

个人信息

姓名: Grant Yu

学历: 博士

性别: 男

专业: 计算机软件

民族: 汉族

住址: 北京市海淀区

电话: 138-1071-7899

邮箱: ZiyueYu@163.com

技能描述

1. 人工智能, 自然语言处理, 图像处理, 机器学习, TensorFlow, SkLearn, KubeFlow 等;
2. CNN, RNN/LSTM, Transformer/Bert 深度模型、优化及应用;
3. Spark/Hadoop/Flink 大数据组件及应用, PB 级系统构建;
4. Java Web 高并发大流量架构, SpringCloud 微服务架构, Kubernetes/docker 组件;
5. python/java/c++, Linux/windows;

教育背景

| | | | |
|----------|-------|----|----------------|
| 北京航空航天大学 | 计算机软件 | 博士 | 2003.9-2008.10 |
| 北京航空航天大学 | 计算机应用 | 硕士 | 1996.9-1999.7 |
| 北京航空航天大学 | 计算机科学 | 学士 | 1992.9-1996.7 |

工作经历

| | |
|------------|-------------------|
| 国家电网 | 2014.07 — 至今 |
| 高级技术专家 | |
| 中铁总公司(铁道部) | 2009.01 — 2014.07 |
| 系统架构师 | |
| 中科院中科网威公司 | 1999.7-2003.1 |
| 软件工程师 | |

项目经历

- 大型企业数据应用和分析系统** 2019.5 - 至今
工作职责: 算法设计, 系统架构, 团队管理
项目描述: 大型企业全业务数据中心数据应用和分析系统的顶层设计和实现。在归集大型企业全业务信息系统数据的基础上, 采用大数据、人工智能、机器学习、图形处理、自然语言处理等技术手段, 围绕用户、项目、人资、电网、财务、物资等业务主题域对全业务数据进行分类梳理。基于主要业务数据线和关键数据场景, 进行业务主题域/业务部门的数据线数据建模、算法设计、实现和调优, 支持企业管理和业务优化提升。
系统和工具: java/python/c++, SKLearn, Tensorflow, Hadoop, Spark,
- 基于深度学习的大型企业舆情和经营风险分析预警系统** 2018.1 - 2019.12
工作职责: 算法设计, 系统架构, 团队管理
项目描述: 基于自然语言处理的舆情风险管控监测技术: 海量 Web (微博、论坛、新闻等) 多区域、跨渠道外部信息的数据结构化内容提取技术; 文档主题生成模型和概率潜在语义分析技术, 获取关于公司的主题结构; 基于自然语言处理的 Web 短文本立场判定技术; 基于深度学习的舆情传播网络溯源分析技术; 基于深度学习的企业内部经营管理风险智能识别技术: 深度学习的经营管理风险特征编码技术; 基于半监督学习的经营管理风险识别技术; 基于特征值扰动的风险成因解释技术。
系统和工具: python/java/c++, NLP 自然语言处理, Tensorflow, Kafka, Hadoop, Spark,
- 企业图像处理智能云** 2017.1 - 2018.10
工作职责: 系统架构, 算法设计, 团队管理
项目描述: 在当前几种典型的目标检测识别方法的基础上, 利用 CNN 等深度学习算法提升小部件识别的准确性和效率, 分析了不同参数对 CNN 检测结果的影响。利用特定的 GPU 计算单元可以实时目

标检测和识别,为无人机巡检图像的智能化处理提供核心支撑。构建精细的识别类别,实现目标部件的分类,实现部件的缺陷识别。

系统和工具: python/java/c++,CNN, 图像处理、Caffe, OpenCV, Hadoop, Spark, Webservice 等。

● **部委业务平台大数据分析平台 (PB 级存储容量)** 2016.3 - 2017.3

工作职责: 系统架构, 算法设计, 团队管理

项目描述: 部委业务平台大数据分析顶层设计和实现。部委业务平台节点数据采集, 数据汇聚, 数据分析机器学习平台。部委业务平台工作节点约百万个, 数据存储为 PB 级别。采用 kafka 分级采集数据, 汇总到部中心数据仓库 Hive, 数据清洗的基础上, 采用机器学习算法, 如分类、聚类等, 实现大数据分析的目标。

