### 高潜用户购买意向预测

**一 小组成员：**

孟玉立 2120161024

谢瑄 2120161065

王龙 2120161054

王欣欣 2120161059

商军英 2120161032

1. **问题描述**

高潜用户购买意向预测问题，首先需要解释的是高潜用户这一概念。所谓高潜用户即在用户群体中有更多可能性下单购买商品的用户，比如用户经常浏览购物网站，用户在购物网站的用户等级比较高的用户。然后从这些用户中预测在给定商品中其可能购买的商品。

网购已经成为现今购物一大主要方式，购物平台一般都积累了海量的真实数据。如何从历史数据中找出规律，去预测用户未来的购买需求，让最合适的商品遇见最需要的人，是大数据应用在精准营销中的关键问题。以真实的用户、商品和行为数据（脱敏后）为基础，构建用户购买商品的预测模型，输出高潜用户和目标商品的匹配结果，为精准营销提供高质量的目标群体。  
三、**数据集**

从京东的算法大赛上获得数据集，然后对数据集进行处理。

数据集的属性：用户编号、商品编号、行为时间、点击模块编号、type类型（6种）、品类ID、品牌ID。

1. **实现步骤**

1、主要思路

1）对于历史记录的用户，绘画出历史购买轨迹，用已购买中评价最高的数据为预测数据

2）对于没有历史记录的用户，用用户数据表中的age、sex、user\_lv\_cd、user\_reg\_tm属性来聚类分析结果。

2、主要特征

第一层特征：使用cate, 行为数据表中的type使用权重处理一下，考虑用户数据的age、sex、user\_lv\_cd、user\_reg\_tm等

第二层特征：在第一层特征基础上加brand，调type的权重

第三层特征：1）在第二层上基础上考虑product

2）考虑comment\_num,bad\_comment\_rate

3）给1）和2）加权整理一下

整体上是线性回归神经网络算法

3、第1层主要步骤

第一步：考虑cate,用权重处理一下type

第二步：根据用户的age、sex等特征聚类出用户群体以及对应的cate最优值。

第三步：加权，0.4x[0.1,0.2,0.3]+0.6X[0.1,0.2,0.3]

1. 第2层主要步骤

第一步：考虑brand,用权重处理一下type

第二步：根据用户的age、sex等特征聚类出用户群体以及对应的brand最优值。

第三步：加权，0.4x[0.1,0.2,0.3]+0.6X[0.1,0.2,0.3]

5、第3层主要步骤

第一步：考虑product,用权重处理一下type

第二步：根据用户的age、sex等特征聚类出用户群体以及对应的product最优值。

第三步：加权，0.4x[0.1,0.2,0.3]+0.6X[0.1,0.2,0.3]

**五、组员分工**

数据收集：王龙

数据处理：谢瑄

数据分析：王欣欣

算法实现：孟玉立

项目总结：商军英