

边圣陶

期望岗位: 算法实习生

Born 1995 in 浙江 诸暨



About me

嗨，我是边圣陶，我对待工作认真严谨，细心负责，善于在工作中提出问题，发现问题、解决问题，有较强的逻辑思维分析能力；喜欢和善于迎接新挑战，喜欢钻研新知识，有创新精神，有较强的探究新知识与学习能力；喜欢计算机相关工作，并热衷于研究计算机方向新技术。

Skills

Python

Java

Deep Learning

Machine Learning

Knowledge Graph

NLP

SQL

Neo4j

Contact

浙江省诸暨市天成锦江苑

bian.shengtao@gmail.com

15557537168

Richard-Bian



Experience

NLP算法实习生（已同意转正）

北京树根互联科技有限公司

since 2021-10-01

一、负责整理清洗问答对，并录入数据库。二、用elasticsearch库建立问答对的倒排索引。用elasticsearch做短语层面的相似性计算。三、用simbert将题库中的标准问题和用户提问转化成句向量，用faiss存储标准问题向量并对用户提问通过余弦相似性计算获得相近语义的标准问题。四、统合两者返回的topk以及score，对超过期待阈值的则返回回答库中的标准问题和回答，并对相似度低的问题录入陌生问题数据库中。五、探索用自产数据对原有simbert预训练模型继续预训练的可能性。



Education

Master of Information Technology

September 2020 - August 2022

Major in Artificial Intelligence and Data Science, University of New South Wales, Sydney, AU.

本科：电子商务 法学（双学位）

September 2014 - August 2018

Major in E-Commerce, 山西财经大学, 太原, 中国.

<> Projects

Cat Breed Classification

Python

这是一个图像识别的多分类任务，由于模型参数限制，项目选择了resnet18和其改进的preact-resnet作为模型，crossentropy作为损失函数，用SGD with Momentum 和 Adam分别作为优化器进行测试。另外，由于项目所给图像集并不大，我对trainset做了数据增强，包括但不限于调整大小，随机裁剪，随机旋转，随机水平翻转以及归一化处理。对于模型的超参，设置了learning-rate, learning rate scheduler, weight decay等参数，并用GridSearch方法来寻找最优参数组合，最终得到0.91的accuracy，在rank中排top3。

剧本角色情感识别

Python

本项目多标签多分类问题。我使用hfl/chinese-bert-wwm-ext 预训练模型和simpletransformers 的MultiLabelClassification 模块来实现。去除其中无角色文本，拼接角色和ta所对应文本，最终训练score为0.67。