**深 圳 大 学 实 验 报 告**

**课程编号：­ 2801000049**

**课程名称：­ 机器学习**

**实验项目名称： 基于机器学习的深圳道路能见度预测**

**学院： 电子与信息工程学院**

**专业： 电子信息工程**

**指导教师： 麦晓春**

古炜

林凡超

卫宏林2022300013

古炜

林凡超

卫宏林

**报告人： 学号： 班级： 文华班**

**实验时间： 2024.5.28 - 2024.6.28**

**实验报告提交时间： 2024.6.28**

**教务部制**

|  |  |
| --- | --- |
| **基于机器学习的深圳道路能见度预测**  **古炜 林凡超 卫宏林**  **Contributing percentage: 33.33% + 33.33% + 33.33%** | |
| **Abstract**  **(250-300 words)** | |
| **1. Introduction**  **Background**  **Motivation**  **The problem our solve**  **Method**  **Contributions and Novelty** | |
| **2. Method (Use at least two pages to illustrate the methodology)**  **2.1**  **2.2**  **2.3** | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **3. Experiments and Results**  **3.1 Data Collection, Preprocessing and Analysis**  Data Collection from the website:  https://opendata.sz.gov.cn/data/dataSet/toDataDetails/29200\_00903518  Preprocessing  Step1: we find the dataset that some characteristic sets have lots of missing value, so we first delete the sets.  Step2: because the dataset that has been processed still have some missing value, we chose to change the missing value to 0.  Step3: we calculate the relationship value for the visibility and other characteristic sets.  The result below:    Figure 1 correlation with VISIBILITY  Step4: we found that the dataset have so many characteristic and many charateritic have low correlation with visibility, so we only choose the characteristic that the correlation value greater than 0.15.  The result below:   |  |  | | --- | --- | | **Feature** | **Correlation** | | VISIBILITY | 1 | | RELHUMIDITY | 0.311094 | | MINRELHUMIDITY | 0.306294 | | INSTANTWINDV | 0.212944 | | HEXMAXWINDV | 0.198315 | | WINDV10MS | 0.162007 | | MAXWINDV10MS | 0.156455 | | AUTOPRECIPAMOUNT | 0.156363 | | GRASSLANDMAXTEMP | 0.152857 |   We also delete the AUTOEVAPGAUGE, because the sets is have so many missing values.  Step5: we explore the distribution for every characteristic.  The result below:            Expect the autoprecipamount, other characterstic similar adapt to normal distribution.  Step6:  We calculate the mean and variance for each characteristic.  The result below.   |  |  |  | | --- | --- | --- | |  | **mean** | **std** | | **VISIBILITY** | 220.496 | 110.7851 | | **RELHUMIDITY** | 72.96193 | 16.2497 | | **MINRELHUMIDITY** | 70.95121 | 16.75499 | | **INSTANTWINDV** | 24.71689 | 15.29268 | | **HEXMAXWINDV** | 49.76139 | 21.20455 | | **WINDV10MS** | 19.82949 | 9.918758 | | **MAXWINDV10MS** | 27.12172 | 11.91181 | | **AUTOPRECIPAMOUNT** | 1.246381 | 9.825435 | | **GRASSLANDMAXTEMP** | 266.4373 | 109.4651 |   **3.2 Evaluation Metrics**  **3.3 Experiments**  **3.3 Experimental Results and Analysis** |

|  |
| --- |
| **4. Discussion**  **5. Conclusions**  **References**  [1]  [2]  [3] |

深圳大学学生实验报告用纸

|  |
| --- |
| **Appendix**  **Appendix A: xxx**  **Appendix B: xxx** |
| 指导教师批阅意见：  成绩评定：  指导教师签字：  年 月 日 |
| 备注： |

注：1、报告内的项目或内容设置，可根据实际情况加以调整和补充。

2、教师批改学生实验报告时间应在学生提交实验报告时间后10日内。