Detection of Risks to Student Mental Health from Educational Data

Introduction:

Mental health issues among students in India are a growing concern. According to a 2022 survey by the National Institute of Mental Health and Neurosciences (NIMHANS), approximately **10-15% of Indian adolescents** experience mental health conditions such as depression and anxiety. The **National Crime Records Bureau (NCRB)** reports that in 2021, over **13,000 student suicides** were recorded, an alarming figure that highlights the urgent need for addressing mental health in the educational context.

The **Indian Council of Medical Research (ICMR)** notes that **23-32% of adolescents** experience symptoms of depression, with many cases going unreported due to social stigma or lack of awareness. Furthermore, a 2019 survey by **Lokniti** revealed that **42% of Indian college students** reported feeling stressed about academics and future career prospects, while **26%** admitted to struggling with loneliness.

Substance abuse is also a significant concern, with data from the **Ministry of Social Justice and Empowerment** indicating that **4.6%** of Indian adolescents have engaged in substance use, often leading to detrimental consequences. Just as with depressive symptoms, these issues often remain unreported until the situation escalates.

Early detection of mental health risks among students can play a crucial role in providing timely intervention and support. Schools and colleges have the potential to offer resources, counselling, and mental health services, but they often lack the tools to identify at-risk students. By leveraging educational data, institutions can detect patterns and behaviours that might indicate mental health risks, enabling educators to intervene before it's too late.

We aim to explore the possibility of predicting student mental health using data that is readily available to schools. This approach would enable schools to identify students at risk of mental health issues and provide necessary interventions or resources. The input data includes educational records such as student grades, absences, parental occupations, health status, and relationships with peers. Since we do not have explicit labels for mental health status, we used a **K-means clustering** algorithm to group students into clusters. These clusters represent different levels of mental health risk, ranging from 0 to 5, allowing us to infer student mental health based on their assigned cluster.

Finally, we deployed the clustering model on [share.streamlit.io](), utilizing the 7 most important features identified through **Principal Component Analysis (PCA)**. This deployment allows users to easily and conveniently interact with the model for insights into student mental health.

Related work:

In one study, two Portuguese higher secondary school students predicted the final exam score (G3) based on their previous two exam scores (G1 and G2) along with other available features from their educational data. For this task, Decision Trees (DT) and Random Forests (RF) performed best in predicting Portuguese grades in a five-way classification task. However, the study did not account for class imbalance, and the hyperparameter tuning was limited. Predicting student health, however, is more challenging since the first and second-period grades were the most significant predictors for the third-period grades, leaving other features less influential in predicting health outcomes.

Additionally, several studies have explored predicting various aspects of student performance or well-being using data collected from administrative and survey sources. For example, researchers at **Stanford University** conducted a study where they predicted student health based on their educational records, providing insights into the broader potential of leveraging school data for mental health risk assessment.

These prior works underscore the importance of feature selection and model tuning when predicting outcomes like academic performance or health, and highlight the growing interest in using educational data for mental health prediction.

Dataset and Features:

We obtained the dataset from the [UC Irvine Machine Learning Repository]() to predict mental health status among students based on their educational records. The dataset consists of 34 features, providing a diverse range of information to analyze. Additionally, we merged the **student_math** and **student_por** datasets to create a larger, more comprehensive dataset.

It is important to note that we do not have any labels for **mental health status** in the dataset. Therefore, our approach involves clustering the data to infer mental health categories. To effectively tackle this problem, we followed these steps:

Data Preprocessing:

a. Data Cleaning:

- We started by inspecting the dataset for missing values, outliers, and inconsistencies. Missing data was handled using imputation or removal techniques, and outliers were addressed appropriately.
- Since some Portuguese students are present in both the **student_math** and **student_por** datasets, we removed duplicate rows where all features, except for the grades (**G1**, **G2**, **G3**), were identical. The cleaned dataset was saved as **df_cleaned**.

b. Data Exploration:

- We explored the dataset to understand the distribution of features and identify relationships between variables. This step helped us in gaining insights into potential patterns within the data.

c. Feature Engineering:

- We identified categorical columns and applied **one-hot encoding** to convert them into numerical representations. The original categorical columns were dropped after encoding.
- To avoid redundancy, we dropped the **nth numerical column** where necessary.
- The one-hot encoded DataFrame was concatenated with the original DataFrame, resulting in the **df_encoded**.
- Finally, we applied **feature scaling** (either standardization or normalization) to **df_encoded** to create **df_scaled**.

Target Variable:

Since the dataset does not contain an explicit mental health label, we used a clustering approach to create a target variable, **"Mental_Health_Status"**. This is a **5-way clustering task**, with clusters ranging from 0 to 5. We explored different numbers of clusters to determine the best separation between groups, which allowed us to predict mental health risk categories based on the educational data.

