

# 魏凡丁 Nick



男 | 1992 年 02 月生 | 26 岁 | 手机: +8618201732014 | 毕业院校: 浙江大学

电子邮箱: [nick\\_fandingwei@outlook.com](mailto:nick_fandingwei@outlook.com) 兴趣: 旅游, 网球, 游泳

LinkedIn: <https://www.linkedin.com/in/fanding-nick-wei-630a7458/>

Github 账号: <https://github.com/nick6918> 技术博客: <http://nickiwei.github.io/>

## 教育背景

- |   |                     |                                 |
|---|---------------------|---------------------------------|
| 08/2010 -06/2014  | 自动化专业<br>主修学位, 工程学士 | 浙江大学信息学院<br>Overall GPA: 3.22/4 |
| 浙江大学控制科学与工程学系(自动化)长期排名全国前两位, 在校期间, 参与了多个实验室研究项目, 主要方向包括控制设备及智能算法研发, 并涉及多种语言及算法研究。                           |                     |                                 |
| 08/2011 -06/2014  | 工业设计专业<br>双学位, 工程学士 | 浙江大学国际设计研究院<br>平均分: 85.6/100    |
| 国际设计研究院设计创新班是浙江大学针对于本科生的一项精英培养项目, 每学年全校招生 40 名左右。其目的在于培养“设计引导创新驱动”的交叉型精英创新人才, 核心课程包括整合与创新设计, 产品系统设计, 服务设计等。 |                     |                                 |
| 07/2013 -12/2013  | 计算机科学与技术专业          | 新加坡科技设计大学 ( 交换生项目 )             |
| 科研项目: 六自由度机械臂远程控制系统设计与开发 课程: 计算机组成, 信息系统设计等   |                     |                                 |

## 工作经历

- |  |                          |
|--|--------------------------|
| 11/2017 – Now  | 百度 (NASDAQ: BIDU)        |
| 展示广告平台部, 高级算法工程师 ( T5 ), 主要负责包括 Feeds 和广告推荐系统内 NLP 相关模型化策略的实现, 包括但不限于: |                          |
| 1, 为 Feeds 和小频道推荐系统实现低俗过滤, 自动化 tag 等 NLP 相关策略。                         |                          |
| 2, 推荐系统优化, 特别是基于 DSSM 的触发策略优化。   |                          |
| 3, 广告改写, 自动化/结构化广告创作等 NLG 项目。  |                          |
| 4, 大搜核心 ASBS 广告触发系统中的部分业务逻辑注入。   |                          |
| 09/2015 – 07/2017  | 微策略软件 (外企, NASDAQ: MSTR) |
| 供职于公司 Infrastructure Team, 团队成员 base 在华盛顿及杭州, 主要职责包括:                  |                          |
| 1, 分布式资源调度平台, 大数据基础平台等工具的开发与测试。  |                          |
| 2, 持续集成系统, 单元测试框架与 IAS(Infrastructure As Code)工具的研发, 优化和维护等。           |                          |
| 3, 百万级核心产品代码的编译系统开发和优化。  |                          |
| 4, 曾兼任一段时间的 team scrum master。   |                          |
| 07/2014 - 08/2015  | 洛哈网络 (创业公司)              |
| 作为 APP 服务器端主程序员, 独立完成或参与了公司多个核心项目的研发, 主要职责包括:                          |                          |
| 1, 服务器后台重难点业务(如弹幕系统, 直播系统等)的系统设计与研发, 核心 API 开发。                        |                          |
| 2, 随着业务量的提升, 参与后台代码的重构, 并引入单元测试等。新技术的测试与讨论等。                           |                          |

## 工作技能

### 编程语言

C++, Python

- 1, 扎实的 C++ 语法基础, 熟悉 C++ 对象模型, OOP 及泛型编程, 熟悉 STL。
- 2, 扎实的 Python 基础, LNMP 网络编程, 熟悉 Django 框架, MVC 模型, ORM 数据模型等。
- 3, 对 Scala 和函数式编程有一定了解, 在一些项目中也使用过 Java, Ruby 等语言进行开发。

### 算法与机器学习

数据结构, 机器学习, 深度学习(NLP) (此部分欢迎查阅我的博客)

- 1, 扎实的数据结构与算法基础, 对各类算法问题有浓厚的兴趣。
- 2, 熟悉各监督(回归, SVM, 贝叶斯等)及非监督(Clustering 等)学习算法, 能独立实现。
- 3, 扎实的深度学习基础, 对 CNN/RNN/GAN/DQN 等有较深刻理解, 能独立实现以上算法。对 BatchNorm, Dropout, FastPath 等常见的模型改进有较深刻的理解。

4, NLP 方向: 熟悉 ngram word embedding 原理, 熟练使用各 RNN 模型(GRU, LSTM 等)及 TextCNN 实现文本分类, 对基于 Seq2Seq 及 Attention 的文本改写与生成任务有一定经验。

