一种模型整合方法及装置

技术领域

[01]本说明书实施例涉及数据处理技术领域，尤其涉及一种模型整合方法及装置。

背景技术

[02]智能化或自动化处理的本质，是由计算机根据给定的输入信息，经计算后得到对应输出信息的过程，模型则是用于描述上述输入-输出关系的一种信息。简单的模型可以是根据人工规则建立，在数据时代，基于大数据及机器学习技术进行数据建模则成为主流，目前在多种应用领域中，都已经广泛使用大数据建模来实现各种智能化处理功能，但是随着模型数量的增加，也相应产生了难于管理及资源消耗过多等问题。

[03]以风控系统为例，目前以模型驱动的风控策略已被广泛应用，特别是针对一些综合的业务平台，可以基于不同场景不同维度的大数据构建出多种模型，从而实现全面的风险控制。然而，在风险预测覆盖不断完善的同时，模型的数量也随之增长，久而久之会导致整体风控体系越来越庞大，不仅令管理维护成本变得不可控，而且由于不同的模型之间经常存在冗余，实际上也造成了系统处理资源的不必要消耗。

发明内容

[04]针对上述技术问题，本说明书实施例提供一种模型整合方法及装置，技术方案如下：

[05]根据本说明书实施例的第一方面，提供一种模型整合方法，该方法包括：

[06]确定初始的整合模型，所述整合模型具有 1 个整合输出值、且以至少 2 个独立模型的输出值作为输入值；

[07]利用以下步骤进行迭代处理，直到满足预设的迭代停止条件，所述迭代停止条件至少包括：当前的整合模型指标满足预设的性能需求：

[08]获得当前的整合模型针对预设的测试事件集输出的预测结果，并对预测结果进行统计得到整合模型的至少 1 个指标值，所述指标值用于评价所述整合模型的性能；根据预设的优化训练事件集对整合模型进行优化处理，所述优化处理包括：对多个独立模型的整合方式进行调整、或者对独立模型进行单独调整；

[09]停止迭代后，若当前的整合模型指标值满足性能需求，则将当前的整合模型确定为

可用整合模型。

[10]根据本说明书实施例的第二方面，提供一种模型整合方法，该方法包括：

[11]确定初始的整合模型，所述整合模型具有 1 个整合输出值、且以至少 2 个独立模型的输出值作为输入值；

[12]获得当前的整合模型针对预设的测试事件集输出的预测结果，并对预测结果进行统计得到至少 1 个指标值，所述指标值用于评价所述整合模型的性能；

[13]如果指标值不满足性能需求，则根据预设的优化训练事件集对整合模型进行优化处理，所述优化处理包括：对多个独立模型的整合方式进行调整、或者对独立模型进行单独调整；

[14]优化处理后，若当前的整合模型指标值满足性能需求，则将当前的整合模型确定为可用整合模型。

[15]根据本说明书实施例的第三方面，提供一种模型整合装置，该装置包括：

[16]初始整合模型确定模块，用于确定初始的整合模型，所述整合模型具有 1 个整合输出值、且以至少 2 个独立模型的输出值作为输入值；

[17]优化处理模块，利用以下子模块配合进行迭代处理，直到满足预设的迭代停止条件，所述迭代停止条件至少包括：当前的整合模型指标满足预设的性能需求；

[18]性能评价子模块，用于获得当前的整合模型针对预设的测试事件集输出的预测结果，并对预测结果进行统计得到整合模型的至少 1 个指标值，所述指标值用于评价所述整合模型的性能；

[19]优化处理子模块，用于根据预设的优化训练事件集对整合模型进行优化处理，所述优化处理包括：对多个独立模型的整合方式进行调整、或者对独立模型进行单独调整；

[20]输出模块，用于在停止迭代后，若当前的整合模型指标值满足性能需求，则将当前的整合模型确定为可用整合模型。

[21]根据本说明书实施例的第四方面，提供一种模型整合装置，该装置包括：

[22]初始整合模型确定模块，用于确定初始的整合模型，所述整合模型具有 1 个整合输出值、且以至少 2 个独立模型的输出值作为输入值；

[23]优化处理模块，用于获得当前的整合模型针对预设的测试事件集输出的预测结果，

并对预测结果进行统计得到至少1个指标值，所述指标值用于评价所述整合模型的性能；如果指标值不满足性能需求，则根据预设的优化训练事件集对整合模型进行优化处理，所述优化处理包括：对多个独立模型的整合方式进行调整、或者对独立模型进行单独调整；

[24]输出模块，用于在优化处理后，若当前的整合模型指标值满足性能需求，则将当前的整合模型确定为可用整合模型。

[25]应用本说明书实施例所提供的技术方案，一方面，可以将多个独立的模型分值整合为一个综合分值，使得输出模型策略更加便于管理，也解决了模型过多所带来的冗余问题，提升系统处理资源的使用效率；另一方面，对于整合得到的模型，可以基于性能监控预警，自动触发整合模型的迭代优化，从而保证综合分值的稳定有效输出。

