*数据库设计文档*

**因子标签库：tags**

1. **数据库概要**

主要用于存储**因子基本信息**（名称、来源、值域等）、**因子有效性信息**（因子在不同模型中的得分）、**标签基本信息**（标签种类+具体值，如：数据处理方式-离散化）以及**因子标签信息**（因子编号+标签编号）。

**二、表设计**

1. **因子信息表：factor\_info**
   1. 描述：存储因子基本信息，便于从来源企业、来源数据库表中快速查找对应的因子信息
   2. 字段设计：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段含义 | 数据类型 | 长度 | 允许为空 | 备注 |
| **factor\_id** | 因子编号 | int | 11 | 否 | 自增 |
| factor\_name | 因子名称 | char | 50 | 否 |  |
| factor\_code | 因子代码 | char | 50 | 否 |  |
| factor\_range | 值域 | char | 50 | 是 |  |
| source\_company | 因子来源企业 | char | 50 | 是 |  |
| source\_database | 因子来源数据库 | char | 50 | 是 |  |
| source\_table | 因子来源表 | char | 50 | 是 |  |
| remark | 备注 | text |  | 是 |  |

1. **因子评分表：factor\_score**

2.1 描述：存储因子在各模型中的表现结果

2.2 字段设计：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段含义 | 数据类型 | 长度 | 允许为空 | 备注 |
| ***factor\_id*** | 因子编号 | int | 11 | 否 |  |
| decision\_tree\_score | 决策树模型评分 | float |  | 是 |  |
| regression\_score | 回归模型评分 | float |  | 是 |  |
| bayes\_score | 贝叶斯模型评分 | float |  | 是 |  |
| final\_score | 综合评分 | float |  | 是 |  |

1. **标签信息表：tag\_info**

3.1 描述：存储标签的基本信息

3.2 字段设计：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段含义 | 数据类型 | 长度 | 允许为空 | 备注 |
| **tag\_id** | 标签编号 | int | 11 | 否 | 自增 |
| tag\_name | 标签值 | char | 50 | 否 |  |
| tag\_type | 标签种类 | char | 50 | 否 |  |
| remark | 备注 | text |  |  |  |

1. **因子标签表：relations**

4.1 描述：存储因子与标签的对应信息  
4.2 字段设计：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 字段含义 | 数据类型 | 长度 | 允许为空 | 备注 |
| **relation\_id** | 关系编号 | int | 11 | 否 | 自增 |
| *factor\_id* | 因子编号 | int | 11 | 否 |  |
| *tag\_id* | 标签编号 | int | 11 | 否 |  |
| remark | 备注 | text |  |  |  |

1. **视图设计**

1. 视图factor\_tag：获得因子名称与标签名称的关系查询结果

视图查询语句：

**select** factor\_info.factor\_id, factor\_info.factor\_name, tag\_info.tag\_type, tag\_info.tag\_name **from** factor\_info **inner** **join** relations **on** relations.factor\_id=factor\_info.factor\_id **inner** **join** tag\_info **where** relations.tag\_id=tag\_info.tag\_id

预期结果示例：

