

# 吴明昊

Beijing, P. R. China

+86 176 0044 2356 | [minghao\\_wu@163.com](mailto:minghao_wu@163.com) | <https://github.com/minghao-wu>

## 教育经历

莫纳什大学

Melbourne, Australia

计算机 (博士)

Expected Mar. 2020 - Expected Mar. 2023

- 研究方向: Neural Machine Translation for Low Resource Languages

墨尔本大学

Melbourne, Australia

计算机 (硕士)

Mar. 2016 - Jul. 2018

- 研究方向: 自然语言处理

悉尼大学

Sydney, Australia

统计学和信息系统 (本科)

Mar. 2013 - Mar. 2016

## 工作经历

京东集团-AI 平台与研究部

北京

算法工程师

Aug. 2018 - Aug. 2019

- 基于业务需要, 利用 NLP 技术, 设计开发线上系统
- 开发、评价、优化机器学习和深度学习模型, 尤其擅长文本匹配、命名实体识别任务, 以及排序系统
- 跟进学术界最新进展, 开展研究发表论文

昆仑智汇数据科技 (北京) 有限公司

北京

实习数据分析师

Apr. 2017 - Dec. 2017

- 分析数据, 提取信息, 并进行数据可视化
- 开发、评价、优化机器学习和深度学习模型
- 跟进学术界最新进展, 开展研究发表论文

## 科研成果

**Evaluating the Utility of Hand-crafted Features in Sequence Labelling**

The University of Melbourne

关键词: 自然语言处理, PyTorch, 命名实体识别

Mar. 2018 - May. 2018

- 首次提出了在深度学习模型中通过 Auto-Encoder 利用人工特征的新方法
- 在 CoNLL 2003 English NER shared task 上取得了  $F_1$  91.89 和 92.29 的当时历史最好成绩, 与 BiLSTM-CRF+ELMo 和 BERT Base 表现相当
- 本工作被 EMNLP 2018 接收, 并做 Oral Presentation
- <https://aclweb.org/anthology/D18-1310>
- <https://github.com/minghao-wu/CRF-AE>

## 项目经历

---

### 京东 AlphaSales 售前智能对话机器人

京东集团

本项目基于京东商城的历史数据开发了抽取式的售前智能对话机器人

Aug. 2018 - Jan. 2019

- 复现 ESIM，计算 Query-Question 相似度，被用于 FAQ 模块和排序模块
- 在排序模块中，构建客服点击特征，提高准确率 6%，构建客服回答特征，提高准确率 3%，综合提高准确率 6.5%，当季度部门 KPI 为准确率提高 8%
- 开发基于 XGBOOST 的置信度模块

### 智慧医疗药学知识图谱

京东集团

京东 AI、京东商城与北医三院达成战略合作协议，共同开发药学知识图谱

Jan. 2019 - Jul. 2019

- 药品说明书的粗粒度结构化解析
- 基于 Protege，协助设计知识图谱本体 Schema
- 训练字符级 NER 模型 BiLSTM-CRF， $F_1$  达到 80.02
- 基于 Stanford CoreNLP 开发规则模板
- 结合规则模板和深度 NER 模型，提取文本中的细粒度结构化信息，准确率达到 80% 以上

### 金风 MD4X 风机健康综合分析系统

昆仑数据

该平台是贯穿于风力发电机组全寿命周期的健康状态评估 + 解决方案综合管理平台

Apr. 2017 - Dec. 2017

- 风机健康评估模型库多个模型的开发：本人协助开发了风力发电机叶片结冰模型，叶片开裂故障模型，风机变桨系统健康评估模型，风机齿形带断裂模型，风机整体性能评估模型等多个模型
- 模型都基于风机 SCADA 秒级数据和录波高频数据，模型精度都到了 90% 以上