5, 图像方向 :对 Image Classification, Captioning, Localization 等主流视觉算法有较深刻理解, 能独立实现以上算法。了解 R-CNN 系列的 Object Detection 的基本原理。

6, 熟练使用 Numpy, **Tensorflow**, Pandas 等常用的 Python 数据编程工具。

软件工程 分布式系统, DevOps, Agile, 测试框架, LNMP 网络编程

1, 熟悉 MapReduce 原理及 Hadoop 的使用, 熟悉 Mesos/ZooKeeper 分布式框架。

2, 扎实的 DevOps 基础, 熟悉 **CI/Jenkins**, Infrastructure As Code, GTest 测试框架等技术。

T 型人才 3, 对 **Git** 有较深入的研究, 熟悉 LNMP 服务器编程及架构。熟悉 Agile, 曾兼任 scrum master。

产品与商业设计, 设计表达

熟练使用 **Axure** 进行产品原型设计, 擅长使用 KeyNote 完成设计表达及 presentation。

熟练掌握 UCD(用户中心设计), 行为相似, 所见即所得等基础设计原理, 对产品有良好的感觉。

## 项目经历 (个人)

02/2017 - 10/2017 深度学习图像算法库的实现 (Github 仓库: MyDeepLearning)  
使用 Numpy 实现了包括 CNN, RNN, DQN 在内的多个算法, 包括反向 pass 梯度计算的实现。  
使用 TensorFlow 实现了包括各类 GAN, Fast(er)-R CNN 在内的多个算法。  
利用上述模型独立实现了 CIFAR-10 Classification, Image Captioning, Style Transfer 等项目。  
独立实现了多个机器学习算法, 包括朴素贝叶斯, SVM 等(仓库: MyMachineLearning)

04/2016 - 10/2016 独立完成邻仓 APP 后台服务器的部署及接口开发与测试  
独立完成服务器后台架构设计, 接口实现, 服务部署及测试。项目基于 LNMP 实现。

## 项目经历 (工作)

11/2017 - 01/2018 Feeds 及广告低俗模型优化  
对 Feeds 标题和广告前链词 (均为短文本) 实现低俗模型识别。文档向量化及低俗识别。  
1, 针对短文本低俗 (分类问题) 测试了 TextCNN 和 BiLSTM 两个模型。模型准召提升至准 6 召 8。并重新设计了数据积累和日志可追踪流程。为后期进一步优化提供了数据保证。  
2, 针对文档长文本训练了 Doc2Vec, Corroption 等多个 embedding 模型。并提供了通用 API。  
3, 针对广告端收入敏感的特性, 优化了带收入权重的模型, 在准召持平的情况下, 将收入跌幅控制在 7% 以内。达到预期目标。

01/2018 - 03/2018 Feeds 推荐模型触发策略模型化测试与优化  
为推荐模型触发阶段策略进行模型化渠道尝试。  
1, 为触发策略模型化测试了 DSSM(MVDSSM/TDDSSM)和 YoutubeDNN 两套模型。  
2, 上线了 SingleView DSSM, 最终的漏出贡献率超过 20%, 整体 CTR 提升 5% 左右。  
3, 对比了直接采用文档向量相似度触发策略, 最终 DSSM 模型效果最优。

03/2018 - Now NLG 结构化广告的生成模型实现与测试  
基于以电商和汽车为主的结构化数据, 实现个性化推荐语自动生成的业务逻辑。  
1, 调研了包括 Data2Seq 在内的多个实验方案, 并独立完成了项目规划。  
2, 采用 Wide & Deep 模型通过挖掘商品详情页构造有标数据。

11/2016 - 06/2017 EcoSystem: 环境管理与研发流程一体化工具的开发, CI 系统的开发与优化  
EcoSystem 集成了环境管理, 编译及测试等多个研发流程, 保证了测试与开发环境的一致性。将开发人员从复杂的编译和测试流程中解脱出来, 提高开发效率。主要负责:  
1, 参与 Gaia 系统, 即 EcoSystem 资源分配与服务管理系统部分功能的开发及测试  
2, 参与 Server Team Ecosystem 开发环境的部署(Linux/Windows), build, test 脚本调试。  
3, 负责了 Post/pre merge build, test 等 job 实现及基于 EcoSystem 的重构和优化。  
4, 长期作为该项目的 who-to-ask person 独立负责帮助 developer 解决 CI/unittest 相关问题。

03/2015 - 08/2015 无线点歌及弹幕系统核心 API 的开发与测试  
系统根据请求客户端的类型自动从 SocketIO(NodeJS 实现), HTTP 1.1 Keep-alive 机制和 HTTP 1.0 轮询中选择最优方案保持连接。显著提高了连接稳定性。

## 语言与沟通

英语	流利听说	日语	基础会话
08/2013	GRE	Overall: 317	V: 152 Q: 165 AW: 3
12/2013	TOEFL	Overall: 102	R: 28 L: 28 W: 26 O: 20