

# 黄孝炜

(+86) 135-4020-5963 · 1446269148@qq.com · AI 算法实习生 · GitHub@kitaharatomoyo

## 个人总结

本人在校成绩优秀、乐观向上，工作负责、自我驱动力强、热爱尝试新事物，致力于自然语言处理核心技术的深入研究与在产品中的落地应用。在校期间长期从事基于强化学习的医疗对话诊断相关研究，对智能对话系统的落地有浓厚兴趣。现就读于哈尔滨工业大学（深圳）。

## 教育背景

哈尔滨工业大学（深圳），计算机科学与技术，学术型硕士研究生 2019.9 - 2022.1  
排名前 10%，预计 2022 年 1 月毕业

四川大学，计算机科学与技术，工学学士 2015.9 - 2019.6  
排名前 5%，国家奖学金

## 专业技能

- 熟悉 Linux 平台, C++, Python
- 熟悉常用机器学习算法理论及 NLP 相关深度学习模型 (word2vec, bert 等)
- 熟悉常用机器学习及深度学习算法包 (scikit-learn, pytorch)
- 具有基本英文专业文献阅读能力 (CET-6)

## 实习经历

深圳网络空间科学与技术广东省实验室（鹏城实验室），PCL, 智慧医疗部门 2020.7-2020.12

- 整体项目：新冠防疫相关的对话系统。召回率达到 TOP1 84%，TOP3 98%，TOP5 100%
- 独立负责数据集的清洗、扩充与负样本的构造。通过短文本聚类，设计模板问题，按照模板问题标注数据，统一模板下的问题可以互相替换；通过短文本聚类，将不同类的样本定义为“伪”负样本对。
- 独立负责候选召回与问题匹配模型的训练。通过 bm25 召回候选问答对；使用 abcnn 模型进行训练。

北京旷视科技有限公司，MEGVII, 物流算法部门 2018.10-2018.12

- 整体项目：物流仓库中机器人的路径规划算法研究
- 负责订单、货架、机器人的模拟器设计实现。使用 C++ 设计业务场景的模拟器，用来检验算法效果。
- 负责探索先进的路径规划算法。主要探索如何针对动态进入的订单，设计灵活的路径规划算法。

## 竞赛获奖/项目作品

- JDDC2020 京东商城多模态多轮对话比赛 **3/53** (<http://jddc.jd.com/rank>) 2020.7-2020.10
  - 独立负责基于检索的问答模块的设计与训练。使用 BM25, DSSM, PolyEncoder 检索出候选问答对。
  - 独立负责基于预训练模型的生成式问答模型的设计与训练。使用 GPT-2 来作为生成模型，并使用熵增益来对生成答案及检索答案进行选择。
  - 负责探索结合检索答案的生成式问答模型。探索利用 COPY-Net 的机制结合检索出的问答对来构建模型。
- ACM-ICPC 区域赛南宁赛区 **银牌**, 2017 年 11 月
- ACM-ICPC 区域赛西安赛区 **银牌**, 2017 年 10 月
- CCF-CCPC 区域赛长春赛区 **银牌**, 2016 年 11 月