会议纪要

**时间**：2016年5月19日15:30至17:30

**地点**：广东电信大数据运营中心研发会议室

**与会人员**：戴志成，金圣，张宇宁（广东电信企业信息化事业部）

梁锐，叶宗刚，马昭（天睿信科技术广州分公司）

**记录整理**：马昭

**会议提要**：进场交流

**会议事项**：

1. 戴志成介绍（PPT）

(1)广东电信数据存储平台：ODS,EDW------>Hadoop(处理)------->GP(对外：脱敏)

(2)广东电信挖掘平台数据字典：移动用户、宽带用户、固话用户、ITV用户、用户行为、移动用户APP、移动宽带上网网站类型、固话移动用户

1. 梁锐介绍（PPT）
2. 大数据机器学习基本流程
3. 项目背景
4. 机器学习平台的应用场景
5. 自动化模型构建流程：整体流程，自动化建模优势
6. 问题讨论与交流

广东电信关心的问题：

1. 机器学习算法是否新颖
2. 是否符合电信情况:基础表字段形式复杂等，如通话清单没有分辨能力
3. 模型输入层的数据：底层数据，以及适用性
4. 如何实现与运营
5. 是否有成功事例及可可应用场景
6. R的单机性能是否满足大数据建模要求
7. 程序运行的平台：Green Plum+TD一体机 (+ASTER)
8. 将模型模块化，提供web接口（JAVA）

Web 接口 (JAVA)

R

系统：open suse

权限

内部网

TD

GP

1. 分公司人员能否掌握操作方法
2. 此模型区别于普通模型的优势
3. SAS,SPSS已有自动挖掘功能：但收费
4. 模型的维护周期及方式

广东电信的愿景：APP Store

1. WEB前台的操作界面及前端设计
2. 模型应用——以APP应用方式呈现
3. 将模型产品化，UI界面设计
4. 附件：详细操作指导