Abstract

This project aims to develop a machine-learning model that can classify an individual's depression state using other variables. We are using a dataset called Depression.csv, which has 16 features. The target feature of our project is the depression state of the individual, and we aim to accurately predict the level of depression (no depression, mild depression, moderate depression, severe depression) based on other features. Our approach involves preprocessing the data to handle missing values, scaling features for optimal model performance, and training and validating the model using robust machine-learning techniques. The successful implementation of this model could enhance mental health monitoring and support, contributing to better mental health outcomes.

Introduction

Nowadays, Depression is one of the significant problems for many people, mainly Adults. It affects a person's mental health, resulting in anxiety, panic attacks, and damage to physical health. Detection of Depression is essential, as well as providing the proper treatment. Machine Learning (ML) is one of the emerging technologies that can help find hidden patterns in data using various techniques. This project aims to apply one of those machine learning (ML) techniques to develop a model that can classify the state of the Depression based on its severity.

Methods

In this project, we used many methodologies to build a reliable model for classifying the depression state. Firstly, we mapped the target variable string values to the numeric values so that 0 for 'no depression,' 1 for 'mild depression,' 2 for 'moderate depression,' and 3 for 'severe depression.' Next, we used StandardScaler to scale the features so that each column contributes equally to the model prediction. Scaling is vital in classification, as it helps arrange the data within the same range. In this project, we used many methodologies to build a reliable model for classifying the depression state. Firstly, we mapped the target variable string values to the numeric values so that 0 for 'no depression,' 1 for 'mild depression,' 2 for 'moderate depression,' and 3 for 'severe depression.' Next, we used StandardScaler to scale the features so that each column contributes equally to the model prediction. Scaling is vital in classification, as it helps arrange the data within the same range. As our model mainly focuses on classification, we utilized a Support Vector Classifier (SVC), a classification model in Python from the sci-kit learn library. This library provides the implementation of this model simply and efficiently. We also used train_test_split from sklearn, which divides the data into train and test data. Finally, we used a confusion matrix to measure the model's accuracy, making detecting errors easy. The rows of this

matrix represent the actual outcomes, whereas the columns represent the predictions of our model.

Data Procurement:

Kaggle is one of the open-source platforms for extensive collections of data sets. We downloaded our data set from Kaggle, which provided relevant and adequate information for training our model.

Data Analytics Life Cycle: (Explain these**)**

I. **Business Case Evaluation**
II. **Data Identification**
III. **Data Acquisition and Filtering**
IV. **Data extraction**
V. **Data validation and cleansing**
VI. **Data aggregation and representation**
VII. **Data Analysis**
VIII. **Data Visualization**
IX. **Utilization of analysis results**

Future work: (Select one**)**

1. Integration with Healthcare Systems

- **Clinical Decision Support Systems (CDSS)**: Integrate the model into CDSS to assist healthcare providers in diagnosing and monitoring depression. This can enhance clinical decision-making by providing additional insights based on patient data.
- **Electronic Health Records (EHR)**: Embed the model within EHR systems to automatically flag potential depression cases based on patient data and trends over time.

2. Mobile and Web Applications

- **Mental Health Apps**: Develop or collaborate with existing mental health apps to provide users with depression risk assessments, personalized tips, and resources.
- **Telehealth Services**: Integrate the model into telehealth platforms to help clinicians assess patients' mental health remotely.

3. Workplace Mental Health Programs

- **Employee Wellness Programs**: Incorporate the model into employee wellness programs to monitor and support employees' mental health. This can help in early detection and intervention.

- **Corporate Partnerships**: Partner with corporations to provide mental health tools that use your model to help employees manage stress and depression.

4. Educational Institutions

- **Student Mental Health Services**: Implement the model in educational institutions to monitor and support students' mental health, potentially integrating with counseling services.
- **Awareness Programs**: Use the model to identify students at risk and provide targeted mental health awareness and intervention programs.