MACHINE LEARNING MODELING FOR STUDENTS' SELF-ESTEEM LEVEL DETECTION

BY

MD. MEHEDI HASSAN MIM
ID: 232-25-027

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science and Engineering.

Supervised By

DR. SHEAK RASHED HAIDER NOORI
Professor & Head
Department of CSE
Daffodil International University

Co-Supervised By

MR. ABDUS SATTAR
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2025

APPROVAL

This Project/Thesis titled **"Machine Learning Modeling for Students' Self-Esteem Level Detection"**, submitted by **Md. Mehedi Hassan Mim**, ID No: **232-25-027** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **11-01-2025**.

BOARD OF EXAMINERS

Chairman

Dr. S.M Aminul Haque
Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Fizar Ahmed
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Md Alamgir Kabir
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Mr. Sadat Hasan
Data Scientist
Risk Management Division,
BRAC Bank Limited

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Dr. Sheak Rashed Haider Noori, Professor & Head,** Department of Computer Science & Engineering (CSE), Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

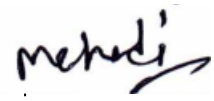
Supervised by:

Dr. Sheak Rashed Haider Noori
Professor & Head
Department of CSE
Daffodil International University

Co-Supervised by:

Mr. Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:



Md. Mehedi Hassan Mim
ID: 232-25-027
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we have to admit that this project could not have been done without the help of adequate people in the respective sectors of our project, proper guidance from our supervisor, and, of course, the grace of the **Almighty Allah**.

We are grateful and wish our profound indebtedness to **Dr. Sheak Rashed Haider Noori**, **Professor & Head**, Department of CSE, Daffodil International University. Deep knowledge & keen interest of our supervisor in the "Machine Learning" field to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Sheak Rashed Haider Noori**, **Professor, and Head**, Department of CSE, Daffodil International University, for his kind help in finishing our project and also to other faculty members and the staff of the CSE department of Daffodil International University.

We would like to thank our entire coursemates at Daffodil International University for participating in this discussion while completing the coursework.

Finally, our deepest gratitude goes towards our parents for their continuous support and utter belief in us which ultimately led us towards the completion of this project.

ABSTRACT

In this report, we aimed at using machine learning approaches to identify whether students' self-esteem was low, normal, or high. Early identification of self-esteem was a strategic intervention in our model because of its key influence on mental health as well as performance outcomes. In questionnaires, we obtained basic demographic data, behavioral characteristics, and self-assessment results from 1,050 students. The results of the dataset showed that participants had low self-esteem in the range of 48.2%, normal self-esteem in the range of 50.9%, and only 1% of participants had high self-esteem. In this paper, we were able to use machine learning algorithms, such as Logistic Regression, Random Forest, k-nearest Neighbors, and Gradient Boosting to build and train the developed model. To compare the effectiveness of each algorithm, we used parameters like precision, recall, F1-score, and accuracy in which Random Forest and Gradient Boosting models showed the best results for the classification of self-esteem categories. More details, hyperparameter tuning further enhanced the models to guarantee reliable predictive accuracy. That the use of machine learning has promising potential in mental health is supported by our studies which indicate how valuable self-esteem assessments in an educational context may prove in terms of learning with early detection of the potential self-esteem issues at hand. Possible future research might provide a greater number of indicators, including social behavior and academic achievements to make the model more precise and suitable for study participants. From this research, it is possible that machine learning will play a significant role in creating practical evidence-based interventions to promote student well-being.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

A person's personal feelings towards themselves that arise from their awareness of how they judge various parts of themselves are included in their self-esteem, defined as the perspectives and values they ascribe to themselves [1]. Self-esteem is what a person thinks, feels, and believes about themselves; it determines how we behave, and experience them. The self refers to our value for ourselves, our abilities, others, and what happens in life. People with strong self-esteem feel good about themselves and their skills. They often are more emotionally tough, can acknowledge their strengths and weaknesses, and look on the bright side even when everything seems bleak. On the other hand, low self-esteem results in self-doubt, insecurity, and a negative sense of oneself leading to focusing on one's weaknesses or feeling inadequate and unworthy of love and success. Life satisfaction, considered a variable that affects teachers' job satisfaction, includes evaluations of individuals' lives [2].

There are a few things that all come into play when it comes to what shapes your self-esteem. One of the biggest factors is likely early life experiences (our upbringing). Being raised in a supportive, caring family environment of love and positivity helps to create an empowered sense of worth. However, criticism, neglect, or emotional abuse will have a lasting impact on self-esteem in the arena of inadequacy, as we get older, social comparisons may also become a factor—constantly competing against others in things like achievements, looks or status can lower your self-esteem too, especially in a time like this where social media shows us everyone in the best light possible. For many teenagers, going from high school to college marks a significant turning point in their lives. Students have the opportunity to study and grow psychologically when they attend university [3].

Studies on self-esteem, break down convincing stats on how they influence everyone at any age, culture, and situation. It found self-esteem varies throughout life for both men and women, following a predictable pattern from adolescence to old age, with the lowest point reaching somewhere around 16-18. Self-esteem peaks in late childhood, goes down among teens (due to social pressure and identity-forming processes), and

tends to rise again through adulthood (American Psychological Association). Midlife is the phase in life when most people are at the peak of their self-esteem, which begins to decline into older adulthood as health problems and other life events such as retirement set in. Low self-esteem has been documented in a variety of antisocial groups, such as conduct-problematic kids and teenagers [4].

Generally speaking, those who have high self-esteem are happier and more likely to have close friendships [4]. Researchers across 48 countries examined data from over 985,000 individuals for a global study of self-esteem published in the Journal of Personality and Social Psychology. Self-esteem tends to be higher in Western cultures, especially in parts of North America and Europe where individualism and self-expression are a dominant part of the cultural landscape. In contrast, countries in East Asia, where group harmony is one of the core collectivist values, report far lower levels of self-esteem. That said, if self-esteem culture plays out in different ways than it does for White people, that is not an argument for interpreting low self-esteem as causally leading to harm.

It has been noted that self-efficacy and self-esteem, two crucial components of fundamental self-evaluations, have a significant implicit impact on mental health [5]. Gender seems to be a factor in self-esteem trends as well. A significant body of research has illuminated differences in self-report of esteem across genders. Men generally report higher levels of self-esteem than women, especially as adolescents and adults. As far as we can discern from the research, published on Thursday in Psychological Bulletin, this gap is not that large – men report modestly third-party monitored self-worth and competence across cultures – but it remains strong. But the difference starts to disappear after an early age as life experience builds personality, and develops shared identity for both males & females. For example, from early adolescence to maturity, self-esteem may increase. It has long been believed that having a strong sense of self-worth can shield people from the damaging effects of bad experiences like failure [6].

The phenomenon of low self-esteem is a risk trigger for mental illness and in most cases is combined with a high incidence of anxiety symptoms. Only a tiny fraction of the current literature focused on student anxiety, but what did exist was clear -anxiety disorders affect huge numbers of students across the globe. Bangladeshi freshman students scored 61% in at least moderate anxiety; it was 15% of undergraduates from Franciscan University, U.S., who presented severe or extremely severe anxiety. A total

of 41. ≤3 % of college students had anxiety problems in China, including approximately 8–13%, mild anxiety, 20% moderate anxiety, and 4–6%), severe anxiousness [6].

Another study revealed how 20% of college students in the container who have reached a clinical degree suffer from abnormal anxiety symptoms that coexist with depression and stress. 8 Similar circumstances also applied to other groups, eg patients with cancer or inflammation, infertile women,11, and women who are experiencing marital conflict. The Covid-19 pandemic has also raised unprecedented levels of angst among college students. For example, in one study of over 70% of undergraduate students at a large US public university, the perception increased anxiety due to other stressors as well, such as no services for mental health help, being engaged in social distancing, and being debarred from academic progress levels. Finding the contributing elements to anxiety issues is crucial given the rising incidence of anxiety and its detrimental impacts; self-esteem has drawn a lot of attention in this regard [6].

Transitioning from high school to a university signifies a significant stage in life that many young adults experience, requiring them to take on a multitude of personal and academic challenges. A university setting offers fresh chances for intellectual development, skill acquisition, and self-reliance training [7]. There are widespread self-esteem development programs in schools and other public institutions that implicitly make the premise that one's subjective belief states are the primary source of one's self-esteem. In particular, the idea behind these programs is that self-esteem can be raised without necessarily referencing objective standards from the outside world [8].

How may children's violence be connected to both high and poor levels of self-esteem? We propose two potential causes. First off, different research has defined "high" and "low" self-esteem differently. A child's self-esteem can be "high" or "low" about peers' self-esteem, but it can also be "high" or "low" about observers' assessments of the child's behavior and abilities, such as peers. According to the latter description, a youngster may overestimate or underestimate their behavior and competencies in comparison to the judgments of observers [9].

Anxiety disorders affect 117 million youth worldwide between the ages of 5 and 19, making them extremely common in children and adolescents. Adolescent anxiety symptoms are linked to negative consequences that last throughout life, such as anxiety, depression, and substance abuse that continues into maturity. Adolescents' ability to go

about their everyday lives is not only severely compromised, but their quality of life is also adversely affected. However, there are important public health ramifications when modifiable risk factors for anxiety symptoms are identified [10].

1.2 Motivation

Self-value is rooted in the internal emotional infrastructure of oneself. A person with low self-esteem might retreat into solitude and avoid interacting socially. Their day-to-day may involve isolating from social events, ignoring family and friends, and declining interest to the point of abandoning pastimes or hobbies entirely. Low self-esteem is becoming an epidemic in today's society, ails us with the feeling of hopelessness.

Another example of low self-esteem is how it might manifest in a man or woman based on their life experience. Diminished self-worth due to various forms of mental stress is also a common issue within the public. But the biggest challenge is that a much larger populace in Bangladesh has yet to even start thinking of matters concerning one's emotional well-being. They often miss the signs indicating low self-esteem. Since they do not ask for assistance from mental health professionals (psychologists, therapists) it may even be worse.

In other words, one of the easiest and most reliable ways to judge whether a person has low self-esteem or not is through their daily activeness in moving forward. And without acknowledging these red flags it is impossible to fathom the mood of an individual. Consequently, we need to create a tool or method so that anyone can evaluate the mood of self-esteem person. Learning systems and technology are also the most effective solutions to offer insights in these cases, that can help individuals improve themselves.

Low Self-esteem often comes out in very subtle but big ways to the individual behaviors, actions, and interactions. This leads to self-exclusion, negative internal chatter, fear of failure, or too much dependency on the opinions of others. The issue however is that such signs are oftentimes dismissed as harmless personality traits or short-lived states of mind, particularly in societies where mental health awareness remains underdeveloped.

A self-esteem test or measure might focus on patterns of behavior and help identify it in things people do each day, also offers a range of knowing about activities pubs choices

work stresses cultural Baring history this must be consulted during the appropriate. If you find that you are avoiding social situations, lacking confidence, or have less motivation for life could be a sign of low self-esteem. With the help of more sophisticated machine learning systems, such a tool could dissect behavioral data and deliver insights into an individual's emotional wellness at deeper levels.

Those systems can do more than just detect, though. They might provide individually tailored responses, and self-help techniques or help guide people towards professional assistance if they do need it. This could also help to highlight health issues in countries like Bangladesh, which still struggle with the idea of emotional well-being as an acceptable condition. Integrating the use of this tool into daily life could also prompt people to become more proactive with their self-esteem and mental health, which would in turn lead to a healthier community.

But those systems are capable of more than just detection. Individualized responses, self-help strategies, or aid for those who potentially need to seek professional support. It may be a good way to improve the visibility of mental health in countries like Bangladesh that have yet to accept even basic emotional well-being as an okay state. Another benefit of the use of this tool in daily life accelerates individuals to make good decisions against their self-respect and mental health which would be a healthy community for itself.

To the best of our knowledge, there are only a very few studies known to us that specifically focus on this topic. There is much more to investigate in this area. There are not enough timely, data-driven answers or solutions to questions surrounding depression detection and emotional health assessment.

So, we are using a very good old approach that has been around for years in an attempt to figure out, how do you build something magical that can detect someone feeling down. We will be able to understand more by using this method as we can learn which methods are used towards recognizing the emotional distress in patterns, behaviors and even expressing creatively like art. By integrating cutting-edge methods with well-established information, the aim is to close the research gap and improve our ability to recognize and comprehend emotional challenges like depression and poor self-esteem.

1.3 Fundamental Principle of the Study

As we have already indicated, the number of full-size paintings from the Bangladeshi perspective that were created in the past and showed signs of desperation has drastically decreased. For this reason, we highly recommend running with despair and utilizing gadget learning strategies. Machine learning, a branch of artificial intelligence, offers probabilistic, statistical, and optimization methods that enable computers to "learn" from patterns and similarities discovered in sizable, complex, noisy, and demanding datasets. Methods for analyzing contemporary electronics are used in the fields of recognition, classification, and detection. Machine learning is used in systems for analysis, disease detection, stock market analysis, and traffic prediction. Many different kinds of detection and classification challenges are resolved by the application of device learning. In this domain, machine learning holds great importance and is frequently employed by scholars, yielding noteworthy outcomes. Consequently, we decided to use device learning for our study.

1.4 Research Questions

We have divided our entire research into a few questions for the convenience of our work. As a result of these questions, we will get to know and get an idea about the whole research and try to find answers to these questions in the next steps of the research.

- Are the machine learning techniques we've used and our dataset valid and dependable?
- Should we apply well-known and popular machine learning techniques, or should we develop new models?
- Which of the machine learning algorithms we used on our dataset had the highest accuracy and why?

1.5 Expected Output

Our paper's goal is to identify human behavior from desperation. One of our main problems is that we can't identify dejection when it happens. Someone can swiftly recollect his intellectual domain by using our strategy. This might mean that the general public is no longer restricted to their intellectual realm. We should talk to a variety of social groupings. We can't help but notice how many of our friends have tied the knot.

They are not engaging in any artistic pursuits inside society. But there's no longer a noticeable decline in these people's diversity throughout society. Finding ideas that represent the intellectual aspect of humanity will be made easier with the help of this essay.

1.6 Project Management & Finance

In the above section, we mentioned detecting self-esteem levels which is done by developing an algorithm to study and classify all our generated data using Machine learning techniques. This project involves gathering data from different sources including surveys and social media text analysis to create a predictive model. The research would not only be on a data collection front but also come with an appropriate detection and validation framework.

We defined an explicit timeline dividing the research into different tasks like data collection, model development/testing, and analysis to make sure it is progressing well. The progress of each phase is monitored in contrast to the pre-defined milestones like finishing data preprocessing, training a model writing down research paper for submission, etc. These milestones help to gauge progress and keep the project on track.

Executing our research needs certain resources such as access to computational tools necessary for executing machine learning algorithms, survey platforms through which the data is collected, and statistical software required while analyzing results. The research team is another valuable asset, with researchers who have a specialization in data science, psychology, and machine learning.

This will include data collection delays, algorithmic bias, and concerns with participant privacy; all risks that have been identified as part of our risk management process. Countermeasures are in place to reduce risk, such as alternate data sources for backup and adherence to ethical guidelines when dealing with sensitive user information.

These key stakeholders in this project are the research funding agency, university ethics committees, and data contributors. In the process, this would also contribute to both academic research as well as public welfare in empowering individuals to comprehend and better their self-esteem.

The outcome of these projects consists of a fully functional model for detecting self-esteem, full datasets and hopefully published accessible research paper including

findings and methodologies used. This output will be disseminated to the scientific community and professionals working in mental health.

The project budget includes funds for data collection tools and software licenses, computational resources, and payments to participants. We performed an itemized cost estimate that included direct equipment, pilot expenses, and indirect costs.

This work is funded by the researcher. This financing will contain all the costs for the project. Regular expenditure tracking to maintain the project within budget margins and financial reports depending on funding will be submitted.

Upon completion of this project, we anticipate a strong return on our investment as the self-esteem detection model enables mental health and psychology professionals to better understand, assess, and enhance self-esteem within various populations both in academia & real-world applications.

1.7 Report Layout

The following are the contents of this research project:

- Chapter 1 provides an overview of the study, including its objectives, guiding concepts, open-ended research questions, and anticipated results.
- The breadth of the problem, the challenges we confront, an overview of the study, and pertinent past publications are all covered in Chapter 2.
- The procedures for data collecting, data preprocessing, statistical analysis, and feature implementation are all described in the flowchart in Chapter 3.
- Chapter 4's experimental analysis includes a summary of accuracy, research outcomes, and other relevant investigations.
- The research's social ramifications are covered in Chapter 5.
- In Chapter 6, the findings, constraints, and future directions of this study are all discussed.

CHAPTER 2

BACKGROUND STUDY

1.1 Introduction

The scope of the problem, the challenges we faced, a summary of previous researchers' findings, and related research will all be covered in this part. In the section on similar work, we include some research, related studies, their applied methodologies, classifiers, and accuracy levels that are relevant to our study. We create a summary of every study and present it as a table in the research summary section to make it easier to read. We talk about how, given the size of the issue, we can support or further this work. The obstacles and challenges we encountered during our research endeavors are discussed in the final part, along with our strategies for overcoming them.

Rather, we discuss how our study either extends or deviates from prior work. Our differences primarily reflect some background issues in the field, as well as just a difference in target readership. Specifically, we should talk about novel or other innovations that set our research apart: maybe new techniques, or algorithms designed in this experiment for example outside frameworks.

We will then discuss the relevance of our findings and their practical implications. In this section, we will describe how our findings can inform real-world settings or advance current systems/policies. We will also discuss how our research could open future work, pointing out the opportunities other researchers can use to build on top of our results.

Then, we will elaborate on the significance of our general results and also implications for practice. In the discussion, we detail how results may be applied in practice, or improve existing systems/policy. We will further talk about how our research might extend to future studies identifying the opportunities that other researchers may leverage upon in order to sequence on top of our findings.