系统和工具: 大流量数据采集、汇聚、融合, 实体关系分析, 分类、聚类等机器学习算法; Kafka, Hadoop, Spark, Hive, Hbase, Mahout, Mllib 等。

● **互联网海量数据面向协议的深度分析和挖掘系统实现** 2015.8-2015.12

工作职责: 系统架构, 团队管理, 算法实现

项目描述: 骨干互联网省级接口 100Gbps~Tbps 带宽全流量数据的数据采集, 数据接入, 数据融合, 关键字搜索, 实体关系分析, 数据分析和挖掘工作。

岗位职责: a.负责骨干互联网省级接口万兆网络大数据平台的架构设计和开发工作; b.负责骨干互联网省级接口网络全流量数据的采集、清洗、整合等工作; c.负责骨干互联网省级接口大数据平台数据分析、数据挖掘等工作。

系统和工具: 大流量数据采集、融合, 实体关系分析, 数据分析和挖掘, Hadoop, Spark, Hive, Hbase, Mahout, ZeroMQ 等。

● **金融网站移动互联项目** 2015.3 - 2015.8

工作职责: 架构和管理, 移动互联功能和技术架构, 系统方案, 移动互联项目总体把控

项目描述:

实现金融网站移动互联项目开发;

实现 IOS, 安卓, 和微信端开发 REST WS, IOS, 安卓 原生开发;

实现基于实时令牌的加固方案; 集成三方支付;

采用负载均衡、缓存技术等机制支持高并发大流量;

● **互联网海量数据复杂实体关联分析系统研究与实现** 2014.7-2015.2

工作职责: 系统架构师, 团队管理, 架构设计, 算法实现

项目描述: 基于系统采集的互联网数据, 对新闻、论坛、博客、微博等多元数据中的人物、组织等实体进行深入的关联分析, 从而为全局性网络舆情分析提供支持。

a.负责实体知识库的构建, 数据规模达到百万级, 提出基于全局特征的感知器模型, 从无结构文本中抽取实体属性; b.作为项目架构师, 研究实体关联分析的关键技术, 协调系统软件各个模块的联合部署、调试与测试; c.面向大规模网络文本抽取特定领域的实体关系, 准确率达到 92%以上;

d.面向领域分类的热点实体发现及推荐。

相关技术工具: 实体和实体属性识别; 文本分类; 关系抽取; 实体消歧; Hadoop, mahout 等

中铁总-高铁票务系统 2009.1-2014.5

● **系统数据挖掘和分析项目** 2012.9-2013.10

工作职责: 系统架构师

针对系统和日志事件分析的需要, 对 SOC 生成的长期数据海量信息进行离线智能分析和高级数据分析 (数据挖掘聚类分析), 寻找有效有用的日志事件分析方法, 调整和优化大数据分析的有效性和效率, 并用于调整系统的运行性能, 加固系统漏洞缺陷, 加强合规性。

相关技术: 海量数据存储, 大规模机器学习和数据挖掘, 数据建模和优化;

系统和工具: CentOS64 Hadoop/Hive/Hbase, Spark, Mahout

- 投资项目可行性分析系统

2011.9-2012.6

工作职责: 项目经理 系统架构师

项目描述: 根据国家发改委可行性报告指南, 开发本软件系统, 本系统为基于 Web 的 JEE 应用, 支持可行性报告生成和分析的全生命周期活动。包括可行性报告的创建, 经济分析报表, 市场经济分析, 市场风险分析, 不确定性分析等, 用户辅助工具等功能。基于 JavaEE 实现了功能全面, 界面友好的报告编辑功能, 经济评价报表 CRUD 功能, 支持功能强大的经济分析和市场风险分析功能。

工具技术和环境: CentOS64 JAVA Spring MyBatis Tomcat Eclipse Svn Oracle PL/SQL Developer ;

