A Comprehensive Study on Mental Health and Resilience of Students using Machine Learning Techniques

Sadman Sakib
Student ID: 1114600
Department of Computer Science
Lakehead University
Thunder Bay, Ontario
ssak2921@lakeheadu.ca

Tahrat Tazrin
Student ID: 1128120
Department of Computer Science
Lakehead University
Thunder Bay, Ontario
ttazrin@lakeheadu.ca

***Abstract*—The ability to handle and overcome adversities, with personal transformation and growth is said to be resilience. It is one of the deciding factors of students' mental health and is vital to recognize the determinants of a positive, developmental reaction in the face of stressful, emotionally demanding states. Mental health is a significant area in medical science however, machine learning techniques has not been applied extensively in this domain. It has been established that resilience is strongly associated with a person's standard of living. Hence, various support groups and programs are initiated to improve one's resilience power in order to achieve a better quality of living. Therefore, to improvise on a person's resilience power, it is especially important to perceive the factors that are directly related to it. Along with that, it is also necessary to know the measure of resilience from those associated factors to have a clear idea of a person's resilience power and how much more they might need to improvise. Therefore, in this research, we focused on exploiting machine learning techniques for predicting resilience levels based on several factors. Furthermore, we have explored the correlations among different factors related to resilience levels and mental health. We also performed clustering to define our own resilience classes and achieved an accuracy of 84.75% with random forest classification model. Therefore, resilience power can be predicted with machine learning models from which people might gain insight of how much to improve upon for a better mental health.**

***Index Terms*—Mental health, Resilience, Machine Learning, Prediction**

I. Introduction

Resilience can be defined as the capability to overcome severe stress and adverse environments. This is an essential concept for health professionals as resilience also defines a more general concept of responding to hurdles that may influence the individual in terms of their health and behavior. It is an important factor in the mental health domain in particular as this domain has an influential role in coping with and recovery from, vital health problems [1].

As a set of personal characteristics, resilience encompasses being able to deal with life events, recognize and observe problems like the opportunity for personal growth, limitations, and personal and collective resources. It also means being able to organize strategies through self-reflection, composure, creativity, enthusiasm, being adaptable and able to act with responsibility and confidence [2]. Thus it can facilitate persons moving on in a positive way from negative, traumatic or stressful experiences.

The majority of the good practices in modern pedagogical methods may support the development of resilience in adulthood however, it has been rarely used as professional development or in education practice and research [3]. It has been suggested that resilience impacts a student's education and professionalism. Resilience has also an instrumental part to play in the clinical presentation of stress disorder. The possibility that resilience can be strengthened by interventions after the early years of life and therefore it is vital for young students to strengthen resilience.

The level of resilience also relates to the quality of life (QoL). [4]. Therefore, to have a better quality of life it is important to improve one's resilience power. Mental health affects all aspects of daily life and therefore consciousness is crucial. It is especially an alarming issue among students as the rate of students dropout affected by mental health problems and difficulties in meeting educational and economic demands are concerning nowadays. As defined by the Canadian Mental Health Association [5], resilience is one of the 5 factors that can describe the mental health of a person. Therefore resilience can be considered as a deciding factor for the mental health and measure of the quality of life for students.

Medical students face various stressors during their clinical years, including difficult clinical events. Academic workload, the consequence of seeing dying patients, as well as external stressors, have all been assumed as reasons for anxiety, leading to larger rates of mental illness than their non-medical students. Encouraging resilience is an assuring way to alleviate the negative consequences of that and help students flourish after stressful experiences [6]. Therefore in this research, we employed a dataset that consists of 1,350 medical students as random samples from 22 Brazilian medical schools [4]. In this work, we implemented machine learning techniques to find out the factors associated with students' resilience level and hence mental health. We intend to predict the resilience level of a student from the dataset which is based on the Dundee Ready Educational Environment Measure (DREEM) [7], World Health Organization Quality of Life questionnaire (WHOQOL) [8], the Beck Depression Inventory (BDI) and the State-Trait Anxiety Inventory (STAI). Here, information from participants includes the Wagnild and Young's resilience scale also.