We also lay out our discussion of the degree to which these limitations accounted for data, and resources as well as describe why we were obligated to grasp till that point., limit or breadth them univariately and multivariate This may involve what future research can do to see beyond the constraints of already collected datasets or how methods and collaboration from other disciplines will address this.

Lastly, the conclusion will recap our findings and what this research adds to the previous one on both text reuse detection as well scientific papers in general. In this way, the section will underscore where our work contributes to a broader field and what may be some of its lasting significance.

2.2 Related Works

We have collected and reviewed many research papers on our topic from various journals. We analyzed and reviewed each paper separately. Among these papers, which are highly related to our work, 10 are mentioned here. What methods have been used in these papers, how much data is there, and what algorithms have been used are discussed below.

K. Solanki [11] do an investigation of self-esteem and general perceived control among first-year college students, for example, revealed a significant positive relationship between these two constructs where GPC accounted for 21.7% of the unexplained variance in SCT. The research was carried out with 99 undergraduate students, and they answered the Rosenberg Self-Esteem Scale (Rosenberg, 1965) together with the Generalized Self-Efficacy Scale - GSE-6 on Google Forms. Supported by the findings, increasing self-efficacy as a way to improve ranged-bullying inhibiting can take its place should be considered an effective intervention argument for those tools such as task-specific interventions particularly in school students experiencing both low teaching background and lack of confidence or at least not confident enough on their ability with somatic video titles. The findings have important implications for how to create a quality educational system and support students in it. However, the results were constrained due to a small sample and some methodologic issues: no control group, and reliance on self-reported data that could be biased. Moreover, the results may not be generalizable to a broader age group than 18–25 years. Validating these results and investigating how broadly applicable they are will require future research.

A. Al-Qahtani et al. [12] highlights the link between health and a child's self-esteem. The piece also looks at gender differences and their stability throughout an academic year. The sample was 351 students at four secondary schools in Trondheim ages 15-21 as of T1 and T2 in the 2016/2017 school year.

The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) measured mental well-being, with the Hopkins Symptom Checklist (HSCL-10) used to measure

depression/anxiety, the Rosenberg Self-Esteem Scale to measure self-esteem, the Adolescent Stress Questionnaire for stress (ASQ-N), and an 11-item pain scale for pain. For statistical analyses, independent samples t-tests were used in addition to multivariate regression analyses, controlling for gender, age, SES (socioeconomic status), stress, and pain. In this study, the girls showed more depression anxiety, and stress, while boys had a higher score for self-concept of mental wellness. These differences were relatively stable: the females ' depression/anxiety Crept up (0.15), while their self-concept improved (1.36). Low self-concept leads to depression/anxiety (β = -.38) and decreased mental well-being (β =.59). Low mental well-being is also correlated with low self-concept (β =.56), supporting both models of prediction. The study concludes that self-concept and mental health are reciprocal relationships, and stresses the importance of gender-specific mental health strategies in school work.

In U.K. Moksnes et al. [13] the author examines based on the Vulnerability model, if self-esteem is low then there may be depression and anxiety; in contrast under the Scar Model poorer moods follow from these disorders finally leading down into perdition without hope of salvation. Altogether, the data consists of 454 subjects. Their self-esteem was measured using the Multidimensional Service Scale, while depression and anxiety were measured with the Psychiatric Scale for Children and Adolescents (SAFA). Two models were tested using path analysis. Model 1 was self-esteem predicting anxiety and depression; Model 2 was anxiety and depression predicting self-esteem while controlling for age and gender. Both models met the criteria of model fit, with Model 1 yielding significant findings, indicating that low self-esteem increases symptoms of depression and anxiety (SBχ^2 = 0.57, CFI = 1.00, RMSEA = 0). A major point is that girls reported higher anxiety levels. Compare Model 2's effect, the more we see the vulnerability model is confirmed. Correlations reflected that self-esteem was significantly negatively related to both depression (r = -.31) and anxiety (r =-.29), whereas depression and anxiety were significantly positively related (r =.83). These findings highlight the importance of promoting self-esteem among adolescents and particularly girls to reduce mental health problems.

In G. Manna et al. [14] there are three separate aspects of self-esteem which are levels, stabilities, and contingency, this research continued to explore how changes in the Psychosocial level will result in changes in academic achievement after being mediated by a general and notable factor: motivation. In Study 1 (N = 600 college students), self-

esteem level (M=3.08, SD= 0.56) and stability (M=3.05, SD= 0.98) had significant negative correlations with self-handicapping (β=-0.22, p<0.01; β =-0.19, p< 0.01). However, contingent self-esteem (M=3.67, SD=0.52) was positively associated with effort management (β=0.11, p<0.05). The indirect effect of contingent self-esteem on performance through effort management was significant (β=-0.04, p<0.05). In study 2 (N=1052 students) it was found that of the three factors tested: worry about tests, contingent self-esteem, and fluctuation in one's feelings of self-worth, would tend to drive up at least one-factor measurement for each such kind of thing. The worry component of test-taking anxiety scores is positively correlated with self-esteem instability (M=2.60, SD=0:89) and contingent self-esteem (M=2.71, SD=0:76). Worry mediated the effect of contingent (B=0.07, P<0.01) and fluctuating self-esteem (B=0.04, P<0.01) on math grades. The model fit in both studies was excellent (CFI = 1.00, RMSEA = 0.00).

In H. Kärchner et al. [15] the authors included a total of 515 college students (258 females and 257 males; 22.54 for their average age years of age with a stD of 2.89) We analyzed how personality traits influence self-esteem; however, whether or not there is a mutual relationship between people´s traits and is still uncertain. Only in the case of extraversion did self-esteem have an appreciable positive correlation. In contrast, although rank orders are low if factually consist of many scores that would also create an appreciable statistical relationship between self-esteem and agreeableness (r = 0.15, p < 0.001). A significant positive correlation was discovered between self-esteem and conscientiousness (r = 0.11, p < 0.05). A significant positive correlation was found between self-esteem and openness (r = 0.38, p < 0.001). However, neuroticism was negatively related to self-esteem (r = -0.35, p < 0.001). Personality traits were found to account for 27 % of the variance in self-esteem (F (5,509) = 37.08, p < 0.01) as evaluated using a regression analysis. Extraversion (β =.29, p <.001) entirely accounted for the difference, while all other variables (− 0.18 <0 n v <1.00) tended not to reach statistical significance. In addition to being a negative predictor of self-esteem (β = 0.52, p < 0.001), neuroticism (β = 0.14, p < 0.0001) was also inversely correlated for all of the five remaining independent variables. There were gender differences in conscientiousness (p < 0.05) and neuroticism (p < 0.05). On both traits, women received higher scores.

In E.Varanarasamma et al. [16] self-esteem (SE) variability and depression are linked by predictors of depression among 139 first-year College students (mean age = 20 years) are examined. SE was indexed twice daily, using the Rosenberg Self-Esteem Scale, for 14 days. SE variability was calculated from these scores' standard deviation. Key predictors included daily happenings, negative generalization, and defectiveness schema. A regression analysis showed that negative generalization (β =.25, p =.001) and adverse relationship events (β =.27, p =.003) were significant variables in predicting SE variability, accounting for 28% of the variance (F (5,130) = 10.12, p =.000). As for interaction effects, relationships between general negativity (β=.17, p=.03) or, equivalently, defectivity (β=.18, p=.02) on negative events were significant predictors of variability in SE. Negative generalization was linked to loss of SE (r=-.19, p=.029) on days with adversities, but on other days led to gain instead. Greater SE variability was found with previously depressed persons, along with mediation analyses indicating that negative generalization (β=.27, p=.008) and for this reason adverse relationship events accounted for. These findings suggest that cognitive styles and stresses in relationships lead to SE levels, which in turn can raise the chances of becoming depressed.

A. Ghartappeh et al. [17] investigated the relationship among self-esteem, humor styles, and emotional affect in 440 undergraduate students (77% female; average age = 20.33 years). The tools used for the data collection were the Rosenberg Self-Esteem Scale, Humor Styles Questionnaire (HSQ), and Positive and Negative Affect Schedule (PANAS). Data were analyzed by using Pearson correlation, also carried out with multiple regression. Experiment 1 thus showed that affiliative humor, as well as positive affect and (to a lesser extent) self-enhancing humor were associated with higher levels of state self-esteem. 001. They found a negative correlation between self-esteem and aggressive humor (r = -0.11, p <. 05), self-defeating humor (r = -0.21, p <. 001), and (3) negative affect r = -0.20, p <. 001). Self-esteem predicted 3.1% of affiliative humor (R^2 =0.031, β=.176), 5.7 %of self-enhancing humor (R^2=0 .057, β =0 .239), 1·2 % of aggressive humor ((R squared nary) =011; Beta=-107) and it explained4 ·19; certain variance in self-defeating humor (R Squared for linear equation corrected for number or predictors adjusted R-squared) (jgnl: RT>|> ('CTE'D f (t (-208)). It additionally accounted for 14.7% of positive effects (R^2 = 0.147, β = 0.384) and by only .2%. Some or other negative effects were mediated (R^2=. =03204204. Findings also

indicate that self-esteem significantly affects adaptive humor styles and emotional well-being, highlighting its importance in ameliorating positive affect and mitigating negative experiences.

A. M. Hayes et al. [18] is conducted by the authors to determine self-esteem and life satisfaction in university students (233 females, 172 males) of Selcuk University/ Turkey within social problem-solving skills. Participants (average age=20.45; SD = 1.52) The Rosenberg Self-Esteem Scale (RSES), Satisfaction with Life Scale, and Social Problem-Solving Inventory-Revised were used to assess self-esteem, life satisfaction, and problem-solving skills respectively. Self-esteem and life satisfaction were positively related to Positive Problem Orientation (PPO) and Rational Problem Solving, whereas they were negatively among Negative Problem Orientation (NPO), Impulsivity/Carelessness Style (ICS), and Avoidance Styles of SPSI-R according to the study results. ($p <. 01$). Results of regression analysis revealed that NPO accounted for the 13% variance in self-esteem ($R^2 = 0.13$, $p <. 001$). Exclusion of PPO reduced the explained variance to 14% ($R^2 = .146$, $p < \times. 001$); adding AS contributed 19 % additionally ($R^2 = .19$, $p <. 05$). RPS explained 10% of the variance in Life Satisfaction (adjusted $R^2 = 0.100$, $p <. 001$), and an AS that accounted for 13% ($R^2 = 0. 001$), with the addition of PPO increasing it to 15% ($R^2 = 0.15$, $p <. 05$). These results specifically demonstrate how increased problem-solving skills, specifically a positive outlook and rational strategies add to self-esteem ad life satisfaction.

The impact of self-esteem and self-efficacy on women's empowerment has been examined by Z. Ozyesi [19] in the survey of 5587 female academic, and administrative staff across 15 governmental universities that are independently affiliated with universities or colleges. A cross-sectional design using multistage cluster sampling was adopted in the study where participants with a mean age of 36.4 (SD = 8.19) were included, mostly married (65%) and residing in urban areas (80.5%). Participants completed the Rosenberg Self-Esteem Scale, General Self-Efficacy, and a Women's Empowerment measure including personal, social/relational empowerment, and environmental/workplace women empowerment. 50.2% of the participants had high self-esteem and 49.8% had medium self-esteem, in addition, 66.9% of the participants had high self-efficacy and only 86.8% showed adequate overall women empowerment. Based on multiple linear regression, we found that self-esteem ($B = 0.521$, $b = 0.127$, t value =13. 001) on the other, and self-efficacy ($B = 2.388$; $\beta = .702$; $t (32533) = 76.049$,

p <. In this context, the Self-Help Group (β = .001) predicts women's empowerment significantly. Self-efficacy was the strongest predictor, accounting for 73.4% of the total variance of empowerment levels. The results indicate the role of self-esteem and creativity in improving women's empowerment within Saudi Arabia's Vision 2030.

E. Hamarta [20] investigated 419 overweight and obese girls living in Kermanshah-Iran. The study was intended to evaluate the association between self-esteem, mental health, and sleep quality with their QOL. Of the students, 58.5% were overweight and 41.5% obese as per BMI which worked out to an average of BMI at. The study, which used Rosenberg's Self-Esteem Questionnaire, the Goldberg General Health Questionnaire, and the Pittsburgh Sleep Quality Index found 11% of the variance in QOL was explained by self-esteem score; sleep quality at 20%; and mental health at a solid quarter. The average self-esteem score was 0.37 ± 0.56, a mental health score of 32.37 ±11.78 and sleep quality was calculated at the total points were added to calculate scores within themselves related statements on similar lifestyle factors (6·05±2·69). The QOL score at an overall level was 82.27 ± 14.47 and MAJOR IMPACT on physical health (25.74 ± 4.88) followed by mental health ((19+.87±4+.71). Mental health and the quality of sleep had a significant influence on all QOL dimensions (P < 0.01), while self-esteem played an important role in social relationship (P > 0,005) and environmental health domains at multivariate level through multiple regression analysis. Mental health was an important factor in improving QOL, which has been validated with a targeted mental health intervention requirement.

2.3 Comparative Analysis & Summary

Some work on the detection and attribution of self-esteem level detection has already been completed with the use of machine learning algorithms and data mining techniques. In the present era, the usage of devices for learning has increased along with the detection of various ailments and self-esteem levels. Table 2.1 provides an appraisal of several research subjects, approaches, and findings.

Table 2.1: An Overview of Related Research Works

Serial Number	Authors Name	Dataset	Methodologies	Description	Result
1	K. Solanki	99	Examine the relationship between self-esteem and self-efficiency using statistical tools like Pearson's correlation and linear regression.	The General Self-Efficacy Scale was utilized to measure self-efficacy and Rosenberg Self-Esteem Scale was utilized to measure self-esteem.	With a moderate accuracy rate of 77.5%, self-efficacy predicts self-esteem.
2	A. Al-Qahtani et al.	5,587	Used Cross-sectional design. SPSS Version 23.0 was used for statistical analysis.	The predictive link between women's empowerment self-esteem and self-efficacy was examined using multiple linear regression.	Self-efficacy and self-esteem taken together could account for 73.4% of the variation in women's empowerment.
3	U.K. Moksnes et al.	351	Used the Vulnerability Model for self-esteem as a predictor of mental health and the Scar Model for mental health as a predictor of self-esteem.	Descriptive statistics were used to present the socio-demographic data and levels of self-esteem, mental health, depression, and stress.	The vulnerability model had higher accuracy with 435 for depression/anxiety and 39% for mental well-being.
4	G. Manna et al.	454	Used Multidimensional Self-Concept Scale for self-esteem and Psychiatric Scale for Children and Adolescents for depression and anxiety.	Two models were used. Model 1 for self-esteem predicts anxiety and depression and Model 2 for anxiety and depression predict self-esteem.	Both models demonstrated a good fit with CFI = 1.00 and RMSEA = 0 indicating the models fit the data very well.
5	H. Kärchner et al.	1652	Used Mplus 7.4 in both studies and nested data structures were taken into account.	Academic performance was mediated by Self-Handicapping and effort management for Study 1, and Test Anxiety for Study 2.	For both studies, the model fit was excellent as suggested by good agreement between data and what would be expected CFI = 1.00 and RMSEA= 0.00.

6	E. Varanarasamma et al.	515	Correlation and regression analyses were used to test the relationships of variables represented by the Big Five personality traits.	The participants were informed that there was no right or wrong answer and their confidentiality would be protected.	Self-esteem was associated positively with extraversion, agreeableness, conscientiousness, and openness on one hand and negatively with neuroticism on the other.
7	A. Ghartappeh et al.	419	Used Linear Regression, multiple regression, and stepwise regression.	Predict the Quality of Life (QOL) of obese and overweight girl students.	The model's accuracy represented by the coefficient of determination was 30%.
8	A. M. Hayes et al.	139	Used Regression Analyses to determine the contributions of hypothesized predictors.	The research was able to look into the ways that various life events and cognitive processes affect changes in self-esteem.	The R2 value of 0.28 reported in this study suggests that these predictors account for about a moderate proportion (i.e., 28%) of variance in self-esteem.
9	Z. Ozyesi	440	Used Pearson correlation coefficients and multiple regression analysis.	Explored how self-esteem correlates with humor style and affect.	For Humor Style the model explained between 1.1% to 5.7% of the variance and for Positive-Negative Affect the model explained 14.7% of the variance.
10	E. Hamarta	405	Pearson correlation coefficient and Stepwise regression analysis.	Try to assess the predictive role of social problem-solving on self-esteem and life satisfaction.	The Stepwise regression analysis model explained 19% of the variance for self-esteem and 15% of the variance for life satisfaction.

These days, deep learning and artificial intelligence (AI) are extensively employed for expectation, classification, and discovery models in all fields of information science. The discovery models make use of numerous widely-known computations, including computed relapse, convolutional neural networks (CNN), artificial neural networks (ANN), SVM, and k-NN. The k-NN, NB, SVM, and DT algorithms have high values

and are frequently employed in forecasting, finding, and acknowledging models, according to the findings of our writing survey. In our study endeavor, we endeavored to perform computations employing k-NN, Decision Tree, Naïve Bayes, Gradient Boosting (GB), Support Vector Machine, Logistics Regression, Extreme Gradient Boosting, and Random Forrest classifiers to identify the individuals in Bangladesh who lack motivation.

2.4 Scope of the Problem

Our investigative work involves employing machine learning algorithms and data analysis to create a display. Our proposed demonstration can recognize sadness. This endeavor will significantly impact society's citizens. People who struggle with mental clutter are prevalent in our culture. Though they are not aware of it, they are discouraged. They don't have the required information and isolate themselves. This leads them to choose to misbehave.

We require a framework, in this case, to help them recognize their issue and assess whether or not they are demoralized. Lastly, give them information so they will know what needs to be done. Recently, the results have been quite good as machine learning and fake insights have been used for various question-finding and sickness projections. Therefore, we decide to use machine learning to create a display of the discouragement location.

