

一种结合门店周围环境的相关指数对门店进行评估的方法

技术领域

本发明涉及门店商业评估技术领域，具体为一种结合门店周围环境的相关指数对门店进行评估的方法。

背景技术

商业用户在进行门店选址、门店分群画像、线下门店对比测试等行动时，通常需要对评估点周围地理环境、商业环境及人群等有客观认识，以便于评估工作，门店大时空指标是通过周边地理信息以及天气数据统计建模后，抽象出来的一系列具有一定业务指导意义的门店大尺度地理环境主题分数以及衡量人群出行便利和舒适度的相关指数，完美符合商业用户的需求，有助于其对门店周围环境有客观清晰的认识，提升其评估工作开展效率；

现有评估工作大多是通过人员实地考察收集地理信息，再采用专家评估投票的方式进行得到区域地理环境评分，人员实地采集信息从人员成本和时间成本来说，消耗巨大，且不能保证信息收集完全，不可作为大范围区域的评估方法，且专家评分得到的地理环境评分主观性过强，评估结果受个人因素影响严重。

发明内容

本发明提供一种结合门店周围环境的相关指数对门店进行评估的方法，可以有效解决上述背景技术中提出人员实地采集信息从人员成本和时间成本来说，消耗巨大，且不能保证信息收集完全，不可作

为大范围区域的评估方法，且专家评分得到的地理环境评分主观性过强，评估结果受个人因素影响严重的问题。

为实现上述目的，本发明提供如下技术方案：一种结合门店周围环境的相关指数对门店进行评估的方法，主要从地理信息中量化出具有业务价值且强解释性的地理环境主题分数，结合地理信息和相关自然特征为指标分数，对门店经营状况或经营需求的能力进行评估，为门店业务场景提供参考依据；

包括如下评估步骤：

S1、收集线下门店周边地理信息数据；

S2、对门店周边主体进行统计建模；

S3、合理提取主题特征指标 SSTI；

S4、线性回归模型评价体系；

S5、指标 SSTI 应用于现实业务场景。

根据上述技术方案，所述 S1 中，线下门店周边地理信息数据在收集时主要通过使用不同数据源来得到目标门店周围的地理信息，该地理信息主要包括门店周围环境信息和相关自然特征；

获取地理信息后，将目标门店附近设施分为交通、居住、购物、美食、旅游、生活、工作和娱乐八个类别的主题，同时针对每一个主题，来得到目标门店附近该主题设施的数量，以及与目标门店的距离信息。

根据上述技术方案，所述 S2 中，对门店周边主体进行统计建模主要通过建立主题强度模型和主题聚集度模型，并综合两个模型的结

果，来生成该主题下的大时空指标；

主要采用某品牌数据为例，结合线性回归模型来验证大时空指标对客流量的影响大小，待验证通过，可将大时空指标应用于品牌的现实业务场景，助力提升品牌对门店周边环境的评估效率，最终作为门店选址等需求的重要参考。

根据上述技术方案，所述 S2 中在具体统计建模时，门店地理位置环境主题分数受到门店周边主题兴趣点数量强度、聚集程度的影响，以门店周边 POI 为样本，对门店周边主题 POI 分布建立以下模型：

$$\text{Loc}\left(\text{POI}_t^{(s)}\right) \stackrel{\text{i.i.d}}{\sim} \mathfrak{B}\left(p_t^{(s)}\right) \times^{\text{indep.}} \mathcal{N}\left(0, \Sigma_t^{(s)}\right)$$

$$p_t^{(s)} \sim \mathbb{B}\left(\alpha_t, \beta_t\right), \sigma_t^{(s)} \sim \mathcal{N}\left(\mu_t, \lambda_t\right)$$

其中：

s 为门店上标；

t 为主题下标，包括交通、居住、购物、美食、旅游、生活、工作、娱乐八类；

$\text{Loc}\left(\text{POI}_t^{(s)}\right)$ 为门店周边主题 POI 分布；

$\mathfrak{B}\left(p_t^{(s)}\right)$ 为门店周边主题 POI 强度模型，服从伯努利分布；

$p_t^{(s)}$ 为门店周边主题 POI 强度模型概率参数，符合贝塔分布；

α_t, β_t 为强度参数分布模型变量；

$\mathcal{N}\left(0, \Sigma_t^{(s)}\right)$ 为门店周边主题 POI 聚集模型，服从二维正态分布；

$\Sigma_t^{(s)}$ 为门店周边主题 POI 聚集模型方差，符合高斯分布，并满足

$$\Sigma_t^{(s)} = \begin{bmatrix} \left(\sigma_t^{(s)}\right)^2 & 0 \\ 0 & \left(\sigma_t^{(s)}\right)^2 \end{bmatrix}, \left(\sigma_t^{(s)}\right)^2 \text{ 为 POI 经度或纬度的方差;}$$

