,及时做出正确有效的决策。

3.2.2 决策过程

著名学者 H. A. Simon 认为决策过程主要由四个阶段组成。

(1) 情报活动:找出存在问题,确定信息决策目标,获取相关信息。

(2) 设计活动:拟订各种备选方案。

(3) 选择活动:从各种备选方案中进行选择。

(4) 评价活动:执行所选方案,对整个过程及其结果进行检查和评价,将所得信息作为下次信息决策的参考,或者提出新问题,启动新一轮信息决策过程。

H. A. Simon 的观点一方面强调了实践的意义,明确了决策的目的在于执行,而执行又反过来检查决策是否正确;另一方面把决策看成是一个不断循环的管理过程,即“决策—执行—再决策—再执行”的循环过程。此外,他还进一步认为:“决策过程中至关重要的因素是信息联系,信息是合理决策的生命线。”也就是说,不只是第一阶段需要情报活动,每个阶段都需要情报活动的参与。

这四个阶段可以分成更详细的九个步骤,即提出问题、确定目标、价值准则、拟订方案、分析评估、选择方案、实验验证、普遍实施和反馈检验。决策过程示意图如图 3.5 所示。

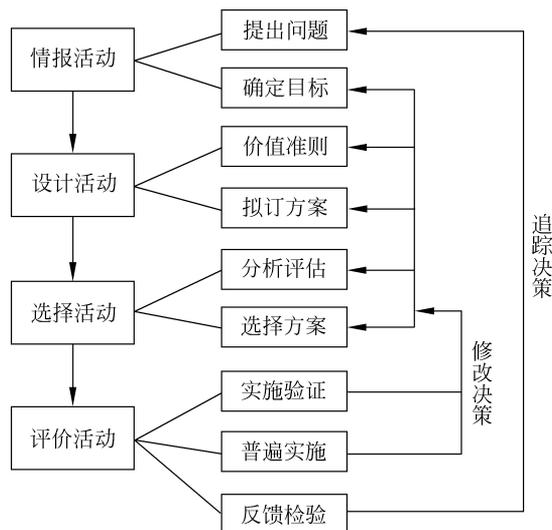


图 3.5 决策过程示意图

1. 提出问题

所有决策工作都从提出问题开始,一般通过寻找实际状况与理想要求(或标准)之间的差距,来发现和提出问题。这一步还需要恰当界定问题,即通过调查研究,分析问题产生的时间、地点、条件和环境,明确问题的性质和特点,确定问题的范围。

2. 确定目标

决策目标的正确与否对信息决策的成败关系极大。为确定决策目标,一般需要通过调

查研究,找出产生问题的原因,还要进行科学预测,以恰当判断未来一段时间内要达到的结果。

决策目标一般有三个特点:

- (1) 目标概念明确或决策目标数量化,这样就能保证各方对目标理解相同;
- (2) 决策目标有时间限制,在规定的时间内完成;
- (3) 决策目标有约束条件限制。

对于多目标的复杂任务,决策常常遵循以下两条原则:

(1) 在满足需要的前提下尽量减少目标个数。因为目标越多,选择方案就越多,会增加选择难度。

(2) 目标要分级落实,协调执行。要分析各个目标的重要性大小,先集中力量实现重要性大的目标。

3. 价值准则

价值准则是落实目标、评价和选择方案的依据。这里所说的价值是指决策目标或方案的作用、效益、收益、意义等,一般通过数量化指标来反映,如产量、产值、成本、质量、效益等。价值准则的设定一般包括如下三项内容:

(1) 价值指标。一般有三类:学术价值、经济价值和社会价值。每类价值又可以分为若干项,每项又可以分为若干条,构成一个价值系统。

(2) 取舍原则。在大多数情况下,要同时达到整个价值系统的指标是困难的,因此,要规定价值的主次、缓急以及在相互矛盾时的取舍原则。

(3) 约束条件。任何决策都有一定的环境,要指明实现这些指标的约束条件。

4. 拟订方案

备选方案的拟订是信息决策过程中关键的一步。这一步主要是根据决策目标和所掌握的信息资料进行的。在拟订备选方案时应注意三条原则:

- (1) 尽可能多地列出集中可行的方案;
- (2) 多种不同方案之间必须有原则上的区别;
- (3) 要依靠专家或专门机构来进行,要广泛运用智囊技术。

5. 分析评估

在拟订一批备选方案后,按价值标准,对各种备选方案进行分析评估。一般有三种方法:经验评价法、数学分析法和实验法。

经验评价法是使用较普遍的评价方法。特别是对复杂的决策问题,只能用经验评价法加以估计。但这种方法局限性较大,科学性较差。

数学分析法是对拟定的备选方案建立相应模型,并利用计算机等工具进行计算。该方法科学性较强,已成为方案评估的基本手段。

实验法能通过实验过程获取其他方法难以获得的评价结果。

6. 选择方案

在分析、评估备选方案后,需要对方案的选择做出决断,这是决策过程中关键的一步。方案优选,必须有合理的选择标准。一般来说,决策的目的是实现一定的决策目标,越是符

合目标的要求就越好,这就是决策方案的价值标准。此外,在理论上人们追求最优标准,但对于实际信息决策,绝对的最优化是不存在的。H. A. Simon 提出一个现实的“满意标准”,就是在现有条件下追求一个满意的结果。

7. 实验验证

当方案选定之后必须进行局部实验,以验证其方案运行的可靠性。在实验验证中,如果实验成功,即可进入普遍实施阶段;否则反馈回去,进行决策修正。

8. 普遍实施

决策方案在局部实验中能稳定地取得较好效果后,就可以加以推广,进行普遍实施。决策方案通过实验验证,可靠程度一般较高,但在实施过程中仍会发生偏离目标的情况。因此,需要加强反馈,不断采取措施加以控制,保证方案的顺利实施。