[26]应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本说明书实施例。

[27]此外，本说明书实施例中的任一实施例并不需要达到上述的全部效果。

附图说明

[28]为了更清楚地说明本说明书实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本说明书实施例中记载的一些实施例，对于本领域普通技术人员来讲，还可以根据这些附图获得其他的附图。

[29]图1是本说明书实施例的整合模型与独立模型的关系示意图；

[30]图2是本说明书实施例的整合模型构建方法的流程图；

[31]图3是本说明书实施例的整合模型优化方法的第一种流程图；

[32]图4是本说明书实施例的整合模型优化方法的第二种流程图；

[33]图5是本说明书实施例的模型整合系统的结构示意图；

[34]图6是本说明书实施例的用于构建整合模型的装置的结构示意图；

[35]图7是本说明书实施例的用于优化整合模型的装置的结构示意图；

[36]图8是用于配置本说明书实施例装置的一种设备的结构示意图。

具体实施方式

[37]为了使本领域技术人员更好地理解本说明书实施例中的技术方案，下面将结合本说明书实施例中的附图，对本说明书实施例中的技术方案进行详细地描述，显然，所描述的实施例仅仅是本说明书的一部分实施例，而不是全部的实施例。基于本说明书中的实施例，本领域普通技术人员所获得的所有其他实施例，都应当属于保护的范围。

[38]以金融领域的风控场景为例，模型驱动的风控系统需要考虑多种风险类型，针对不同的风险类型，需要构建不同的风控模型，例如反欺诈模型、反盗用模型、反作弊模型等。而在每种风险类型下又会针对不同场景不同维度设计多种专项模型。以反欺诈模型为例，可以进一步细分为离线欺诈者模型、离线欺诈卡模型、在线 FP（转账到账户）模型、在线 FZ（转账到卡）模型、兼职被骗模型等等。为了防止风险漏过，模型策略大部分采用多个模型并用（即每个模型分值阈值加一定规则并行）的方式，例如以下两条策略并行使用：

策略 1 为：模型 A 分值大于 0.5 且交易金额大于 10 元；

策略 2 为：模型 B 分值大于 0.6 且排除双方关系强且交易金额大于 100 元；

[39]长期来看，每新增一个模型，就需要部署新的模型策略，一定时间累积后会造成模型策略冗余且成本增加不可控，针对上述问题，本说明书提出：将不同模型的输出分值进行标准处理后，整合输出一个综合分值。从而使得输出模型策略更加便于管理，解决模型过多所带来的冗余问题，提升系统处理资源的使用效率；另一方面，对于整合得到的模型，可以基于性能监控预警，自动触发整合模型的迭代优化，从而保证综合分值的稳定有效输出。

[40]图 1 所示为整合模型与独立模型的关系示意图：将多个独立模型的输出进行整合，得到 1 个整合模型，其中每个独立模型具有 1 个输出值，多个独立模型的输出值经整合模型后统一输出 1 个整合输出值。

[41]下面将分别从“整合模型构建”以及“整合模型优化”两方面，对本说明书所提供的模型整合方案进行说明。

[42]图 2 所示，为本说明书提供的一种整合模型构建方法的流程图，该方法可以包括以下步骤：

[43]S101，确定待整合的至少 2 个独立模型。

[44]首先确定需要参与整合的多个独立模型，例如离线欺诈者模型、离线欺诈卡模型、在线 FP（转账到账户）模型、在线 FZ（转账到卡）模型等等。其中每个独立模型都具有 1 个专项风险的预测输出分值、若干个输入特征，且不同独立模型之间可能会复用相同的输入特征。

[45]为了保证整合模型的性能，这里可以选择性能较好的独立模型，这里的“性能较好”的标准可以是预测覆盖率较高、准确率较高、打扰率较低等等，最好可以具有一定的额外输出增益。例如在一种具体实施方式中，可以选择以下独立模型参与整合：端模型（指部署在手机客户端的风险识别模型）、低风险识别模型、在线风险识别模型、离线风险识别模型。当然，本说明书并不需要对选择独立模型的具体方案进行限定，另外，这里的独立模型也可以是基于规则建立的模型，并不局限于狭义的“利用机器学习建立的模型”。

[46]S102，从历史事件记录中获取事件数据集；

[47]针对 S101 中所确定的参与整合的独立模型，从历史事件记录中获取事件数据。这里可以选择某个时间窗口内（例如过去一周，过去一个月等）的发生过的事件记录作为数据源，对于任意事件 j，需要获取的数据包括：

a）每个独立模型对于事件 j 的预测输出值；

b）事件 j 的综合风险标签值；

如表 1 所示：

事件 id	模型 1 分值	模型 2 分值	模型 3 分值	综合风险标签
1	0.1	0.01	0.3	0
2	0.55	0.3	0.6	1
3	0.9	0.8	0.9	1
……				

表 1

[48]其中“风险标签”是指某事件是否实际构成案件的判定结果，需要说明的是，该结果与各模型的预测结果并不一定完全一致，该标签值可以根据人工标注确定，也可以根据用户是否报案确定，本说明书对标签值的具体来源不需要进行限定。

