姓名： 郭子谆
手机： 159 8208 7339
微信： william-awesome

个人网址： https://zizhunguo.com/
领　　英： https://linkedin.com/in/zizhguo/
邮　　箱： zizhunguo@gmail.com

工作经历

机器学习算法工程师 | 竞技世界网络有限公司 (北京)

北京 | 2021年12月 - 现在

基于规则和机器学习模型的匹配推荐系统

(Python, Java, Shell, PySpark, Spark, Hadoop, Hue, Impala SQL, Hive SQL, Pandas, Scikit-Learn, XGboost)

- 负责人机对战中匹配推荐规则的构建，挖掘用户的走棋风格如进攻性和速度等级等，筛选出目标用户群，为其匹配不同类型的AI对手，提升用户游戏体验
- 构建XGBoost模型，训练优化用户再来一局和胜率的预测模型，对比深度学习模型的表现，协助部署模型服务
- 跨部门合作，推动和协助打通匹配系统工程链路，校验和实现离线/实时特征提取模块，评估AI干预效果和用户数据表现
- 构建并维护离线特征库，提取游戏录像中的对局上下文信息，统计用户历史特征信息

实时风控异常用户识别及处理

(Python, SQL, Imapla, Hive, Hue, Zeppelin, PySpark, Spark, Hadoop)

- 独立负责构建异常用户识别策略和进行数据分析，负责项目上线后的效果评估工作，协助打通实时特征处理的工程链路体系，负责与其他5个部门就相关工作沟通与推进
- 构建运行于Spark集群的程序，从提取用户行为特征标签，用于风控规则迭代和计算风控效果评估指标
- 通过对规则引擎命中的异常用户实施干预手段，实现将受异常用户影响的活跃玩家比例，从10%降至1%

AI训练样本的分布式处理

- 负责构建基于PySpark训练样本转换工具，增大了转换效率，将原来算法团队与工程团队对接需要的沟通等待成本和单枳离线样本转换时间，由4天天降为30分钟，工具推广到团队4名同事使用，被应用在了10种以上该产品的不同模式中
- 提升了团队数据传输的安全性，数仓表和集群数据处理全程未出库，避免敏感数据如AI玩家清单泄露的风险
- 主持召开一次组群内的关于pyspark工具使用的分享会

项目经历

SLG游戏用户画像标签

北京 | 2022年

参与SLG项目AI智能体研发团队，负责设计AI “画像标签体系”，AI将具备多种性格和行为习惯，在游戏发行时替代低质量用户，减少产品买量营销成本，在中后期增强用户体验

基于文本挖掘的Yelp商户种类预测

纽约 | 2020年

一个对美国Yelp商户点评网站的商户数据和用户评论数据的种类预测任务，实现了利用多种NLP，机器学习和深度学习框架对文本内容进行特征提取建模训练调参优化的过程（TF-IDF, LDA, Word2Vec和GloVe文本模型建模，CNN,GRU,LSTM,encoder-decoder神经网络建模）

教育背景

罗切斯特理工大学

纽约 | 2016年8月 - 2021年5月

计算机科学/理学硕士（GPA 3.50/4.00）「全美专业排名52名」

- 相关课程：大数据导论，大数据分析（数据挖掘），数据库系统实现，自然语言处理（文本挖掘），数据结构和算法，面向对象编程等
- 获得学院大数据分析高级证书文凭
- 连续三年获得学院最高等奖学金

游戏设计与开发/理学硕士（GPA 3.28/4.00）「全美专业排名7名」

成都信息工程大学 | 电子信息工程/工学学士

成都 | 2012年9月 - 2016年6月

专业技能及证书

语言： 熟悉SQL，Python, 了解Java, C/C++, C#, Clojure, HTML, CSS, JS, XML
框架： 熟悉Spark SQL, MLlib, Scikit-learn, Pandas, NumPy, 了解 Xgboost, TensorFlow, Keras, Matplotlib, NLTK, Genism
工具和概念： Spark, Hive, Imapla, Zeppelin, Hadoop, Linux, Shell, AWS EC2, AWS S3，MySQL, MongoDB, Git, GitHub, REST, OOP, API
英语口语流利