μ_t, λ_t 为聚集参数分布模型变量。

根据上述技术方案，在 S2 中，基于模型的计算流程包括输入、输出和计算过程，具体如下：

输入： t ，评估主题，包括交通、居住、购物、美食、旅游、生活、工作、娱乐共计八类；

m ：总评估门店数；

$n_t^{(s)}$ ： t 主题 s 家门店查询到的 POI 数量；

$N_t^{(s)}$ ： t 主题 s 家门店查询到的 POI 次数；

$p_t^{(s)}$ ： t 主题 s 家门店的 POI 强度模型初始概率参数；

$\sigma_t^{(s)}$ ： t 主题 s 家门店所有 POI 与门店距离的标准差；

$(x_i^{(s)}, y_i^{(s)})$ ： s 门店附近 i POI 点相对于门店的坐标；

α, β ：门店间强度模型先验的服从贝塔分布， α, β 为模型变量；

μ, λ ：门店间聚集度模型先验的服从高斯分布， μ, λ 为模型变量；

输出： $SSTI_t^{(s)}$ ： s 家门店 t 主题下的得分。

根据上述技术方案，所述 S2 中计算过程如下：

1)、在每个 t 中执行；

2)、循环开始；

3)、强度模型中，门店周边主题 POI 强度服从伯努利分布，其概率参数服从贝塔分布，强度模型的联合概率密度函数为：

$$\mathfrak{B}(\alpha, \beta, p) = \left(\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \right)^m \left(\prod_{s=1}^m \left(p_t^{(s)} \right)^{(\alpha - 1 + n_t^{(s)})} \left(1 - p_t^{(s)} \right)^{(\beta - 1 + N_t^{(s)} - n_t^{(s)})} \right);$$

4)、根据极大似然估计, 令 $\mathfrak{B}(\alpha, \beta, p)$ 关于 p 的偏导数为 0, 求其 p 为 p_t , 其中每家门店的强度模型概率参数为 $p_t^{(s)}$;

5)、聚集度模型中, 门店周边主题 POI 服从二维正态分布, POI 聚集模型方差符合高斯分布, 聚集度联合概率密度函数为:

$$\mathcal{N}(\lambda, \mu, \sigma) = \prod_{s=1}^m \left(\frac{1}{\lambda \sqrt{2\pi}} e^{-\frac{(o_t^{(s)} - \mu)^2}{2\lambda^2}} \left(\prod_{i=1}^{n_t^{(s)}} \frac{1}{2\pi (o_t^{(s)})^2} e^{-\frac{(x_i^{(s)})^2 + (y_i^{(s)})^2}{2(o_t^{(s)})^2}} \right) \right);$$

6)、根据极大似然估计, 令 $\mathcal{N}(\lambda, \mu, \sigma)$ 关于 σ 的偏导数为 0, 求其 σ 为 σ_t , 其中每家门店的集中度模型标准差为 $\sigma_t^{(s)}$;

7)、得到每家门店 t 主题的主题强度模型概率参数为 $p_t^{(s)}$ 和集中度模型标准差为 $\sigma_t^{(s)}$;

8)、循环结束;

9)、门店信息、实时天气数据作为自变量, 客流数据作为因变量, 构建线性回归模型;

10)、 R^2 —模型决定系数;

11)、AIC—赤池信息量准则;

12)、BIC—贝叶斯信息准则;

13)、定义主题得分计算方式 $SSTI_t^{(s)} = \frac{p_t^{(s)}}{\sigma_t^{(s)}}$;

- 14)、在每个t中执行;
- 15)、循环开始;
- 16)、每家门店t主题的主题得分 $SSTI_t^{(s)}$;
- 17)、在线性回归模型加入主题得分特征 $SSTI_t^{(s)}$;
- 18)、得到加入主题得分后的 R_t^2 、 AIC_t 和 BIC_t ;
- 19)、循环结束;
- 20)、当各t中 R_t^2 效果无明显提升或 AIC_t 过大或 BIC_t 过大时;
- 21)、循环执行;
- 22)、重新定义主题得分计算方式, 执行 14 至 19 步;
- 23)、观察 R_t^2 、 AIC_t 、 BIC_t ;
- 24)、直到各t中 R_t^2 效果提升明显且 AIC_t 、 BIC_t 数值均合理;
- 25)、循环结束;
- 26)、保留对应的t的得分 $SSTI_t^{(s)}$;
- 27)、 $SSTI_t^{(s)}$ 可应用至品牌门店选址和分群等真实业务场景, 为评估人员选择最优地理环境提供各维度主题评分参考依据。

