



谢虹宇的简历

(' - ') 干杯~

Contact. 联系方式

☎ 电话: 156-0051-0487; 156-01042995

✉ 邮箱: 1990xhy@gmail.com

☎ 微信: xhy1990xhy

👤 QQ: 489632554

Application. 应聘岗位

机器学习工程师【优先】
Java Web研发工程师

我是一个对黑科技由衷热爱、甚至沉迷前沿技术的幽默感十足的软件工程师。从事前沿科技相关的工作能让我热血沸腾，干劲十足。

我目前正在寻找机器学习工程师岗位的工作机会，希望借此为贵司献上我的一点绵薄之力，快招我到碗里来吧！

Basic info. 基本信息

个人信息: 谢虹宇 / 女

毕业院校: 东北大学(2010-2014)

院校性质: 教育部直属国家首批985、211重点建设院校

工作年限: 5年+【实际4年+(含大四期间实习),多出来的是加班加出来的,你懂得】

业余爱好: 科幻电影,推理电影,Geek盛会,球类运动

Linkedin: www.linkedin.com

CSDN博客: blog.csdn.net/loverain90

GitHub: github.com/hongyuXie Star 23

Skill. 技能清单

总览



Experience. 项目与工作经验

个人项目

机器学习相关【开始】

个人博客-AI【目前仅有：机器学习MachineLearning】

源代码

Follow @hongyuXie 8

简要介绍：机器学习是一门研究在非特定编程条件下让计算机采取行动的学科。最近二十年，机器学习为我们带来了自动驾驶汽车、实用的语音识别、高效的网络搜索，让我们对人类基因的解读能力大大提高。当今机器学习技术已经非常普遍，您很可能在毫无察觉情况下每天使用几十次。本项目涵盖机器学习、数据挖掘和统计模式识别。相关主题包括：(i) 监督式学习（参数和非参数算法、支持向量机、核函数和神经网络）。(ii) 无监督学习（集群、降维、推荐系统和深度学习）。(iii) 机器学习实例（偏见/方差理论；机器学习和AI领域的创新）。在不同领域应用学习算法，例如智能机器人（感知和控制）、文本理解（网络搜索和垃圾邮件过滤）、计算机视觉、医学信息学、音频、数据库挖掘等领域。

资格证书：



涵盖内容:

- Linear Regression 线性回归
- Logistic Regression 逻辑回归
- Neural Networks 神经网络
- Support Vector Machines 支持向量机
- K-means Clustering K-means聚类算法
- Principal Components 主成分分析
- Analysis 分析
- Anomaly Detection 异常检测
- Collaborative Filtering 协同过滤
- Object Recognition 模式/目标识别
- Matlab 数学分析软件Matlab使用
- Octave 数学分析软件Octave使用

最初的最初：是个人对人工智能领域的热爱；热爱各类科幻电影，并认为只有想不到的，没有做不到的；科幻最终一定会走出银幕，成为现实。

后来的后来：与Andrew Ng的清华大学的一面之缘，从此埋下了一颗种子；后来，基于个人对🔗[Andrew Ng](#)的崇拜,当年还没毕业，参加百度在清华大学的招聘宣讲和笔试，和Andrew Ng有过一面之缘，从此特别关注人工智能领域。无意中，发现Andrew Ng原来曾创建过MOOC的平台，从此开始了自学机器学习的漫漫长路。

现在的现在：经过一年的坚持，终于修完Andrew Ng在Coursera的课程。

以后的以后：继续跟着Andrew Ng学习，Ng从百度离职以后，创办了🔗[deeplearning.ai](#)致力于让全世界的人都能通过自学课程，成为AI领域的专家。

深度學習工程師

deeplearning.ai

已開課

01. 神經網絡和深度學習

開始學習

02. 改善深層神經網絡：超參數調試、正則化以及優化

開始學習

03. 結構化機器學習項目

開始學習

04. 卷積神經網絡

開課時間待定

05. 序列模型

開課時間待定

机器学习相关【结束】

其他项目【开始】

个人简历

GitHub项目地址: [MyResume_CN.git](#)

GitHub访问地址: [MyResume_CN](#)