9. 反馈检验

这一步骤也称为“后评价”,是指决策实施后,应检验和评价实施的结果,检验是否达到预期的目标,回顾整个信息决策过程,总结经验教训,为今后的决策提供信息和借鉴,或者提出新问题,启动新一轮决策。

3.3 人工智能

人工智能(Artificial Intelligence, AI)是知识工程学,是一门研究机器智能的学科,即用人工的方法和技术,研制智能机器或智能系统来模仿、延伸和扩展人的智能,实现智能行为。

3.3.1 人工智能的基本概念及历史

1. 人工智能的不同定义

定义 1: 人工智能之父 John McCarthy 认为,人工智能(知识)就是制造智能的机器,更特指制作人工智能的程序。人工智能模仿人类的思考方式使计算机能智能地思考问题;人工智能通过研究人类大脑的思考、学习和工作方式,然后将研究知识的表示方法、知识的运用和知识的获取作为开发智能软件和系统的基础。

定义 2: 人工智能(学科)是一门基于计算机科学、生物学、心理学、神经科学、数学和哲学等学科的科学和技术。人工智能的一个主要推动力是开发与人类智能相关的计算机功能,例如推理,学习和解决问题的能力。

定义 3: 人工智能(能力)是智能机器所执行的通常与人类智能有关的智能行为,如推理、证明、识别、感知、理解、思考、学习和问题求解等思维活动。

2. 人工智能的发展历史

1940—1950年:一些来自数学、心理学、工程学、经济学和政治学领域的科学家在一起讨论人工智能的可能性,当时已经研究出了人脑的工作原理是神经元电脉冲工作。

1950—1956年:艾伦·图灵(Alan Turing)发表了一篇具有里程碑意义的论文,其中他预见创造了思考机器的可能性。

重要事件:曼彻斯特大学的 Christopher Strachey 使用 Ferranti Mark 1 机器编写了一

个跳棋程序, Dietrich Prinz 编写了一个国际象棋程序。

1956年: 达特茅斯会议, 人工智能诞生。约翰·麦卡锡创造了“人工智能”一词并且演示了卡内基-梅隆大学首个人工智能程序。

1956—1974年: 推理研究, 主要使用推理算法, 应用在棋类等游戏中; 自然语言研究, 目的是让计算机能够理解人的语言。日本早稻田大学于1967年启动 WABOT 项目, 并于1972年完成世界上第一个全尺寸智能人形机器人 WABOT1。

1974—1980年: 由于当时的计算机技术限制, 很多研究迟迟不能得到预期的成就, 这时候 AI 处于研究低潮。

1980—1987年: 在20世纪80年代, 世界各地的企业采用了一种称为“专家系统”的人工智能程序, 知识表达系统成为主流人工智能研究的焦点。在同一年, 日本政府通过其第五代计算机项目积极资助人工智能。1982年, 物理学家 John Hopfield 发明了一种神经网络可以以全新的方式学习和处理信息。

1987—1993年: 第二次 AI 研究低潮。

1993—2011年: 出现智能代理, 它是感知周围环境, 并采取最大限度提高成功的机会的系统。这个时期自然语言理解和翻译、数据挖掘、Web 爬虫出现了较大的发展。

里程碑的事件: 1997年深蓝击败了当时的世界象棋冠军 Garry Kasparov。2005年, 斯坦福大学的机器人在一条没有走过的沙漠小路上自动驾驶 131 mile (1 mile = 1609.344 m)。

2011年至今: 在深度学习、大数据和强人工智能方面发展迅速。

3.3.2 人工智能的主要技术

人工智能技术分成两大类: 一类是传统人工智能; 另一类是计算智能。传统人工智能是以知识为基础, 通过推理进行问题求解; 计算智能是以数据为基础, 通过训练建立联系进行问题求解, 包括人工神经网络、遗传算法、模糊系统、机器学习等方法。

1. 专家系统

专家系统是一类具有专门知识和经验的计算机智能程序系统, 通过对人类专家的问题求解能力的建模, 采用人工智能中的知识表示和知识推理技术来模拟通常由专家才能解决的复杂问题, 达到具有与专家同等解决问题能力的水平。这种基于知识的系统设计方法是以知识库和推理机为中心而展开的, 即

$$\text{专家系统} = \text{知识库} + \text{推理机}$$

它把知识从系统中与其他部分分离开来。专家系统强调的是知识而不是方法。很多问题没有基于传统的计算机程序算法的解决方案(如表 3.1 所示), 采用专家系统可以利用人类专家拥有的丰富知识, 因此专家系统也称为知识系统(Knowledge-Based Systems)。一个专家系统一般应该具备以下三个要素:

- (1) 具备某个应用领域的专家级知识;
- (2) 能模拟专家的思维;
- (3) 能达到专家级的解题水平。

建造一个专家系统的过程可以称为“知识工程”, 它是把软件工程的思想应用于设计基于知识的系统。知识工程包括以下四个方面:

- (1) 从专家那里获取系统所用的知识(即知识获取);

表 3.1 专家系统与传统的计算机程序的主要区别

列 项	传统的计算机程序	专家系统
处理对象	数字	符号
处理方法	算法	启发式
处理方式	批处理	交互式
系统结构	数据和控制集成	知识和控制分离
系统修改	难	易
信息类型	确定性	不确定性
处理结果	最优解	可接受解
适用范围	无限制	封闭世界假设