[49]S103，根据所获得的事件数据集，生成整合模型训练样本；

[50]理想情况下，可以直接利用 S102 所获得的事件数据集作为训练样本，即对于任意事件 j，直接以各独立模型 A 的分值作为整合模型的输入特征值、以综合风险标签值作

为整合模型的标签值。而实际应用中，可能需要对所获得的事件数据集做进一步处理，以供后续训练使用。

[51]根据模型训练的需求，这里的处理方式可以包括各独立模型的输出值进行标准化（standardization）、归一化（normalization）等常规处理方式，从而将各独立模型的输出值映射至一定的数值区间。

[52]为了令整合模型的特征值表达更加精细，还可以将独立模型的值域划分为若干个子区间，对于同一模型，当模型输出值处于不同子区间时，将被映射为不同的数值，具体处理如下:

[53]S103a，子区间划分:

[54]假设参与整合的独立模型数量为 M，则对于任一独立模型 i，可以将 i 的输出值的值域划分为 L_i 个子区间（$L_i \geqslant 2$， $i=1,2,3\cdots M$）。其中，根据不同模型输出值的实际含义或特性，可以对不同模型采用相同或不同的值域子区间划分方式（例如子区间的划分规则、子区间数量等等）。

[55]在本说明书提供的一种具体实施方式中，可以针对 S102 所获得的事件数据集，根据独立模型 i 对各事件的输出值的分布情况，将各独立模型的输出值的值域划分为多个子区间。

[56]例如，假设风险预测输出分值在[0,1]区间，分值越高代表事件风险越高。在风控场景，综合考虑高准确率、高覆盖率、低打扰率的性能指标需求，在选取风险判决临界值时，一般对越高的分段越敏感，相应地对高分段的划分粒度也应该更为精细，而对于低分段分值，则可以采取相对较为粗糙的划分粒度。基于上述需求，可以采用划分策略举例如下:

[57]首先对所有数据的独立模型得分进行排序，然后根据排序结果做以下划分:

将排序在前 0.01%的分值对应的区间段，划分为 10 个子区间;

将排序在前 0.01%~1%的分值对应的区间段，划分为 100 个子区间;

将排序在前 1%~5%的分值对应的区间段，划分为 20 个子区间;

将排序在前 5%~100%的分值对应的区间段，划分为 1 个子区间;

即将值域区间[0,1]共计划分为 131 个子区间。

[58]下面以一个简化的例子对上述划分方式进行说明，以表 1 所示数据为例，假设 3 个

模型的输出分值都在[0,1]区间，子区间的划分数量均为 3，划分原则是：分别将得分排序在前 0~15%、15%~50%、50%~100%的分值所对应的区段划分为 1 个子区间。假设通过统计所有数据（图 1 中仅示出 3 条）在各个模型输出值的分布占比，可以得到以下值域子区间划分结果：

值域子区间	模型 1	模型 2	模型 3
子区间 1	（0.85，1]	（0.3，1]	（0.6，1]
子区间 2	（0.5，85]	（0.005，0.3]	（0.3，0.6]
子区间 3	[0，0.5]	[0，0.005]	[0，0.3]

表 2

[59]可以看出，由于采用的是实际数据的分布占比作为值域子区间的划分依据，因此即便不同独立模型的子区间划分数量相同、子区间对应的输出值分布占比相同，其子区间对应的端点也不一定相同。

[60]可以理解的是，上述划分方式以及实例仅用于示意性说明，不应理解为对本说明书方案的限定。另外在实际应用中，不同独立模型所选用的值域子区间划分方式也可以是不同的，本领域技术人员可以根据实际需求进行选择。

[61]S103b，数值映射：

[62]在已划分子区间的基础上，针对于独立模型 i 的任意输出值，可以根据该输出值所处的子区间，确定该输出值对应的映射结果。其中，对于独立模型 i 的任意输出值，可以对应一个或多个映射值，举例说明如下：

[63]对于独立模型 i 的任意输出值，可以根据该输出值所处的子区间，将该输出值转换为 One-Hot 编码值。

[64]One-Hot 即独热编码，也称一位有效编码，其方法是使用 N 位状态寄存器来对 N 个状态进行编码，每个寄存器位之间相互独立，并且在任意时刻，其中只有一位有效。在本说明书中，由于独立模型输出值必然属于 L_i 个子区间中的任意一个，因此用长度为 L_i 的 One-Hot 编码值表示独立模型的输出值：将输出值所处的区间对应编码为 1，其他（L_i-1）个区间对应编码为 0。结合前述的区间划分方案，可以将原本的一个特征值进一步稀疏化，从而在一定程度上起到扩充特征的作用，以便为后续训练分类模型提供更多的信息。

[65]以表 1 所示的数据集合以及表 2 所示的子区间划分结果为例，其 One-Hot 编码结果如表 3 所示：

事件id	模型 1			模型 2			模型 3		
	onehot 1	onehot 2	onehot 3	onehot 1	onehot 2	onehot 3	onehot 1	onehot 2	onehot 3
1	0	0	1	0	0	1	0	0	1
2	0	1	0	0	1	0	0	1	0
3	1	0	0	1	0	0	1	0	0
… …									