介绍：这是个人的中文在线简历，为了实时更新个人的经历，一劳永逸，特别将此简历做成在线的。后期会在github上新开一个分支，做中英文简历。欢迎star。

启明星辰（2015.11 - 至今）

机器学习相关【开始】

产品-日志自动范式化系统

介绍：Syslog是一种被广泛应用于日志系统的日志描述格式，SOC作为日志采集和分析系统，需要从各种不同的系统和设备上采集并分析日志。但大多数系统并没有完全遵守syslog的格式标准，并且具有各种自身系统所定义的日志格式。这使得SOC在对采集的日志进行解析和抽取所需要的字段（范式化）时，无法做到对所有日志格式的兼容。目前的解决方案是采用人工编写正则表达式的方法来对不同系统和设备的日志格式进行适配。该方法存在有一定的弊端，当用户环境中的产生日志的设备因为版本升级日志格式发生调整或者增加了新的未识别过的设备后，都会使得SOC当前的正则表达式对日志的解析功能失效，这时就需要前线技术人员现场增加可以解析新的日志格式的正则表达式，但是，由于前线技术人员个人技术的差别加上对SOC所定义的字段和用户日志的理解的不同，会使得写出的正则表达式错误较多且解析不准确，因此需要设计一种能够在不需要人工干预或者少量人工干预的情况下，完成对未识别过的日志格式进行自动解析和识别的方法，来增加SOC对不同日志的兼容性和适配能力。

负责：

- (0) 机器学习的研究
- (1) 方案预研究

机器学习相关【结束】

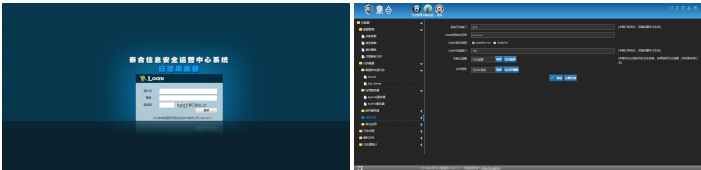
Java Web相关【开始】

产品研发-日志采集系统

介绍：针对不同的应用通过授权，能够采用不同的采集方式，同时，对已部署的系统，应该能够不断的添加新开发出来的采集方式，因此插件模式也就应运而生。在采集器(模块)中，每一种采集方式都被定义为一个插件，这个插件在自己的代码控件中实现本身特有的采集功能，并自己控制插件配置的保存和读取。

负责：各采集插件服务的研发工作，目前实现的主要的采集插件有:apache、iis、jdbc(区分不同数据库类型，主要有oracle;sqlserver;mysql;db2等)、mysql、oracle、mysql、textfile、textfilepath等。

技术栈：quartz框架，实时多线程任务调度；nutz框架



产品研发-日志审计系统

介绍：系统总13个一级模块，若干二三级模块。系统能实时采集企业和组织中各种不同厂商的安全设备、网络设备、主机、操作系统、以及各种应用系统产生的日志、事件、报警等信息，并将数据信息汇集到审计中心，进行集中存储、展现、查询和审计。产品解决安全行业中：各种安全产品及设备的日志数据通常杂乱无序，同时也无法体现它们之间的相互关系。因此日志审计系统就是为解决这些问题建立起一个信息交换、信息存储、信息处理的平台，通过该平台，可以对各类产品的日志、事件进行统一管理、分析。

负责：目前主要是规则、审计、告警三个一级模块的研发和维护等工作。其中，规则指的是审计的前提，告警是审计的结果。此三个模块是日志审计的核心所在。规则作为最基础的部件，规则配置的条件面板采用了插件式的设计。实时采集的syslog日志通过某种机制从采集器转发过来后，通过日志范式化后形成可读性较高的结构化数据，即：事件，针对匹配规则的某些事件便会产生告警通知客户，是否进行处理。

技术栈：

前端栈：Jquery；统计分析图表展示框架：echarts。

后端栈：elasticsearch用于快速检索，kibana用于分析展示，nutz框架等。

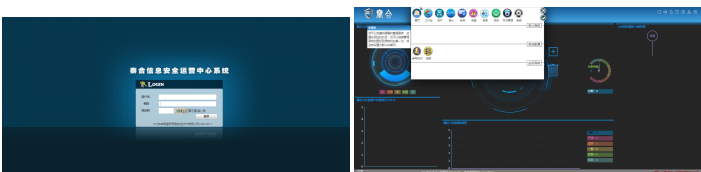
软技能：

原型设计：Axure RP Pro 7.0

设计文档：Word2010;Visio2010;Project2010

分析工具：思维导图Mindjet

进度规划：甘特图



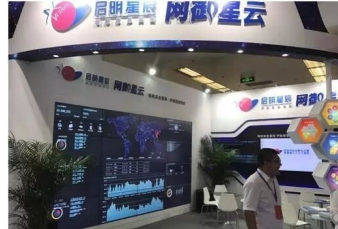
产品研发-SOC安全管理系统

介绍：全面管理，综合分析平台。SOC安全管理系统以 IT 资产为基础，以业务信息系统为核心，以用户体验为指引，从监控、审计、风险、运维四个维度建立一套可度量的统一业务支撑平台，使得各种用户能够对业务信息系统进行可用性、性能与服务水平监控，事件分析、审计、预警与响应、风险及态势的度量与评估，标准化、例行化、常态化的安全流程管控，从而最终实现业务信息系统的持续安全运营。