根据上述技术方案, 所述 S3 中, 合理提取主题特征指标 SSTI 主要指利用最大似然估计来估计每家门店每个主题的 $\hat{p}_t^{(s)}$ 和 $\hat{\sigma}_t^{(s)}$;

具体的, $\hat{p}_t^{(s)}$ 越大, 门店周边主题 POI 数量强度越大, $\hat{\sigma}_t^{(s)}$ 越大, 门店周边主题 POI 聚集程度越高, 因此, 定义以下的 $SSTI_t^{(s)}$ 来描述门店主题下的 POI 分布:

$$SSTI_t^{(s)} = \frac{p_t^{(s)}}{\sigma_t^{(s)}}$$

其中： $\hat{p}_t^{(s)}$ 为由 MLE 估算出的s门店周边t主题的 POI 强度模型概率参数；

$\widehat{\sigma}_t^{(s)}$ 为由 MLE 估算出的s门店周边t主题的 POI 聚集模型标准差，门店 $SSTI_t^{(s)}$ 分数越高，主题特征越明显。

根据上述技术方案，所述 S4 中，线性回归模型评价体系具体指通过选取一个可用于监测的且能够衡量门店经营状况或经营目标的指标，如客流量、销量，使用地理信息指标以及其他相关自然特征建立模型，通过模型拟合度增益来体现地理信息指标的合理性；

其中，线性回归模型具备良好的线性解释性和特征兼容性，拥有完备的数值化特征评估和推断机制，且能够简单描述特征叠加的增益，对评估线下门店大时空特征有天然的适配性，以及它不需要训练集和测试集的前提条件，因此，由线性回归模型来评估自变量特征对因变量的影响大小及正负。

根据上述技术方案，在 S4 中，衡量模型拟合度的指标定为决定系数，表示在目标变量的总变异中回归关系所能解释的百分比，数值越接近于 1 则拟合的效果越佳，赤池信息量准则AIC和贝叶斯信息准则BIC，可权衡所估计模型的复杂度和模型拟合数据的优良性，数值越小模型效果越佳；

门店信息、实时天气及 8 类地理环境指标等数据作为自变量，客流数据作为因变量，构建线性回归模型，将 8 类地理环境特征分次单个加入评估模型后， R^2 有所提升，其中交通指数的效果最为明显，依次是娱乐指数和旅游指数，对目标变量-客流量变化的可解释性增强，

各指标的AIC和BIC均有所下降，表示模型拟合良好，未出现过拟合现象。

根据上述技术方案，所述S5中，指标SSTI应用于现实业务场景，主要结合以上检验结果，为品牌提供一定的门店选址及门店分群方面的业务指导和参考价值；

在具有其他相似条件下，新门店地址可首选参考交通便利处并具备一定的娱乐设施，如临近地铁口的商业综合体，同时也可考虑旅游胜地等场所，根据现实场景需求，门店大时空指标为商业用户量化出周边环境的客观数字依据，快速提升地理评估的工作效率。

与现有技术相比，本发明的有益效果：本发明结构科学合理，使用安全方便：

1、通过将目标区域点周边的地理信息及自然环境量化成包含通、居住、购物、美食、旅游、生活、工作和娱乐八个维度的数值，方便使用者快速高效地掌握周围环境状态，结合地理信息和相关自然特征为指标分数，对门店经营状况或经营需求的评估能力，对门店选址、门店分群、门店管理等真实场景提供参考依据，此外，八大指标还可作为门店经营情况的预测特征，运用回归模型观察周边环境是否对客流量等关注指标产生显著性影响，以及产生的影响大小，因此对业务也具有指导意义；

同时，使用统计建模这种高效稳定的方式，能够极大降低人员成本，并消除人为主观因素的影响，且大时空指标与专家评审并不是对立的，可以使用大时空指标来辅助专家做出更加科学、客观的评审，

便于取得更好的结果，实现了从地理信息中量化出具有业务价值且强解释性的地理环境主题分数

2、线性回归模型具备良好的线性解释性和特征兼容性，拥有完备的数值化特征评估和推断机制，且能够简单描述特征叠加的增益，对评估线下门店大时空特征有天然的适配性，以及它不需要训练集和测试集的前提条件，以便更加快速的提升地理评估的工作效率。