- (2) 选择合适的知识表示形式(即知识表示);
- (3) 进行软件设计;
- (4) 以合适的计算机编程语言实现。

专家系统通常由人机交互界面、知识库、推理机、解释器、综合数据库、知识获取六部分构成。专家系统的基本结构如图 3.6 所示,其中箭头方向为信息流动的方向。

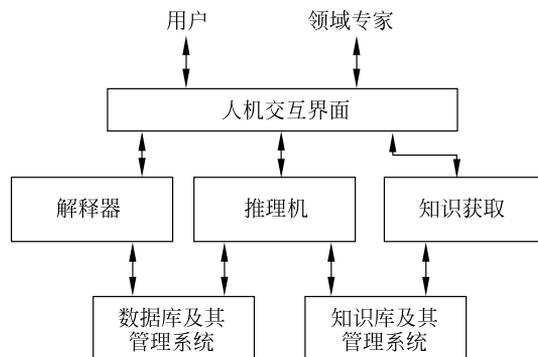


图 3.6 专家系统的基本结构

知识库是问题求解所需要的领域知识的集合,包括基本事实、规则和其他有关信息。知识的表示形式可以是多种多样的,包括框架、规则、语义网络等。知识库中的知识源于领域专家,是决定专家系统能力的关键,即知识库中知识的质量和数量决定着专家系统的质量水平。知识库是专家系统的核心组成部分。一般来说,专家系统中的知识库与专家系统程序是相互独立的,用户可以通过改变、完善知识库中的知识内容来提高专家系统的性能。

推理机是实施问题求解的核心执行机构,它实际上是对知识进行解释的程序,根据知识的语义,对按一定策略找到的知识进行解释执行,并把结果记录到数据库的适当空间中。推理机的程序与知识库的具体内容无关,即推理机和知识库是分离的,这是专家系统的重要特征。它的优点是对知识库的修改无须改动推理机,但是纯粹的形式推理会降低问题求解的效率。将推理机和知识库相结合也不失为一种可选方法。

知识获取负责建立、修改和扩充知识库,是专家系统中把问题求解的各种专门知识从人类专家的头脑中或其他知识源那里转换到知识库中的一个重要机构。知识获取可以是手工的,也可以采用半自动知识获取方法或自动知识获取方法。

人机交互界面是系统与用户进行交流时的界面。通过该界面,用户输入基本信息、回答系统提出相关问题。系统输出推理结果及相关解释也是通过人机交互界面进行的。

综合数据库也称为动态库或工作存储器,是反映当前问题求解状态的集合,用于存放系统运行过程中所产生的所有信息,以及所需要的原始数据,包括用户输入的信息、推理的中间结果、推理过程的记录等。综合数据库中有各种事实、命题和关系组成的状态,既是推理机选用知识的依据,也是解释机制获得推理路径的来源。

解释器用于对求解过程做出说明,并回答用户的提问。两个最基本的问题是 Why 和 How。解释机制涉及程序的透明性,它让用户理解程序正在做什么和为什么这样做,向用户提供了关于系统的一个认识窗口。在很多情况下,解释机制是非常重要的。为了回答“为什么”得到某个结论的询问,系统通常需要反向跟踪动态库中保存的推理路径,并把它翻译成用户能接受的自然语言的表达方式。

2. 人工神经网络

人工神经网络(Artificial Neural Network, ANN)是一种具有非线性适应性信息处理能力的算法,可克服传统人工智能方法对于直觉,如模式、语音识别等非结构化信息处理方面的缺陷。早在 20 世纪 40 年代人工神经网络已经受到关注,并随后得到迅速发展。人工神经网络从信息处理角度对人脑神经元网络进行抽象,建立某种简单模型,按不同的连接方式组成不同的网络。在工程与学术界也常直接将其简称神经网络或类神经网络。

神经网络是一种运算模型(MP 模型),由大量的结点(或称神经元)相互连接构成,利用神经元的信息传播模型进行学习得到训练结果,并用于解决各类问题。MP 模型如图 3.7 所示。每个结点代表一种特定的输出函数,称为激励函数(Activation Function)。每两个结点间的连接都代表一个对于通过该连接信号的加权值,称为权重,这相当于人工神经网络的记忆。网络的输出则依网络的连接方式、权重值和激励函数的不同而不同。而网络自身通常都是对自然界某种算法或者函数的逼近,也可能是对一种逻辑策略的表达。

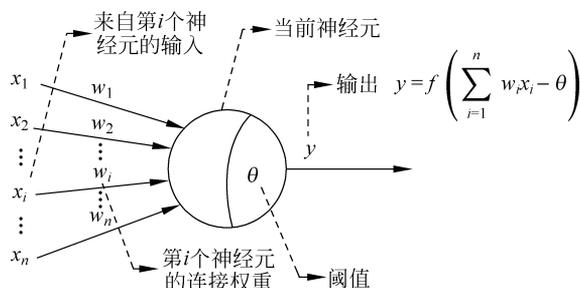


图 3.7 神经元的信息传播模型(MP 模型)

图 3.7 中, x_1, x_2, \dots, x_n 为输入; y 为该神经元 i 的输出; w_i 为外面神经元与该神经元连接强度(称为权值); θ_i 为阈值; $f(x)$ 为该神经元的作用函数。

神经元的信息传播是一个多输入、单输出的结构,神经元之间的连接强度通过权值来表示。神经元之间的连接权值就是神经网络的知识,它是通过大量样本的学习而获得的。神经网络的推理就是信息传播模型。神经网络主要有前馈式网络、反馈式网络和自组织网络。使用最多的是前馈式网络。图 3.8 所示的反向传播模型(Back Propagation, BP)网络结构就是目前使用最多的前馈式神经网络模型。BP 网络是多层网络结构,在输入层与输

