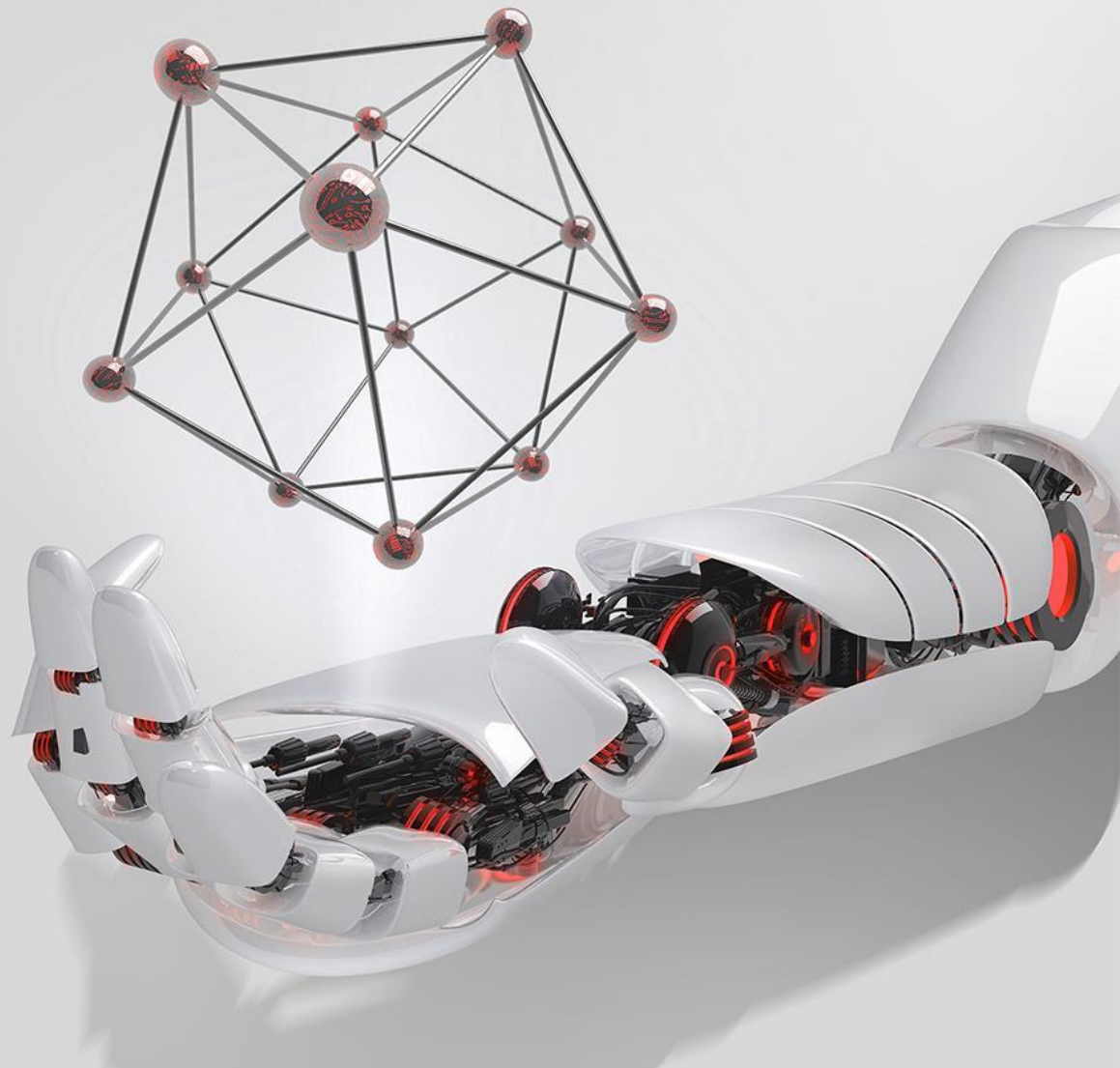
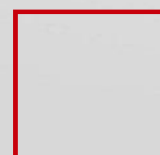


# 大数据创新训练营方案

部门：人才生态发展部

日期：2019.6



Security Level:



# 大数据创新训练营概况

创新训练营共历时5天，含理论精讲、3个关卡的闯关赛和创新优化实践；旨在让学生短时间内建立对大数据的综合认知并灵活运用取得分析成果。通过提前预习、集中讲解和实训指导，学生不仅能掌握大数据的采集、预处理、存储、分析挖掘等理论知识，并且能通过实例进行演练并可视化展示。



# 目录

---

1. 阶段一：理论精讲+闯关赛
2. 阶段二：创新实践

# 阶段一 理论集训营+闯关赛

课前	4天	
《大数据技术基础》在线视频	线上预习	1天
《Python基础》在线视频	线上预习	1天
《大数据分析与挖掘》在线视频	线上预习	2天
正式赛段	4天	
《数据处理流程概览及KNIME》课程	理论精讲	0.75天
关卡1: 大数据分析挖掘思维实践: KNIME平台实现	闯关赛	0.25天
《大数据分析与挖掘进阶》课程	理论精讲	1天
关卡2: 信用风控领域的大数据综合实训	闯关赛	1天
《大数据可视化实现: DLV平台》课程	理论精讲	0.5天
关卡3: 大数据大屏可视化实战	闯关赛	0.5天

说明: 每关卡后设置抽组成果发表环节  
考核: 配套《综合评分表》, 以完成度和时间为依据。



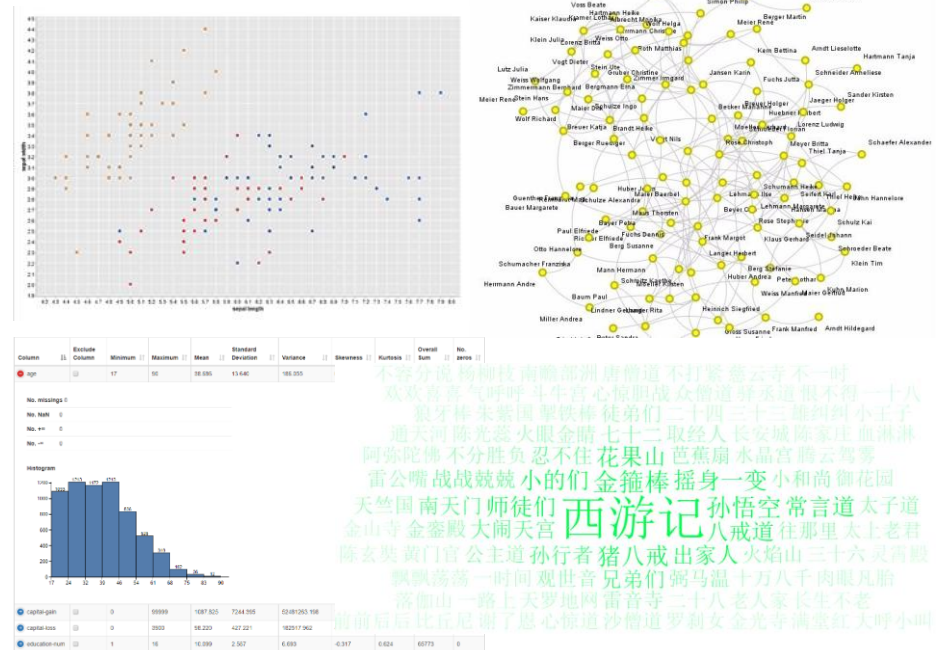
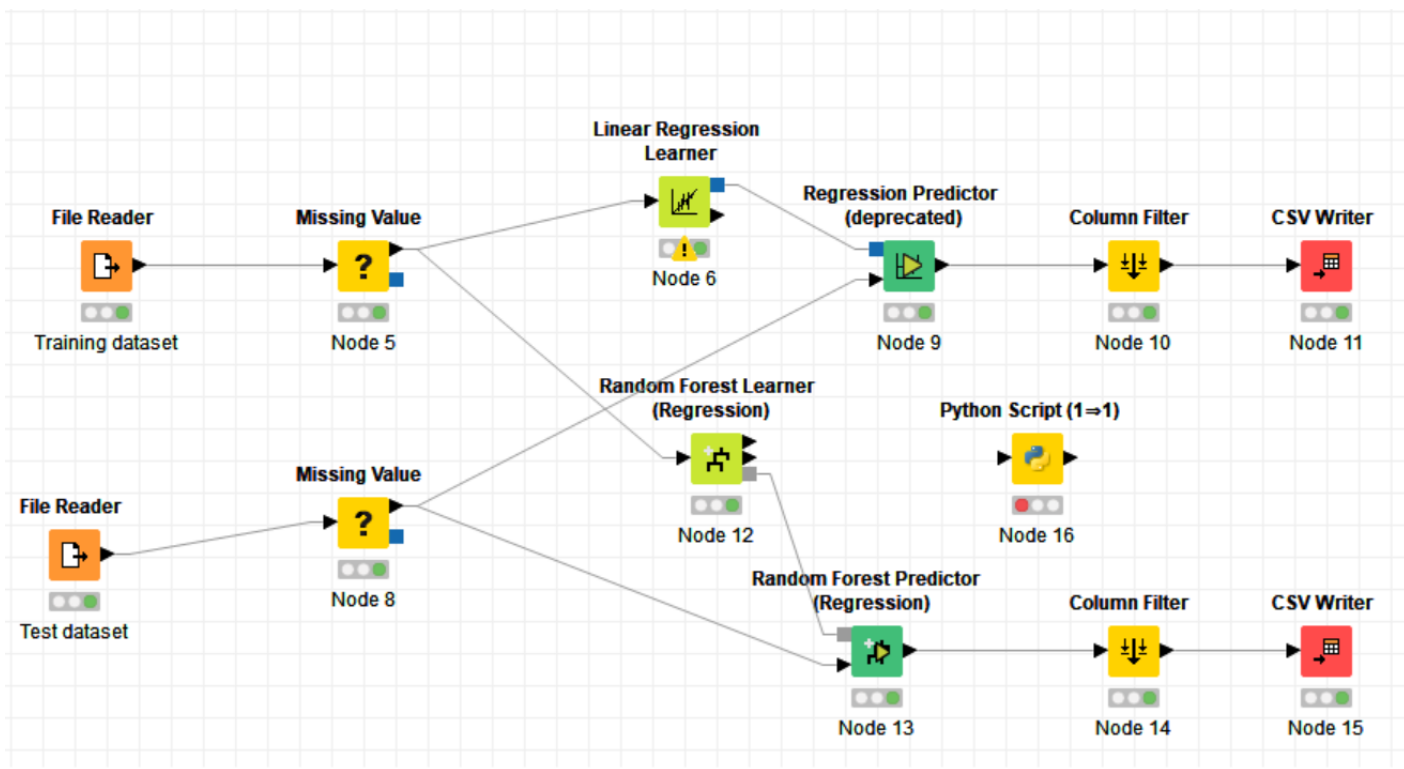
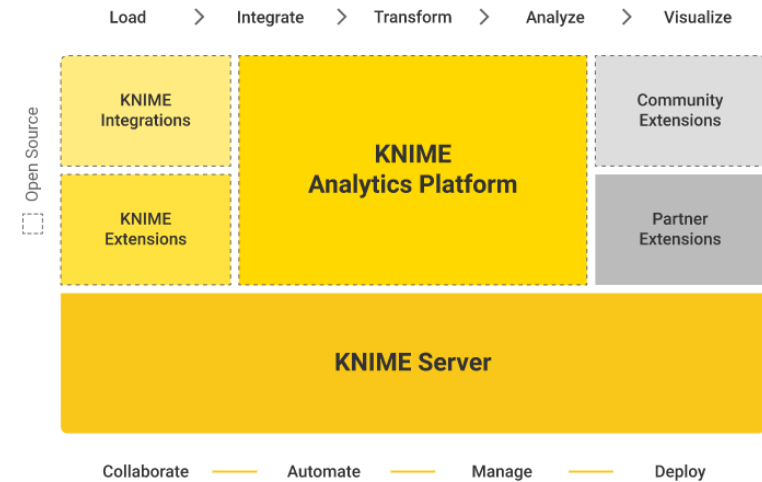