To that end, we suggest flagging patients with signs of depression using machine learning models based on data. These models can use massive amounts of data about someone — for example, the text they input into a computer or the way their voice sounds on a recording to pick up faint traces of emotional distress that people might not be aware of in themselves. It acts as an early warning system, tracking major symptoms benchmarks for depression — mood swings, loneliness, and conversational shifts among them. Once determined, the frame will not only alert users of likely depression but also provide personalized advice and interventions.

The impact of this project is transformative in that using a simple tool can make someone aware of his/her condition, and actions could be taken before the worker becomes sicker. We need a proactive solution to mental health management and integrating machine learning allows for that, providing the knowledge of when help should be sought out.

2.5 Challenges

During our investigation, we often encountered challenges and the confidence of both experienced examinations, just as frequently stepping on changing points. Collecting authentic data was one of the most challenging parts. It bored down to the point that it was a battle; who is “healthy” and whose depression is acting up. Those (who disclosed) mental health issues often declined to participate in the data collection process. After trying to connect with any group, multiple articles, and consults from different doctors we couldn't get a straight solution. It was difficult to tailor the survey tools because of limitations associated with depressive symptom variability in this age range and there were fewer responses than expected due to privacy concerns.

The engineering side of things, as well as some machine learning algorithms, were complete unknowns at the beginning and we encountered a lot more problems here. Some days, we felt like complete failures not able to understand any of the tough concepts. Nevertheless, under the patient instruction of our supervisor and after some trial-and-error we managed to overcome these challenges together. Even if it was sometimes a tough process that strengthened my nerves and also often challenged me to some extent, in the end, every challenge brought us forward in our knowledge and guts as we had one more part of this project successfully overcome.

We certainly did not go away; however, along the way, every challenge became a lesson. We quickly adapted and were more resilient in our new approach. But those data collection challenges ended up serving as useful nudges to think more inventively about workarounds. Our starting point was to consider public data sets, anonymized datasets, and literature around detecting or inferring depression indirectly. By examining social media content or behavior-based models. Moving forward with this new strategy helped us to take pride in the fact that we had recommitted ourselves to a worthy cause but not at an unsustainable cost of compromising individual privacy.

Initially, the machine learning algorithms seemed insurmountable to us and felt like something that only a select few could comprehend, but as we worked harder on it, things started slowly making sense. The learning curve was hard, but every step of the way it felt like getting a success to tell me that I could indeed learn and do something. In due course, we reaped the fruits of our labor, and that in itself made us believe more.

Overcoming these hurdles is not only how we managed to advance our research but also how, grew as people. Every problem solved was a living tribute to endurance and the empowerment of overcoming obstacles. What we learned is that they were not limits on what was possible, but limitations in our imagination of where forward lay. In hindsight, every time we doubted ourselves, each of those moments added to the foundation for more confidence and self-respect in terms of our technical expertise and what it is that we want from this project.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This study aims to create a model for self-esteem detection in people. It is built on common behaviors and reactions at different levels of self–esteem. We use Logistic Regression (LR), Random Forest (RF), k-nearest Neighbors (k-NN), Gradient Boosting (GB), Decision Trees (DT) and Naive Bayes (NB) algorithms. The algorithms will classify people as High, Moderate, or Low self-esteem. We found twenty-four characteristics corresponding with self-esteem, either directly or indirectly. Before we could get to the final implementation of the model, we had some necessary dataset preprocessing to do. We also computed precision, recall, and F1−score, and plotted Receiver Operating Characteristics (ROC) curves, to determine the best-performing model.

The model itself refined the dataset towards a wide array of behaviors and psychological signals believed by literature to be correlated with self-esteem. Traits were chosen for analysis for their relevance to self-esteem from the fields of psychology and behavioral science. For example, characteristics including confidence, frequency of social interaction, risk-taking behaviors, and self-reflective responses were combined as possible proxies.

After pre-processing the data by establishing procedures to normalize and clean the data to further eliminate inconsistencies or bias, feature selection techniques were performed to determine significant predictors of self-esteem levels. The purpose of this step was to simplify the model without sacrificing performance to make an optimized model.

The finalization of the dataset led us to run a few algorithms to check how useful they are for classifying the self-esteem level. To ensure robustness and persistence on new data (avoiding overfitting), we employed k-fold cross-validation for each model.

Our first results with each algorithm revealed significant differences in their performance, both for sensitivity, specificity, and accuracy. However, Random Forest and Gradient Boosting are great models to be able to capture small differences between Moderate and High self-esteem. Even though Logistic Regression and k-NN are simpler algorithms, they were stable as well and achieved an F1-score above 0.90 at all

points in the tuning process. We tuned parameters which include learning rates, tree depths, and regularization terms with grid search and randomized search to get the best performance. Last, ROC curve analysis was used to characterize each model's output.

To summarize, this study used an iterative process, explored various machine learning methods, and gave a detailed feature-extraction analysis to create a generalizable model capable of accurately identifying self-esteem levels. These findings could help power mental health, personalized therapy, and self-development applications, providing a data-driven approach to comprehend and improve your self-esteem.

3.2 Data Collection

The collection of information includes many different elements or facets that are either directly or indirectly connected to self-esteem. Since the government had informed them that sharing patient information could violate patients' privacy, affect their policies, and occasionally not be available as a prepared dataset, we were unable to gather the necessary data while we were traveling to the clinic. As a result, we decided to create our dataset utilizing data collected through online government papers, in-person surveys, and a paper with a list of questions. With any luck, we were able to compile 1050 human facts based only on 13 elements that are common to human behavior. The primary difficulty emerged once all the data was gathered: the data might be categorized as low, normal, and high.

As we seek guidance and spot trends to solve problems, we refer to a range of papers. With the assistance of a doctor, a physiatrist, and a psychology student, the facts were ultimately leveled independently, with an emphasis on consensus, and their three alternatives were included in the final leveling outputs. Among the one fact, there are 10 people with high self-esteem, 534 with normal self-esteem, and 506 with poor self-esteem. There are differences between secondary and higher secondary schools, colleges, and universities, according to the data we collected via an online survey.

We collected the data based on the below topics:

1. Age.
2. Gender.
3. Highest education level.
4. Satisfaction level.
5. State of mind.

6. No. of good qualities.
7. Confidence level.
8. Proud level.
9. Performance level.
10. Compare to others.
11. Respect level.
12. Success or failure.
13. Positive attitude.

We came up with the questions by reading studies and articles and speaking with physicians and physiatrists. To observe the differences, we separated self-esteem levels into three groups. Details of the procedure are given below.

3.3 Research Subject and Instrumentation

These days, deep learning, data mining, and system mastering algorithms are quite suitable and well-liked for all types of detection, recognition, and prediction. To ascertain which system-mastering algorithms would best meet our requirements and function well, we tried testing a range of them on the dataset we had obtained. Some of the system mastering methods we used were LR, RF, k-NN, GB, DT, and NB.

In recent years, "Python" has emerged as one of the most well-liked and extensively used programming languages, and scholars regularly employ it for research purposes. As a result, we used Microsoft Excel as our dataset, Google Colab, Google Forms, and Python as our programming language.

Expanding on this, we were able to build on this foundation using Python's framework of data science libraries such as Pandas for data manipulation, Sk-learn for the implementation of machine learning algorithms as well as Matplotlib and Seaborn for data visualization. This helped us identify the patterns and relationships associated with low or high self-esteem. we were restricted to these particular tools for in-depth analysis and some sort of visualization.

Once again, Google Colab specifically was great for this as they gave us interactive environments to code and visualize results on the fly. These can also be easily shared with other researchers, which allows for collaboration using these platforms as well as for sharing results and visualizations. Using Google Forms, we gathered responses

quickly from various types of participants, which was especially useful for collecting data in real-time.

In the preprocessing stage where data was gathered and cleaned the input needed to be sufficient data with its quality and consistency maintained. This includes dealing with missing values, normalizing the numerical features, and encoding the categorical values. Once this preprocessing was done, exploratory data analysis (EDA) helped us learn more about what the shapes of our data look like, where correlations might be found, and feature distributions. Feature selection using this step was important as it was aimed towards determining the best features that could increase the model accuracy and model interpretability.

After preprocessing and analyzing the data, we ran each machine learning algorithm on the dataset to validate its suitability. To prevent overfitting and have a model that works on unseen data, each model was trained and validated with cross-validation techniques. Relevant metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve are used to evaluate each algorithm, resulting in a comprehensive understanding of the strengths and weaknesses of each model.

Based on these initial evaluations, we were able to uncover a few algorithms that performed above and beyond the rest with regards to high accuracy along with meeting the metric stabilities with again a couple of methods. Next, hyperparameter tuning was performed to further tune the performance of each algorithm. As an example, for Random Forest and Gradient Boosting, the number of estimators was optimized.

By utilizing a variety of programming tools and platforms, along with systematic machine-learning routines, we were able to analyze, process, and model our data. Knowledge generated by this research may open the door to more advanced self-esteem classifiers which could benefit mental health and general fields related to psychology and personal development.

3.3.1 Proposed Methodology

In Figure 3.1 our steps of the proposed methodology are shown:

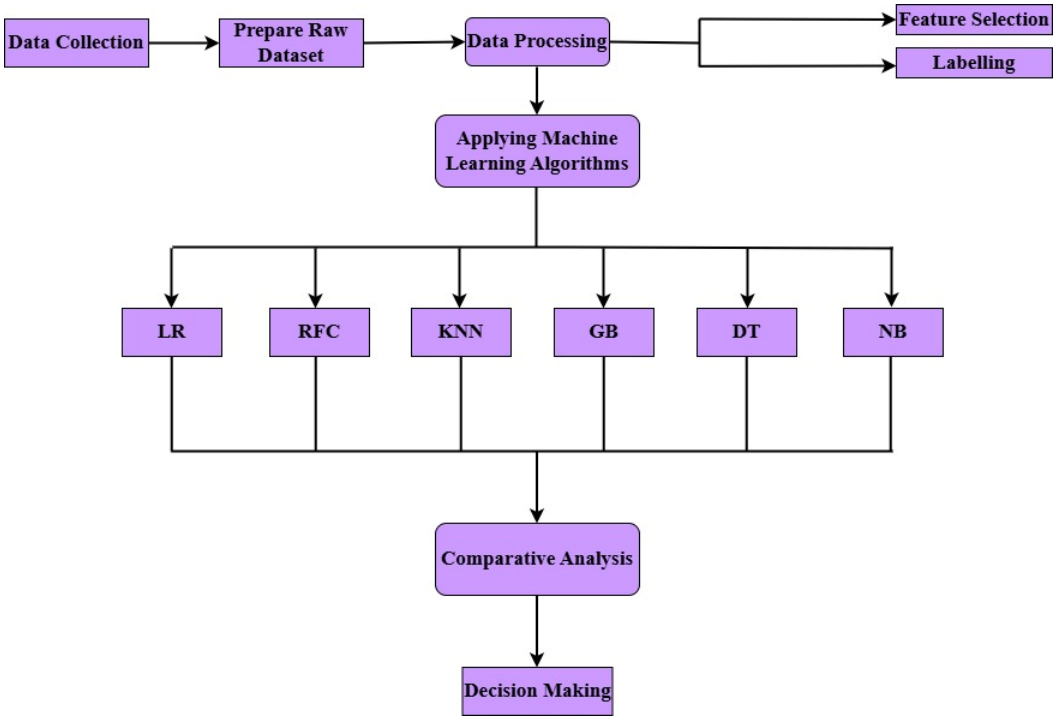


Figure 3.1: Steps of Proposed Methodology

A structured, multi-pronged approach towards developing a machine learning model for self-esteem detection that is being proposed is given below:

It all starts with Data Collection, where it collects whatever data is relevant. It could be responses to surveys, aspect of questionnaires, or other strengths-based assessments that are meant to be pointers to self-esteem. We will call this first dataset data and it becomes the basis of the model.

The next step after data collection is to Prepare the Raw Dataset. This stage is concerned with organizing and formatting raw data and making sure that it is structured and ready for processing. This could be preparing the data by eliminating any errors or inconsistencies.

After preparing the raw dataset, it is then subjected to Data Processing. The operations in this phase can get really detailed including handling missing values, normalizing numeric data, encoding categorical data, etc. Data processing is the main step for

assuring that the data is correct and that the data goes into a clean state that makes the data ready for correct modeling.

After we process the data, Feature Selection is done. The identified features/traits which are contributing or are directly related to predicting self-esteem levels are the most relevant features, in this step. Such features can be selected through statistical significance or domain knowledge because they are the best predictors of high (non), and (moderate) low self-esteem respectively.

Once the features are selected, we perform labeling which means to differentiate between data. Each of the instances in the dataset is categorized according to the degree of self-esteem reflected from it, namely High, Moderate, and Low. The labels are necessary for the supervised learning algorithms, which learn from examples that have been labeled. Now that we have the labeled and preprocessed data, the next step involves Applying Machine Learning algorithms. Predictive models are created using several machine-learning techniques.

Once we process this data, we perform Feature Selection. At this step, the most relevant features are those identified features/traits that are contributing to or are directly related to the prediction of self-esteem levels. Features like these can be chosen based on the statistical significance or domain knowledge as these are the best predictors of high (non), (normal) low self-esteem respectively.

After selecting the features, we come to the labeling process that is we have to label the data. Each of the instances of the dataset falls into three classes by reflecting the measure of self-esteem from it i.e., High, Normal, and Low. These labeled examples are required for the supervised learning algorithms to learn from.

Having the labeled and preprocessed data set, now we will proceed to Applying Machine Learning algorithms. Several machine learning methods generate predictive models.

3.3.2 Data Preprocessing Process

When we successfully compile an abundance of data, we usually find that there are multiple data types, including numerical and categorical data, and that some of the data contains missing values. This type of knowledge is not appropriate for machine learning approaches. Consequently, we frequently arrange our knowledge to fit our tastes and

make it algorithm-compatible. Processing can convert knowledge into acceptable representations after it has been collected. data that has undergone more precise processing to help generate the greatest outcomes as soon as possible.

The data preprocessing process is shown in Figure 3.2 below.

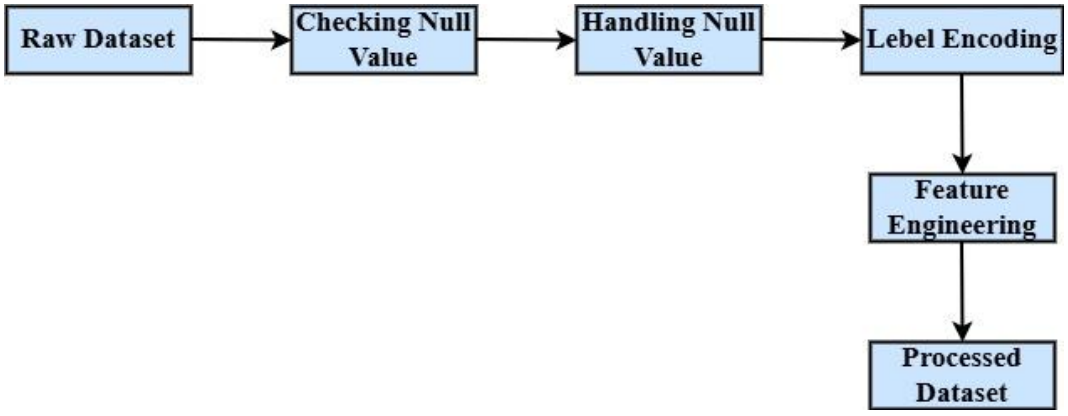


Figure 3.2: Data Preprocessing Process

To begin our job, we first gathered and produced our raw dataset. After that, we concentrated on tidying up the data. We checked to see whether there were any null or missing data in the statistics set. The section of the text or specialized statistical data that transforms everything into pertinent numerical statistics was then encoded. Standardization was then used to finish the statistical transformation. We then determined and investigated the relationship between several aspects to achieve the function engineering aim. Then, retaining the other two features—gender and age—as suggested, we determined which six out of the thirteen features were the best. Consequently, we were able to acquire the final processed dataset of 18 characteristics, which included both age and gender features. All of the work required to process the data was done in Google Colab.

3.4 Statistical Exploration

We gathered data from our poll of 1050 students across various levels and categories. The information we collected included students from schools, colleges, and universities. People's levels of low, normal, and high self-esteem are depicted in Figure 3.3. We constructed our project, worked on it, and completed all other procedures. Our research revealed that 10 people had high self-esteem, 534 had normal self-esteem, and

506 had low self-esteem. About 48.2% of persons have low self-esteem, 50.9% have normal self-esteem, and 1.0% have high self-esteem.

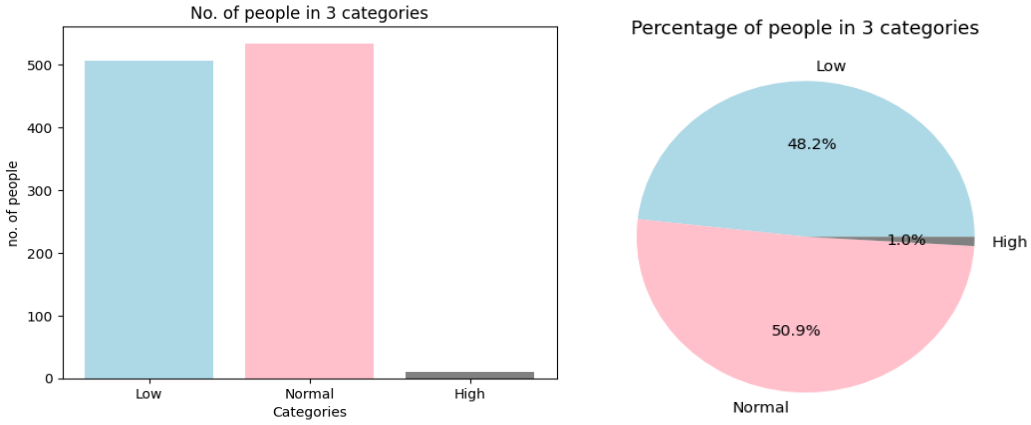


Figure 3.3: Low, Normal & High Self-Esteem Cases

The number of males in each of the three categories is displayed in Figure 3.4. There were 610 men in all. Of them, 280 had low self-esteem, 324 had normal self-esteem, and 6 had high self-esteem. In terms of proportion, 1.0% was high, 53.1% was normal, and 45.9% was low. The figure below is 3.4.