表 3

[66]可见，对于任一条数据样本 j，原本 M 个独立模型的 M 个输出值，将值域统一划分为 L 个子区间、并且经 One-Hot 编码处理后，变成了共 M*L 个数值。根据表 3 示例，M*L=9，这 9 个数值可以在后续的训练过程中作为条数据样本 j 的 9 个特征值使用。

[67]对于独立模型 i 的任意输出值，还可以根据该输出值所处的子区间，将输出值所处的子区间的 WOE 分值确定为该输出值的映射结果。

[68]WOE（Weight of Evidence，证据权重），也是一种将数值做离散化处理的方式，WOE 值表达的是变量取某个值时对事件结果（例如风险事件）的一种影响。

[69]在本说明书方案中，对于任意数值子区间 k，该子区间的对应的 WOE 值定义为:

$$\mathrm{WOE_k} = \ln \frac{\text{子区间 k 中的黑样本数量/全部黑样本数量}}{\text{子区间 k 中的白样本数量/全部白样本数量}}$$

[70]上述公式中，“黑样本”表示综合风险标签为 1（有风险）的数据样本，“白样本”表示综合风险标签为 0（无风险）的数据样本。

[71]如表 4 所示，假设对某个独立模型 i，值域子区间划分数量为 3，总样本数量是 1000，其中白样本数量为 900、黑样本数量为 100，根据黑/白样本在每个子区间的分布情况，计算每个子区间的 WOE 分值如表 4 所示:

值域子区间	黑样本数量	白样本数量	WOE
子区间 1	80	50	=ln[(80/100) /(50/900)] ≈ 2.67
子区间 2	15	350	=ln[(15/100)/(350/900)] ≈ -0.95
子区间 3	5	500	=ln[(5/100)/(500/900)] ≈ -2.14
总计	100	900	

表 4

[72]可见，WOE 的数值是也是基于各个子区间的样本分布占比情况统计得到的，每个子区间的 WOE 值反映的是每个分段的白/黑样本比例与总体白/黑样本比例的差异。根据本说明书方案，首先针对各个独立模型的值域子区间划分情况，分别计算出各个独立模

型各子区间对应的 WOE 分值，进而对于独立模型 i 的任意输出值，该输出值所处子区间的对应 WOE 分值即为该输出值的映射值。

[73]可见，对于任一条数据 j，原本 M 个模型的 M 个输出值，经 WOE 编码处理后仍然是 M 个数值，这 M 个数值可以在后续的训练过程中作为数据样本 j 的 M 个特征值使用。

[74]可以理解的是，上述数值映射方式以及实例仅用于示意性说明，不应理解为对本说明书方案的限定。另外在实际应用中，不同独立模型所选用的数值映射方式也可以是不同的，本领域技术人员可以根据实际需求进行选择。

[75]S104，利用有监督学习算法对样本进行训练得到整合模型。

[76]根据 S103 得到的训练样本，利用有监督学习算法进行训练得到整合模型。整合模型具有 1 个整合输出值，且以多个独立模型的输出值作为输入值。

[77]在训练过程中，每条数据样本训练标签值即为该条数据样本的综合风险标签，而每条数据样本的训练特征值则根据 S103 的处理方式确定，例如可以是多个独立模型输出值本身，也可以是多个独立模型输出值经处理后的结果。

[78]整合模型的形式可以根据实际训练需求选择，例如线性回归模型（linear regression model）、逻辑斯谛回归模型（logistic regression model）等等。本说明书对模型的选择及具体的训练算法均不需要进行限定。

[79]由于预测模型是基于一定的历史数据样本训练得到，而随着时间的推移，可能会有多种因素导致模型的适用性下降，因此需要适时对模型进行优化调整。对于整合模型而言也同样存在优化调整的需求，基于整合模型的特性，本说明书进一步提供整合模型的优化方法。

[80]图 3 所示，为本说明书提供的一种整合模型优化方法的流程图，该方法可以包括以下步骤：

[81]S201，确定初始的整合模型；

[82]本说明书中以“初始整合模型”指代优化处理前的整合模型，根据前面实施例的描述，该整合模型是由至少 2 个独立模型整合得到，具有 1 个整合输出值、且以至少 2 个独立模型的输出值作为输入值。

[83]在本说明书中，“初始整合模型”用于指代模型的某个待优化状态，既可以是训练后完全未经优化的模型，也可以是上一次优化后的结果。另外，该“初始整合模型”可以

利用前述的整合模型训练方法得到，也可以从其他途径得到，本说明书对初始整合模型的获得方式并不进行限定。

[84]S202，获得当前的整合模型针对预设的测试事件集输出的预测结果，并对预测结果进行统计得到至少 1 个指标值；

[85]用于评价风控预测模型的性能指标，一般包括预测覆盖率、预测命中率、用户打扰率等，这些指标同样适用于整合模型，而这些指标是否表现足够好，则是决定是否要对整合模型进行优化的重要依据。

