

客户推荐系统说明文档

| 文档作者 | 联系方式 | 版本 | 日期 |
|------|---|------|--------------------|
| 刘凌海 | Liulh@fxiaoke.com 13726288907 | V0.1 | 2016 年 3 月 29 日星期二 |

文档目录

- 客户推荐系统说明文档.....1
- 1. 核心内容.....3
- 2. 前言.....3
- 3. 市场调研.....4
 - 3.1. Everstring.....4
- 4. 我们的客户推荐系统.....5
 - 4.1. 应用场景.....5
 - 4.2. 实现路径.....6
 - 4.3. 可行性分析.....7
 - 4.4. 算法.....7
 - 4.5. 数据来源.....8
- 5. 挑战.....8

1.核心内容

推荐系统可以带来：

- 推荐高质量客户资源
- 提供辅助销售信息
- 提高效率，降低成本

推荐系统使用的技术：

- 分类、聚类、模式匹配算法
- 有可能会用到深度学习将多源异构信息整合

推荐系统挑战：

- 模型设计
- 数据来源以及清理

2.前言

中国大概有四千万家中小型企业，这对于 2B 企业来说是一个很大的市场，但是如果以最大效率去扩大市场份额以及推销产品是一个很大的问题。从一线销售人员的角度来说，有如下几个问题需要回答：

- 需要把产品卖给哪家公司？有哪些潜在的买家？
- 那些买家最有可能成交？
- 谁更适合去联系这个买家？
- 如何制定相应的销售与推广策略？

推荐系统就是为了解决这些问题而生，以帮助企业客户更高效的联系到高质量销售目标，同时合理安排人员，进行精准推广，降低运营成本。

3.市场调研

3.1. Everstring

EverString 是北美的一家初创企业，其通过全网用户大数据分析来为客户建立用户分析模型进而提供智能用户分析和推荐服务。Everstring 主要通过大数据和人工智能学习客户的特征。根据内外部数据建立模型，这个模型的本质就是用来帮企业客户针对用户进行画像的，判断什么样的用户转化率比较高，从而给出一个打分。根据这个打分，企业就可以优先选择分数高的用户进行营销，这样会比之前的平均效率提高 3-4 倍，加速了企业的销售流程，节省了营销费用。以下是其服务架构。

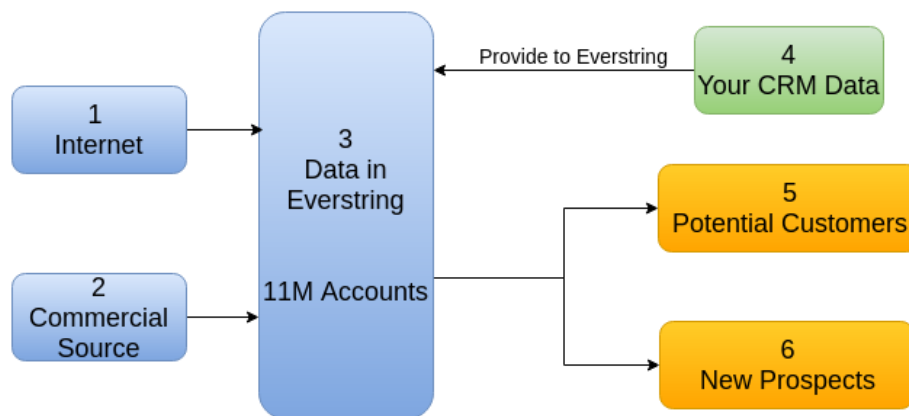


Figure 1 Everstring 的服务架构

1)和2)是 Everstring 收集数据的过程,通过全网数据爬取和商业数据来源,现在 Everstring 大概有 11M 企业数据。通过对这些数据建模挖掘,形成结构化的企业画像,以待推荐。4)是当 Everstring 的客户寻求其服务时,需要提供其自身的 CRM 等数据给 Everstring, Everstring 通过特定算法和匹配,在 5)和 6)中返回高质量的潜在客户,同时带来更多的机会。