# 关卡1：KNIME平台数据分析挖掘思维实践

旨在熟悉数据分析与挖掘的端到端流程，掌握数据分析的关键步骤。

评估点：

- 1- 完整的KNIME流程结果
- 2- XMind思维导图，梳理完整的数据分析挖掘的流程



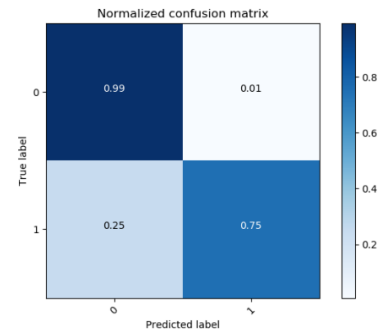
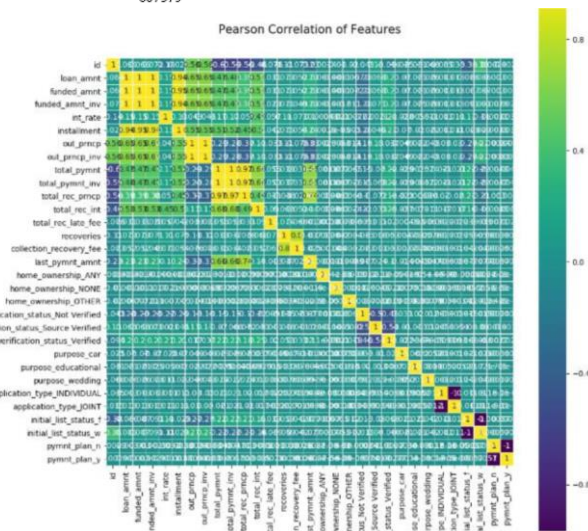
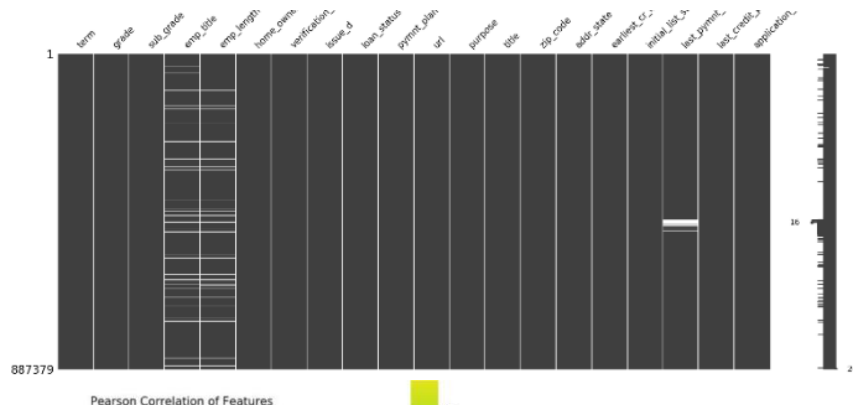
# 关卡2：信用风控领域的大数据综合实训

## 综合实训介绍：

互联网金融90%的业务是信贷，如何在无抵押的情况下甄选高信用的用户和提前预测可能发生的违约风险，成为互联网公司的核心竞争力。

本实训采用Python对某P2P平台数据进行分析和数据预处理，并进行相关的属性挑选，结合表征信息、行为信息、状态信息和利益信息等多个维度的信息进行机器学习、建立数学模型，从而构造一个客户行为信用评级模型，对客户贷款是否会违约进行预测，最终决定是否放款。

通过本实训，学员可掌握特征选择等数据分析预处理技能，运用有监督学习算法进行建模和评估，体验分布式数据库和大数据内存计算引擎的使用。同时掌握Python/Scala等多种语言，以及相关开发工具的使用，是一个强调全知识面、贴近真实案例的综合实训。



PCA SVM

# 关卡3： 大数据大屏可视化实战

认证全景



 <p><b>适合人群</b></p> <p>对实时流计算和可视化感兴趣的从业人员，社会大众和高校师生</p>	 <p><b>培训方案</b></p> <p>结合华为云服务搭建基于流计算的可视化平台</p>	 <p><b>技术能力</b></p> <p>了解流计算的关键技术，掌握华为云基于流计算的可视化解决方案</p>	 <p><b>认证价值</b></p> <p>掌握基于流计算的可视化平台搭建，实时展现业务成果，帮助企业办公效率的快速提升</p>
--	---	---	--



# 示例:



实时体验:

<https://console.huaweicloud.com/dlvapp/screen.html?share=3956f355cda940e5a579acfada4b9f53&locale=zh-cn&region=cn-north-1>