[86]在本步骤中，利用当前的整合模型，对某个已知结果（即具有风险标注结果）的目标事件集进行预测，将预测结果与风险标签进行对比，并且统计准确率、打扰率等指标，以便在后续步骤中根据这些指标确定是否开始进行优化。实际应用中，该测试事件集可以取自某个时间窗口（例如最近一周、最近一个月等等），从而实现整合模型性能的按需优化。

[87]另外，根据图 3 可知，一次完整的优化过程，可能需要迭代若干个轮次才能够优化到位，因此这些指标值同样可以作为一次完整优化中“是否开启下一轮次的优化”的判断依据。实际应用中，对于不同的迭代轮次而言，每次使用的测试事件集可以是完全相同的，也可以有一定的差别。例如将测试事件集划分为多个子集，每次迭代针对其中一个子集进行预测输出；或者每次从测试事件集中随机抽取一部分用于本次迭代的预测输出。

[88]S203，判断是否满足预设的迭代停止条件，如果是则进行执行 S205，否则执行 S204；

[89]根据前面的描述可知，“指标是否表现足够好”是决定是否要对整合模型进行优化的重要依据，因此这里的迭代停止条件应至少包括：当前整合模型的指标满足预设的性能需求。例如“用户打扰率在 1%以下”、“预测覆盖率在 90%以上”等等。当存在多个指标值时，可以对各个指标值分别设置判断标准，也可以根据一定的策略综合判断，本说明书对具体的判断标准不需要进行限定。

[90]可以理解的是，“当前整合模型”在第 1 轮迭代中指代的是“初始的整合模型”，在第 2 轮及以后的迭代轮次中则指代是上一轮优化后得到的整合模型。

[91]实际应用中，对于初始的整合模型，可以周期性（例如每周一次、每月一次等）获得整合模型预测结果，进而根据统计得到的指标值确定是否需要开始优化。也可以持续对整合模型预测结果进行监测，发现性能指标不满足需求后进行告警提示，并且触发优

化流程。

[92]此外，根据实际需求，这里的迭代停止条件还可以是其他形式，例如：

[93]可以在达到某个预设的迭代次数后控制迭代停止；

[94]可以在连续多次采用同样的优化策略、但却没有明显性能提升的情况下，控制迭代停止；

[95]可以在执行完所有可用的优化策略后控制迭代停止；

[96]当然，以上迭代停止条件仅用于示意性说明，不应理解为对本说明书方案的限定。

[97]S204，对整合模型进行优化处理；

[98]根据整合模型的特性，本说明书提供多种整合模型优化策略，根据策略优化对象的不同，整体上可分为两大类：对多个独立模型的整合方式进行调整、或者对独立模型进行单独调整，其中每一大类策略还可以进一步包含多种实现方式，并且在迭代优化的过程中，不同轮次的优化策略安排还可以产生更为丰富的策略形式。本步骤的具体实现方式将在后面的实施例中详细说明。

[99]S205，若当前的整合模型指标值满足性能需求，则将当前的整合模型确定为可用整合模型。

[100] 对于初始整合模型、或者任意轮次优化后得到的整合模型，如果能够满足性能需求，则可以作为最终的优化结果投入使用。如果在停止迭代处理后，当前的模型仍未满足性能需求，则可用转入人工介入其他处理流程，具体处理方式与本说明书方案无关，这里不再详细说明。

[101] 以上介绍了模型优化处理的整体流程，下面将对 S204 所涉及的优化策略做进一步展开说明：

[102] 根据前面的说明，从“优化对象”的维度划分，可以将优化策略分为两类：

a）对多个独立模型的整合方式进行调整；

b）对独立模型进行单独调整。

[103] 此外，如果从“优化手段”维度划分，也可以将优化策略分为两类：

a）不改变模型的输入特征、仅对特征参数进行调整（refit）；

b）增加或减少模型的输入特征，重新训练得到新的模型（retrain）。

[104] 根据上述两个维度的划分，排列组合后可的优化策略如表 5 所示：

	独立模型	整合模型
refit	独立模型 refit	整合模型 refit
retrain	独立模型 retrain	整合模型 retrain

表 5

[105] 表 5 所示的 4 种优化策略具体解释如下：

[106] 整合模型 refit：不改变参与整合的独立模型，对独立模型在整合模型中的整合权重进行调整，这里的“整合权重”即整合模型的各个特征值（也就是各个独立模型的输出）的参数；

[107] 整合模型 retrain：增加或减少用于提供整合模型输入值的独立模型，重新训练得到新的整合模型；

[108] 独立模型 refit：不改变独立模型的输入特征，对独立模型的各个特征参数进行调整；

[109] 独立模型 retrain：增加或减少独立模型的输入特征，重新训练得到新的独立模型。

[110] 当然，以上仅列出了 4 中典型的优化策略，不应理解为对本说明书方案的限定，例如，对于基于人工规则建立的独立模型，可以直接通过修改或替换规则的方式实现调整。