附图说明

附图用来提供对本发明的进一步理解，并且构成说明书的一部分，与本发明的实施例一起用于解释本发明，并不构成对本发明的限制。在附图中：

图 1 是本发明评估方法的步骤流程图。

具体实施方式

以下结合附图对本发明的优选实施例进行说明，应当理解，此处所描述的优选实施例仅用于说明和解释本发明，并不用于限定本发明。

实施例：如图 1 所示，本发明提供一种技术方案，一种结合门店周围环境的相关指数对门店进行评估的方法，主要从地理信息中量化出具有业务价值且强解释性的地理环境主题分数，结合地理信息和相关自然特征为指标分数，对门店经营状况或经营需求的能力进行评估，为门店业务场景提供参考依据；

包括如下评估步骤：

S1、收集线下门店周边地理信息数据；

S2、对门店周边主体进行统计建模；

S3、合理提取主题特征指标 SSTI；

S4、线性回归模型评价体系；

S5、指标 SSTI 应用于现实业务场景。

基于上述技术方案，S1 中，线下门店周边地理信息数据在收集时主要通过使用不同数据源来得到目标门店周围的地理信息，该地理信息主要包括门店周围环境信息和相关自然特征；

获取地理信息后，将目标门店附近设施分为交通、居住、购物、美食、旅游、生活、工作和娱乐八个类别的主题，同时针对每一个主题，来得到目标门店附近该主题设施的数量，以及与目标门店的距离信息。

基于上述技术方案，S2 中，对门店周边主体进行统计建模主要通过建立主题强度模型和主题聚集度模型，并综合两个模型的结果，来生成该主题下的大时空指标；

主要采用某品牌数据为例，结合线性回归模型来验证大时空指标对客流量的影响大小，待验证通过，可将大时空指标应用于品牌的现实业务场景，助力提升品牌对门店周边环境的评估效率，最终作为门店选址等需求的重要参考。

基于上述技术方案，S2 中在具体统计建模时，门店地理位置环境主题分数受到门店周边主题兴趣点数量强度、聚集程度的影响，以门店周边 POI 为样本，对门店周边主题 POI 分布建立以下模型：

$$\text{Loc}\left(\text{POI}_t^{(s)}\right) \stackrel{\text{i.i.d.}}{\sim} \mathfrak{B}\left(p_t^{(s)}\right)^{\text{indep.}} \times \mathcal{N}\left(0, \Sigma_t^{(s)}\right)$$

$$p_t^{(s)} \sim \mathbb{B}(\alpha_t, \beta_t), \sigma_t^{(s)} \sim \mathcal{N}(\mu_t, \lambda_t)$$

其中:

s 为门店上标;

t 为主题下标, 包括交通、居住、购物、美食、旅游、生活、工作、娱乐八类;

$\text{Loc}(\text{POI}_t^{(s)})$ 为门店周边主题 POI 分布;

$\mathfrak{B}(p_t^{(s)})$ 为门店周边主题 POI 强度模型, 服从伯努利分布;

$p_t^{(s)}$ 为门店周边主题 POI 强度模型概率参数, 符合贝塔分布;

α_t, β_t 为强度参数分布模型变量;

$\mathcal{N}(0, \Sigma_t^{(s)})$ 为门店周边主题 POI 聚集模型, 服从二维正态分布;

$\Sigma_t^{(s)}$ 为门店周边主题 POI 聚集模型方差, 符合高斯分布, 并满足

$$\Sigma_t^{(s)} = \begin{bmatrix} \left(\sigma_t^{(s)}\right)^2 & 0 \\ 0 & \left(\sigma_t^{(s)}\right)^2 \end{bmatrix}, \left(\sigma_t^{(s)}\right)^2 \text{ 为 POI 经度或纬度的方差};$$

μ_t, λ_t 为聚集参数分布模型变量。

基于上述技术方案, 在 S2 中, 基于模型的计算流程包括输入、输出和计算过程, 具体如下:

输入: t , 评估主题, 包括交通、居住、购物、美食、旅游、生活、工作、娱乐共计八类;

m : 总评估门店数;

$n_t^{(s)}$: t 主题 s 家门店查询到的 POI 数量;

$N_t^{(s)}$: t 主题 s 家门店查询到的 POI 次数;

$p_t^{(s)}$: t 主题 s 家门店的 POI 强度模型初始概率参数;

$\sigma_t^{(s)}$: t主题s家门店所有 POI 与门店距离的标准差;

$(x_i^{(s)}, y_i^{(s)})$: s门店附近i POI 点相对于门店的坐标;

α, β : 门店间强度模型先验的服从贝塔分布, α, β 为模型变量;

μ, λ : 门店间聚集度模型先验的服从高斯分布, μ, λ 为模型变量;

输出: $SSTI_t^{(s)}$: s家门店t主题下的得分。

基于上述技术方案, S2 中计算过程如下:

1)、在每个t中执行;

2)、循环开始;

3)、强度模型中, 门店周边主题 POI 强度服从伯努利分布, 其概率参数服从贝塔分布, 强度模型的联合概率密度函数为:

$$\mathfrak{B}(\alpha, \beta, p) = \left(\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \right)^m \left(\prod_{s=1}^m (p_t^{(s)})^{\alpha - 1 + n_t^{(s)}} (1 - p_t^{(s)})^{\beta - 1 + N_t^{(s)} - n_t^{(s)}} \right);$$

4)、根据极大似然估计, 令 $\mathfrak{B}(\alpha, \beta, p)$ 关于p的偏导数为0, 求其 p为 p_t , 其中每家门店的强度模型概率参数为 $p_t^{(s)}$;

5)、聚集度模型中, 门店周边主题 POI 服从二维正态分布, POI 聚集模型方差符合高斯分布, 聚集度联合概率密度函数为:

$$\mathcal{N}(\lambda, \mu, \sigma) = \prod_{s=1}^m \left(\frac{1}{\lambda \sqrt{2\pi}} e^{-\frac{(\sigma_t^{(s)} - \mu)^2}{2\lambda^2}} \left(\prod_{i=1}^{n_t^{(s)}} \frac{1}{2\pi (\sigma_t^{(s)})^2} e^{-\frac{(x_i^{(s)})^2 + (y_i^{(s)})^2}{2(\sigma_t^{(s)})^2}} \right) \right);$$

- 6)、根据极大似然估计, 令 $\mathcal{N}(\lambda, \mu, \sigma)$ 关于 σ 的偏导数为 0, 求其 σ 为 σ_t , 其中每家门店的集中度模型标准差为 $\sigma_t^{(s)}$;
- 7)、得到每家门店 t 主题的主题强度模型概率参数为 $p_t^{(s)}$ 和集中度模型标准差为 $\sigma_t^{(s)}$;
- 8)、循环结束;
- 9)、门店信息、实时天气数据作为自变量, 客流数据作为因变量, 构建线性回归模型;
- 10)、 R^2 -模型决定系数;
- 11)、AIC-赤池信息量准则;
- 12)、BIC-贝叶斯信息准则;
- 13)、定义主题得分计算方式 $SSTI_t^{(s)} = \frac{p_t^{(s)}}{\sigma_t^{(s)}}$;
- 14)、在每个 t 中执行;
- 15)、循环开始;
- 16)、每家门店 t 主题的主题得分 $SSTI_t^{(s)}$;
- 17)、在线性回归模型加入主题得分特征 $SSTI_t^{(s)}$;
- 18)、得到加入主题得分后的 R_t^2 、 AIC_t 和 BIC_t ;
- 19)、循环结束;
- 20)、当各 t 中 R_t^2 效果无明显提升或 AIC_t 过大或 BIC_t 过大时;
- 21)、循环执行;
- 22)、重新定义主题得分计算方式, 执行 14 至 19 步;
- 23)、观察 R_t^2 、 AIC_t 、 BIC_t ;
- 24)、直到各 t 中 R_t^2 效果提升明显且 AIC_t 、 BIC_t 数值均合理;

25)、循环结束;

26)、保留对应的t的得分 $SSTI_t^{(s)}$;

27)、 $SSTI_t^{(s)}$ 可应用至品牌门店选址和分群等真实业务场景,为评估人员选择最优地理环境提供各维度主题评分参考依据。

基于上述技术方案, S3 中, 合理提取主题特征指标 SSTI 主要指利用最大似然估计来估计每家门店每个主题的 $\hat{p}_t^{(s)}$ 和 $\hat{\sigma}_t^{(s)}$;

具体的, $\hat{p}_t^{(s)}$ 越大, 门店周边主题 POI 数量强度越大, $\hat{\sigma}_t^{(s)}$ 越大, 门店周边主题 POI 聚集程度越高, 因此, 定义以下的 $SSTI_t^{(s)}$ 来描述门店主题下的 POI 分布:

$$SSTI_t^{(s)} = \frac{p_t^{(s)}}{\sigma_t^{(s)}}$$

其中: $\hat{p}_t^{(s)}$ 为由 MLE 估算出的s门店周边t主题的 POI 强度模型概率参数;

$\hat{\sigma}_t^{(s)}$ 为由 MLE 估算出的s门店周边t主题的 POI 聚集模型标准差, 门店 $SSTI_t^{(s)}$ 分数越高, 主题特征越明显。