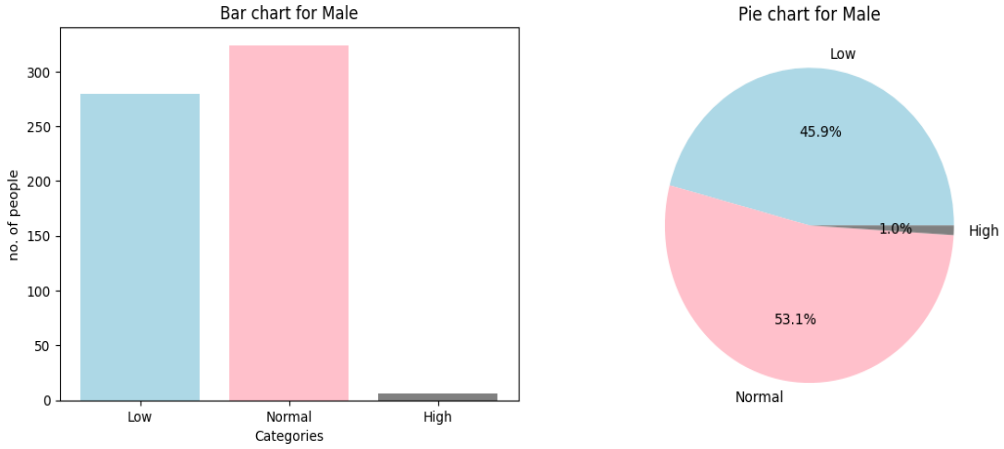


Figure 3.4: Male's Self-Esteem Levels

The number of females in each of the three categories is displayed in Figure 3.5. There were 423 women in all. Out of them, 216 had low self-esteem, 203 had normal self-esteem, and 4 had high self-esteem. 51.1% of the population was low, 48.0% was normal, and 0.9% was high. The figure below is 3.5.

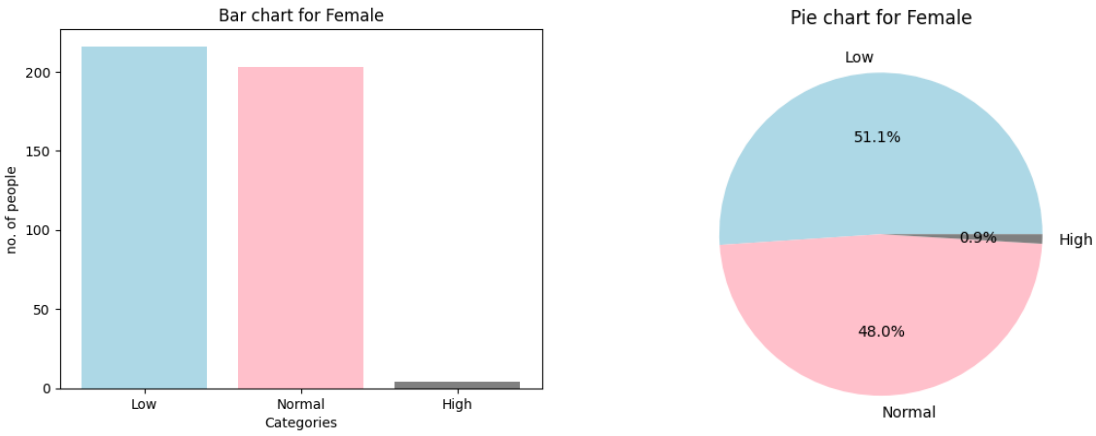


Figure 3.5: Female Self-Esteem Level

Figure 3.6 shows the number of prefer not to say in three categories. There was a total of 17 prefer not to say. Of them 10 were low, and 7 were normal self-esteemed persons. In percentage 58.8% were low, and 41.2% were normal Self-esteemed persons. Figure 3.6 is shown below.

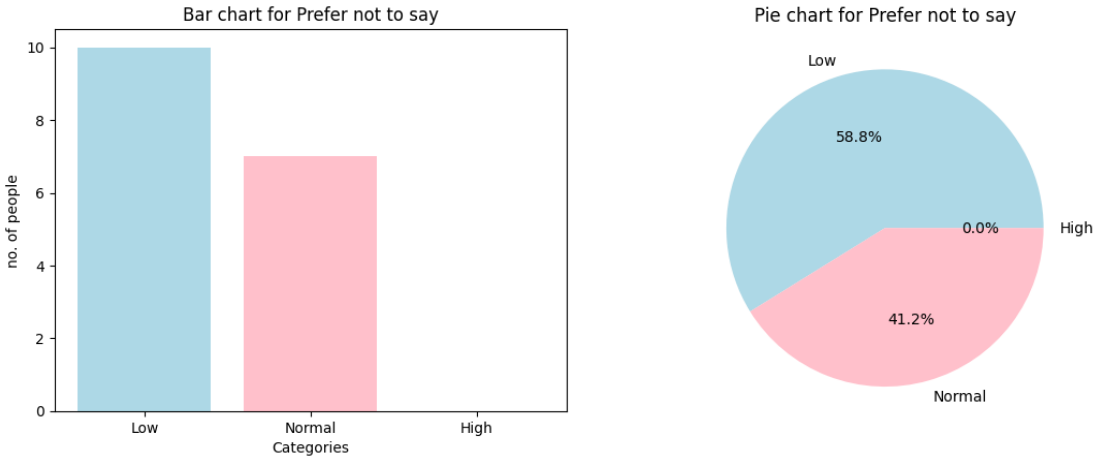


Figure 3.6: Prefer Not to Say's Self-Esteem Levels

Figure 3.7 shows the number of less than 18 in three categories. There was a total of 49 less than 18. From them 21 were low and 28 were normal self-esteemed people. In

percentage 42.9% were low and 57.1% were normal self-esteemed persons. Figure 3.7 is shown below.

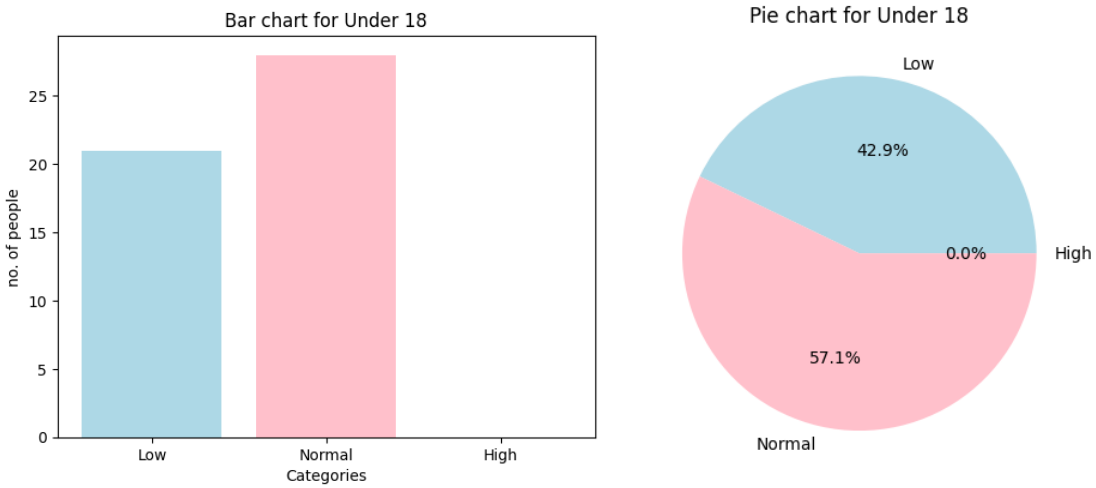


Figure 3.7: Less than 18 Self-Esteem Levels

Figure 3.8 shows the number of ages 18-30 in three categories. There was a total of 575 ages 18-30. Of them 279 have low, 286 have normal and 10 have high self-esteem levels. In percentage 48.5% were low, 49.7% were normal and 1.7% were high self-esteemed individuals. Figure 3.8 is shown below.

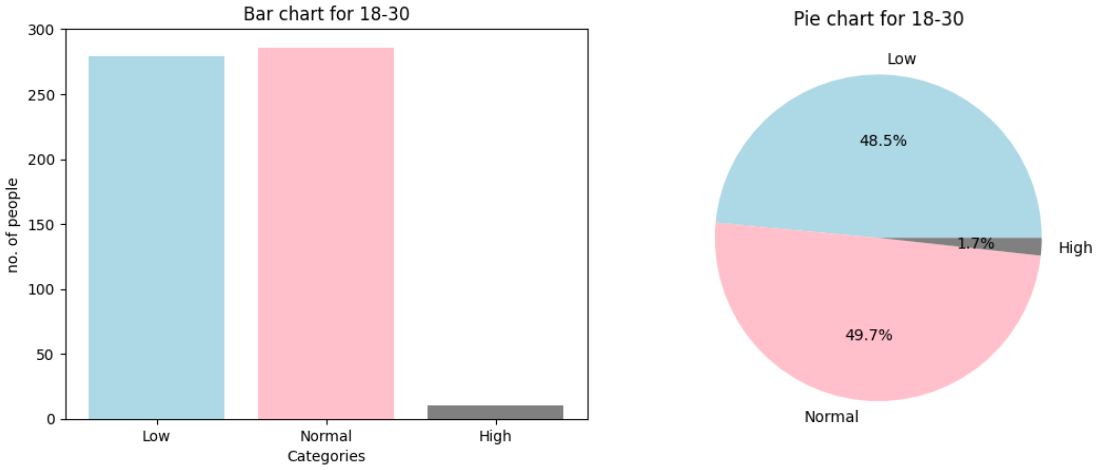


Figure 3.8: 18-30 Self-Esteem Levels

Figure 9 shows the number of ages 31-45 in three categories. There was a total of 416 ages 31-45. Of them 201 have low, and 215 have normal self-esteemed persons. In percentage 48.3% were low, and 51.7% were normal self-esteemed individuals. Figure 3.9 is shown below.

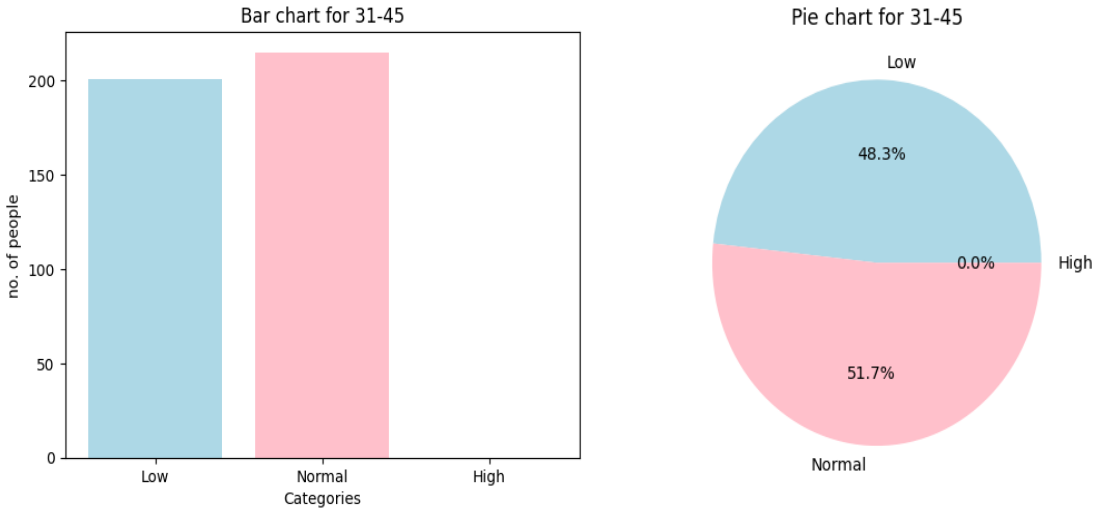


Figure 3.9: 31-45 Self-Esteem Levels

Figure 3.10 shows the number of ages more than 45 in three categories. There was a total of 10 persons ages more than 45. Of them 5 have low, and 5 have normal self-esteemed persons. In percentage 50.0% were low, and 50.0% were normal self-esteemed individuals. Figure 3.10 is shown below.

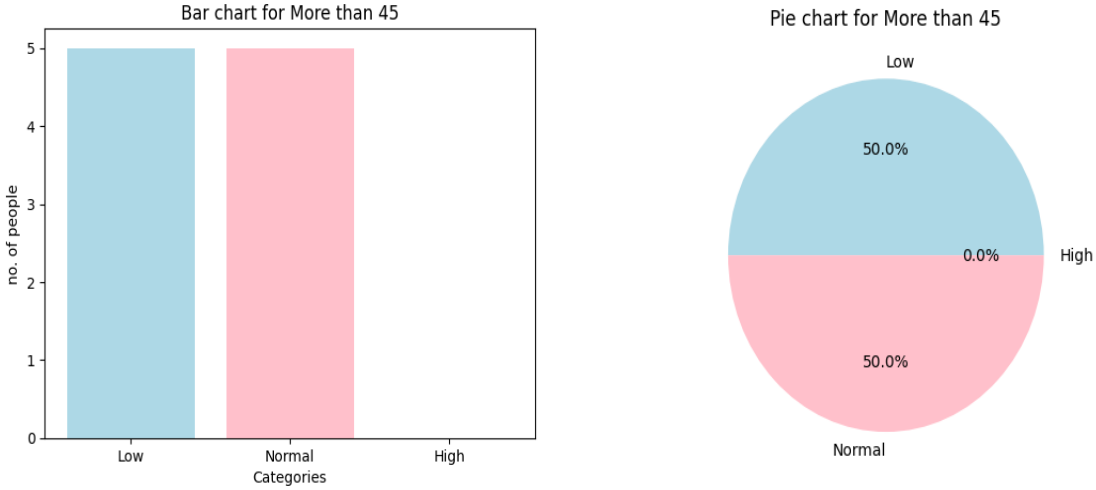


Figure 3.10: More than 45 Self-Esteem Levels

Figure 3.11 shows the number of people with high self-esteem according to gender. There was a total of 10 high self-esteem people. Six were males, and four were females. In percentage, 60% were males, and 40% were females. Figure 3.11 is shown below.

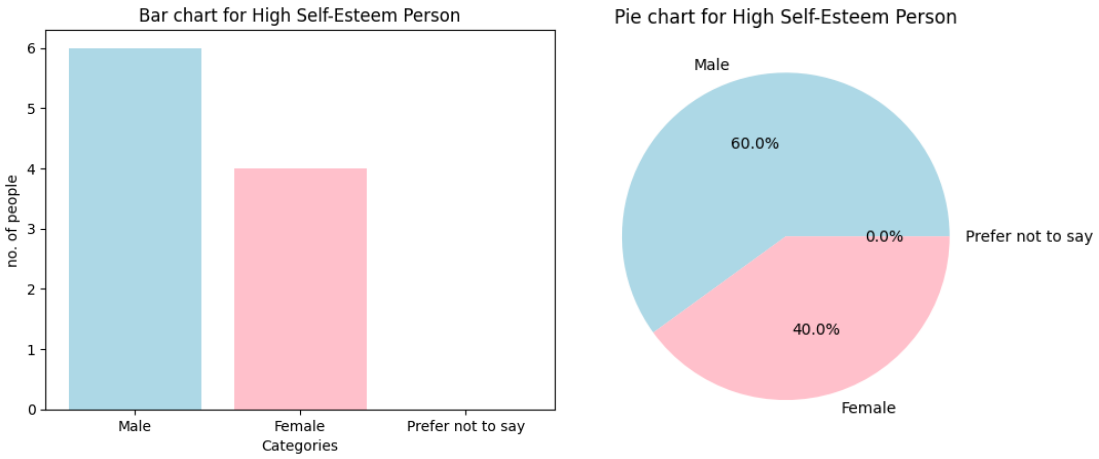


Figure 3.11: High Self-Esteem According to Gender

Figure 3.12 shows the number of people with normal self-esteem according to gender. There were 534 people with normal self-esteem in total. Of them, 324 were males, 203 were females, and 7 were prefer not to say. In percentage, 60.7% were males, 38% were females, and 1.3% were prefer not to say. Figure 3.12 is shown below.

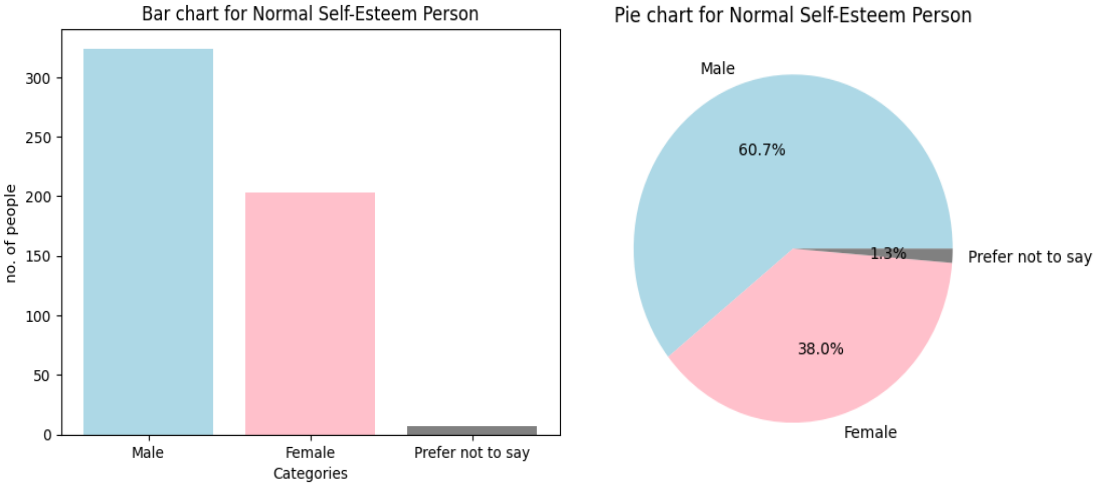


Figure 3.12: Normal Self-Esteem According to Gender

Figure 3.13 shows the number of low self-esteem persons according to gender. There was a total of 506 low self-esteem people. Of them 280 were males, 216 were females and 10 were prefer not to say. In percentage 55.3% were males, 42.7% were females and 2.0% were prefer not to say. Figure 3.13 is shown below.