[111] 无论是 refit 还是 retrain，本质上都需要通过训练完成，由于模型性能不达标的原因很可能是过去可用的模型已经无法满足最新的情况，因此可以选用最近一段时间（例如最近一周、最近一个月等等）发生的事件作为优化训练事件集，进而根据优化训练事件集进行有监督训练以实现 refit 或 retrain。

[112] 实际应用中，“测试事件集”和“优化训练事件集”可以从同一事件集中获取，例如将最近一段时间发生过的事件划分为两部分，其中一部分用于统计指标值、另一部分用于对整合模型进行优化处理。

[113] 在一次完整的模型优化过程中，可以选用表 5 所示出的一种或多种优化策略，如果从训练成本方面考虑，应优先对整合模型进行调整、然后再对独立模型进行调整；并且优先对模型进行 refit、然后再对模型进行 retrain。实际应用中，客观情况可能并不允许随意增减模型或特征的数量，因此 retrain 的策略使用频率也相对较小。

[114] 例如，假设整合模型是由 3 个独立模型整合得到，模型表达式如下：

$$y_{整}=h(-0.1+0.7\ y_1+\ 0.2\ y_2+0.5\ y_3)$$

[115] 其中 y_1，y_2，y_3 分别表示 3 个独立模型的输出，0.7、0.2、0.5 为整合权重。

[116] 如果发生性能告警，则在第一个迭代轮次中，优先利用优化训练事件集对整合进行 refit，假设调整后的整合权重如下：

$$y_{整\ refit}=h(-0.2+0.8\ y_1+\ 0.1\ y_2+0.4\ y_3)$$

[117] 在本例中，不考虑“对整合模型进行 retrain”这一策略，因此如果 refit 后的整合模型仍然不满足性能需求，则需要进行整合因子拆解，以便分别对各个独立模型进行优化。

[118] 对于多个独立模型的优化处理，可以分别独立进行，也可以按照一定的顺序逐个执行。例如，可以各个独立模型在整合模型中的整合权重，优先对整合权重较高的独立模型进行调整。

[119] 以上述的 $y_{整\ refit}$ 表达式为例，根据整合权重可知，独立模型 1 对整合输出值的贡献最大，其次是独立模型 3、最后是独立模型 2，因此独立模型的调整优先级为 1→3→2。

[120] 对每个独立模型的优化策略优先采用 refit、然后是 retrain。对于同一独立模型，由于可能包含多种增减参数方案，因此可能包含多个轮次的 retrain，另外根据实际情况，对于某些独立模型也可以不进行 retrain。

[121] 每个轮次优化完成后，均可以利用测试事件集获得性能指标值，如果满足性能需求则停止，否则继续开启下一轮优化。如果对所有独立模型均进行优化后仍不满足性能需求，可以转为人工方式做进一步诊断。

[122] 图 3 示出了一种基于循环流程的整合模型优化方案，其特点是按需优化，即每次优化之后即对优化效果进行检测，如果满足需求则不再继续进行优化。可以理解的是，在某些情况下也可以不使用迭代方式实现整合模型的优化，例如图 4 所示的顺序优化流程，与图 3 所示循环流程的区别在于，在 S303 可以一次性执行全部的优化策略，例如仅使用一种优化策略、或按照固定规则执行多种优化策略。顺序流程中其他部分的实现均可参见循环流程中相应部分的说明，这里不再重复说明。

[123] 以上分别从“模型构建”以及“模型优化”两方面对模型整合方案进行说明，

应用模型构建方案，可以将多个独立的模型分值整合为一个综合分值，使得输出模型策略更加便于管理，也解决了模型过多所带来的冗余问题，提升系统处理资源的使用效率；应用模型优化方案，对于整合得到的模型，可以基于性能监控预警，自动触发整合模型的迭代优化，从而保证综合分值的稳定有效输出。可见，本说明书提供的“模型构建”以及“模型优化”方案，既可以共同构成一个整体，也可以分别构成独立的方案以解决不同的技术问题。

[124] 可以理解的是，除风控场景之外，本说明书方案在其他的应用场景依然能够有效适用，例如：在文本语义识别场景中，存在多个模型利用不同的文本特征维度来识别语义；在图像内容识别场景中，存在多个模型就不同的图像特征判断该图像是否为某一事物；在严格身份认证场景中，存在面部特征模型、指纹特征模型等多个模型共同判断用户身份；……。事实上，对于针对同一目的、使用不同特征构建多种模型的应用场景，都可以应用本说明书方案进行模型整合及优化，因此本说明书实施例中的使用的场景举例不应理解为对本说明书方案的限定。

[125] 相应于上述方法实施例，本说明书实施例还提供一种模型整合系统，参见图 5 所示，该系统可以包括两部分：用于构建整合模型的装置以及用于优化整合模型的装置，可以理解的是，上述两种装置也可以分别独立运行以解决不同的技术问题。以下分别进行说明：