出层之间增加若干层(称为隐层)神经元,这些神经元称为隐单元,它们与外界没有直接的联系,但其状态的改变能影响输入与输出之间的关系,每一层可以有若干个结点。

前馈式神经网络是利用大量标准样本(已知样本的输入信息和输出信息)进行学习,获得网络的权值(知识)的。这些知识可以用来对新实例(已知输入信息)进行神经网络的推理完成识别,求出该实例的输出信息。

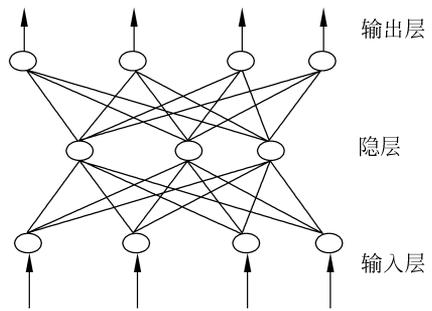


图 3.8 BP 网络结构

3. 遗传算法

遗传算法(Genetic Algorithm,GA)是模拟生物进化的自然选择和遗传机制的一种寻优算法。它模拟了生物的繁殖、交配和变异现象,从任意一初始种群出发,产生一群新的更适应环境的后代。这样一代一代不断繁殖、进化,最后收敛到一个最适应环境的个体上。遗传算法的工作过程如图 3.9 所示。

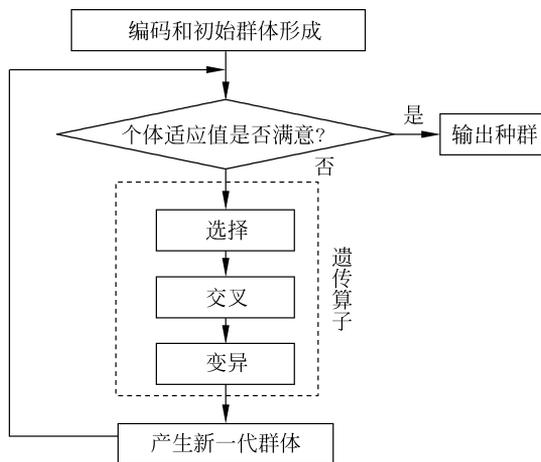


图 3.9 遗传算法的工作过程

其中,选择、交叉和变异是遗传算法的三个主要操作算子,它们构成了遗传操作,使遗传算法具有了其他传统方法所没有的特性。

选择算子又称复制、繁殖算子,其作用是从种群中选择生命力强的染色体产生新种群。依据每个染色体的适应值大小选择,适应值越大,被选中的概率就越大,其子孙在下一代产生的个数就越多。

交叉算子又称重组、配对算子。遗传个体重组是分两步骤进行的,首先在新复制的群体中随机选取两个个体,然后沿着这两个个体随机地取一个位置,二者互换从该位置起的末尾部分。

通过上面的选择和交叉算子操作,基本上完成了遗传算法的大部分搜索功能,而变异算子则增加了使用遗传算法找到接近最优解的能力。变异就是以很小的概率,随机地改变字符串某个位置上的值。变异操作是按位进行的,即把某一位的内容进行变异。

遗传算法对于复杂的优化问题无须建模和进行复杂运算,只需要利用遗传算法的算子就能找到问题的最优解或满意解,因此已经广泛地应用于各类优化问题和分类学习问题。

4. 机器学习

机器学习是人工智能及模式识别领域的共同研究热点,其理论和方法已被广泛应用于解决工程应用和科学领域的复杂问题。机器学习是研究怎样使用计算机模拟或实现人类学习活动的科学,是人工智能中最具智能特征、最前沿的研究领域之一。自 20 世纪 80 年代以来,机器学习作为实现人工智能的途径,在人工智能界引起了广泛的兴趣,特别是近十几年来,机器学习领域的研究工作发展很快,它已成为人工智能的重要课题之一。

机器学习不仅在基于知识的系统中得到应用,而且在自然语言理解、非单调推理、机器视觉、模式识别等许多领域也得到了广泛应用。一个系统是否具有学习能力已成为是否具有“智能”的一个标志。

机器学习的研究主要分为两类。第一类是传统机器学习的研究,该类研究主要是研究学习机制,注重探索模拟人的学习机制;传统机器学习的研究方向主要包括决策树、随机森林、人工神经网络、贝叶斯学习等方面研究。

第二类是大数据环境下机器学习的研究,该类研究主要是研究如何有效利用信息,注重从巨量数据中获取隐藏的、有效的、可理解的知识。该类研究以深度学习为代表,借鉴人脑的多分层结构、神经元的连接交互信息的逐层分析处理机制,自适应、自学习的强大并行信息处理能力,在很多方面收获了突破性进展,其中最具有代表性的是图像识别领域。机器学习算法的类型结构示意图如图 3.10 所示。