II. Literature Review

The importance of resilience power of a human being is contemplated by the vast amount of studies conducted on this particular topic. It has been established that resilience has a high effect on a person's quality of living. Therefore, various programs are conducted for people to realize their resilience power and ways of improving upon it for a successful living. Past research has focused on using surveys to associate resilience with various factors that affect itself and gets affected by it. However, applying machine learning tools to such data has not received much attention. Thus, in this paper, we will try to implement the deduced statements in terms of machine learning. The dynamics of resilience should be studied more thoroughly because it is directly related to human psychology. Thus, understanding the concepts in detail will make people more prone to healthy aging and loss management [9]

Resilience is a very significant factor in the field of medical science study. It has been suggested that resilience influences medical student's learning and medical professionalism [3].In a previous research, it was found that a majority of medical students suffer from anxiety (around 85%) and depression (41%) [10]. Coping up with stress and other difficulties in their program might be aggravated with increased resilience power. In another recent study, Tempski et. al discovered that resilience has a strong association with one's quality of life[4]. It was concluded that, the more resilient a person is, the better is their quality of life. Linear regression was performed to back up their statement along with calculation of beta coefficient and confidence interval, and they discovered a dose-effect relationship between these two factors.

A similar study was conducted regarding mental health of engineering students using classification and regression approaches [11]. In the paper, Deziel et. al. found some interesting relationships between mental health condition with respect to various academic factors. They concluded that students having more flexibility in their program had a better mental health score than others. A research conducted by Canadian institute of health and research stated that resilience is a process and the factors associated with resilience may vary [12]. The study tried to highlight the case of aboriginals as to how their history and upbringing also contributed towards them having comparatively lower resilience power. While there might be some cases having multiple variable factors, there are some core factors that always are vitally related to resilience. With this study of resilience with medical students, we will try to figure out those universal factors which can be applied to all cases.

III. Methodology

A. Dataset

The dataset that was used for our analysis is collected from Data Dryad. It contains a random sample of 1,350 medical students from 22 Brazilian medical schools. Information from participants included the Wagnild and Young's resilience scale, the Dundee Ready Educational Environment Measure (DREEM), the World Health Organization Quality of Life measure (WHOQOL), the Beck Depression Inventory (BDI) and the State-Trait Anxiety Inventory (STAI) as well as some other personal features of each students like their gender, age, school status and school location. There were a total of 22 features with around 1350 instances. The Wagnild and Yong's resilience scale is an overall measure of resilience, consisting of 14 items that was clustered in 5 domains: self-reliance, meaning, equanimity, perseverance and existential aloneness. The students' quality of life was recorded both in terms of their medical aspect as well as the overall aspect. The overall quality of life was evaluated as global self-assessment and a set of questionnaires validated by world health organization quality of life (WHOQOL-BREF). In WHOQOL-BREF, 26 items were clustered in four fields: environment, psychological, social relationships and physical health that ranged from 0 – 100. The higher the scores are, the better is their quality of life. Another set of features are the educational environment in terms 5 domains - learning, teachers, social self-perception, academic self-perception and atmosphere which was measured by Dundee Ready Educational Environment questionnaire. Beck Depression Inventory was used to get an approximation of depression and state trait anxiety inventory were used to get a measure of state anxiety and trait anxiety of the students.

B. Dataset Pre-processing

The dataset required some pre-processing steps. The feature DREEM Global score was just the summation of other DREEM features and did not have any significance on its own. Thus the feature was excluded from the dataset while building our models. The categorical features gender, school legal status & school location had two values for each, so, they were encoded using label encoder, where one of the categories of the feature was assigned to 0 and another to 1. Other categorical features that had multiple values like group, overall QoL and medical QoL were encoded using one-hot encoding. This caused all categories of each feature to be assigned to a column and the column had a value of 1 in rows where it was present. The age feature was used as both numerical and categorical features. The raw values were used for age when used as numerical feature. For converting it to a categorical feature, binning method was used. The values of age were distributed in bins and each bin was considered as a category. Several number of bins were tried out and it was seen that for 3 bins, the instances were distributed almost equally. Therefore three bins were used to categorize the values, one from 17 to 21 (493 instances),another from 22 to 24(542 instances) and lastly from 25 to 40 (315 instances).