[126] 参见图 6 所示，用于构建整合模型的装置可以包括：

[127] 独立模型确定模块 110，用于确定待整合的至少 2 个独立模型，每个独立模型具有 1 个输出值；

[128] 事件数据集获取模块 120，用于从历史事件记录中获取事件数据集，每条事件数据中可以包括：每个独立模型针对该事件的预测输出值、该事件的综合标签值；

[129] 训练样本生成模块 130，用于根据所获得的事件数据集，生成整合模型训练样本；

[130] 训练模块 140，用于利用有监督学习算法对样本进行训练得到整合模型，整合模型具有 1 个整合输出值、且以多个独立模型的输出值作为输入值。

[131] 在本说明书提供的一种具体实施方式中，训练样本生成模块可以具体用于：

[132] 根据预设的映射规则，对各独立模型的输出值进行映射处理，将映射结果确定为用于训练整合模型的训练特征值。

[133] 在本说明书提供的一种具体实施方式中，训练样本生成模块，可以利用以下方式对各独立模型的输出值进行映射处理：

[134] 将独立模型 i 的输出值的值域划分为 L_i 个子区间；其中 $L_i \geqslant 2$， i=1,2,3…M，M 为参与整合的独立模型数量；

[135] 对于独立模型 i 的任意输出值，根据该输出值所处的子区间，确定该输出值对应的映射结果。

[136] 在本说明书提供的一种具体实施方式中，训练样本生成模块可以利用以下方式将独立模型 i 的输出值的值域划分为 L_i 个子区间：

[137] 针对所获得的事件数据集，根据独立模型 i 对各事件的输出值的分布情况，将独立模型 i 的输出值的值域划分为 L_i 个子区间。

[138] 在本说明书提供的一种具体实施方式中，不同独立模型 i 所对应的子区间划分数量 Li 相同，且不同独立模型间的相应子区间对应的输出值分布占比相同。

[139] 在本说明书提供的一种具体实施方式中，训练样本生成模块可以利用以下方式确定输出值的映射结果：

[140] 对于独立模型 i 的任意输出值，将输出值所处的子区间的 WOE 分值确定为该输出值的映射结果；独立模型 i 的每个输出值对应 1 个训练特征值。或者

[141] 对于独立模型 i 的任意输出值，根据该输出值所处的子区间，将该输出值转换为长度为 L_i 的 One-Hot 编码值；独立模型 i 的每个输出值对应 L_i 个训练特征值。

[142] 上述整合模型构建装置，还可以进一步配置优化处理模块，用于在确定当前的整合模型的指标值不满足性能需求的情况下，根据预设的优化训练事件集对整合模型进行优化处理。例如对多个独立模型的整合方式进行调整、或者对独立模型进行单独调整。具体的优化策略使用可以参见本说明书中的其他实施例，这里不再重复说明。

[143] 参见图 7 所示，用于优化整合模型的装置可以包括：

[144] 初始整合模型确定模块 210，用于确定初始的整合模型，整合模型具有 1 个整合输出值、且以至少 2 个独立模型的输出值作为输入值；

[145] 优化处理模块 220，利用以下子模块配合进行迭代处理，直到满足预设的迭代停止条件，迭代停止条件至少可以包括：当前的整合模型指标满足预设的性能需求；

[146] 性能评价子模块，用于获得当前的整合模型针对预设的测试事件集输出的预测

结果，并对预测结果进行统计得到整合模型的至少 1 个指标值，指标值用于评价整合模型的性能；

[147] 优化处理子模块，用于根据预设的优化训练事件集对整合模型进行优化处理，其中优化处理可以包括：对多个独立模型的整合方式进行调整、或者对独立模型进行单独调整；

[148] 输出模块 230，用于在停止迭代后，若当前的整合模型指标值满足性能需求，则将当前的整合模型确定为可用整合模型。

[149] 在本说明书提供的一种具体实施方式中，优化处理子模块可以具体用于：

[150] 在迭代处理过程中，优先对多个独立模型的整合方式进行调整，如果无法满足性能需求，则在后续的迭代轮次中对独立模型进行单独调整。

[151] 在本说明书提供的一种具体实施方式中，优化处理子模块可以具体用于利用以下方式对多个独立模型的整合方式进行调整：

[152] 不改变参与整合的独立模型，对独立模型在整合模型中的整合权重进行调整；或者

[153] 增加或减少用于提供整合模型输入值的独立模型，重新训练得到新的整合模型。

[154] 在本说明书提供的一种具体实施方式中，优化处理子模块可以具体用于利用以下方式对独立模型进行单独调整：

[155] 在迭代处理过程中，根据各个独立模型在整合模型中的整合权重，优先对整合权重较高的独立模型进行调整，如果无法满足性能需求，则在后续的迭代轮次中对整合权重较低的独立模型进行调整。