决策树是机器学习中的一个预测模型,是对象属性与对象值之间的一种映射关系。每个决策树都表述了一种树状结构,树中每个结点表示某个对象,而每个分叉路径则代表的

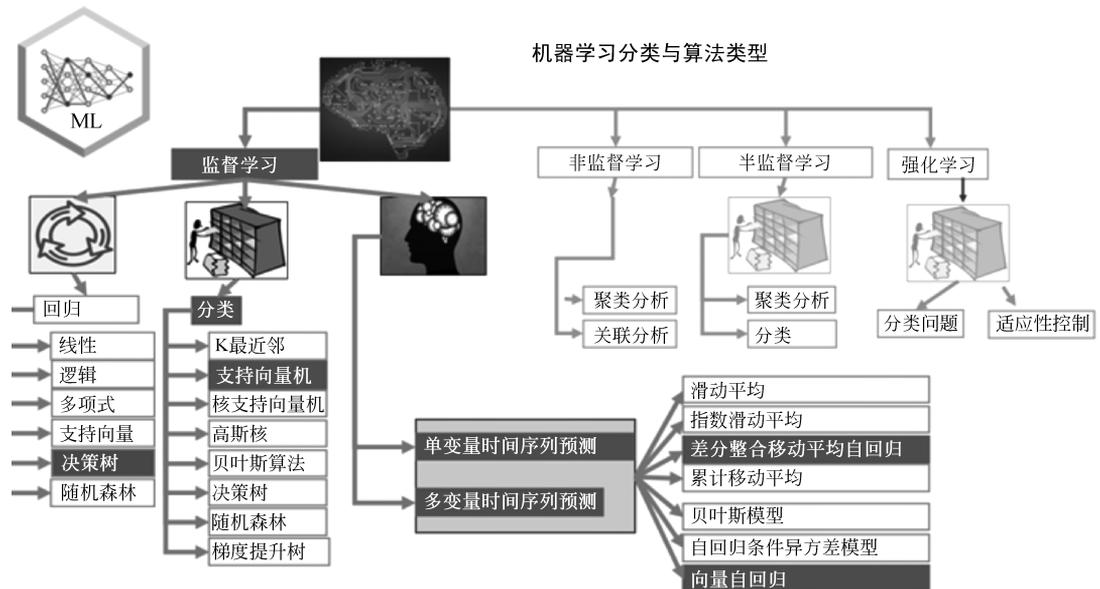


图 3.10 机器学习算法的类型结构示意图

某个可能的属性值,而每个叶结点则对应从根结点到该叶结点所经历的路径所表示的对象的值。决策树仅有单一输出,若欲有多路输出,可以建立独立的决策树以处理不同输出。

每个决策树可以依靠对源数据库的分割进行数据测试。这个过程可以递归式地对树进行修剪。当不能再进行分割或一个单独的类可以被应用于某一分支时,递归过程就完成了。另外,随机森林分类器将许多决策树结合起来以提升分类的正确率。随机森林(RF)作为机器学习重要算法之一,是一种利用多个树分类器进行分类和预测的方法。近年来,随机森林算法的发展十分迅速,已经在生物信息学、生态学、医学、遗传学、遥感地理学等多领域开展应用性研究。

贝叶斯学习是机器学习较早的研究方向,其方法最早起源于英国数学家托马斯·贝叶斯(Thomas Bayes)在1763年所证明的一个关于贝叶斯定理的一个特例。经过多位统计学家的共同努力,贝叶斯统计在20世纪50年代之后逐步建立起来,成为统计学中一个重要的组成部分。贝叶斯分类器的分类原理是通过某对象的先验概率,利用贝叶斯公式计算出其后验概率,即该对象属于某一类的概率,选择具有最大后验概率的类作为该对象所属的类。目前研究较多的贝叶斯分类器主要有四种,分别是朴素贝叶斯算法(Naive Bayes)、TAN、BAN和GBN。

贝叶斯网络是一个带有概率注释的有向无环图,图中的每一个结点均表示一个随机变量,图中两结点间若存在着一条弧,则表示这两结点相对应的随机变量是概率相依的,反之则说明这两个随机变量是条件独立的。网络中任意一个结点 X 均有一个相应的条件概率表(Conditional Probability Table, CPT),用以表示结点 X 在其父结点取所有可能值时的条件概率。若结点 X 无父结点,则 X 的CPT为其先验概率分布。贝叶斯网络的结构及各结点的CPT定义了网络中各变量的概率分布。

朴素贝叶斯算法是以贝叶斯网络为基础的分类算法总称,对于给定的待分类项,利用贝叶斯函数计算出此项条件下各种类别的概率值,其中概率值最大的类别就是待分类项所属的类别。其基本思想如下:

假设 $D = \{X_1, X_2, \dots, X_n\}$ 是输入数据集,设 A_1, A_2, \dots, A_n 表示 n 个属性,有 m 个类的集合 $\{C_1, C_2, \dots, C_m\}$ 和一个未知类归属的样本 $X = \{x_1, x_2, \dots, x_m\}$, x_i 是属性 A_i 的值,计算 X 属于类别 C_k 的概率。由贝叶斯定理可得

$$P(C_k | X) = \frac{P(X | C_k)P(C_k)}{P(X)} \quad (3.1)$$

但是因为计算 $P(X | C_k)$ 很复杂,所以假定 n 个属性变量是相互独立的,则

$$P(X | C_k) = \prod_{i=1}^n P(x_i | C_k) \quad (3.2)$$

式中 $P(X)$ 对于所有类都为常数,因此只要 $P(X | C_i)$ 、 $P(C_i)$ 值最大的时候, $P(C_i | X)$ 为最大,因此朴素贝叶斯分类模型为:

$$C(X) = \operatorname{argmax} P(C_k) \prod_{i=1}^n P(x_i | C_k) \quad (3.3)$$

支持向量机(Support Vector Machine, SVM)是一种监督式学习方法,它广泛应用于统计分类以及回归分析中。支持向量机属于一般化线性分类器,可以认为是提克洛夫规范化(Tikhonov Regularization)方法的一个特例。这种分类器的特点是能够同时最小化经验误