基于上述技术方案, S4 中, 线性回归模型评价体系具体指通过选取一个可用于监测的且能够衡量门店经营状况或经营目标的指标, 如客流量、销量, 使用地理信息指标以及其他相关自然特征建立模型, 通过模型拟合度增益来体现地理信息指标的合理性;

其中, 线性回归模型具备良好的线性解释性和特征兼容性, 拥有完备的数值化特征评估和推断机制, 且能够简单描述特征叠加的增益, 对评估线下门店大时空特征有天然的适配性, 以及它不需要训练集和

测试集的前提条件，因此，由线性回归模型来评估自变量特征对因变量的影响大小及正负。

基于上述技术方案，在 S4 中，衡量模型拟合度的指标定为决定系数，表示在目标变量的总变异中回归关系所能解释的百分比，数值越接近于 1 则拟合的效果越佳，赤池信息量准则AIC和贝叶斯信息准则BIC，可权衡所估计模型的复杂度和模型拟合数据的优良性，数值越小模型效果越佳；

门店信息、实时天气及 8 类地理环境指标等数据作为自变量，客流数据作为因变量，构建线性回归模型，将 8 类地理环境特征分次单个加入评估模型后， R^2 有所提升，其中交通指数的效果最为明显，依次是娱乐指数和旅游指数，对目标变量-客流量变化的可解释性增强，各指标的AIC和BIC均有所下降，表示模型拟合良好，未出现过拟合现象。

基于上述技术方案，S5 中，指标 SSTI 应用于现实业务场景，主要结合以上检验结果，为品牌提供一定的门店选址及门店分群方面的业务指导和参考价值；

在具有其他相似条件下，新门店地址可首选参考交通便利处并具备一定的娱乐设施，如临近地铁口的商业综合体，也可考虑旅游胜地等场所，根据现实场景需求，门店大时空指标为商业用户量化出周边环境的客观数字依据，快速提升地理评估的工作效率。

最后应说明的是：以上所述仅为本发明的优选实例而已，并不用于限制本发明，尽管参照前述实施例对本发明进行了详细的说明，对

于本领域的技术人员来说，其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权利要求

1. 一种结合门店周围环境的相关指数对门店进行评估的方法，其特征在于：主要从地理信息中量化出具有业务价值且强解释性的地理环境主题分数，结合地理信息和相关自然特征为指标分数，对门店经营状况或经营需求的能力进行评估，为门店业务场景提供参考依据；

包括如下评估步骤：

S1、收集线下门店周边地理信息数据；

S2、对门店周边主体进行统计建模；

S3、合理提取主题特征指标 SSTI；

S4、线性回归模型评价体系；

S5、指标 SSTI 应用于现实业务场景。

2. 根据权利要求 1 所述的一种结合门店周围环境的相关指数对门店进行评估的方法，其特征在于：所述 S1 中，线下门店周边地理信息数据在收集时主要通过使用不同数据源来得到目标门店周围的地理信息，该地理信息主要包括门店周围环境信息和相关自然特征；

获取地理信息后，将目标门店附近设施分为交通、居住、购物、美食、旅游、生活、工作和娱乐八个类别的主题，同时针对每一个主题，来得到目标门店附近该主题设施的数量，以及与目标门店的距离信息。

3. 根据权利要求 1 所述的一种结合门店周围环境的相关指数对门店进行评估的方法，其特征在于：所述 S2 中，对门店周边主体进行统计建模主要通过建立主题强度模型和主题聚集度模型，并综合两个模型的结果，来生成该主题下的大时空指标；

主要结合线性回归模型来验证大时空指标对客流量的影响大小，待验证通过，可将大时空指标应用于品牌的现实业务场景，助力提升品牌对门店周边环境的评估效率，最终作为门店选址等需求的重要参考。

4. 根据权利要求 3 所述的一种结合门店周围环境的相关指数对门店进行评估的方法，其特征在于：所述 S2 中在具体统计建模时，门店地理位置环境主题分数受到门店周边主题兴趣点数量强度、聚集程度的影响，以门店周边 POI 为样本，对门店周边主题 POI 分布建立以下模型：

$$\text{Loc}\left(\text{POI}_t^{(s)}\right) \stackrel{\text{i.i.d.}}{\sim} \mathfrak{B}\left(\mathbf{p}_t^{(s)}\right) \times^{\text{indep.}} \mathcal{N}\left(0, \Sigma_t^{(s)}\right)$$

$$\mathbf{p}_t^{(s)} \sim \mathbb{B}\left(\alpha_t, \beta_t\right), \sigma_t^{(s)} \sim \mathcal{N}\left(\mu_t, \lambda_t\right)$$