- 票务系统管理运营中心 SOC 研发

2009.6-2011.6

工作职责: 系统架构师

高铁票务系统(12306 后台, 亿级 PV 大流量)管理运营中心采集所有软硬件设备生产的信息, 包括日志文件、交易记录、网络消息、传感器采集的数据等, 这些数据包含了关于用户行为、网络系统风险、服务水平、欺诈活动和客户体验的关键信息。采用分布式系统架构, 对票务系统的内分散的网络设备和系统进行集中管理。数据管理采用三级模式: 实时分析、短期存储和长期存储, 根据业务和应用的需要进行分析处理。

参与管理系统 SOC 架构设计和核心算法代码的编写。

相关技术: 分布式计算架构; 负载均衡技术; 大数据存储处理和存储技术;

系统和工具: UNIX/Windows/Linux、C++/Java、大数据 Hadoop, Scribe 框架, 数据库 MySQL;

- 高铁 12306 票务系统

2009.1-2011.12

工作职责: 系统架构

高铁票务系统(12306 后台, 亿级 PV 大流量)管理子系统的架构分析和设计: 根据高铁票务系统管理系统的特点和互联网发展的趋势, 为系统的研发提供前瞻的技术方向; 评估和设计管理系统的整体技术架构(需求、方案、设计和部署)。高铁票务系统是一个重要接口是面向 Web 的大规模订票系统, 具备处理大规模数据和多用户高并发的能力, 高铁票务系统具备实时业务处理能力。高铁票务系统采用集中式架构, 总体上采用三级架构, 部署在全路网区域中心和车站, 并在部中心设置集中管理平台。高铁票务系统处理全路网的售票业务, 设置综合服务平台, 为顾客提供互联网、手机、电话、短信、邮件等客票销售和信息服务功能。

北京航空航天大学

- 无线自组网络协议和算法

2006.3-2008.6

工作职责: 项目负责人

系统描述: 基于当前的 Ad hoc 仿真模型, 采用虚拟现实的碰撞检测算法, 实现了一种新型 Ad hoc 仿真模型运动模型; 和其他模型相比, 本模型具有较多创新, 能够使网络节点的运动更加平滑、真实和智能; 仿真实验表明, 本模型对路由协议性能有较大影响。

系统和工具: 网络仿真、面向对象的建模、设计和开发、Linux, C/C++;

- 无线自组网络系统架构设计

2004.1 – 2005.1

工作职责: 项目负责人

系统描述: 提出一种完备的无线自组网络系统架构, 系统架构包含预防式和反应式两类系统策略, 集成三类系统机制; 相较于其他系统方案, 该架构是基于网络协议栈的分层系统方案, 同时强调系统和性能的平衡, 可用于指导设计和实施无线自组网络系统方案。

- Good Neighbors 分布式文件系统

2003.2 – 2003.8

工作职责：项目负责人

系统特征：根据分片的算法，系统中需要存储的文件，被分割发布到网络中的节点上。每个网络节点可以检索其他节点的文件分片，从而回复出来一个完整的文件。系统容错的原理在于，当部分节点离线时，在线节点仍然可以从其他在线节点中检索并回复完整的文件；

服务节点管理和负载均衡；冗余存储技术；节点失败恢复；

系统和工具：Linux、C/C++、TCP/IP 协议，P2P 算法协议；

中科院中科网威公司

1999.7- 2003.1

● 网络安全管理系统产品研发

工作职责：软件工程师

作为成熟的网络管理产品在市场上推广多年，有多个行业的成功部署案例；采用的基于规则和状态迁移的网络管理检测技术；根据各种网络具体的协议分析结果进行检测，具备更强的识别能力；

支持外部联动；系统和工具：FreeBSD, C/C++, MySQL，面向对象的建模、设计和开发；

● 上海证券交易所网站建设

工作职责：技术顾问

负责为上海证券交易所提供网络建设方面的技术咨询，作为技术专家，对该网站进行了性能评估，找出了网站的性能隐患和缺陷，并切合实际的对网站的网络架构进行改造与升级，使网络性能服务始终保持国内领先；

系统和工具：Solaris OS、Apache、Load Balancing；

英语能力

● 英语等级：CET-6；

● 较强的英语读写能力，较好的英语口语和听力。