差与最大化几何边缘区,支持向量机也称为最大边缘区分类器。在统计计算中最大期望(EM)算法是在概率(Probabilistic)模型中寻找参数最大似然估计算法,其中概率模型依赖于无法观测隐藏变量(Latent Variable)。最大期望算法经常用在机器学习和计算机视觉的数据集聚(Data Clustering)领域。最大期望算法经过两个步骤交替进行计算,第一步是计算期望,也就是将隐藏变量像能够观测到的一样包含在内从而计算最大似然估计的期望值;第二步是最大化,也就是最大化在第一步中找到的最大似然估计的期望值从而计算参数的最大似然估计。在第二步中找到参数后用于另一个第一步的计算,这个过程不断交替进行。

支持向量机的主要思想可以概括为两点:①它是针对线性可分情况进行分析,对于线性不可分情况,通过使用非线性映射算法将低维输入空间线性不可分的样本转换为高维特征空间使其线性可分,从而使得高维特征空间采用线性算法对样本的非线性特征进行线性分析成为可能;②它基于结构风险最小化理论,在特征空间中建构最优分割超平面,使得学习器得到全局最优化,并且在整个样本空间的期望风险以某个概率满足一定上界。

3.4 基于人工智能技术的智能决策支持系统

3.4.1 智能决策支持系统的定义及体系结构

智能决策支持系统(Intelligent Decision Support System, IDSS)是决策支持系统与人工智能技术相结合的系统,其中既包括决策支持系统所拥有的组件,例如数据库系统、模型库系统和人机交互系统,同时集成最新发展的人工智能技术,如专家系统、神经网络和遗传算法等。IDSS 是以信息技术为手段,应用管理科学、计算机科学及有关学科的理论和方法,针对半结构化和非结构化的决策问题,通过协助明确问题、修改完善模型、列举可能方案等方式,为管理者做出正确决策提供帮助的智能人机交互信息系统。智能决策支持系统的广义结构如图 3.11 所示。

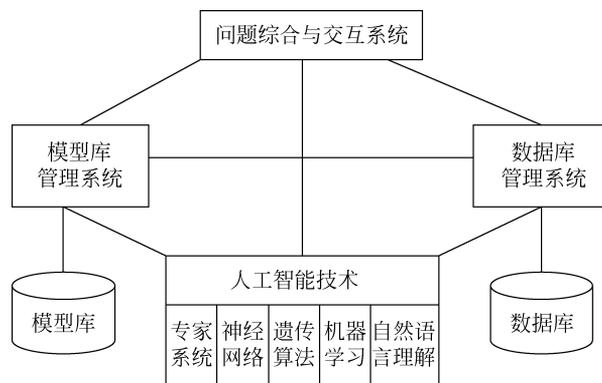


图 3.11 智能决策支持系统的广义结构

智能决策支持系统(IDSS)=决策支持系统(DSS)+人工智能(AI)技术。人工智能技术包括专家系统、神经网络、遗传算法、机器学习和自然语言理解等。其中,专家系统的核心是知识库和推理机;神经网络涉及样本库和网络权值库(知识库),神经网络推理机是MP模型;遗传算法的核心是“选择、交叉、变异”三个算子;机器学习包括各种推理算法库;

自然语言理解需要语言文法库(知识库),处理对象是语言文本,对语言文本的推理采用推导和归约两种方式。因此,智能决策支持系统的统一结构如图 3.12 所示。

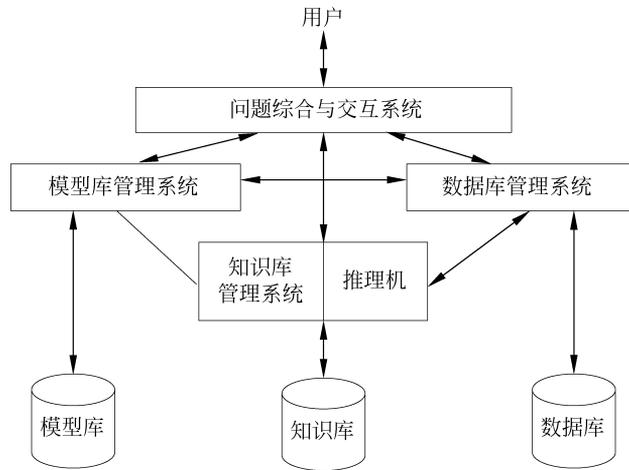


图 3.12 智能决策支持系统的统一结构

3.4.2 基于专家系统的智能决策支持系统

专家系统(ES)与决策支持系统(DSS)结合的智能决策支持系统充分发挥了专家系统以知识推理形式解决定性分析问题的特点,又发挥了决策支持系统以模型计算为核心解决定量分析问题的特点,充分做到定性分析和定量分析的有机结合,使得解决问题的能力 and 范围得到一个大的发展。专家系统与决策支持系统的具体集成结构如图 3.13 所示。

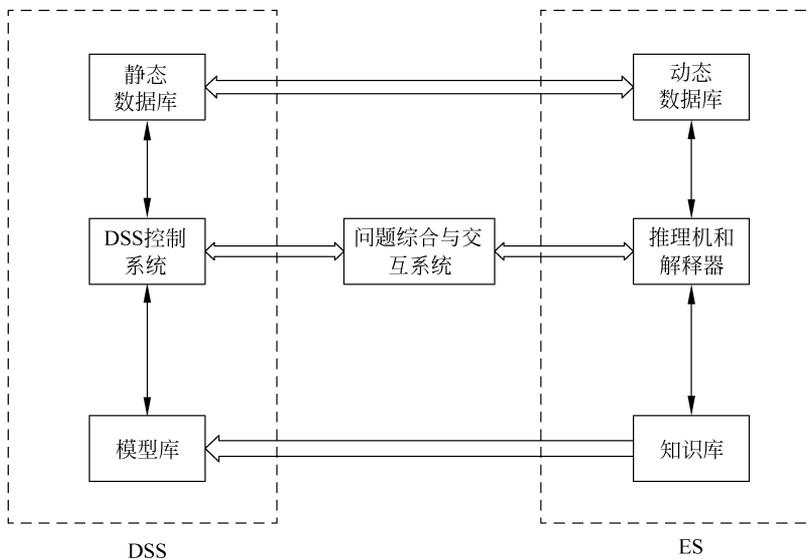


图 3.13 专家系统与决策支持系统的具体集成结构

DSS 与 ES 结合主要体现在以下三个方面:

- (1) DSS 与 ES 的总体结合。由集成系统把 DSS 与 ES 有机结合起来。
- (2) 知识库和模型库的结合。模型库中的数学模型和数据处理模型作为知识的一种形

式,即过程性知识,加入知识推理中。也可以把知识库和推理机作为智能模型加入模型库中。

(3) 静态数据库和动态数据库的结合。静态数据库为动态数据库提供初始数据,ES推理结束后,动态数据库中的结果再送回到DSS中的静态数据库中。

3.5 智能决策支持系统案例——厦门航空航班智能恢复系统

3.5.1 项目概述

航班恢复是困扰民航发展的世界性难题。长期以来,航班智能恢复技术被少数欧美公司掌握,但国内外运行条件差异较大,这些国外的航班智能恢复技术在国内遭遇“水土不服”,无法有效满足国内民航发展的需要。

2019年,厦门航空联合阿里云、同济大学共同研发的航班智能恢复系统正式投产,经过不断优化完善算法,进行工程化应用,建成国内首个航班智能恢复系统,并于2019年1月在厦门航空有限公司(简称厦航)正式上线试运行,2019年6月正式运行。航班智能恢复系统创造性地运用回旋时空网络模型、列生成等人工智能技术,从航空运行实际出发,引入旅客、机场、维修计划、运行限制等更多的复杂变量,可在30min内实现2000班航班的大规模恢复问题,在实际运行中取得了良好效果。

航班智能恢复系统的上线运行标志着国内民航业运用运筹规划和人工智能解决大面积航班调整问题从零到一的突破,打破该技术被欧美少数国家垄断的局面,实现了关键核心技术的中国自主可控,为减少航班延误、保障旅客出行贡献了中国智慧,获得了广泛的社会好评。

3.5.2 项目背景

2016年9月15日,“莫兰蒂”台风袭击厦门,厦航取消了168个航班,调整406个航班,数万名旅客行程受到影响。航班延误极大地浪费航空公司的资源和旅客的时间,据统计,全国每年有约20%的航班发生延误,由此带来的损失高达数十亿元。如何在航班延误后快速地恢复航班,并尽可能减少航班变化对旅客出行的影响,是一个长期困扰航空运行的难题。恢复航班需要同时考虑飞机、航线、机组、旅客等诸多因素限制,决策难度大、效率低、耗时长,如果调整决策不当或不及时,会直接给旅客的出行带来诸多不便,也影响到公司运行的安全、正常、服务与效益。

3.5.3 项目介绍

1. 基本概念

航班智能恢复系统应用运筹优化、人工智能和数据挖掘等技术,实现恶劣天气如台风、雾霾、大雪、飞机故障、机场关闭、流量控制等条件下的不正常航班快速恢复与旅客恢复。

航班智能恢复系统能够突破人工方案仅考虑少数因素的局限性,在综合考虑旅客、航班、飞机等多项复杂因素后做出更优决策,大大提高航班调整决策效率,并让旅客在最短的时间知晓未来航班动态、调整行程计划、减少延误等待时间,尽可能减少航班变化对旅客出行的影响,提升旅客服务体验。

2. 业务流程

航班恢复的主要业务流程是:设置运行业务规则及目标函数,输入运行相关基础数据,

选择航班恢复场景、运行政策、决策偏好,调用算法运算,输出航班恢复结果。最后,再传输到 FOC 签派调整系统中进行规则检查,并由签派、客运、机务进行三方协调评估,通过后生效,后续各运行保障部门按照此调整方案开展相关航班保障及旅客通知工作。航班智能恢复业务流程如图 3.14 所示。