目前, Everstring 的用户包括英特尔 McAfee 和 Salesforce RelateIQ 等世界 500 强企业。Everstring 向他们的企业客户收取每年 5 万美金到 300 万美金不等的年费。

4.我们的客户推荐系统

我们的客户推荐系统与 Everstring 的思路比较类似，然而它不仅为我们自己带来市场占有率和效率上的提高，还会部署到我们的分享平台上，为客户服务，届时将成为最为重要的一个增值功能之一。下面将详细介绍推荐系统的应用场景、实现系统的可能性、实现路径、具体算法与数据情况。

4.1. 应用场景

下图是推荐系统的应用场景之一。系统从 CRM，internet 和 government 处获取企业画像信息，首先对企业内部客户数据进行画像，然后将 internet 和 government 处获取的数据进行画像与内部数据进行匹配。将高分潜在客户名单以及相应数据返回。另外，对销售的过往客户进行画像，也就是对销售画像。然后将销售与潜在客户匹配，进一步提高成功率。

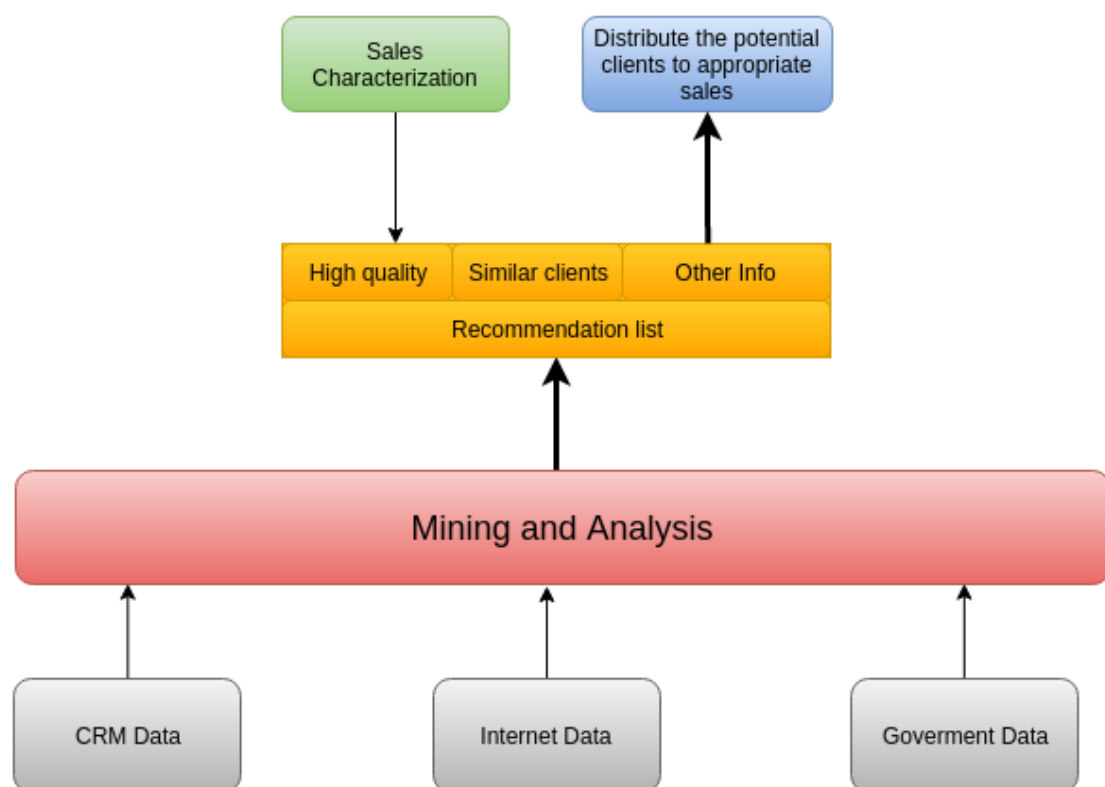


Figure 2 客户推荐系统应用场景一

下图是客户推荐系统的应用之二。系统从 CRM，internet 和 government 处获取企业画像信息，当销售人员获取到新的客户资源时，输入系统查询，系统返回客户的评分（评价评价