[156] 在本说明书提供的一种具体实施方式中，优化处理子模块可以具体用于利用以下方式对独立模型进行单独调整：

[157] 不改变独立模型的输入特征，对独立模型的参数进行调整；或者

[158] 增加或减少独立模型的输入特征，重新训练得到新的独立模型。

[159] 在本说明书提供的一种具体实施方式中，用于统计指标值的测试事件集、以及用于对整合模型进行优化处理的优化训练事件集，为取自同一事件集的测试集和验证集。

[160] 在图 7 所示的用于优化整合模型的装置独立工作的情况下，初始整合模型确定

模块可以用于实现整合模型的构建:

[161] 从历史事件记录中获取事件数据集，每条事件数据中可以包括：每个独立模型针对该事件的预测输出值、该事件的综合标签值;

[162] 以所获得的事件数据集作为样本，利用有监督学习算法对样本进行训练得到整合模型。

[163] 另外，相应于方法实施例中的顺序优化流程，优化处理模块 220 还可以采用非循环的方式执行优化，例如仅使用一种优化策略、或按照固定规则执行多种优化策略等等，具体实现可参见前面实施例，这里不再重复说明。

[164] 本说明书实施例还提供一种计算机设备，其至少包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其中，处理器执行所述程序时实现前述的模型整合方法，具体可以是在一台设备中同时实现前述的整合模型构建方法及整合模型优化方法，也可以是在不同设备中分别实现前述的整合模型构建方法及整合模型优化方法。

[165] 图 8 示出了本说明书实施例所提供的一种更为具体的计算设备硬件结构示意图，该设备可以包括：处理器 1010、存储器 1020、输入/输出接口 1030、通信接口 1040 和总线 1050。其中处理器 1010、存储器 1020、输入/输出接口 1030 和通信接口 1040 通过总线 1050 实现彼此之间在设备内部的通信连接。

[166] 处理器 1010 可以采用通用的 CPU（Central Processing Unit，中央处理器）、微处理器、应用专用集成电路（Application Specific Integrated Circuit，ASIC）、或者一个或多个集成电路等方式实现，用于执行相关程序，以实现本说明书实施例所提供的技术方案。

[167] 存储器 1020 可以采用 ROM（Read Only Memory，只读存储器）、RAM（Random Access Memory，随机存取存储器）、静态存储设备，动态存储设备等形式实现。存储器 1020 可以存储操作系统和其他应用程序，在通过软件或者固件来实现本说明书实施例所提供的技术方案时，相关的程序代码保存在存储器 1020 中，并由处理器 1010 来调用执行。

[168] 输入/输出接口 1030 用于连接输入/输出模块，以实现信息输入及输出。输入输出/模块可以作为组件配置在设备中（图中未示出），也可以外接于设备以提供相应功能。其中输入设备可以包括键盘、鼠标、触摸屏、麦克风、各类传感器等，输出设备可以包

括显示器、扬声器、振动器、指示灯等。

[169] 通信接口 1040 用于连接通信模块（图中未示出），以实现本设备与其他设备的通信交互。其中通信模块可以通过有线方式（例如 USB、网线等）实现通信，也可以通过无线方式（例如移动网络、WIFI、蓝牙等）实现通信。

[170] 总线 1050 包括一通路，在设备的各个组件（例如处理器 1010、存储器 1020、输入/输出接口 1030 和通信接口 1040）之间传输信息。

[171] 需要说明的是，尽管上述设备仅示出了处理器 1010、存储器 1020、输入/输出接口 1030、通信接口 1040 以及总线 1050，但是在具体实施过程中，该设备还可以包括实现正常运行所必需的其他组件。此外，本领域的技术人员可以理解的是，上述设备中也可以仅包含实现本说明书实施例方案所必需的组件，而不必包含图中所示的全部组件。

[172] 本说明书实施例还提供一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现前述的模型整合方法，具体可以是在一台设备中同时实现前述的整合模型构建方法及整合模型优化方法，也可以是在不同设备中分别实现前述的整合模型构建方法及整合模型优化方法。

[173] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存（PRAM）、静态随机存取存储器（SRAM）、动态随机存取存储器（DRAM）、其他类型的随机存取存储器（RAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、快闪记忆体或其他内存技术、只读光盘只读存储器（CD-ROM）、数字多功能光盘（DVD）或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体（transitory media），如调制的数据信号和载波。

[174] 通过以上的实施方式的描述可知，本领域的技术人员可以清楚地了解到本说明书实施例可借助软件加必需的通用硬件平台的方式来实现。基于这样的理解，本说明书实施例的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在存储介质中，如 ROM/RAM、磁碟、光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本说明书实施例各个实施例或者实施例的某些部分所述的方法。

[175] 上述实施例阐明的系统、装置、模块或单元，具体可以由计算机芯片或实体实现，或者由具有某种功能的产品来实现。一种典型的实现设备为计算机，计算机的具体形式可以是个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件收发设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任意几种设备的组合。

[176] 本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于装置实施例而言，由于其基本相似于方法实施例，所以描述得比较简单，相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，在实施本说明书实施例方案时可以把各模块的功能在同一个或多个软件和/或硬件中实现。也可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

[177] 以上所述仅是本说明书实施例的具体实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本说明书实施例原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本说明书实施例的保护范围。