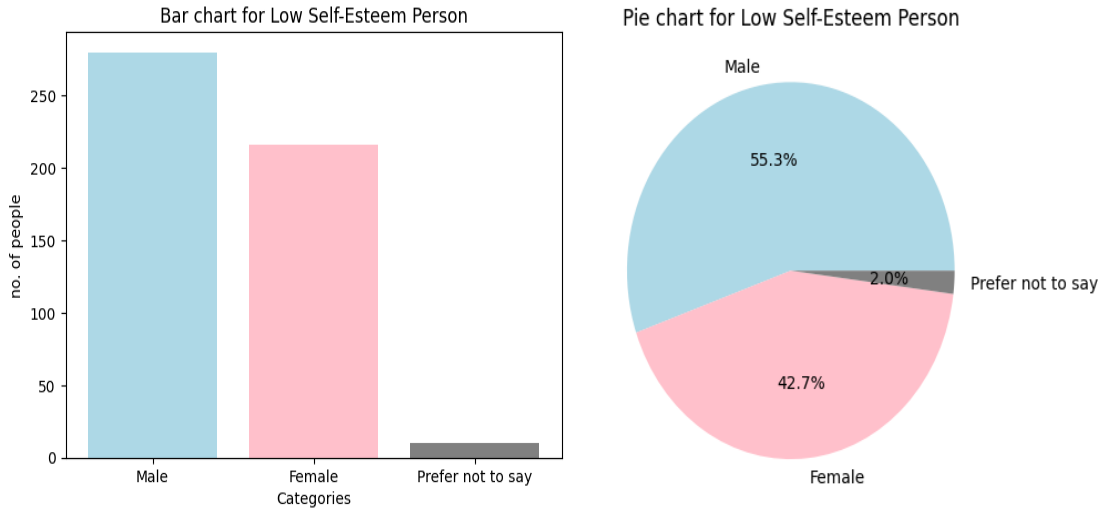


Figure 3.13: Low Self-Esteem According to Gender

Figure 3.14 shows the number of high self-esteem persons according to age. There was a total of 10 high self-esteem people. Of them, 10 were less than 18-30. In percentage, it is 100%. Figure 3.14 is shown below.

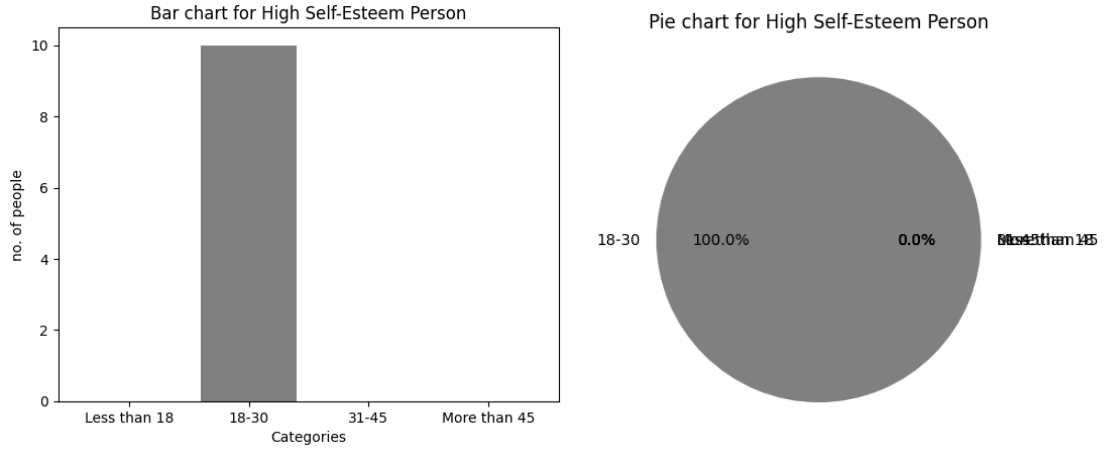


Figure 3.14: High Self-Esteem According to Age

Figure 3.15 shows the number of people with normal self-esteem according to age. There was a total of 534 normal self-esteem people. From them 28 were less than 18, 286 were 18-30, 215 were 31-45 and 5 were more than 45. In percentage 5.2% were less than 18, 53.6% were 18-30 and 40.3% were 31-45, and 0.9% were more than 45. Figure 3.15 is shown below.

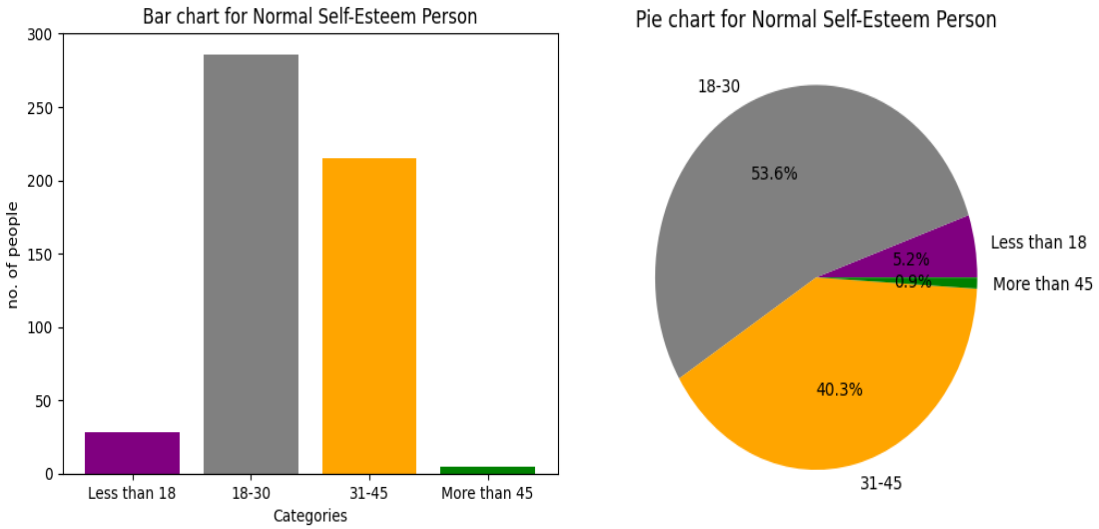


Figure 3.15: Normal Self-Esteem According to Age

Figure 3.16 shows the number of low self-esteem persons according to age. There was a total of 506 low self-esteem people. From them 21 were less than 18, 279 were 18-30, 201 were 31-45, and 5 were more than 45. In percentage 4.2% were less than 18, 55.1% were 18-30, 39.7 were 31-45, and 1.0% were more than 45. Figure 3.16 is shown below.

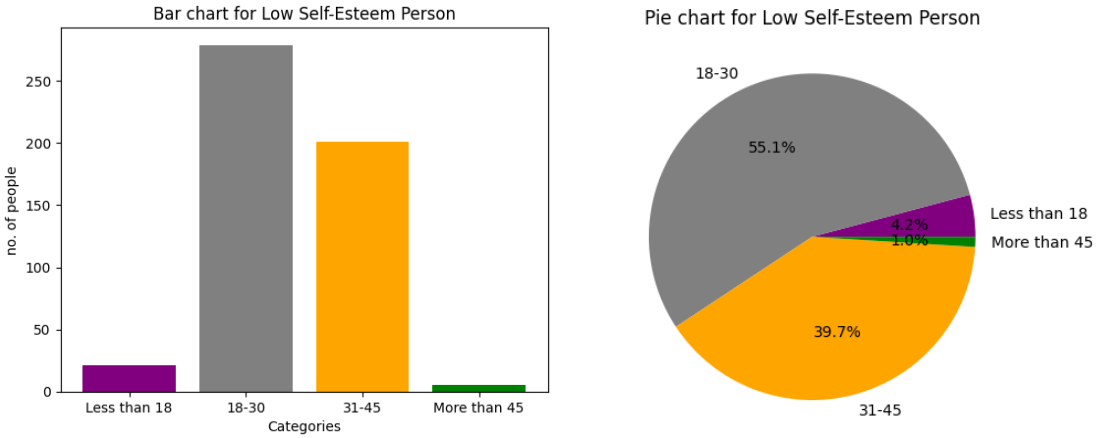


Figure 3.16: Low Self-Esteem According to Age

Figure 3.17 shows the correlation of multiple features. Which is used to find out the similarity rates among them. Figure 3.17 is shown below.

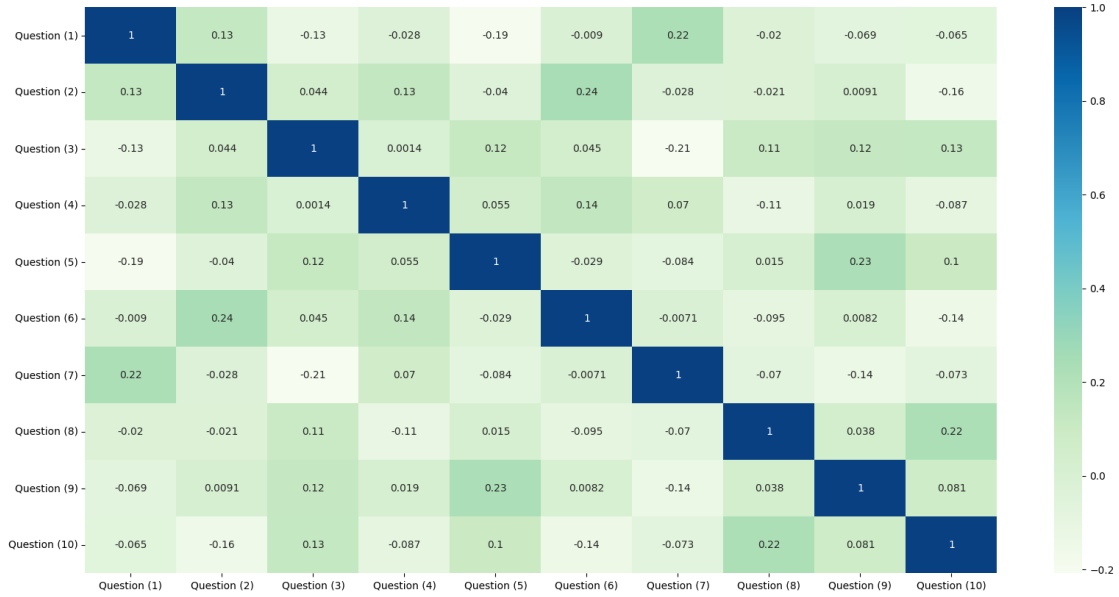


Figure 3.17: Correlation Matrix with Multiple Feature

Correlation heatmap indicating the correlation of responses to 10 questions. Each cell indicates the correlation coefficient for a specific pair of questions. More vivid dark blue squares indicate a strong positive correlation, while progressively paler shades of eco-friendly to light great indicate a much more delicate connection. Light green indicates a negative value, and white areas suggest the correlation is near 0. On the diagonal, each question has a self-correlation of 1. This leads to a heatmap that shows how respondents tend or more likely to answer together across questions.

In this example the heatmap is used to highlight correlations between pairs of ten survey questions, making it easier for us to determine if or not they are correlated visually. The color scale goes from dark blue to light green and allows for quick identification of which questions are more closely related. STATA Example If, for a pair of response alternatives that should correlate with one another, e.g., measuring different aspects of the same thing, Responses with values close to 1 indicate that respondents replied similarly as they responded to both questions. On the other hand, negative correlations indicate that a subject scores high on one question and low on another, suggesting potentially opposite dimensions. It helps eliminate survey questions that are similar to, or the opposite of, each other a common tool used in practice for optimizing and refining survey designs.

Table 3.1: Feature for Self-Esteem Level Detection

Feature Name	Evidence-based on	Feature Name	Evidence-based on
Age	[21]	Proud level	[21]
Gender	[21]	Performance level	[21]
Highest education level	[21]	Compare to others	[21]
Satisfaction level	[21]	Respect level	[21]
State of mind	[21]	Success or failure	[21]
No. of good qualities	[21]	Positive attitude	[21]
Confidence level	[21]		

Age, sex, education level, and other items such as the degree of scooter, and degree of feeling. Moreover, certain categories appear to evaluate relationships or forms of comparison. The other listed characteristics, such as performance level, favorable attitude, and success or failure imply a measure of output or achievement. This table seems to be designed for weighing or judging people on these various measures. It may even be an assessment of traits used to measure personal development, success factors, or comparison. These characteristics may be useful in propagating evidence-based evaluations, wherein each feature can reflect an updated report for global individual growth, performance, or mood.

This table provides personal and professional characteristics for the evaluation of individual development. Demographics, psychological state, and relational characteristics. Positive attitude and success or failure hint at performance metrics for productivity and accomplishment. These sets of features individually and collectively allow assessments of individuals to be made along the lines of evidence-based comparisons for attributes that promote personality development, relationships, romantic interest, and success or failure in verbal expression likely not dissimilar to self-assessments or comparative performance measures.

Figure 3.18 shows the scatter plot matrix of four features. The features are decision-making, inattention or forgetfulness, insecurity by family, and educational life happiness. Figure 3.18 is shown below.

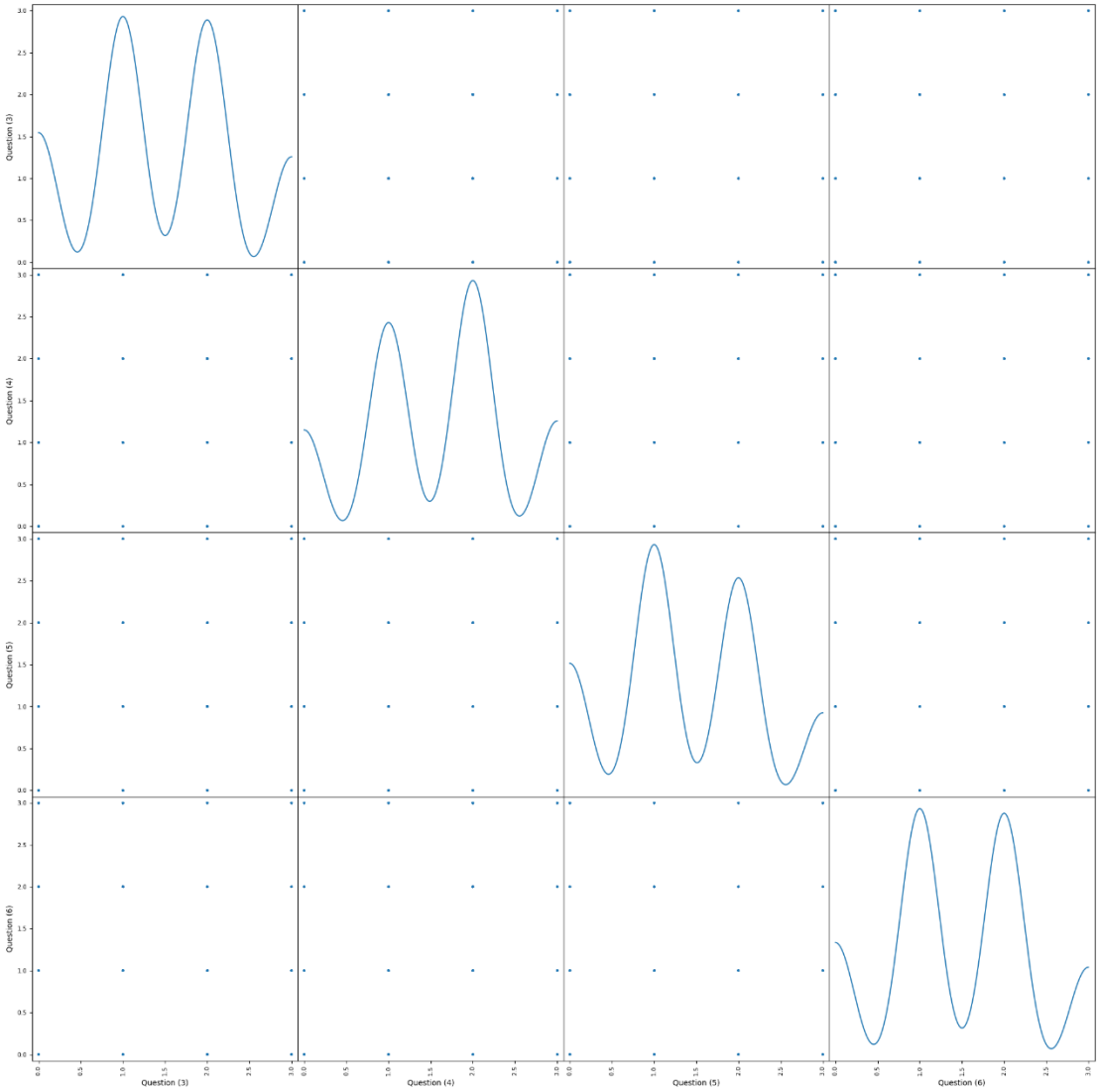


Figure 3.18: Scatter Plot Matrix

As you can see in the image, above is a scatter plot matrix displaying the pairwise relationships between all of the variables. Asymmetry of some diagonals appears, such as sinusoidal curves, suggesting a periodic signal in the data for specific variables. In off-diagonal plots, we show the relationships between pairs of questions. The sparse scatter of points in these plots indicates that, for the most part, pairs of variables are uncorrelated with points dispersed widely instead of following any detectable linear or non-linear trend. Now you would see this type of visualization helps us in looking at individual variables' distributions or relationships with other time series variables.