客户成为买家的可能性)，以及辅助销售信息（类似已成交的企业、相应销售销售人员以及相关行业信息），销售可以根据返回的信息制定销售策略。

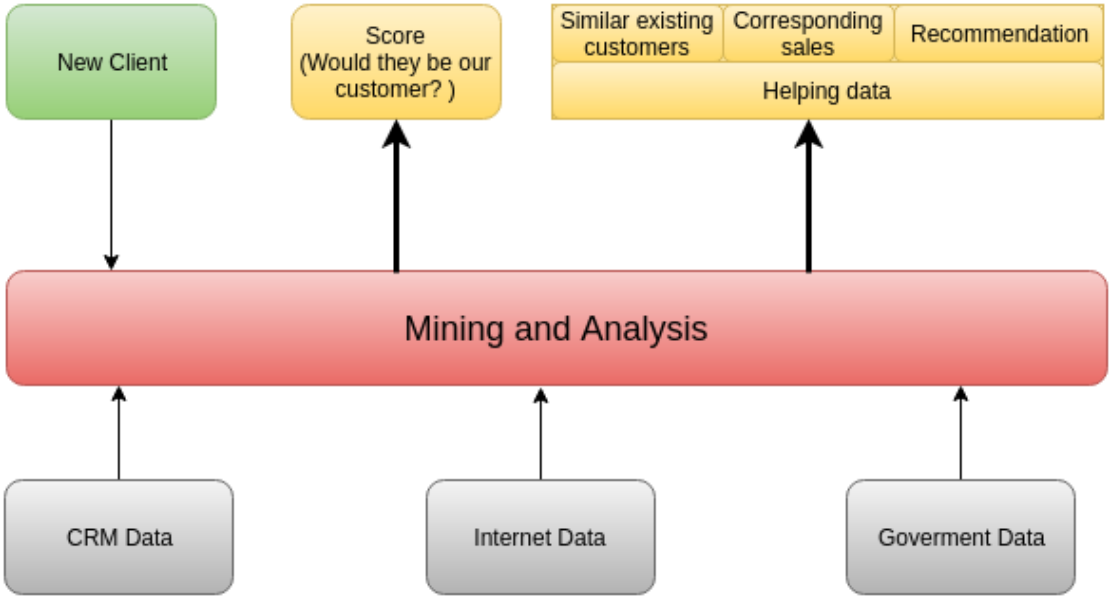


Figure 3 客户推荐系统应用场景二

4.2. 实现路径

推荐系统不是一个简单的系统，其中涉及算法和模型设计、数据采集和工程实现等。



需要将大的目标拆分成可以逐步迭代提高的阶段目标。

第一步： 利用现在客户在 CRM 中的 feature 对其进行画像和聚类/分类分析，将客户中的 CRM 数据（也就是客户的客户）作为推荐对象，挖掘高质量客户推荐给销售，同时提供辅助信息（例如签单的同类型企业，签单的销售）。

第二步：在实现第一步的同时，会编写爬虫爬取互联网上的企业信息，同时确定内部外部数据统一 **feature**。将 CRM 数据之外的高质量企业推荐给销售。

第三步：将成熟的经过验证的推荐系统部署到产品上。也就是说，购买我们产品的客户，都有一个推荐系统为其推荐潜在高质量客户。这将成为最重要的增值功能之一。

4.3. 可行性分析

- 第一步：目前平台上的付费客户大概有 7000 家，使用客户有 20000 家，客户的 CRM 数据综合将比这个要大一到两个数量级。因此第一步是可以实现的。
- 第二步：其中如何获取全网信息将是难点，但是以目前的爬虫，NLP 设计和异构数据机器学习算法的情况，也是可以实现的。
- 第三步：将相同的算法应用到客户的 CRM 数据中，为客户带来转化率的提高。这是我们平台的优势。
- 从算法的角度来说，目前聚类分析和推荐算法等都比较成熟，需要根据具体情况做相应调整，以获得最高的准确率。从算法上来说是可行的。

