**蒋文君**

**出生年月：**1998.3.16 **现居住地：**湖南省祁阳县

**电 话：**15377038183 **邮 箱：**[15033529591@163.com](mailto:15033529591@163.com)

**意向岗位：**机器学习 **微 信：**jwj0024



**教育背景**

**2019.09-2022.06 武汉大学 软件工程学硕**

主要课程：高级算法设计与分析、高级人工智能、数据仓库与数据挖掘等

**2015.09-2019.06 东北大学 市场营销**

**荣誉奖项**

* 2020年获得发明专利：基于变量加权的软子空间聚类方法；
* 2018年获得实用型专利：一种辅助行李箱爬楼梯的装置；
* 2017年获创新创业奖学金；2015-2016年两次获得校综合三等奖学金；
* 2017年第六届大学生节能减排社会实践与科学竞赛中获三等奖；
* 2016年获全国大学生电子商务“创新，创意，创业”大赛获省三等奖；

**专业技能**

* 熟悉Python，有良好的编码能力，熟悉面向对象编程。
* 熟悉机器学习常用算法，掌握K-means、谱聚类、层次聚类、AP聚类等聚类算法。
* 熟悉深度学习常用算法，掌握LSTM、Seq2Seq+attention等模型。
* 熟练使用sklearn、tensorflow等工具框架。
* 熟悉Linux常用命令，掌握Linux环境下的项目部署。

**科研项目**

**1、独立研究基于多元时间序列数据的软子空间集成聚类**

* 提出一种基于变量加权和K-means的软子空间聚类方法，将变量权重嵌入基于形状的距离度量的目标函数中，并且通过不同的更新策略对聚类结果，变量权重和聚类中心分别进行迭代更新。
* 提出基于反近邻的密度峰值算法来选取K-means的初始聚类中心。
* 为了进一步提高算法性能，将多个软子空间聚类算法结果整合到一起，提出了软子空间聚类集成。
* 在十个真实数据集上的实验结果表明，我们提出的方法相比四个最新的方法在两个常用聚类指标上分别提高了约5%和14%。

**2、参与国家重点研发课题：“群体协作的组织设计与调控”**

* 提出了一种基于信息性和代表性的样本选择策略，从未标记样本中选择最典型的样本进行标注。
* 引入了快速相似性度量策略，评估样本间的近似相似性距离，达到快速且准确的度量效果。
* 提出了一种基于图的成对约束传播方法，以充分利用专家提供的少量成对约束监督信息。
* 协助撰写科技报告和论文，并在国际著名学术期刊IEEE Transactions on Systems, Man, and Cybernetics上发表论文一篇。

**3、基于内容的小说推荐**

* 使用requests、BeautifulSoup获取奇书网上一万多本小说的基本信息和正文，存放在Mysql数据库中
* 使用结巴分词对每本小说进行分词，并构建词典。
* 对词典里每个词计算TF-IDF值，对每本小说进行向量化表示。
* 通过余弦相似度计算小说之间的相似性，并保存在KD-Tree树中，方便查询每本小说的近邻（相似）小说。