南京邮电大学

毕业设计(论文)任务书

|  |  |
| --- | --- |
| 题 目 | 面向手机传播网络的用户识别研究与实现 |
| 专 业 | 软件工程（NIIT） |
| 学生姓名 | 陈志聪 |
| 班级学号 | B12041120 |
| 指导教师签字 | 陈兴国 |
| 指导教师职称 | 讲师 |
| 指导单位 | 南京邮电大学计算机学院 |
| 单位负责人签字 |  |
| 学院领导签字 |  |
| 日 期 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 题　　目 | | 面向手机传播网络的用户识别研究与实现 | | |
| 课题类型 | | 理论研究■  工程设计□ | 课题类别 | 理论与实验■　 硬软结合（含硬件）□  软件□ 其他□ |
| 设计内容与技术要求  、  成  果形式 | 1. 设计内容 2. 基于手机用户数据，构建社会网络模型，从中挖掘有价值的属性。 3. 利用这些属性采用机器学习方法进行用户画像，预测用户性别、年龄等。 4. 记录分析实验结果，与相关研究成果进行对比，比较预测的准确性。 5. 技术要求 6. 用pandas+shell处理GB级原始txt数据，并进行数据预处理和数据清洗 7. 用networkx进行手机通话和短信网络的建模和实现 8. 用scikit-learn实现机器学习算法 9. 用git进行版本控制和管理 10. 实验开发环境为CentOS 5.6 + IPython Notebook   三、成果形式  1、设计并实现能提高预测效果的用户识别方法；  2、提交毕业设计报告及其电子文档。 | | | |
| 设计进度 | 1. 学习社会网络分析、机器学习相关知识，搜索课题相关文献 3周 2. 数据清洗和数据预处理，抽样，完成开题报告 3周 3. 构建手机用户ego network，提取合适的特征 6周 4. 结合用户属性，构建数据集，实现机器学习算法 6周 5. 结果误差分析，与他人研究成果进行对比 3周 6. 整理资料，论文写作，准备答辩 3 周 | | | |
| 参考资料 | [1] Onnela, J. P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., ... & Barabási, A. L. (2007). Structure and tie strengths in mobile communication networks. Proceedings of the National Academy of Sciences, 104(18), 7332-7336.  [2] Dashun Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabasi. 2011. Human mobility, social ties, and link prediction. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '11). ACM, New York, NY, USA, 1100-1108.  [3] Newman M E J. Assortative mixing in networks[J]. Physical review letters, 2002  [4] Eagle N, Pentland A S, Lazer D. Inferring friendship network structure by using mobile phone data[J]. Proceedings of the national academy of sciences, 2009, 106(36): 15274-15278.  [5] de Montjoye Y A, Quoidbach J, Robic F, et al. Predicting personality using novel mobile phone-based metrics[M]//Social computing, behavioral-cultural modeling and prediction. Springer Berlin Heidelberg, 2013: 48-55.  [6] 崔瀚文,栾石圳南,李远帆,蔡志杰,曹沅. 空巢老人手机用户的精准识别[J]. 数学建模及其应用,2014,01:49-62. | | | |