负责：威胁情报功能开发，以及采用elasticsearch默认支持的脚本groovy进行威胁情报的配置，匹配查询等。

技术栈：基本同上-->产品研发-日志审计系统；新增groovy实现威胁情报脚本匹配查询。

软技能：完全同上-->产品研发-日志审计系统



产品研发-通用产品打包虚拟机研发

介绍：打包，万里长征的最后一步，如何把系统部署到测试人员或最终用户的机器上？为了便捷快速的实现部署到相关服务器上，我们需要形成一个产品包。这就需要有一个便捷的工具，实现产品源码的编译，打包，针对不同操作系统，如linux是一个tar包，而windows是一个exe的可执行安装包。

负责：1、针对windows系统：

- (0) 自动化代码编译、打包的shell脚本实现；
- (1) 采用innosetup进行压制成exe，通过iss脚本实现；
- (2) iss脚本和bat脚本做安装接口：针对安装包安装，需启动各类服务，如：tomcat，es，db.java等服务；
- (3) iss脚本和bat脚本做卸载接口：针对安装包卸载，需清空安装目录；
- (4) iss脚本和bat脚本做elasticsearch服务内存分配接口，根据es官网给出的建议es内存分配机制进行合理的分配jvm内存。

2、针对linux系统：

- (0) 同上；
- (1) 不需要第三方工具压制，直接采用linux系统的tar -czvf命令；

技术栈：

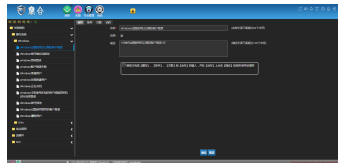
linux
shell
innosetup(iss)
bat
ssh
xmanager

产品研发-分布式关联分析系统

介绍：关联分析中心建设思路旨在南方电网项目中独立出关联分析中心，从而减轻系对CPU的消耗，提升产品运行速度。关联分析中心是未来承载TSOC管理平台的事件规则告警的分析中心，实现事件分析独立分布式的处理建设目标。

负责：目前项目刚启动,已经完成的主要功能有TMS消息服务。TMS目前支持两种传输方式：MessageClient和RemoteClient，可以按照实际需要选择合适的传输类型。MessageClient：支持点对点单向传输，由客户端发送到服务器，传输数据支持缓存，能够实现断点续传。RemoteClient：支持多点双向传输，由客户端发送到服务器，服务器可以返回数据，也可以不返回，直接作为单向传输，但是都不支持缓存，无法实现断点续传。

技术栈：消息服务中间件：Kafka、RocketMQ、Redis等【尚未定论】;数据转发器



定制项目-海关大数据日志审计系统

介绍：针对海关总署和下级海关,总署海关需要获得子署海关的告警信息，同时针对并开发告警需要消息提示等。因此，总署海关和子署海关采用级联方式进行管理和查看，同时针对子署海关开发的事件，

需要一个事件补全，以了解 子署海关中心的健康运行健康状况等。

负责：

并发告警

事件补全

技术栈：

同上-->产品研发-日志审计系统

并发告警webservice

级联

定制项目-保密日志审计系统

介绍：针对日志审计产品保密资质申请，其它不便多说。

负责：

(0) 日志审计产品保密版本功能研发

(1) 日志审计产品保密资质申请跟踪

技术栈：同上-->产品研发-日志审计系统

航天宏图 航天宏图 (2014.4 - 2015.10)

定制项目-航遥中心数据分发系统

介绍：

(0) 数据查询检索：实现用户对遥感影像产品、专题产品及其它相关数据检索

(1) 服务门户模块：在此系统提供了新闻发布功能、产品介绍功能、工具下载功能、信息共享功能

(2) 订单管理：用来实现用户对已有订单、再加工订单、数据需求单、常规需求单的管理，包括订单查询、订单状态跟踪与统计分析、订单内容修改删除、订单取消等功能

(3) 数据同步：同步基础数据支撑平台数据到信息共享与服务系统数据库；实现将内网经过筛选后的元数据和快视图数据同步到外网系统的功能；产品订单进行内网导出，可以为外网订单进行导入做好数据支撑

负责：订单管理模块的开发与维护

技术栈：前端栈：

DWR框架

统计分析图表展示框架：highcharts

后端栈：

基本环境：jdk1.6+tomcat6.0+oracle11g+Arcgis

SSH框架

软技能：

需求文档编写

概要设计文档

功能界面原型设计：Axure RP Pro 7.0

定制项目-军工数据分发系统和应用调度系统

介绍：项目包含5个子系统：

(0) 资料存档管理子系统

(1) 查询检索子系统

(2) 产品分发子系统

(3) 产品定制服务子系统

(4) 应用调度子系统

负责：

(0) 产品分发子系统

(1) 应用调度子系统

技术栈：同上-->定制项目-航遥中心数据分发系统

Java Web相关【结束】