CHAPTER 4

RESULT AND DISCUSSION

4.1 Introduction

We used a few machine-learning techniques on our dataset and came up with some conclusions. In this section, we shall discuss the methods and the results. Many methods were used, such as LR, RF, k-NN, GB, DT, and NB. After putting them into practice, we were able to ascertain which approach gave us the most accuracy. Following the application of preprocessing and feature engineering approaches to the processed data, the accuracy is calculated. 1050 data points were gathered from individuals with low, normal, and high self-esteem. To calculate data as accurately as possible, the following actions were taken: The data was first split into training and testing, with 80% of the data going towards training and the remaining 20% going towards testing. We assessed the algorithms' accuracy after splitting the entire data set into training and testing sets.

We used LR, RF, k-NN, GB, DT, and NB for all of the machine learning models on the training set we got after splitting the data. Once we tuned the hyperparameters to their optimal values, each model was scored on the test set for predictive accuracy. By applying certain preprocessing techniques, we increased the performance of each model regarding the variations present in the dataset. The highest accuracy was recorded by one of the models, showing that it is capable of robustly predicting self-esteem scores. These are results that will inform upon the best model to use for future analyses.

4.2 Experimental Analysis

There were six machine learning algorithms used. A few classifications, including precision, recall, F1-score, and accuracy, were then computed using confusion matrices, and we compared them with each other. The training set and testing set were the two groups into which the entire dataset was separated for computation.

Confusion matrices were used to calculate precision, recall, F1-score, and accuracy for each algorithm which we then compared to choose the optimal model. Comparing these metrics, we were able to understand the strengths and weaknesses of each model in predicting self-esteem levels, which will help us explain the effectiveness of our models and gain an understanding of their reliability.

4.2.1 Experiment for Evaluation

We used a few machine-learning techniques on our dataset and obtained interesting findings. In this section, we'll review the methods and results. Many methods were used, such as LR, RF, k-NN, GB, DT, and NB. After implementing them, we were able to identify which approach gave us the best accuracy.

The sigmoid function, which accepts input as independent variables and generates a probability value between 0 and 1, is used in logistic regression for binary classification. Class 0 and Class 1 are two examples of our classes. An input is classified as Class 1 if its logistic function value is greater than the threshold value of 0.5; otherwise, it is classified as Class 0. Although it is mostly employed for classification problems, it is called regression since it is an extension of linear regression.

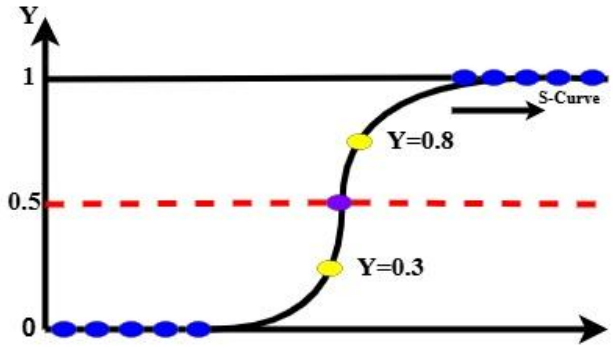


Figure 4.1: Logistic Regression

The Random Forest algorithm is a successful tree-learning technique in machine learning. It creates certain Decision Trees throughout the training phase. Each tree is constructed using a random subset of the data set to measure a random subset of characteristics in each partition. This unpredictability reduces the likelihood of overfitting and improves prediction performance overall by adding variation among individual trees.

By voting or averaging, the algorithm aggregates each prediction tree's output. This collaborative decision-making method, which benefits from the insights of multiple trees, offers an example of precise and dependable results. Random forests are widely used for classification and regression problems due to their ability to handle complex data, reduce overfitting, and generate accurate forecasts in a range of circumstances.

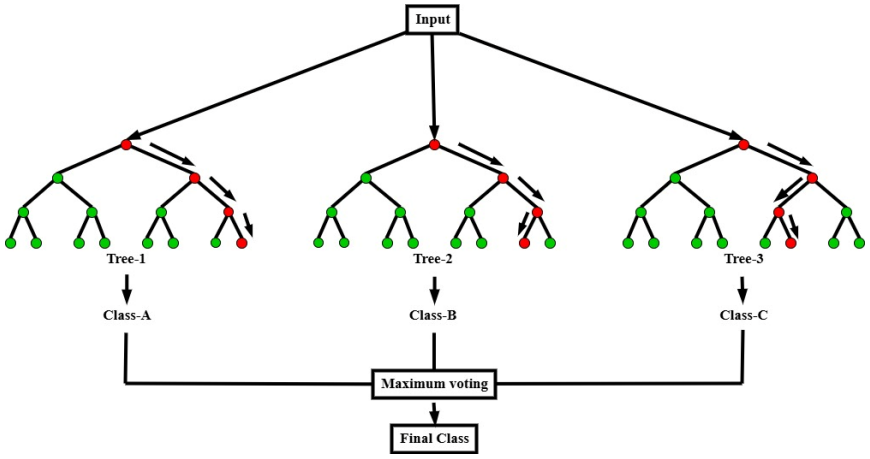


Figure 4.2: Random Forest

KNN is one of the most basic yet important machine-learning classification techniques. It belongs to the class of supervised learning and finds extensive application in pattern recognition, data mining, and intrusion detection.

This method is frequently used in real-world situations because it is non-parametric, which means it makes no underlying assumptions about the distribution of data, in contrast to other algorithms like GMM, which assume a Gaussian distribution of the provided data. We are given some historical data that arranges coordinates based on specific characteristics; this is commonly referred to as training data.

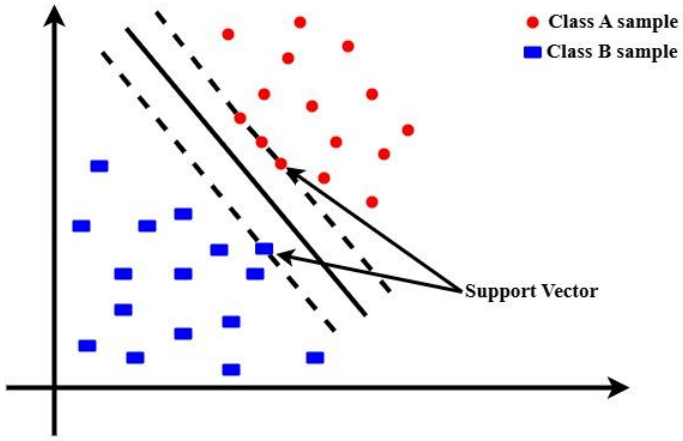


Figure 4.3: K-Nearest Neighbor (KNN)

A key component in coordinating the complex dance of model optimization is gradient descent. Fundamentally, it is a numerical optimization procedure that seeks to minimize a specified cost function to determine the ideal weights and biases for a neural network.

A popular optimization technique in deep learning and machine learning, gradient descent (GD) minimizes a neural network model's cost function while it is being trained. Until the cost function's minimum is attained, iteratively modifying the model's weights or parameters in the direction of the cost function's negative gradient is how it operates.

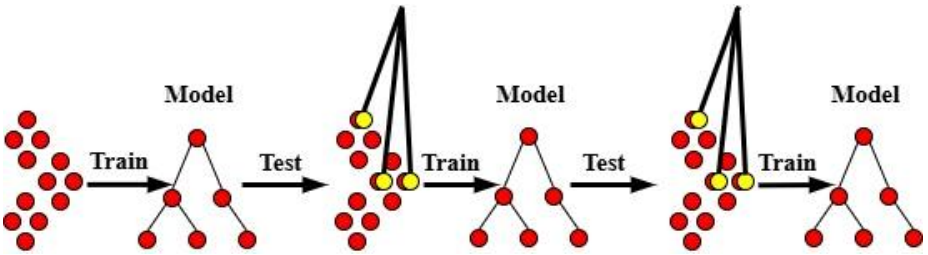


Figure 4.4: Gradient Boosting

Decision trees and other supervised learning algorithms are widely employed in machine learning to model and predict outcomes based on input data. Each internal node tests a property, each branch represents an attribute value, and each leaf node represents the final decision or prediction, giving the structure a tree-like appearance. One type of supervised learning method is the decision tree algorithm. They can be used to address both regression and classification problems.

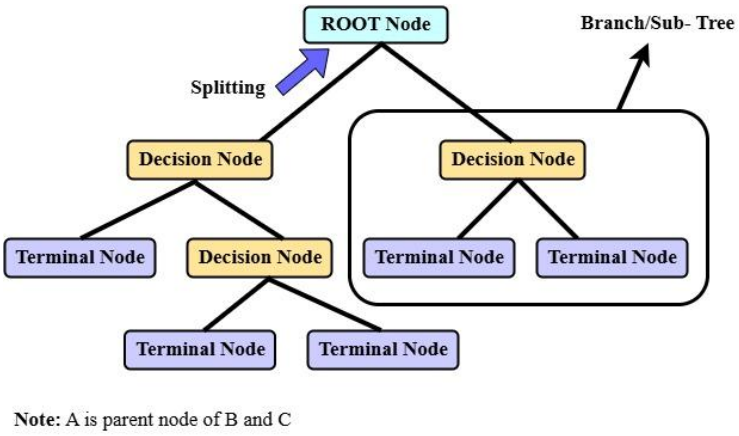


Figure 4.5: Decision Tree

A group of classification algorithms based on Bayes' Theorem are known as naive Bayes classifiers. It is a family of algorithms rather than a single method, and they are all based on the same idea—that is, each pair of features being classified is independent of the others. Let's start by looking at a dataset.

The Naïve Bayes classifier, one of the simplest and most efficient classification algorithms, facilitates the quick creation of machine learning models with quick prediction capabilities.

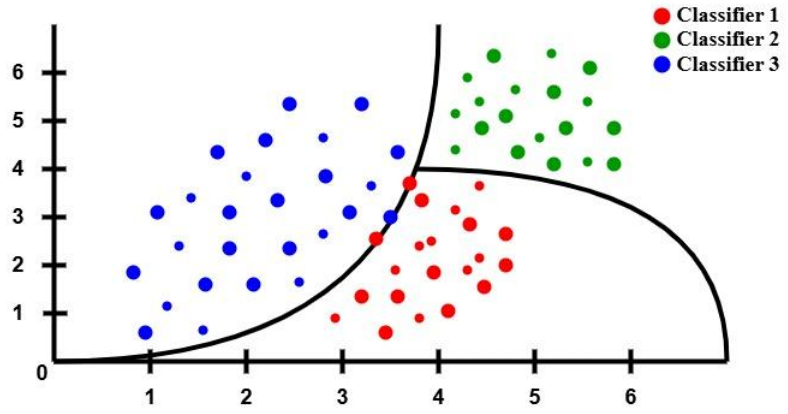


Figure 4.6: Naïve Bayes classifier

Figure 25 Bar chart shows the accuracies of different machine learning algorithms as a percentage. The algorithms on the x-axis are Logistic Regression, Random Forest Classifier, K-Nearest Neighbors (K-NN), Gradient Boosting, Decision Tree, and Naive Bayes Accuracy is plotted on the y-axis, with a scale from 0 to 100%. The light blue bar indicates the removal of noise, while darker and grey bars represent accuracy levels for each algorithm You can see that Gradient Boosting has with highest accuracy, followed by Logistic Regression and Random Forest. Though K-NN shows lower accuracy than others, Naive Bayes and Decision Tree also validate well into accuracy.

This bar chart presents a visual comparison of the six machine-learning algorithms in terms of accuracy. Logistic Regression (light blue) and Random Forest Classifier (pink) are performing quite well too. Compared to these models we observe slightly reduced performance in K-NN, shown in gray. The highest accuracy is given to Gradient Boosting (the bright yellow bar), which indicates it may be the best model for this data. As we see the Decision Tree (orange) and Naive Bayes (green) also give a very competitive accuracy where Naive Bayes is quite near to Gradient Boosting. Understanding which algorithms might be more stable and accurate across models allows one to choose wisely when selecting which is the best-fitting model. Figure 4.7 is given below.

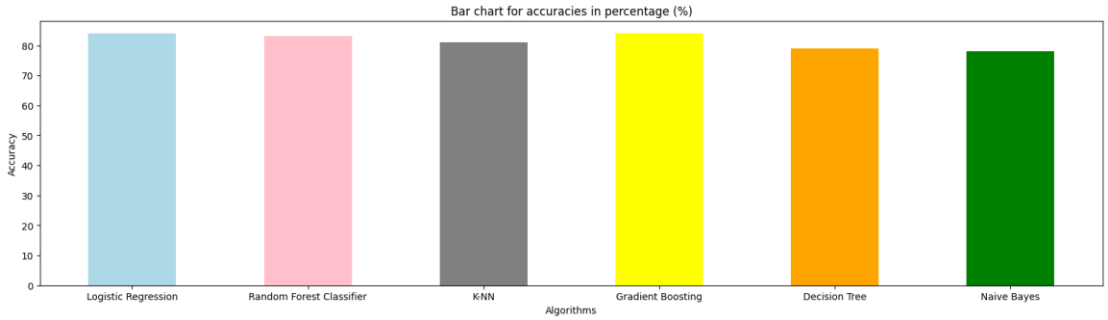


Figure 4.7: Accuracy of Dataset

Table 3 contains the accuracy percentages of six machine learning algorithms. Logistic Regression (LR) and Gradient Boosting (GB) have the best accuracy of 84.3%. Random Forest (RF) is closely behind at 82.9% and K-nearest Neighbors (KNN) at 81.4%. The accuracies of Decision Tree (DT) and Naive Bayes (NB) have lower results, respectively, 78.6% and 77.6%. It helps in discovering the best models for any predictive tasks. Table 4.1 is given below.

Table 4.1: Algorithm's Accuracy for Dataset

Algorithm	Accuracy
LR	84.3%
RF	82.9%
KNN	81.4%
GB	84.3%
DT	78.6%
NB	77.6%

4.2.2 Expressive Analysis

We analyzed the precision, recall, F1-score, accuracy, confusion matrix, and ROC curve of each method in addition to evaluating the accuracy of a large number of algorithms. A model evaluation is necessary for any model selection. In the case of model evaluation, measurements of particular categories are necessary. The test data sets serve as the basis for categorization measurements for advanced measures.

In conjunction with accuracy, additional statistical metrics such as precision, recall, F1-score, confusion matrix, and ROC curve examine model performance in a detailed manner. On the first hand, precision and recall show how well the model could detect

positive cases while the F1-score balances those metrics which helps to analyze the usage of imbalanced data. The confusion matrix gives us information about true and false predictions, while the ROC curve shows how well a classification has performed at various thresholds. With these metrics calculated on test datasets, a holistic assessment can be carried out, allowing for better model selection about the specific task of interest.

The number of true positives that were discovered is called recall. For this reason, it is the ratio of the expected negative value to the true positive value.

$$Recall = \frac{TP}{\text{TP+FN}} \times 100\% \quad \text{(i)}$$

Precision is the number of positively classified cases that were pertinent. The ratio of the expected positive value to the true positive value is crucial. By just producing a positive result for the one outcome it is most sure of, a test can optimize this.

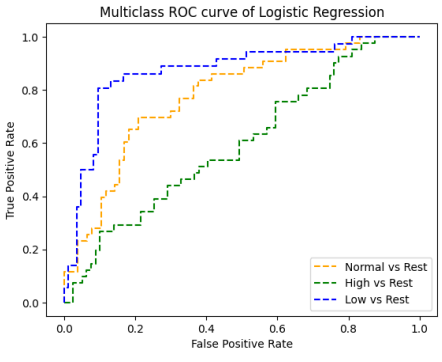
$$Precision = \frac{\text{TP}}{\text{TP+FP}} \times 100\% \quad \text{(ii)}$$

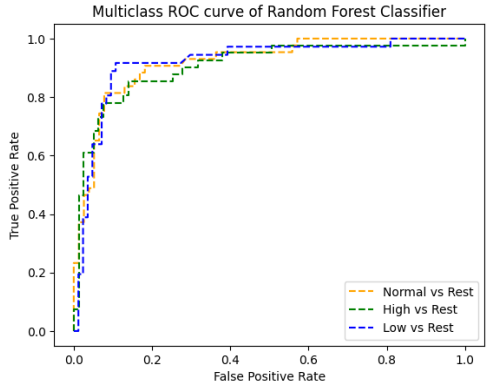
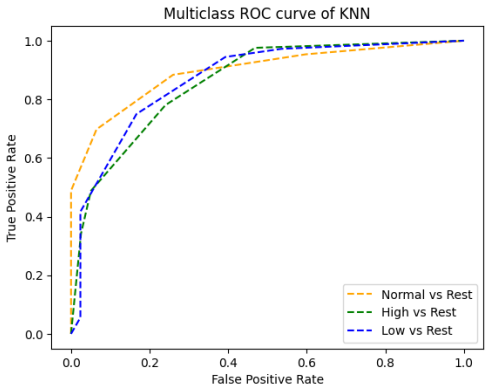
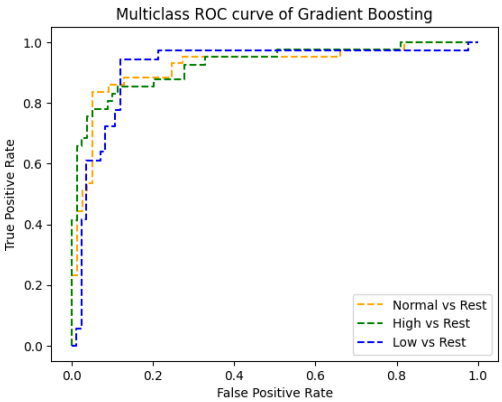
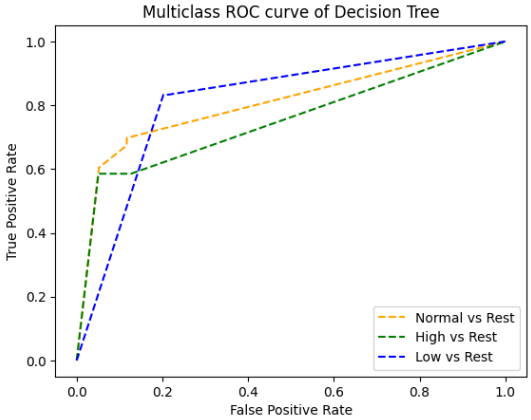
The weighted average of precision and recall is known as the *F1 score*. This score then takes into account both false positives and false negatives.