其中：

s 为门店上标；

t 为主题下标，包括交通、居住、购物、美食、旅游、生活、工作、娱乐八类；

$\text{Loc}\left(\text{POI}_t^{(s)}\right)$ 为门店周边主题 POI 分布；

$\mathfrak{B}\left(\mathbf{p}_t^{(s)}\right)$ 为门店周边主题 POI 强度模型，服从伯努利分布；

$\mathbf{p}_t^{(s)}$ 为门店周边主题 POI 强度模型概率参数，符合贝塔分布；

α_t, β_t 为强度参数分布模型变量；

$\mathcal{N}\left(0, \Sigma_t^{(s)}\right)$ 为门店周边主题 POI 聚集模型，服从二维正态分布；

$\Sigma_t^{(s)}$ 为门店周边主题 POI 聚集模型方差，符合高斯分布，并满足

$$\Sigma_t^{(s)} = \begin{bmatrix} \left(\sigma_t^{(s)}\right)^2 & 0 \\ 0 & \left(\sigma_t^{(s)}\right)^2 \end{bmatrix}, \left(\sigma_t^{(s)}\right)^2 \text{ 为 POI 经度或纬度的方差};$$

μ_t, λ_t 为聚集参数分布模型变量。

5. 根据权利要求 3 所述的一种结合门店周围环境的相关指数对门店进行评估的方法，其特征在于：在 S2 中，基于模型的计算流程包括输入、输出和计算过程，具体如下：

输入： t ，评估主题，包括交通、居住、购物、美食、旅游、生活、工作、娱乐共计八类；

m ：总评估门店数；

$n_t^{(s)}$ ： t 主题 s 家门店查询到的 POI 数量；

$N_t^{(s)}$ ： t 主题 s 家门店查询到的 POI 次数；

$p_t^{(s)}$ ： t 主题 s 家门店的 POI 强度模型初始概率参数；

$\sigma_t^{(s)}$ ： t 主题 s 家门店所有 POI 与门店距离的标准差；

$(x_i^{(s)}, y_i^{(s)})$ ： s 门店附近 i POI 点相对于门店的坐标；

α, β ：门店间强度模型先验的服从贝塔分布， α, β 为模型变量；

μ, λ ：门店间聚集度模型先验的服从高斯分布， μ, λ 为模型变量；

输出： $SSTI_t^{(s)}$ ： s 家门店 t 主题下的得分。

6. 根据权利要求 5 所述的一种结合门店周围环境的相关指数对门店进行评估的方法，其特征在于：所述 S2 中计算过程如下：

1)、在每个 t 中执行；

2)、循环开始；

3)、强度模型中, 门店周边主题 POI 强度服从伯努利分布, 其概率参数服从贝塔分布, 强度模型的联合概率密度函数为:

$$\mathfrak{B}(\alpha, \beta, p) = \left(\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \right)^m \left(\prod_{s=1}^m \left(p_t^{(s)} \right)^{\alpha - 1 + n_t^{(s)}} \left(1 - p_t^{(s)} \right)^{\beta - 1 + N_t^{(s)} - n_t^{(s)}} \right);$$

4)、根据极大似然估计, 令 $\mathfrak{B}(\alpha, \beta, p)$ 关于 p 的偏导数为 0, 求其 p 为 p_t , 其中每家门店的强度模型概率参数为 $p_t^{(s)}$;

5)、聚集度模型中, 门店周边主题 POI 服从二维正态分布, POI 聚集模型方差符合高斯分布, 聚集度联合概率密度函数为:

$$\mathcal{N}(\lambda, \mu, \sigma) = \prod_{s=1}^m \left(\frac{1}{\lambda \sqrt{2\pi}} e^{-\frac{(\mu_t^{(s)} - \mu)^2}{2\lambda^2}} \left(\prod_{i=1}^{n_t^{(s)}} \frac{1}{2\pi (\sigma_t^{(s)})^2} e^{-\frac{(x_i^{(s)})^2 + (y_i^{(s)})^2}{2(\sigma_t^{(s)})^2}} \right) \right);$$

6)、根据极大似然估计, 令 $\mathcal{N}(\lambda, \mu, \sigma)$ 关于 σ 的偏导数为 0, 求其 σ 为 σ_t , 其中每家门店的集中度模型标准差为 $\sigma_t^{(s)}$;

7)、得到每家门店 t 主题的主题强度模型概率参数为 $p_t^{(s)}$ 和集中度模型标准差为 $\sigma_t^{(s)}$;

8)、循环结束;

9)、门店信息、实时天气数据作为自变量, 客流数据作为因变量, 构建线性回归模型;

10)、 R^2 ←模型决定系数;

11)、AIC←赤池信息量准则;