# 目录

---

1. 阶段一：理论精讲+闯关赛
2. 阶段二：创新实践

# 阶段二 创新实践 + 成果答辩

## 目标:

通过1天的创新实践活动，学员在闯关赛的基础上优化反欺诈实训模型，或自主选择开源数据集开发创新应用。

阶段二	1天	
创新实践	实验操作	0.75天
成果展示	答辩	0.25天
结班仪式	证书颁发	

## 考核形式:

高校&华为联合组织创新实践成果展示。

## 评价维度:

- 1- 闯关赛总分
- 2- 创新性
- 3- 可推广性
- 4- 呈现效果



# 参考资料：辅导材料

## 参考材料：

1-大数据技术基础 (HCIA-Big Data):

<https://ilearningx.huawei.com/portal/#/courses/course-v1:HuaweiX+EBGTC00000163+2018.7/about>

2-《Python基础和进阶》：<https://ilearningx.huawei.com/portal/#/courses?q=python>

3-《Python--numpy和pandas库》：<https://www.bilibili.com/video/av16378934/?p=1>

4-大数据分析挖掘 (HCIE-Big Data Mining):7月底推出、后续更新

## 拓展链接：

KNIME大数据Spark组件

<https://www.knime.com/nodeguide/big-data>

<https://www.knime.com/knime-extension-for-apache-spark>

华为云数据可视化DLV服务

<https://www.huaweicloud.com/product/dlv.html>

华为云ModelArts AI/大数据开发平台

<https://www.huaweicloud.com/product/mrs.html>

<https://www.huaweicloud.com/product/modelarts.html>

# 参考资料：华为e学云免费在线视频

序号	课程名称	课程描述
1	<a href="#">融以致远-带你走进大数据</a>	从“大数据是什么”开始，到华为大数据解决方案介绍，接着分享华为大数据的应用案例。最后，讲解大数据技术的学习认证指南。
2	<a href="#">HCIA-Big Data 华为认证大数据工程师V2.0</a>	本课程主要介绍了大数据技术前沿趋势，业界常用且重要大数据组件技术原理，华为大数据 FusionInsight HD解决方案以及大数据组件的基础操作和应用场景综合实践。
3	<a href="#">Fusion Insight二次开发--实时检索</a>	通过公安业务实时检索的场景，介绍了实时检索场景中涉及的Kafka、SparkStreaming、ES和HBase等组件以及各组件的使用方法。
4	<a href="#">Fusion Insight二次开发--实时流处理</a>	通过电商业务实时流处理的场景，介绍实时流处理中涉及的Kafka、Flink和Redis等组件以及各组件的使用方法。
5	<a href="#">Fusion Insight二次开发--交互查询</a>	通过公安系统分析场景介绍了交互查询的开发思路。大致介绍了交互查询所需要的组件以及各组件的使用方法。
6	<a href="#">Fusion Insight二次开发--融合数仓</a>	通过电商数据存储场景，介绍了融合数仓中涉及的Gauss DB 200组件以及组件的使用方法。
7	<a href="#">Fusion Insight二次开发--离线批处理</a>	通过电商推广离线处理的场景，介绍了离线批处理场景中涉及的HDFS、Spark和SparkSQL等组件以及各组件的使用方法。
8	<a href="#">洞见数据的秘密：金融大数据采集实战</a>	本课程是金融大数据采集训战课程的先导课，旨在利用“120分钟”的在线学习，帮助目标学员全面了解： <ul style="list-style-type: none"> <li>• 大数据概念；</li> <li>• 大数据技术生态；</li> <li>• 生活中哪些事情是大数据技术在发挥作用；</li> <li>• 金融业务的大数据技术应用。</li> </ul>





# 参考资料：行业案例

## 华为企业业务行业成功案例

<https://e.huawei.com/cn/case-studies>



华为助力许昌市政府，  
打造河南政务云中心标  
杆



有一种邂逅叫：华为SAP  
HANA遇上“冷酸灵”



华为助力光大银行构建  
全国分行二级资源池



基因说：最懂我的人，  
原来是你！

# Thank you.

把数字世界带入每个人、每个家庭、  
每个组织，构建万物互联的智能世界。

Bring digital to every person, home, and  
organization for a fully connected,  
intelligent world.

**Copyright©2019 Huawei Technologies Co., Ltd.  
All Rights Reserved.**

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