4.4. 算法

第一是手动设计 **feature**，第二是利用 **deep learning** 来将原始数据转化为更有效的 **feature**。

下图是利用手动 **feature** 来对企业进行画像、分类、聚类。将外部数据内部数据匹配，获取高质量潜在客户进行推荐。

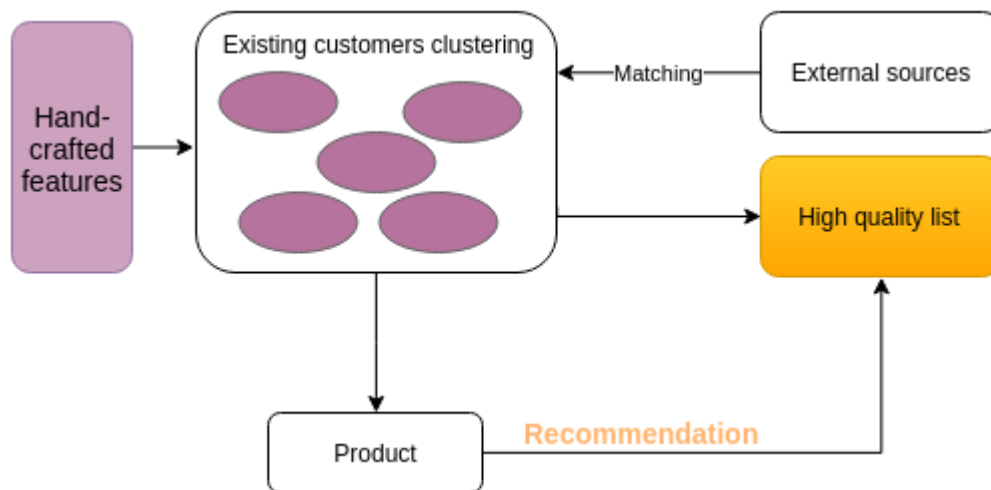


Figure 4 推荐算法图示

深度学习算法与上面结构类似，但是采用深度学习作为特征提取器。下图是一个深度神经网络的示意图。

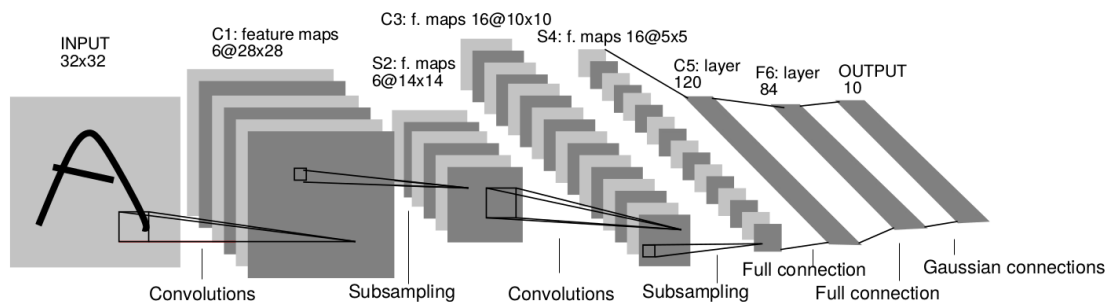


Figure 5 名为 LeNet 的深度学习神经网络图示

利用神经网络，可以将数字、图像和文本异构信息结合起来，提取更有效的分类聚类 features。

4.5. 数据来源

- 平台的 CRM 数据是非常重要的一部分。
- CRM 之外的企业数据可以从互联网爬取以及从工商局或者其他第三方机构购买。

5. 挑战

挑战主要在于模型的设计，数据来源。这两点需要结合具体情况深入研究。

- Everstring 采用了 automatic feature selection and model building，这对多源异构数据的模型设计比较重要。我们的数据也分为内部数据和外部数据，如何统一两者 feature 以及

处理 feature 缺失的问题是最重要的一個难点。

- 除了 CRM 数据之外，从互联网获取有效的企业数据也是非常重要的一点。互联网企业数据都是非结构化的，如何提取结构化的信息，转化为可用的特征需要更多的工作。