



智行酒店排序模型实践

陈晓龙
携程旅行网

提纲

智行酒店排序的场景

SVM rank

Factorization Machine

Pointwise, Pairwise方法

智行酒店排序的场景：全城推荐、地标附近

下午1:42 北京 住 05-10 离 05-11 关键字/位置/... 地图 看过

位置区域 价格/星级 欢迎度排序 筛选

促销优惠 火车专享 机场车站 国贸CBD 中

北京燕山大酒店 高档型
4.7分 "餐厅很好, 离地铁近" 8700+人住过
距市中心10.3公里 中关村/五道口
¥720起 ¥750 半小时前有人预订
黄金会员减30元

维也纳国际酒店(北京首体店) 舒适型
5.0分 "前台热情, 离地铁近" 12000+人...
距市中心8.2公里 西直门/北京展览馆
¥500起 ¥534 半小时前有人预订
返现31元 返现 预订立减

飘HOME连锁酒店(北京王府井步行街店) 经济型
4.6分 "前台热情, 靠近商业区" 26000+...
距市中心1.3公里 天安门/王府井
¥327起 半小时前有人预订
精选推荐

海友酒店(北京前门店) 经济型
4.6分 "前台热情, 离地铁近" 3800+人住过
距市中心1.4公里 前门/崇文门
¥279起 ¥290 1小时前有人预订
黄金会员减11元

秋果酒店(北京华贸店) 高档型
5.0分 "设计不错, 位置好" 11000+人住过
距市中心7公里 国贸CBD

北京 住 05-10 离 05-11 北京爱琴... 地图 看过

4公里以内 价格/星级 欢迎度排序 筛选

促销优惠 火车专享 经济连锁 含早 免费取消

北京译夫宾馆 经济型
4.0分 558人住过
距北京爱琴海购物中心600米 国展中心
¥203起 ¥228 半小时前有人预订
多重减25元 会员专享 预订立减

智行金牌 贯通现代酒店(北京国展店) 舒适型
4.6分 "前台热情, 离地铁近" 2700+人住过
距北京爱琴海购物中心1公里 国展中心
¥389起 半小时前有人预订
精选推荐

色彩连锁酒店(北京国展二店) 经济型
4.7分 "前台热情, 离地铁近" 1200+人住过
距北京爱琴海购物中心700米 国展中心
¥246起 ¥269 仅剩1间
多重减23元 预订立减

北京黎明园宾馆 经济型
4.2分 "很安静, 老板热情" 618人住过
距北京爱琴海购物中心950米 国展中心
¥326起 ¥368 仅剩1间
多重减42元 会员专享 天天特价 预订立减

如家酒店(北京芍药居对外经贸大学店) 经济型
4.6分 "停车方便, 离地铁近" 1800+人住过
距北京爱琴海购物中心1.1公里 亚运村/奥...

Pointwise learning to rank

- 用分类的方法判断商品的相关度 (label: 购买/未购买、点击/未点击)
- 例如, 线性得分函数
- $\text{Score}(d, q) = \text{Score}(\alpha, \omega) = a\alpha + b\omega + c$
- 数据不平衡问题: 下采样

Pairwise learning to rank : SVM

- Pair定义：同一个搜索词下任意两个搜索结果
- 目标：将pair的排序关系分类正确
- 希望得到一个函数 f 使
 - $c_i > c_k$ iff $f(\psi_i) > f(\psi_k)$
- 其中， c_i 是第 i 个样本的相关度得分， $f(\psi_i) = \mathbf{w} \bullet \psi_i$
- 应用于智行酒店排序的地标搜场景

SVMranking的优缺点：

- 很好的平衡价格，酒店点评分，酒店设施的影响
- 可解释性强（配合样条）
- 交互特征学习能力弱，特征稀疏的样本学习能力差

The screenshot displays a mobile app interface for hotel searches. At the top, there are navigation icons and filters for location (北京), dates (05-10 to 05-11), and search terms (北京爱琴...). Below this, there are tabs for different hotel types: 促销优惠, 火车专享, 经济连锁, 含早, and 免费取消. The main content area shows a list of hotel results, each with a thumbnail image, hotel name, rating, price, and location details.

Hotel Name	Rating	Price (Start)	Location
北京译夫宾馆 经济型	4.0分	¥203起	距北京爱琴海购物中心600米 国展中心
贯通现代酒店(北京国展店) 舒适型	4.6分	¥389起	距北京爱琴海购物中心1公里 国展中心
色彩连锁酒店(北京国展二店) 经济型	4.7分	¥246起	距北京爱琴海购物中心700米 国展中心
北京黎明园宾馆 经济型	4.2分	¥326起	距北京爱琴海购物中心950米 国展中心
如家酒店(北京芍药居对外经贸大学店) 经济型	4.6分	-	距北京爱琴海购物中心1.1公里 亚运村/奥...

