|  |  |
| --- | --- |
| **Table of Contents** | **Page No** |
| Declaration | i |
| Certificate(2) | ii- iii |
| Acknowledgement (2) | iv- v |
| Abstract | vi |
| 1. Introduction | 1-2 |
| 1.1 Background | 2-4 |
| 1.2 Research Objectives | 4 |
| 2. Literature Review | 5 |
| 2.1 Introduction | 5 |
| 2.1.1 About National Air Quality Index | 5 |
| 2.2 Machine Learning Techniques | 6 |
| 2.3 Regression Models | 6 |
| 2.3.1 What is Regression Model? | 6 |
| 2.3.2 Why do we use Regression Model? | 6-7 |
| 2.2.3 How many types of regression techniques do we have?  A. Multiple Linear Regression  B. Support Vector Regression  C. Decision Tree Regression | 7  8  8-9  9-10 |
| 3. Data | 10-11 |
| 3.1 Data Collection | 11 |
| 3.2 Creating Data Set | 12 |
| 3.2.1 Calculate sub index   * PM2.5 * PM10 * NO2 * NH3 * SO2 * CO * Ozone | 12-20  13-14  14-15  15-16  16  17  18  19 |
| 3. 3 Calculating AQI | 20 |
| 3.3.1 How to calculate the | 20-21 |
| 3.4 Data Correlation | 22-23 |
| 3.5 Training Testing Splitting | 23 |
| 4. Methods and Models Set up | 24 |
| 4.1 Multiple Linear Regression(MLR) | 24-25 |
| 4.2 Support Vector Regression | 25 |
| 4.3 Decision Tree Regression | 25-26 |
| 4.4 Pre-processing and Feature Selection | 26-27 |
| 4.5 Result | 27-28 |
| 4.6 Algorithm Comparison | 28-29 |
| 5. Conclusion | 30 |
| 6. References | 30 |