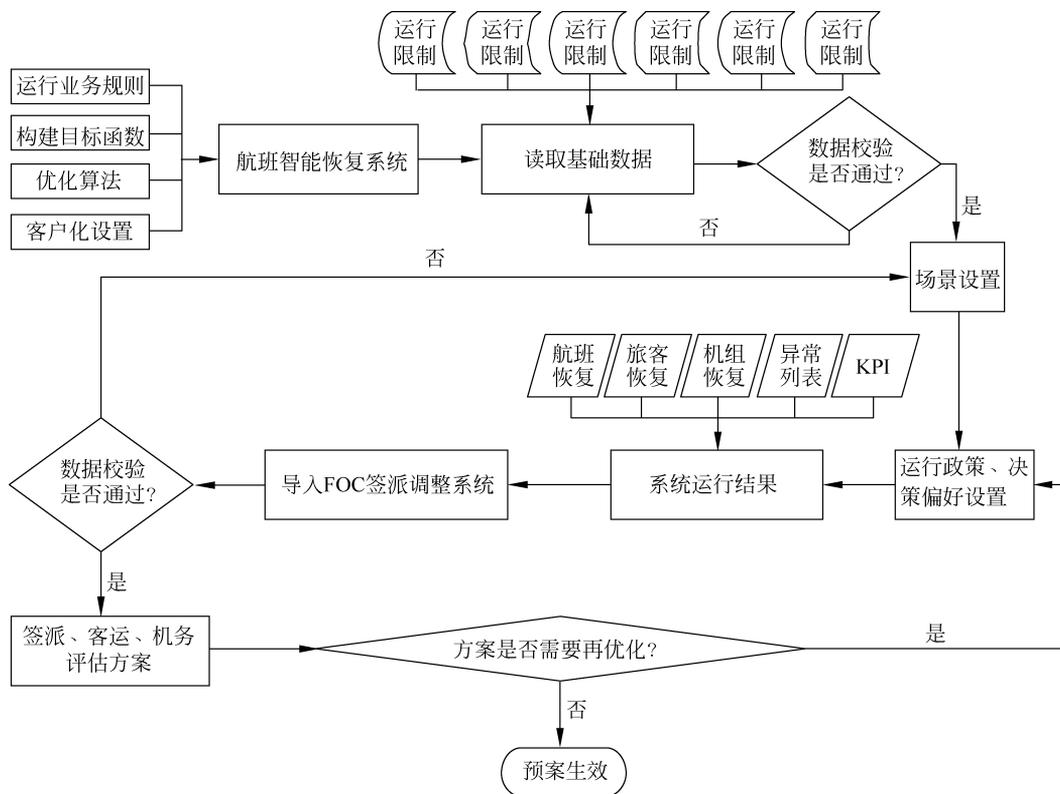


图 3.14 航班智能恢复业务流程

3. 系统构架

系统由视图层、控制层、算法层、外部资源层组成。系统构建了前端交互式组件化 UI 界面展示,并结合 honeycomb 脚手架集成应用托管、集群管理、UI 控制台、应用开发框架,实现微应用模式开发、自动路由规划,达成业务功能间解耦的同时,不增加服务的运维成本;基于 redis 集群的独特数据缓存设计,对 100 多个接口数据进行清洗融合,提升系统访问速度;采用 encryptor 技术对核心数据进行加密,确保数据访问安全;使用 Java 本地调用(Java Native Interface,JNI)、消息队列等技术对算法任务进行编排和跨平台调用。

4. 系统功能

系统实现了对航班的动态监控和异常告警,满足不同业务场景、多种恢复目标的设置,支持不同方案优劣评价指标对比。

在场景支持上,系统针对大规模台风天气、大规模雨雪雾造成的机场关闭等大场景,2~11 天航班计划、飞机维修计划恢复、正常性优化等中场景,当日航班过站时间不足优化、临时飞机故障、临时航班调减、优化指定航班的过站时间等小场景,均能快速给出合理方

案,方案的业务水准显著超过经过长期专业训练的熟练业务人员的水准。

在方案目标设置上,可以对数百项参数进行调整,以满足航空公司实际决策中选择“少延误航班”或“运送更多的旅客”等各种灵活需要。

为进一步提升方案效果,系统设计了航班价值评估、航班延误预估、航班损失回归三个数据挖掘模型。从海量历史航班的大数据分析中,找到季节、天气、航班量、流量控制等众多因素对航班正常性的影响规律,形成智能规则固化到系统中,用于预测未来可能延误的航班,及早进行人工干预,优化航班衔接与过站时间,优化机组、飞机任务连线,实现航班计划的预先管理,减少航班延误班次。在方案优劣对比上,系统生成航班总收入、航班总成本、正常率、客座率、飞机利用率等几十项指标,综合评估得到一个总分,既支持单个方案调整前后的对比,又支持不同方案的对比,还提供了关键指标的图形化对比,满足不同层级管理者的决策需求。

3.5.4 项目应用

2019年8月,超强台风“利奇马”与“白鹿”来袭,厦航对台风密切跟踪,分段决策,用航班智能恢复最优方案替代人工方案,台风场景的调整时间从6~8h缩短到15min,延误和取消航班数可分别减少7%和2%,交出了圆满的答卷。航班智能恢复系统的上线运行,意味着航班控制人员通宵达旦调整航班将成为历史,在面临恶劣天气、飞机故障、流量控制等复杂运行环境中可以更加从容和游刃有余,做到安全、运行、服务与效益的全面结合。

3.5.5 项目影响

厦航航班智能恢复系统,凝集了我国优秀高校、互联网公司以及航空企业的集体智慧。整个系统的设计、建设、运行全过程均自主完成,实现了航班恢复领域核心技术与产品的中国自主创新、自主可控。

国务院国有资产监督管理委员会(简称国资委)网站在2019年1月23日发布新闻《厦航航班大面积调整实现智能恢复,延误减少7%》,民航资源网等多家行业媒体相继发布新闻报道,厦视新闻台在2019年1月24日新闻栏目的黄金时间档进行报道。2019年5月16~5月17日,在北京举办的“2019中国民航发展论坛”会议上,阿里巴巴集团分享了与厦航合作的航班智能恢复系统,该系统成为近年来人工智能与航空业深度融合的最佳实践案例之一。

本章小结

决策支持系统是以信息技术为手段,应用决策技术及有关学科的理论和方法,针对某一类型的半结构化和非结构化的决策问题,通过提供背景材料、协助明确问题、修改完善模型、列举可能方案、进行比较等方式,为管理者做出正确的决策提供帮助的信息系统。决策过程本身是人类的一种智能活动,决策支持系统能够极大程度地模拟人类智能,进而能够科学地做出决策方案。同时,要使决策支持系统具有更多智能,必须提供大量高性能的关于某个问题领域的知识,智能决策支持系统的研究和开发正是向这种基于知识的智能决策过程转变的结果。

课后练习

- 3.1 什么是决策支持系统？
- 3.2 决策支持系统体系结构主要有哪几种？分别画出结构示意图。
- 3.3 决策可以分为哪些类别？决策过程由哪些步骤组成？
- 3.4 人工智能技术包括哪些内容？
- 3.5 试述智能决策支持系统的体系结构？