**TABLE OF CONTENT**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CHAPTER NO** | **TITLE** | | | | **PAGE NO** |
|  | **ABSTRACT** | | | | iv |
| **1** | **CHAPTER - 1**  **INTRODUCTION** | | | | 1 |
|  | 1.1 | | OVERVIEW OF THE PROJECT | | 1 |
|  | 1.2 | | OBJECTIVE OF THE PROJECT | | 2 |
|  | 1.3 | | WHY MACHINE LEARNING? | | 3 |
|  | 1.4 | | USES OF MACHINE LEARNING | | 4 |
| **2** | **LITERATURE REVIEW** | | | | 5 |
| **3** | **SYSTEM DESIGN** | | | |  |
|  | 3.1 | | EXISTING SYSTEM | | 7 |
|  | 3.1.1 | | | DISADVANTAGES | 7 |
|  | 3.2 | | PROPOSED SYSTEM | | 8 |
|  | 3.2.1 | | | ADVANTAGES | 8 |
| **4** | **BACKGROUND STUDY** | | | | 9 |
|  | 4.1 | CLASSIFICATION OF MACHINE LEARNING | | | 9 |
|  | 4.2 | SUPERVISED LEARNING | | | 10 |
|  | 4.3 | UNSUPERVISED LEARNING | | | 11 |
|  | 4.4 | CLUSTERING ALGORITHMS | | | 13 |
|  | 4.5 | APPLICATION OF MACHINE LEARNING | | | 14 |
|  | 4.5.1 | | | MACHINE LEARNING IN HEALTHCARE | 14 |
|  | 4.6 | PRINCIPAL COMPONENT ANALYSIS | | | 16 |
| **5** | **METHODOLOGY** | | | | 17 |
|  | 5.1 | | DATA PREPROCESSING | | 17 |
|  | 5.2 | | DATA COLLECTION | | 17 |
|  | 5.3 | | TRAINING DATA AND TEST DATA | | 17 |
|  | 5.4 | | MODEL CREATION | | 17 |
|  | 5.5 | | MODEL PREDICTION | | 17 |
| **6** | **MODULES** | | | | 18 |
|  | 6.1 | DATA PREPARATION AND ENGINEERING | | | 18 |
|  | 6.2 | MODEL TRAINING AND EVALUATION | | | 19 |
|  | 6.3 | MODEL PERSISTENCE | | | 20 |
|  | 6.4 | WEB APPLICATION DEVELOPMENT | | | 21 |
| **7** | **ALGORITHM IMPLEMENTATION** | | | | 22 |
|  | 7.1 | MULTINOMIAL NAIVE BAYES ALGORITHM | | | 22 |
|  | 7.2 | COUNTVECTORIZER | | | 24 |
|  | 7.3 | FLOW CHART | | | 25 |
|  | 7.4 | SYSTEM REQUIREMENTS | | | 26 |
|  | 7.4.1 | | | HARDWARE REQUIREMENTS | 26 |
|  | 7.4.2 | | | SOFTWARE REQUIREMENTS | 26 |
|  | 7.5 | SOFTWARE DESCRIPTION | | | 30 |
|  | 7.5.1 | | | PYTHON | 27 |
|  | 7.5.2 | | | VERSION 3 | 28 |
|  | 7.5.3 | | | PACKAGES | 28 |
|  | 7.5.4 | | | NUMPY | 29 |
|  | 7.5.5 | | | PANDAS | 29 |
|  | 7.5.6 | | | SKLEARN | 29 |
|  | 7.5.7 | | | FLASK | 30 |
| **8** | **APPENDIX** | | | | 31 |
|  | 8.1 | SOURCE CODE | | | 31 |
|  | 8.2 | SCREENSHOTS | | | 33 |
| **9** | **CONCLUSION & FUTURE WORK** | | | | 37 |
|  | 9.1 | CONCLUSION | | | 37 |
|  | 9.2 | FUTURE WORK | | | 38 |
| **10** | **REFERENCE** | | | | 39 |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **FIG NO** | **TITLE** | **PAGE NO** |
| 1.3 | PROCESS OF MACHINE LEARNING | 3 |
| 4.2 | SUPERVISED LEARNING | 10 |
| 4.3.1 | UNSUPERVISED LEARNING | 11 |
| 4.3.2 | CLUSTERING IMAGE | 12 |
| 4.4 | K-MEANS CLUSTERING | 13 |
| 4.5.1 | APPLICATIONS OF MACHINE LEARNING | 14 |
| 4.5.2 | HEALTHCARE | 14 |
| 6.1 | DATA PREPARATION AND FEATURE ENGINEERING | 18 |
| 6.2 | CONFUSION MATRIX | 19 |
| 6.3 | MODEL PERSISTENCE | 20 |
| 6.4 | WEB APPLICATION DEPLOYMENT RESULT | 21 |
| 7.3 | FLOW CHART | 25 |
| 8.2.1 | Python Code - 1 | 33 |
| 8.2.2 | Python Code - 2 | 33 |
| 8.2.3 | Python Code - 3 | 34 |
| 8.2.4 | Python Code – 4 | 34 |
| 8.2.5 | Output Console | 35 |
| 8.2.6 | Web Application XSS | 35 |
| 8.2.7 | Web Application Result | 36 |
| 8.2.8 | Web Application Result | 36 |

**LIST OF ABBREVATIONS**

|  |  |  |
| --- | --- | --- |
| XSS | - | Cross-Site Scripting |
| OWASP | - | Open Web Application Security Project |
| CSRF | - | Cross-Site Request Forgery |
| HTTP | - | Hypertext Transfer Protocol |
| NLP | - | Natural Language Processing |
| MNB | - | Multinomial Naive Bayes |
| ML | - | Machine Learning |
| DoS | - | Denial of Service |
| ID | - | Identifier |
| ACID | - | Atomicity, Consistency, Isolation, And Durability |
| API | - | Application Programming Interface |