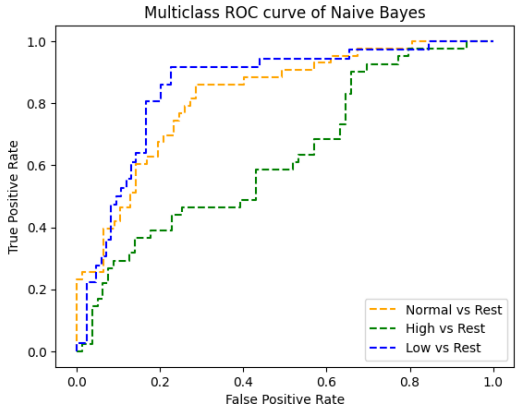
$$F-1\ Score = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \times 100\% \quad \text{(iii)}$$

ROC curves are especially useful for visually comparing curve classification models. The ROC curve was constructed using true positive and false positive rates. The diagonal line represents a guess. A model that has a curve that looks like a random estimation is less accurate. A few ROC curves produced using our approach are shown here in 4.2 table.

Table 4.2: ROC Curve

Algorithm	Roc Curve
Logistic Regression	

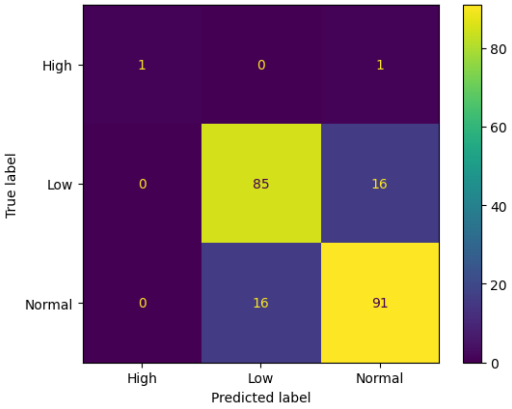
Random Forest	
K-Nearest Neighbor	
Gradient Boosting	
Decision Tree	

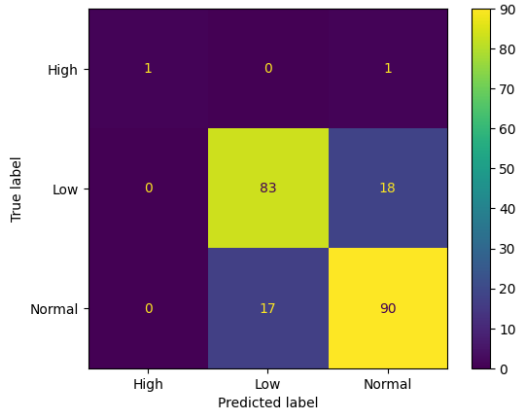
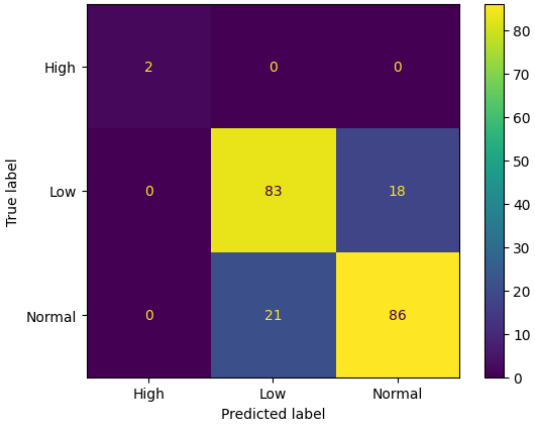
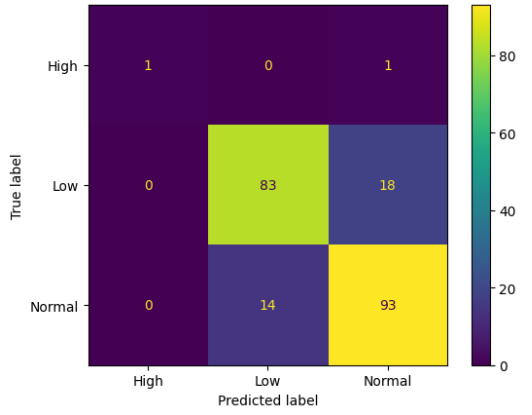
Naïve Bayes	

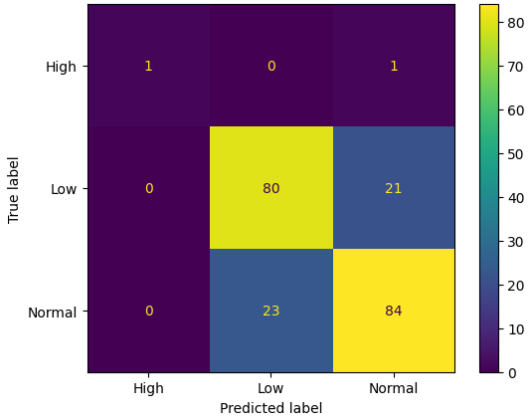
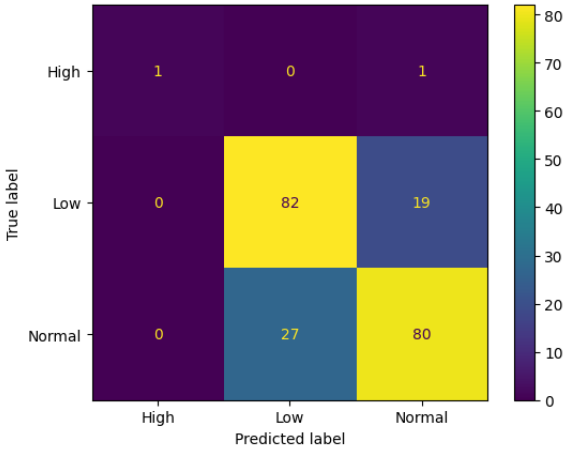
The confusion matrix is a machine-learning technique for forecasting results on a classification task. This is contrasted with the actual goal values projected by the machine learning model. We can quickly assess the effectiveness of our categorization model. We can also identify the kinds of mistakes we make. It is essential for figuring out a classifier's effectiveness.

The confusion matrix of the methods we utilized is shown in Table 4.3. The table below provides an accurate description of each classification.

Table 4.3: Confusion Matrix of All Classifiers

Algorithm	Confusion Matrix
Logistic Regression	

Random Forest	
K-Nearest Neighbor	
Gradient Boosting	

Decision Tree	
Naïve Bayes	

Key metrics for evaluating classifier performance such as True Positives, True Negatives, False Positives, and False Negatives are presented in the confusion matrix. Using these values, we can derive some important evaluation metrics such as accuracy, precision, recall, and F1-score. While accuracy describes the percentage of correctly predicted cases across the two categories, precision and recall seek to measure the model's classification ability for each category. The F1-score is the harmonic mean of precision and recall, which helps to balance these measurements and provides a better indication of the test's performance on imbalanced datasets. The confusion matrix also helps us identify specific kinds of errors; it allows us to target the improvement on these particular aspects of performance which is another plus point. The confusion matrix presents key metrics for evaluating classifier performance, such as true positives, true negatives, false positives, and false Negatives. We can compute useful evaluation metrics from these values like accuracy, precision, recall, and F1-score. Accuracy summarizes the proportion of correctly predicted cases in the two categories, while

precision and recall attempt to quantify the performance of a classifier on all data examples corresponding to one category. The F1 score is the harmonic mean of precision and recall which helps combine these metrics to get a sense of how well your test performed overall on data where one class may be more prominent than others. Another advantage of the confusion matrix is that it gives us an idea about the types of errors, we can specifically try to improve on these aspects of performance.

Table 4.4 displays each algorithm's performance. Based on the algorithms' performance and correctness, we will choose the best approach for our model. When it comes to accuracy, specificity, and precision, LR is unquestionably the best.

Table 4.4: Performance of Classifier

Algorithms	Classes	Precision (%)	Recall (%)	*F*-1 Score (%)
LR	Low	84	84	84
	Normal	84	85	85
	High	100	50	67
RF	Low	82	83	83
	Normal	83	83	83
	High	100	50	67
k-NN	Low	80	82	81
	Normal	83	80	82
	High	100	100	100
GB	Low	86	82	84
	Normal	83	87	85
	High	100	50	67
DT	Low	77	81	79
	Normal	81	78	79
	High	100	50	67
NB	Low	75	81	78
	Normal	80	75	77
	High	100	50	67

4.3 Comparative Performance Analysis

Table 4.5 provides an overview of various qualitative studies identifying self-esteem and its correlates. While these studies span several domains, the majority are focused on self-esteem with additional topics including mental health, anxiety, depression, and associated personality traits as well as life satisfaction. The research sample sizes are disparate from 99 participants (K. Solanki) to 5,587 participants (A. Al-Qahtani et al.).

Survey-based text data are the primary methodological domain, with approaches ranging from machine learning algorithms and statistical techniques to descriptive-analytical and longitudinal survey designs. These studies mainly aim to establish patterns and relationships, perform factor analysis, or predict some aspects of self-esteem. For instance, studies by A. Ghartappeh et al. and A.M. Hayes et al. Focus more on predictive analysis, and implications of self-esteem on mental health and disorders.

Table 4.5: Comparative Analysis with Other Works

Authors	Objects	Data Type	Problem Domain	Sample Size	Method
Our work	Self-esteem	Survey-based text	Detection	1050	ML Algorithms
K. Solanki [11]	Self-esteem	Text, Question	Recognition	99	Statistical Method
A. Al-Qahtani et al. [12]	Self-esteem	Survey-based text	Recognition	5,587	Cross-Sectional Survey Design
U.K. Moksnes et al. [13]	Self-esteem & Mental health	Text	Recognition	351	Longitudinal Survey Design
G. Manna et al. [14]	Self-esteem, anxiety & depression	Survey-based text	Recognition	454	Cross-Sectional Survey Design
H. Kärchner et al. [15]	Self-esteem	Survey-based text	Recognition	1652	Factor Relationship
E. Varanarasamma et al. [16]	Self-esteem & personality traits	Survey-based text	Recognition	515	Factor Relationship
A. Ghartappeh et al. [17]	Self-esteem, mental health & sleep quality	Survey-based text	Prediction	419	Descriptive-Analytical Study
A. M. Hayes et al. [18]	Self-esteem & depression	Survey-based text	Prediction	139	Longitudinal Survey Design
Z. Ozyesi [19]	Self-esteem	Survey-based text	Prediction	440	Factor Relationship
E. Hamarta [20]	Self-esteem & life satisfaction	Survey-based text	Prediction	405	Factor Relationship

4.4 Discussion

The ROC curve, correlation matrix, F1-score, accuracy, recall, precision, and method representation are all looked at in this section. The mathematical operation and efficacy of evolutionary models are also discussed here. We can see that the LR and GB approach has a maximum accuracy of 84.3%. Lastly, we discover that LR and GB may yield the best results for our self-esteem level detecting algorithm.

Next, we focus on the metrics and explain what this means for the models (LR & GB) of our self-esteem detection algorithm. This precision and recall tell us how well these models balance detections of true positives versus false positives. Both LR and GB have a high area under the curve (AUC) values, as shown on the ROC curve, confirming their robustness. These estimates could be further optimized with additional cross-validation or fine-tuning of the models. Taking into account accuracy and interpretability, LR and GB seemed to be good candidates for our use case.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

Mental health is crucial for us all and the way we see people with mental disorders can affect their self-esteem a lot. A lot of people think when he/she is having trouble with mental health they are also dangerous people but this misunderstanding makes the people who have bad mental disorders isolated so miserably. Most at-risk people who encounter these challenges are major assets to society. This gives them confidence and recognition of their abilities, thus making them feel like valued members of the community.

Individuals suffering from mental health issues are vulnerable to low self-esteem, as the society we live in imposes high demands and brings about bullying or struggles like money problems. These external issues make people feel either less worthy or not so capable. Still, some positive self-care practices on moving your body and eating well put the power back in people to control their mental health. For instance, exercising not only diminishes anxiety but provides a sense of accomplishment as well. A well-nutrition body has all the necessary energy for physical and mental health. These self-care steps also brand the option of having better self-esteem as they demonstrate to people that one is capable of looking after their well-being.

People who come with mental health problems are targeted and the society we reside in expects a lot of things from people without even thinking that they can also be bullied having low confidence or struggling to get money. Such external problems lower the self-esteem of a person. But not even your simplest forms of self-care on things like moving your body or eating good food put the power back in people to manage their mental health. For example, exercise reduces anxiety and also brings a sense of achievement. Fueling your body with nutrients provides adequate energy needed for both physical and mental health. This self-care, in turn, sells the idea of having improved self-esteem by showing peers what it appears like to be able to care for your health.

The piece advocates such a concept of human potential, that we can succeed if stress does not paralyze us. This reinforces a healthy level of self-worth and ensures that everyone knows they are valuable to the organization in some form or another. By working to stave off stress, provide for themselves, and find the good in others, those with mental health issues can gain higher self-esteem and tackle a positive rule within their lives as well as society.

5.2 Impact on Environment

Identifying self-esteem levels can affect our neighbors and environmental world around us greatly. The ability to identify low self-esteem in people and families may create a response on the part of communities creating a more supportive environment. Through compassion and mental health, we ensure people have spaces to feel seen, heard, and accepted which lessens isolation in people and elevates the community from each other.

Self-esteem affects the way people relate to the surrounding environment. Those who view themselves as valuable are more likely to feel empowered by their environment and the natural world. The top prompt is a sense of confidence and worth that can only lead to higher levels of participation in environmental-related community-led projects or action towards sustainability when enabled. On the contrary, low self-esteem can make us feel alienated or impotent; and thus, more disengaged from environmental stewardship.

Moreover, having a lack of self-esteem may result in the use of substance abuse or neglect of surroundings. Less littering, pollution, and waste of resources can be resolved by addressing self-esteem issues. When a person feels better about themselves, they are more likely to make better choices in life and for the environment around them.

It can also detect and promote a positive self-understanding which can lead to better lifestyle choices as well. Individuals with high self-esteem may feel better about exercising and going into nature, leading to more positive feelings toward parks and natural spaces. This bond supports an urge to protect these places and benefits both human well-being and the natural world.

Putting it all together the self-esteem levels data could also give one an idea to develop policies that ensure psychological well-being, inclusiveness, and environmental sustainability. When communities and leaders realize the connection between self-image and positive social constructs through mental health, they will begin to make it

a priority amongst programs that are designed for environmental action. Thus, self-esteem detection does not only benefit people; it also has an echo effect that leads to healthier, greener, and more unified communities.

5.3 Ethical Aspect

Concerning self-esteem levels, detecting these could provide important insight for predicting treatment pathways but also raise crucial ethical issues. Privacy and consent are paramount, of course. Persons must consent, only after being informed what the data are going to be measured and how such data is going to be used, stored, and protected. Preserving privacy helps to keep the mental stability of people private and safe.

Data security is another important issue. Due to the sensitivity of self-esteem data, it must be heavily guarded against finding nonlegitimate access which finds breaches and unveils target mental state stats. These measures are crucial as people may feel exposed or suspicious when sharing this type of sensitive information.

It is also crucial not to take averages of self-esteem and stigmatize people. Seeing the trait of low self-esteem as a defect opens them up to negative labeling or stereotypes. Data on self-esteem, rather than being constraining should be constructive support, not condemn individuals, and make them feel accepted and appreciated in their groups.

Nondiscrimination and fair use are also important. The self-esteem information should never be used by employers, insurers, or the like to unfairly assess individuals, especially in a way that would negatively affect an individual. It is unethical to use this information to prevent access to resources or benefits, and it is unfair discrimination.

Such interventions are not easy for the practitioners or patients, if self-esteem detection-based interventions have to be made, they should be planned carefully and compassionately. Should assistance or resources be made available, they need to take into consideration the individual dignity and self-determination of those who require financial assistance. Interventions are not designed to “fix” people or allow outsiders to dictate what someone should want, need or desire. In your search for autonomy, you provide individuals an option about whether to engage with self-esteem detection tools and interventions. As much as it is nice that some people can share their experience, there should never be a pressure to share an experience and there should never be any pressure put on somebody to seek help.

Self-esteem detection organizations should be transparent and accountable. Transparency will maintain trust institutions should clarify what they plan to do, why, and how. Transparent and accountable, such organizations can use ethical practices that tend to the welfare of actual persons living in communities.

To wrap it up, ethical self-esteem detection respects the privacy and consent of the users, protects their data, does not stigmatize them, prevents discrimination, and is transparent about its use. When grounded in these ethics, self-esteem detection can serve to support and not hinder encouraging well-being while honoring dignity and autonomy.

5.4 Sustainability Plan

A self-esteem detection system that is sustainable necessitates careful planning to ensure that what it produces is both ethical and inclusive. It also needs to be beneficial in the long run. First, ethical data management is fundamental Data collected should be kept secure, and individuals need to give informed consent. They deserve to know just how information about them is going to be used and protected. This is a way of establishing trust, and also keeping one's privacy safe.

