**Table of contents**

|  |  |
| --- | --- |
| Topic | Page Number |
|  |  |
| Abstract | 1 |
|  |  |
| 1 Introduction | 2 |
| 1.1 Problem definition | 2 |
| 1.2 Objective | 2 |
|  |  |
| 2 *K-*means algorithm | 3 |
| 2.1 Clustering by *K*-means algorithm | 3 |
| 2.2 Results | 3 |
| 2.3 Applications of clustering | 4 |
|  |  |
| 3 Principal Component Analysis | 6 |
| 3.1 Data preprocessing | 6 |
| 3.2 Principal Component Analysis (PCA) algorithm | 6 |
| 3.3 Results | 7 |
| 3.4 Applications of Principal Component Analysis | 8 |
|  |  |
| 4 One-vs-all logistic regression | 9 |
| 4.1 One-vs-all Algorithm | 9 |
|  |  |
| 5 Results | 10 |
| 5.1 Supervised Learning on dataset as Input | 10 |
| 5.2 Supervised Learning on Results of *K* Means | 10 |
| 5.3 Supervised Learning on Results of Principal Component Analysis | 11 |
|  |  |
| 6 Conclusion and future work | 12 |
|  |  |
| References | 13 |