

高可靠特征库的使用说明书

版本说明

版本	时间	说明
V0.90	2023.12.20	第一个版本发布。
V0.91	2024.01.15	增加特征表的使用方法举例。
V0.92	2024.01.18	增加交流联系方式

特征表的说明

名称	说明
varname	特征名称
buy_sell	多空方向，0 表示空方向相关的特征，1 表示多方式相关特征，0 和 1 方向特征有可能重叠
absprofit	衡量收益能力的指标，数值越高，该特征的收益能力越强。
corr	和收益目标的相关性，相关性的 size 只是一种参考，有些相关性小的，但是有可能收益会比较高。
period_base	数据的基础时间框架，包括 M5,M15,M30,H1,H3,H12,D1 七种时间框架，同时间框架的特征之间的相关性进行了一定的控制，一般相关性不超过 0.6，不同时间框架之间的特征没有做相关性控制，使用者可以根据自己的策略进行长短时间周期的特征进行组合配对使用。

特征的使用说明

名称	说明
带 201 的特征	预测第二天早上开盘到晚上的收益能力，比如 935-2245

带 202 的特征	带 202 的特征：预测第二天早上开盘到下午收盘的收益能力，比如 935-1445
带 101 的特征	预测第二天早上开盘到第三天早上开盘的收益能力，比如第二天 935 到第三天 935

日内的特征暂时不开放，计算量巨大并且实时要求太高，可以做定制化开发。

数据特征都是日级别，101，201，202 分别适合相同不同的场景。

使用举例：

Target 格式如下：

Index	ID	Before_date	date	Time_start	Time_end	profit_loss
1	CZCE.SA	2023.03.04	2023.03.05	09:45	22:45	1.234
2	CZCE.TA	2023.03.04	2023.03.05	09:45	22:45	2.22
3	DCE.FB	2023.03.04	2023.03.05	09:45	22:45	-1.233
...

品种的命名方式 市场.品种，对齐的时候请使用大写字母对齐。

上面的数据中预测某天开盘 9:45 到收盘 22:45 的收益的情况 profit_loss，profit_loss 可以作为 target。

Feature 特征表格式如下：

Index	ID	date	N223	N224	N226	...
1	CZCE.SA	2023.03.04	1.2313	2.232	3.232	...
2	CZCE.TA	2023.03.04	2.323	5.3434	2.223	...
3	DCE.FB	2023.03.04	2.3134	2.343	2.323	...
...

那么 Target 数据和特征数据对应的由 ID，date 决定，

On Target.ID=Feature.ID and Target. Before_date=Feature. Date

所有的特征表都有 ID，date 两个字段，理论上可以把所有的特征都是用起来。

特征的统计数据

201 特征统计								
行标签	D1	H1	H12	H3	M15	M30	M5	总计
0	399	1022	742	880	1055	1053	1070	6221
1	118	225	302	206	272	207	215	1545
总计	517	1247	1044	1086	1327	1260	1285	7766

202 特征统计								
行标签	D1	H1	H12	H3	M15	M30	M5	总计
0	1005	1522	2025	1190	1593	1647	1386	10368
1		17		18	32	24	19	110
总计	1005	1539	2025	1208	1625	1671	1405	10478

101 特征统计								
行标签	D1	H1	H12	H3	M15	M30	M5	总计
0	38	213	148	225	307	239	295	1465
1	869	2061	1585	1576	2294	2197	2250	12832
总计	907	2274	1733	1801	2601	2436	2545	14297

数据更新频率

每天夜跑 3 点之后开始更新，一般 4 点前完成更新。

联系方式

微信：446860105

联系暗号：quant