Factorization Machine

- FM, \mathbf{x} 非常稀疏

$$\hat{y}(\mathbf{x}) := w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j$$

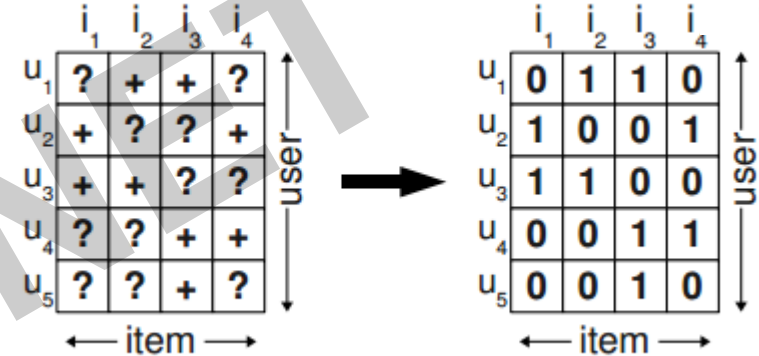
- 建模了二阶交互特征, \mathbf{v} 作为潜向量

Feature vector \mathbf{x}															Target y							
$\mathbf{x}^{(1)}$	1	0	0	...	1	0	0	0	...	0.3	0.3	0.3	0	...	13	0	0	0	0	...	5	$y^{(1)}$
$\mathbf{x}^{(2)}$	1	0	0	...	0	1	0	0	...	0.3	0.3	0.3	0	...	14	1	0	0	0	...	3	$y^{(2)}$
$\mathbf{x}^{(3)}$	1	0	0	...	0	0	1	0	...	0.3	0.3	0.3	0	...	16	0	1	0	0	...	1	$y^{(2)}$
$\mathbf{x}^{(4)}$	0	1	0	...	0	0	1	0	...	0	0	0.5	0.5	...	5	0	0	0	0	...	4	$y^{(3)}$
$\mathbf{x}^{(5)}$	0	1	0	...	0	0	0	1	...	0	0	0.5	0.5	...	8	0	0	1	0	...	5	$y^{(4)}$
$\mathbf{x}^{(6)}$	0	0	1	...	1	0	0	0	...	0.5	0	0.5	0	...	9	0	0	0	0	...	1	$y^{(5)}$
$\mathbf{x}^{(7)}$	0	0	1	...	0	0	1	0	...	0.5	0	0.5	0	...	12	1	0	0	0	...	5	$y^{(6)}$
	A	B	C	...	TI	NH	SW	ST	...	TI	NH	SW	ST	...	Time	TI	NH	SW	ST	...		
	User				Movie					Other Movies rated						Last Movie rated						

BPR : Bayesian Personalized Ranking

$$\text{BPR-OPT} := \ln p(\Theta | \mathcal{D}_S)$$

$$= \sum_{(u,i,j) \in \mathcal{D}_S} \ln \sigma(\hat{x}_{uij}) - \lambda_{\Theta} \|\Theta\|^2$$



说明：

$$\hat{x}_{uij} := \hat{x}_{ui} - \hat{x}_{uj}$$

把FM变成了 pairwise learn to rank

$$\sigma(x) := \frac{1}{1 + e^{-x}}$$

$$\mathcal{D}_S := \{(u, i, j) | i \in I_u^+ \wedge j \in I \setminus I_u^+\}$$

FM的优点：

- 实现个性化排序
- 能处理非常稀疏的特征向量

下午1:42 88

北京 住 05-10 离 05-11 关键字/位置/... 地图 看过

位置区域 价格/星级 欢迎度排序 筛选

促销优惠 火车专享 机场车站 国贸CBD 中

北京燕山大酒店 高档型
4.7分 "餐厅很好, 离地铁近" 8700+人住过
距市中心10.3公里 中关村/五道口
¥720起 ¥750 半小时前有人预订
您曾住过 黄金会员减30元

维也纳国际酒店(北京首体店) 舒适型
5.0分 "前台热情, 离地铁近" 12000+人...
距市中心8.2公里 西直门/北京展览馆
¥500起 ¥531 半小时前有人预订
返现31元 返现 预订立减

飘HOME连锁酒店(北京王府井步行街店) 经济型
4.6分 "前台热情, 靠近商业区" 26000+...
距市中心1.3公里 天安门/王府井
¥327起 半小时前有人预订
智行金牌 本月人气酒店 精选推荐

海友酒店(北京前门店) 经济型
4.6分 "前台热情, 离地铁近" 3800+人住过
距市中心1.4公里 前门/崇文门
¥279起 ¥290 1小时前有人预订
黄金会员减11元

秋果酒店(北京华贸店) 高档型
5.0分 "设计不错, 位置好" 11000+人住过
距市中心7公里 国贸CBD
500

总结

- Ranking SVM作为酒店以POI为搜索条件的模型
- FM-BPR 作为酒店没有指定POI的城市默认排序模型

Q&A

ITPUB.NET