- 12)、BIC<-贝叶斯信息准则;
- 13)、定义主题得分计算方式 $SSTI_t^{(s)} = \frac{p_t^{(s)}}{\sigma_t^{(s)}}$;
- 14)、在每个t中执行;
- 15)、循环开始;
- 16)、每家门店t主题的主题得分 $SSTI_t^{(s)}$;
- 17)、在线性回归模型加入主题得分特征 $SSTI_t^{(s)}$;
- 18)、得到加入主题得分后的 R_t^2 、 AIC_t 和 BIC_t ;
- 19)、循环结束;
- 20)、当各t中 R_t^2 效果无明显提升或 AIC_t 过大或 BIC_t 过大时;
- 21)、循环执行;
- 22)、重新定义主题得分计算方式, 执行 14 至 19 步;
- 23)、观察 R_t^2 、 AIC_t 、 BIC_t ;
- 24)、直到各t中 R_t^2 效果提升明显且 AIC_t 、 BIC_t 数值均合理;
- 25)、循环结束;
- 26)、保留对应的t的得分 $SSTI_t^{(s)}$;
- 27)、 $SSTI_t^{(s)}$ 可应用至品牌门店选址和分群等真实业务场景, 为评估人员选择最优地理环境提供各维度主题评分参考依据。

7. 根据权利要求 1 所述的一种结合门店周围环境的相关指数对门店进行评估的方法, 其特征在于: 所述 S3 中, 合理提取主题特征指标 SSTI 主要指利用最大似然估计来估计每家门店每个主题的 $\hat{p}_t^{(s)}$ 和 $\hat{\sigma}_t^{(s)}$;

具体的, $\hat{p}_t^{(s)}$ 越大, 门店周边主题 POI 数量强度越大, $\hat{\sigma}_t^{(s)}$ 越大, 门店周边主题 POI 聚集程度越高, 因此, 定义以下的 $SSTI_t^{(s)}$ 来描述门店主题下的 POI 分布:

$$SSTI_t^{(s)} = \frac{p_t^{(s)}}{\sigma_t^{(s)}}$$

其中: $\hat{p}_t^{(s)}$ 为由 MLE 估算出的 s 门店周边 t 主题的 POI 强度模型概率参数;

$\hat{\sigma}_t^{(s)}$ 为由 MLE 估算出的 s 门店周边 t 主题的 POI 聚集模型标准差, 门店 $SSTI_t^{(s)}$ 分数越高, 主题特征越明显。

8. 根据权利要求 1 所述的一种结合门店周围环境的相关指数对门店进行评估的方法, 其特征在于: 所述 S4 中, 线性回归模型评价体系具体指通过选取一个可用于监测的且能够衡量门店经营状况或经营目标的指标, 再使用地理信息指标以及其他相关自然特征建立模型, 通过模型拟合度增益来体现地理信息指标的合理性;

其中, 线性回归模型具备良好的线性解释性和特征兼容性, 拥有完备的数值化特征评估和推断机制, 且能够简单描述特征叠加的增益, 对评估线下门店大时空特征有天然的适配性, 以及它不需要训练集和测试集的前提条件, 因此, 由线性回归模型来评估自变量特征对因变量的影响大小及正负。

9. 根据权利要求 8 所述的一种结合门店周围环境的相关指数对门店进行评估的方法, 其特征在于: 在 S4 中, 衡量模型拟合度的指标定为决定系数, 表示在目标变量的总变异中回归关系所能解释的百

分比，数值越接近于 1 则拟合的效果越佳，赤池信息量准则AIC和贝叶斯信息准则BIC，可权衡所估计模型的复杂度和模型拟合数据的优良性，数值越小模型效果越佳；

门店信息、实时天气及 8 类地理环境指标等数据作为自变量，客流数据作为因变量，构建线性回归模型，将 8 类地理环境特征分次单个加入评估模型后， R^2 有所提升，其中交通指数的效果最为明显，依次是娱乐指数和旅游指数，对目标变量-客流量变化的可解释性增强，各指标的AIC和BIC均有所下降，表示模型拟合良好，未出现过拟合现象。

10. 根据权利要求 1 所述的一种结合门店周围环境的相关指数对门店进行评估的方法，其特征在于：所述 S5 中，指标 SSTI 应用于现实业务场景，主要结合以上检验结果，为品牌提供一定的门店选址及门店分群方面的业务指导和参考价值；

在具有其他相似条件下，新门店地址可首选参考交通便利处并具备一定的娱乐设施，根据现实场景需求，门店大时空指标为商业用户量化出周边环境的客观数字依据，快速提升地理评估的工作效率。