Name
ZiZhun Guo

Phone: 159 8208 7339
Email: zizhunguo@gmail.com
Website: https://zizhunguo.com/

Github: https://github.com/zizhGuo/
Linkedin: https://linkedin.com/in/zizhguo/

Experience

Machine Learning Engineer | JJWORLD (Beijing) NETWORK TECHNOLOGY CO.LTD — BEIJING | 12.2021 – Present

A Matchmaking System Based on Rule and ML Model
(Python, Java, Shell, PySpark, Spark, Hadoop, Hue, Impala SQL, Hive SQL, Pandas, scikit-learn, XGboost)

- Forged the human-vs-AI matchmaking rules based on users' personal behavior preferences mined from statistics and machine learning algorithms. Improved the users' number of games per day by 2 percent
- Managed and maintained the features databases daily. Executed feature extraction on user essential attributes, game context attributes, and user's historical attributes
- Collaborated with ETL engineers for real-time feature engineering, implemented Spark programs updating Hive feature tables, and assisted carry forward T+1 task scheduling functions

Real-Time Risk Control Abnormal User Identification and Processing
(Python, SQL, Imapla, Hive, Hue, Zeppelin, PySpark, Spark, Hadoop)

- Initiated the core rule strategy of risk management to identify the users with abnormal game behaviors. The inter-vention reduced the affecting regular players from 10% to 1%
- Built Spark programs to extract user behavior feature tags for formulating risk control rules and conducted the post-mortem assessment of intervention effects
- Collaborated with five other departments to ensure the construction of the real-time ETL feature extraction pipe-line

Distributed Processing of AI Training Samples

- Created a PySpark-based tool for conversion of Mahjong training samples, increasing the conversion efficiency by reducing the cross-department communication and waiting for cost from 4-5 days to 30 minutes
- Improved the security of team data transmission. The Hive tables are processed between clusters to avoid the risk of leak of sensitive data such as AI player user lists
- Hosted a group sharing session on the use of Pyspark tools

Projects

SLG Game User Profile System — BEIJING | 2022

Enlisted in the AI R&D team of the strategy game project. Designed the AI character profile system, which diversifies AI's characters and behaviors for replacing churn users, reducing marketing costs and enhancing the gaming experience

Text Mining towards Yelp Business Category Prediction — New York | 2020

A business category prediction task based on Yelp Dataset with vendor data and user reviews. Conducted data preprocessing, feature extraction, modeling, training, and parameter tuning using a variety of NLP, machine learning, deep learning frameworks (including TF-IDF, LDA, Word2Vec and GloVe text model modeling, LR, SVM, NB, Random Forest, CNN, GRU, LSTM, encoder-decoder)

Education

Rochester Institute of Technology — NewYork • USA | 8.2016 – 5.2021

M.S. in Computer Science (GPA 3.50/4.00)

- Advanced Graduate Certificate in Big Data Analytics
- **Relavant courses: Data Structures and Algorithms, Object-Oriented Programming, Intro to Big Data, Data Mining, Database System Implementation, Web Service & Service Oriented Computing, NLP Text Mining**

M.S. in Game Design & Development (GPA 3.28/4.00)

Chengdu University of Information Technology — CHINA | 9.2012 – 6.2016

B.Eng. in Electronic and Information Engineering

Skills & Certificates

- **Language:** SQL, Python, Java, C/C++, C#, Clojure, HTML, CSS, JS, XML
- **Framework:** Spark SQL, MLlib, Scikit-learn, Pandas, NumPy, Xgboost, TensorFlow, Keras, Matplotlib, NLTK, Genism
- **Tool & Concept:** Hive, Spark, Imapla, Zeppelin, Hadoop, Linux, Shell, AWS EC2, AWS S3, MySQL, MongoDB, Git, GitHub, REST, OOP, API
- **Fluent in Spoken English**