Accessibility to self-esteem detection technology is another priority. Programs ought to be a source of liberation for people from diverse groups who have different backgrounds experiences and lifestyles. Making sure that everyone can avail themselves not only equally but without difficulty helps all individuals, especially those in compromising circumstances, to use these tools.

Sustainable self-esteem detection should also emphasize interventions that empower, rather than brand, individuals. The program should provide support to those with low self-esteem, such as counseling, workshops on confidence-building skills, or relay work of other sorts. These services should meet the individual's unique needs.

Community involvement is essential for self-sustaining development. By linking up with local leaders, educators, mental health workers, and community organizations, the program can be fine-tuned to meet the needs of those it serves. By building in points of entry for local input, self-esteem support becomes more meaningful and effective in reality.

In order to enhance knowledge and reduce stigma, etiquette education about high self-esteem is also crucial. Public campaigns, community events, and classes can promote understanding and inspire mutual encouragement so that everyone has a positive and supportive living environment.

Monitoring and evaluation are essential to any long-term project. Through constant review we can learn what is working and what isn't; feedback from those involved will enable us to make the support resources even better, so they will keep being useful. If good information continues to come out from the centers that produce it will help in reforming government policies on drug abuse.

Financial sustainability and effective resource management are equally crucial. A sustainable project needs funds that are stable over time for operating expenses, such as those from grants, donations, or partnerships with organizations concerned about mental health. This stable budget comes from a clear and realistic plan; only money spent on intervention works to make the most impact. Numerous reviews also help ensure that our funds are used cost-effectively.

Lastly, transparency and accountability help establish confidence. With program goals, results, and problems made public, it acts of its own accord to earn credibility and leaves participants the opportunity to contribute a few comments. This form of transparency is to make sure that the program stays on track, promoting well-being in a straightforward and responsible manner.

Digest brings us to say if you aim to set up a sustainable self-esteem detection program: ethical data management methods, universal access, community-based supportive treatment, education, regular review of program performance, financial soundness, and terms of openness. The program could continue to provide meaningful, long-term assistance for self-esteem and mental health in the region.

CHAPTER 6

SUMMARY, CONCLUSION AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

We develop a machine learning program to evaluate student's own feelings situation. Our system can grade self-respect into low, normal, and high grades. Based on the critical role self-esteem plays in mental health, interpersonal relationships, and personal development across students, we designed this model to promote not only educational but also health awareness of mentality with a focus specifically on Bangladesh.

To build our model, first, we collected data from surveys that have demographics such as age and sex as well as measurable behaviors represented in people's confidence levels or self-reflection scales. We pre-processed the collected data and kept our accuracy as high for use by machine learning algorithms.

To classify the self-esteem levels of people in the data sample used in this study, we utilized machine learning models such as Logistic Regression, Random Forest, k-nearest Neighbors, and Gradient Boosting. To further improve the accuracy of our model, feature selection was carried out, which allowed us to determine this set of important indicators that characterize self-esteem. While assessing using basic measures of accuracy, precision, recall, and F1-score, we noted that Random Forest and Gradient Boosting were the most accurate in classifying moderate and high self-esteem.

In the study, there were challenges most especially in data gathering and data sensitivity which were overcome by using self-administered anonymous questionnaires. In the future, we propose to develop this model to further expand our database variety, and work on increasing the model's prediction scores. In general, this paper demonstrates the possibilities of using machine learning in the early assessment and intervention of emergent mental health concerns, providing theoretical and pragmatic alternatives that are also directly relevant to school contexts.

6.2 Conclusion

In this paper, we show how it is possible to use machine learning to identify the level of self-esteem among students and enrich the knowledge of mental health assessment

in educational contexts. Through the proposed model that categorized students according to their self-esteem level as low, normal, and high self-esteem, we have filled a gap in the early detection of low self-esteem, which affects students' social, emotional, and academic lives. Our strategy entailed first collecting survey data on both behavioral and demographic status and then running an ML decision tree with Random Forest and gradient-boosting algorithms that presented the greatest efficacy in distinguishing between various levels of self-esteem.

Thus, our results stress the need to apply the predictive means in mental health models, especially in communities where people have low mental health literacy. Real-world issues required to be faced in terms of privacy and data aggregation point towards the importance of having more ethically built tools into play along with more volume and variety in sample data that should eventually help better the model. To improve the outcomes, we continue to optimize these approaches and enlarge the pool of materials we work with, thus the further enhancement of the model's effectiveness can be potentially used in the creation of individualized student mental health support.

Finally, this research offers an empirical base for developing the field of mental health detection using machine learning as highly applicable to the educational context. Further development of such models may allow their incorporation into the support of concepts used by teachers and psychologists in student affairs.

6.3 Implication for Further Study

The outcomes of the present research bring some exciting opportunities for primary research about the use of machine learning in mental health assessment in a learning environment. Subsequent research could extend our present investigation by utilizing more extensive and heterogeneous samples, which would enhance external validity and enhance the precision of the model regarding variability of the self-esteem at population and cultural levels. Further, including Longitudinal data could also improve the model's ability to explore how student's self-esteem changes over time to detect changing behaviors that may be a sign of developing mental health issues.

The scope of classification could also be enriched by additional behavioral and environmental parameters to make further decisions in the model. Such additional data may be related to the behavior of students in social networks, their academic results, and physical activity levels which can give a better understanding of self-estimation

and other aspects of mental health. It is possible to develop algorithms with the help of deep learning and natural language processing that will detect the beginnings of self-esteem problems in real-time based on text entries or voice including the need for the intervention of a psychologist.

Last, more research needs to consider how the use of machine learning solutions concerns ethical and privacy concerns in sensitive mental health spheres and how guidelines may help students and their data to be secured and processed appropriately. Such developments may contribute to the development of positive, school-related assets that may be employed by teachers and other personnel in the field of mental health to ensure students are provided with the appropriate kind of help as soon as possible, in a preventive measure manner.

REFERENCES

[1] M. Fatima, S. Niazi, and S. Ghayas, "Relationship between Self-Esteem and Social Anxiety: Role of Social Connectedness as a Mediator," 2017. Available: https://www.gcu.edu.pk/pages/gcupress/pjscp/volumes/pjscp20172-2.pdf.

[2] G. Büyükşahin Çevik, "The Roles of Life Satisfaction, Teaching Efficacy, and Self-esteem in Predicting Teachers' Job Satisfaction," *Universal Journal of Educational Research*, vol. 5, no. 3, pp. 338–346, Mar. 2017, doi: https://doi.org/10.13189/ujer.2017.050306.

[3] L. J. Friedlander, G. J. Reid, N. Shupak, and R. Cribbie, "Social Support, Self-Esteem, and Stress as Predictors of Adjustment to University Among First-Year Undergraduates," *Journal of College Student Development*, vol. 48, no. 3, pp. 259–274, 2007, doi: https://doi.org/10.1353/csd.2007.0024.

[4] A. E. Wells, L. M. Hunnikin, D. P. Ash, and S. H. M. van Goozen, "Low Self-Esteem and Impairments in Emotion Recognition Predict Behavioural Problems in Children," *Journal of Psychopathology and Behavioral Assessment*, Jun. 2020, doi: https://doi.org/10.1007/s10862-020-09814-7.

[5] C. M. Fekry, E. M. Mahfouz, N. N. Kamal, T. A. Abd-El Rahman, and E. E. Hassan, "Contribution of low-level Self-esteem and self-efficacy in adverse mental outcome among secondary school adolescents in Minia city," *Minia Journal of Medical Research*, vol. 34, no. 1, pp. 15–24, Jan. 2023, doi: https://doi.org/10.21608/mjmr.2023.168523.1204.

[6] X. Liu, X. Cao, and W. Gao, "Does Low Self-Esteem Predict Anxiety Among Chinese College Students?," *Psychology Research and Behavior Management*, vol. Volume 15, pp. 1481–1487, Jun. 2022, doi: https://doi.org/10.2147/prbm.s361807.

[7] G. D. Valenti and P. Faraci, "Predicting University Adjustment from Coping-Styles, Self-Esteem, Self-Efficacy, and Personality: Findings from a Survey in a Sample of Italian Students," *European Journal of Investigation in Health, Psychology and Education*, vol. 11, no. 3, pp. 894–907, Aug. 2021, doi: https://doi.org/10.3390/ejihpe11030066.

[8] G. L. Brase and E. C. Guy, "The demographics of mate value and self-esteem," *Personality and Individual Differences*, vol. 36, no. 2, pp. 471–484, Jan. 2004, doi: https://doi.org/10.1016/s0191-8869(03)00117-x.

[9] S. Diamantopoulou, A.-M. Rydell, and L. Henricsson, "Can Both Low and High Self-esteem Be Related to Aggression in Children?," *Social Development*, vol. 17, no. 3, pp. 682–698, Aug. 2008, doi: https://doi.org/10.1111/j.1467-9507.2007.00444.x.

[10] W. Li *et al.*, "Reciprocal relationships between self-esteem, coping styles and anxiety symptoms among adolescents: between-person and within-person

effects," *Child and Adolescent Psychiatry and Mental Health*, vol. 17, no. 1, Feb. 2023, doi: https://doi.org/10.1186/s13034-023-00564-4.

[11] K. Solanki, "RELATIONSHIP BETWEEN SELF ESTEEM AND SELF EFFICACY," *INTERNATIONAL JOURNAL OF RESEARCH IN SOCIAL SCIENCES & HUMANITIES*, vol. 12, no. 04, pp. 926–941, 2022, doi: https://doi.org/10.37648/ijrssh.v12i04.049.

[12] U. K. Moksnes and R. J. Reidunsdatter, "Self-esteem and mental health in adolescents – level and stability during a school year," *Norsk Epidemiologi*, vol. 28, no. 1–2, May 2019, doi: https://doi.org/10.5324/nje.v28i1-2.3052.

[13] G. Manna, Giorgio Falgares, S. Ingoglia, Maria Rosaria Como, and Sandro De Santis, "The Relationship between Self-Esteem, Depression and Anxiety: Comparing Vulnerability and Scar Model in the Italian Context," *Mediterranean Journal of Clinical Psychology*, vol. 4, no. 3, 2016, doi: https://doi.org/10.6092/2282-1619/2016.4.1328.

[14] H. Kärchner, C. Schöne, and M. Schwinger, "Beyond level of self-esteem: exploring the interplay of level, stability, and contingency of self-esteem, mediating factors, and academic achievement," *Social Psychology of Education*, vol. 24, no. 2, pp. 319–341, Feb. 2021, doi: https://doi.org/10.1007/s11218-021-09610-5.

[15] E. Varanarasamma, A. Kaur, G. Singh, and N. Muthu, "The Relationship between Personality and Self-Esteem towards University Students in Malaysia," 2018.

[16] A. M. Hayes, M. S. Harris, and C. S. Carver, "Predictors of Self-Esteem Variability," *Cognitive Therapy and Research*, vol. 28, no. 3, pp. 369–385, Jun. 2004, doi: https://doi.org/10.1023/b:cotr.0000031807.64718.b9.

[17] Z. Ozyesil, "The Prediction Level of Self-Esteem on Humor Style and Positive-Negative Affect," *Psychology*, vol. 03, no. 08, pp. 638–641, 2012, doi: https://doi.org/10.4236/psych.2012.38098.

[18] E. Hamarta, "A prediction of self-esteem and life satisfaction by social problem solving," *Social Behavior and Personality: an international journal*, vol. 37, no. 1, pp. 73–82, Feb. 2009, doi: https://doi.org/10.2224/sbp.2009.37.1.73.

[19] A. Al-Qahtani, H. Ibrahim, W. Elgzar, H. Sayed, and R. Essa, "Self-esteem and self-efficacy in women empowerment The role of self-esteem and self-efficacy in women empowerment in the Kingdom of Saudi Arabia: A cross-sectional study," *African Journal of Reproductive Health*, vol. 25, no. 1s, 2021, doi: https://doi.org/10.29063/ajrh2021/v25i1s.7.

[20] A. Ghartappeh, M. Arefi, S. Faramarzi, and Y. Pasdar, "The Prediction of Quality of Life Based on Self-esteem, Mental Health and Sleep Quality Variables in Obese and Overweight Girl Students," *Journal of Clinical*

Research in Paramedical Sciences, vol. 9, no. 2, Aug. 2020, doi: https://doi.org/10.5812/jcrps.96810.

[21] "Rosenberg's Self-Esteem Scale," Wwnorton.com, 2019. https://www.wwnorton.com/college/psych/psychsci/media/rosenberg.htm

[22] J. D. Brown and M. A. Marshall, "Self-Esteem and Emotion: Some Thoughts about Feelings," Personality and Social Psychology Bulletin, vol. 27, no. 5, pp. 575–584, May 2001, doi: https://doi.org/10.1177/0146167201275006.

[23] A. Pedersen, J. Attwell, and D. Heveli, "Prediction of negative attitudes toward Australian asylum seekers: False beliefs, nationalism, and self-esteem," Australian Journal of Psychology, vol. 57, no. 3, pp. 148–160, Jan. 2005, doi: https://doi.org/10.1080/00049530500125157.

[24] A. Mokady and N. Reggev, "The Role of Predictions, Their Confirmation, and Reward in Maintaining the Self-Concept," Frontiers in Human Neuroscience, vol. 16, Mar. 2022, doi: https://doi.org/10.3389/fnhum.2022.824085.

[25] M. Mann, "Self-esteem in a broad-spectrum approach for mental health promotion," Health Education Research, vol. 19, no. 4, pp. 357–372, May 2004, doi: https://doi.org/10.1093/her/cyg041.

[26] X. Liu, X. Cao, and W. Gao, "Does Low Self-Esteem Predict Anxiety Among Chinese College Students?," Psychology Research and Behavior Management, vol. Volume 15, pp. 1481–1487, Jun. 2022, doi: https://doi.org/10.2147/prbm.s361807.

[27] F. Talaeizadeh, "Prediction of family self-esteem," 2019.

[28] T. Heatherton and N. Ambady, "SELF-ESTEEM, SELF- PREDICTION, AND LIVING UP TO COMMITMENTS."

[29] A. Martens, J. Greenberg, J. J. B. Allen, J. Hayes, J. Schimel, and M. Johns, "Self-esteem and autonomic physiology: Self-esteem levels predict cardiac vagal tone," Journal of Research in Personality, vol. 44, no. 5, pp. 573–584, Oct. 2010, doi: https://doi.org/10.1016/j.jrp.2010.07.001.

[30] M. Roth and T. Altmann, "A Comparison of the Predictive Validity of Self-Esteem Level and Directly Measured Self-Esteem Stability in the Temporal Prediction of Psychological Distress," Frontiers in Psychology, vol. 11, Jul. 2020, doi: https://doi.org/10.3389/fpsyg.2020.01770.

[31] S. Sharma, G. Singh, and M. Sharma, "A comprehensive review and analysis of supervised-learning and soft computing techniques for stress diagnosis in humans," Computers in Biology and Medicine, vol. 134, p. 104450, Jul. 2021, doi: https://doi.org/10.1016/j.compbiomed.2021.104450.

[32] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, and L. Pecchia, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis," Biomedical Signal Processing and

Control, vol. 18, pp. 370–377, Apr. 2015, doi: https://doi.org/10.1016/j.bspc.2015.02.012.

[33] E. Labbé, N. Schmidt, J. Babin, and M. Pharr, “Coping with Stress: The Effectiveness of Different Types of Music,” Applied Psychophysiology and Biofeedback, vol. 32, no. 3–4, pp. 163–168, Oct. 2007, doi: https://doi.org/10.1007/s10484-007-9043-9.

[34] A. Orabi, P. Buddhitha, M. Orabi, and D. Inkpen, “Deep Learning for Depression Detection of Twitter Users,” Association for Computational Linguistics, 2018.

[35] Md. Z. Islam, Z. Jannat, Md. T. Habib, Md. S. Rahman, and G. Z. Islam, “Detection of Facebook Addiction Using Machine Learning,” Lecture Notes in Networks and Systems, pp. 625–638, 2022, doi: https://doi.org/10.1007/978-3-031-12413-6_49.

[36] Jongyoon Choi, B. Ahmed, and R. Gutierrez-Osuna, “Development and Evaluation of an Ambulatory Stress Monitor Based on Wearable Sensors,” IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 2, pp. 279–286, Mar. 2012, doi: https://doi.org/10.1109/titb.2011.2169804.

[37] M. Stankevich, V. Isakov, D. Devyatkin, and I. Smirnov, “Feature Engineering for Depression Detection in Social Media,” Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods, 2018, doi: https://doi.org/10.5220/0006598604260431.

[38] D. M. Shukla, K. Sharma, and S. Gupta, “Identifying Depression in a Person Using Speech Signals by Extracting Energy and Statistical Features,” 2020 IEEE International Students’ Conference on Electrical,Electronics and Computer Science (SCEECS), pp. 1–4, Feb. 2020, doi: https://doi.org/10.1109/sceecs48394.2020.60.

[39] A. R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel, and A. S. Malik, “Machine Learning Framework for the Detection of Mental Stress at Multiple Levels,” IEEE Access, vol. 5, pp. 13545–13556, 2017, doi: https://doi.org/10.1109/access.2017.2723622.

[40] A. Hossain, “Mental Health Disorders Status in Bangladesh: A Systematic Review,” JOJ Nursing & Health Care, vol. 7, no. 2, Apr. 2018, doi: https://doi.org/10.19080/jojnhc.2018.07.555708.

Report.pdf

ORIGINALITY REPORT

15%	13%	4%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	6%
2	Submitted to Daffodil International University Student Paper	1%
3	ajrh.info Internet Source	<1%
4	pdfs.semanticscholar.org Internet Source	<1%
5	www.mdpi.com Internet Source	<1%
6	Submitted to Queen's College Student Paper	<1%
7	www.researchgate.net Internet Source	<1%
8	mdpi-res.com Internet Source	<1%
9	Submitted to University of North Texas Student Paper	<1%