**《机器学习》课程实验报告**

|  |  |
| --- | --- |
| **实验名称:** | Short-term prediction with linear regression model |
| **姓 名:** | 李若昊 |
| **学 号：** | 5120180269 |
| **教 师:** | 赵旭剑 |
| **完成时间:** | 2020/6/20 |

|  |
| --- |
| **一、实验目的和内容**  **利用「怀卡托智能分析环境」实践理论课中学习过的线性回归，体验它的实际应用。** |
| **二、实验过程**  **加载数据：**  **用WEKA创建线性回归模型：**    **这个模型表明生育率、老龄人口占比、其他人口死亡占比、人口流动比率不会影响实际人口数，随着年份增长、老龄人口死亡加剧、女性育龄人口增多，实际人口数均会减少（女性育龄人口占比增大实际人口会减少有点奇怪）**  **所以我把女性育龄人口占比这一列去掉，再生成一个模型：**  **这个模型解释起来就相对合乎情理一些了。**  **接着我用2012、2013年的实际数据依次验证两个模型，结果均列在了实验结果部分。** |
| **三、实验结果**  **用第一次生成的模型预测2012、2013年实际人口数（以万计），得到：**  **2012年：（实际为2069 ）**  **2013年：（实际为2115 ）**  **用第二次生成的模型再次预测2012、2013年实际人口数（以万计），得到：**  **2012年：（实际为2069 ）**    **2013年：（实际为2115 ）** |
| **四、实验总结**  **第一次用所有的数据列让WEKA自动以线性回归方式预测数学模型，发现模型解释起来有些不符合常理，所以我通过去除解释起来牵强的变量修改模型，再次预测，可以通过结果看到显然经过修改后的第二次的模型预测值更加符合真实值。** |