Model Selection & Training:

Cluster Model Assignment

For our prediction task, we selected **K-Means Clustering** as the primary clustering algorithm due to its efficiency and effectiveness in grouping similar data points. The selection process involved the following steps:

1. **Basic 2D Visualization**:
    - Initially, we performed basic 2D visualizations to understand the data distribution and observe preliminary cluster formations.
2. **Principal Component Analysis (PCA)**:
    - To manage high-dimensional data, we applied **PCA** for dimensionality reduction while retaining the most significant variance in the data. This step facilitated more effective clustering and clearer visualizations.
3. **t-Distributed Stochastic Neighbor Embedding (t-SNE)**:
    - For non-linear dimensionality reduction and handling more complex data structures, we employed **t-SNE**. This method is particularly useful for visualizing clusters in a lower-dimensional space, providing a more intuitive separation between clusters.
4. **Visualization and Saving Results**:
    - We visualized the clusters using t-SNE, which offered an intuitive and clear depiction of cluster separations.
    - The t-SNE results and cluster assignments were saved to a CSV file named tsne_clusters.csv.
5. **Model Persistence**:
    - The trained K-Means model was saved using the **joblib** library as **kmeans_model.pkl** for future use and predictions.

Model Evaluation:

To evaluate the performance of our clustering model, we utilized the saved **K-Means** model (**kmeans_model.pkl**) to predict the cluster for the first row (**zeroth tuple**) of the scaled dataset (**df_scaled**). This step ensures that the model can correctly assign clusters to new or existing data points based on learned patterns.

Feature Selection and Reduction:

For enhanced usability and to simplify the user input process, we limited the model to use only the top **7 key features** based on **PCA principal components**. The steps involved are:

- **Identifying Key Features**:
  - We conducted a feature importance analysis using PCA to identify the 7 most significant features contributing to the clustering process.
- **Data Subsetting**:
  - The scaled data was subset to include only these top 7 features, resulting in a new DataFrame **df_reduced**.
- **Training on Reduced Features**:
  - **K-Means Clustering** was applied to **df_reduced** to train a new clustering model.
- **Enhanced Visualization with t-SNE**:
  - We applied t-SNE on the reduced data to achieve better visualization of the clusters in a two-dimensional space.
- **Saving Reduced Model and Results**:
  - The t-SNE results and corresponding cluster assignments for the reduced feature set were saved to **tsne_clusters1.csv**.
  - The K-Means model trained on the reduced dataset was saved as **kmeans_model1.pkl**.
- **Prediction on Reduced Features**:
  - We predicted the cluster for the first row (zeroth tuple) of **df_reduced** using the newly trained K-Means model. This prediction aligns with the earlier model that used the full feature set, ensuring consistency in cluster assignments.

Deployment:

To facilitate user interaction and accessibility, the project was deployed on **Streamlit** at **share.streamlit.io**. The deployment includes:

- **User Interface**:
  - Users can input values for the selected 7 features to predict the cluster assignment for a concerned student.
- **Visualization**:
  - The t-SNE visualizations allow users to see cluster separations and understand where their input data point falls within the existing clusters.
- **Ease of Use**:
  - The streamlined interface ensures that users can easily and comfortably interact with the model without requiring deep technical knowledge.

Interpretability:

Understanding the factors that influence cluster assignments is crucial for actionable insights. To enhance model interpretability, we employed the following techniques:

- **Principal Component Analysis (PCA)**:
    - PCA was used to identify the principal components that capture the most variance in the data. This helps in understanding which features contribute most to the clustering.
- **Feature Importance**:
    - By analyzing the loadings of PCA components, we determined the importance of each feature in defining the clusters.
- **Cluster Profiling**:
    - We examined the characteristics of each cluster by analyzing the mean values of the key features within each cluster. This profiling helps in identifying common traits and potential mental health risk factors associated with each cluster.
- **Actionable Insights**:
    - The interpretability analysis provides educators and administrators with insights into which areas (e.g., academic performance, attendance, parental occupation) are most indicative of mental health risks, enabling targeted interventions.

Conclusion:

The clustering model developed in this project demonstrates the potential for predicting student mental health status based on educational data. By leveraging **K-Means Clustering** and dimensionality reduction techniques like **PCA** and **t-SNE**, we were able to group students into distinct clusters without requiring explicit labels for mental health status. These clusters can serve as proxies for identifying students who may share similar risk factors, allowing for early interventions.

Through feature selection, we identified the **7 most significant features** that influence cluster formation, simplifying the input process for users without sacrificing predictive accuracy. The deployment of **Streamlit** makes this tool accessible to educators and administrators, enabling them to predict student cluster assignments and visualize clusters with ease.

The interpretability of the model, achieved through PCA and cluster profiling, provides actionable insights into the factors most associated with mental health risks. These insights can guide targeted support strategies, helping schools proactively address mental health challenges among students.

In conclusion, this clustering approach is a promising method for educational institutions to utilize existing data to identify students who may be at risk for mental health issues, ultimately improving student well-being through timely interventions and resource allocation.

References:

1. National Institute of Mental Health and Neurosciences (NIMHANS), 2022 report on adolescent mental health.
2. National Crime Records Bureau (NCRB), 2021 report on student suicides.
3. Indian Council of Medical Research (ICMR), 2022 study on adolescent depression rates.
4. Lokniti-CSDS Youth Study, 2019 survey on college student stress and loneliness.
5. Ministry of Social Justice and Empowerment, 2022 report on adolescent substance abuse in India.