Before constructing the models, the data was scaled using standard scaling method z score normalization also known as standardization. It normalized the data within 0 to 1 by excluding the mean and balancing to unit variance. The standard score of sample x can be calculated as:

$$z = (x - \mu)/\sigma$$

here, μ = the mean of the training samples
σ = standard deviation

Our class variable, resilience score, contained numerical values initially. In order to classify resilience, the resilience score was discretized into 6 classes according to a previous study [2]. There, a resilience score of 14 - 56 was considered as very low, 57 - 64 as low, 65-73 as moderately low, 74 - 81 as moderately high, 82 - 90 as high and 91 - 98 as very high. This resilience class was used as our class variable initially. While using the resilience as 6 categories as mentioned, an issue of class imbalance arose. For example, the number of data in each category were 87, 89, 188, 348, 435 & 203 respectively as shown in Table II. Therefore Synthetic minority oversampling technique (SMOTE) was applied to the minority classes in order to balance the data. SMOTE is as oversampling technique, that generates synthetic data points for minority class. Here, a line is generated between two data points of a minority class and a new data point is synthesized by picking a point in the line. This is continued until all the classes become balanced. The technique was validated using 10 fold cross validation, where smote was applied on the training folds and classifiers were trained based on the SMOTE applied data. The models were then tested on the original test fold.

C. Clustering

Apart from implementing the resilience classes as stated in the previous study, we defined our own resilience classes. In order to divide the classes, K- means clustering was performed. In clustering, data points are divided into k number of clusters (groups), where each point in a cluster has minimum difference with other points in the same cluster. A silhouette width is calculated in order to measure the fitness of the clusters. For implementing clustering, a range of clusters (2-10) was constructed and their respective silhouette width was calculated. It was seen that for k = 2, the clusters are the most compact (figure 1), with a high silhouette score. Therefore, two classes for the two clusters was considered and the problem became a binary classification problem with class 0 (16- 73) & class 1 (73- 98). When k was 3, the cluster fitness was the second best. Thus, resilience was also divided into 3 categories according to cluster assignments resulting class 0 (16-65), class 1 (66-82) and class 2 (83-98). The cluster arrangements when k was 2 and 3 are shown in figure 2 and 3 respectively.

D. Correlation

We attempted to figure out if a correlation exist between the features and resilience. For this, we applied different measures

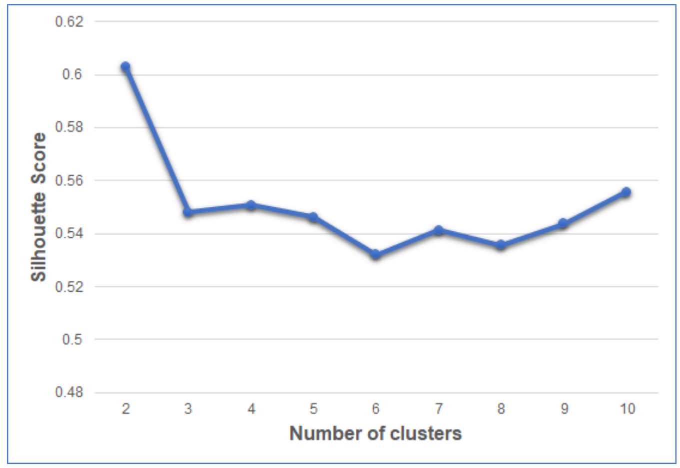


Fig. 1. Values of Silhouette Width for Various Clusters

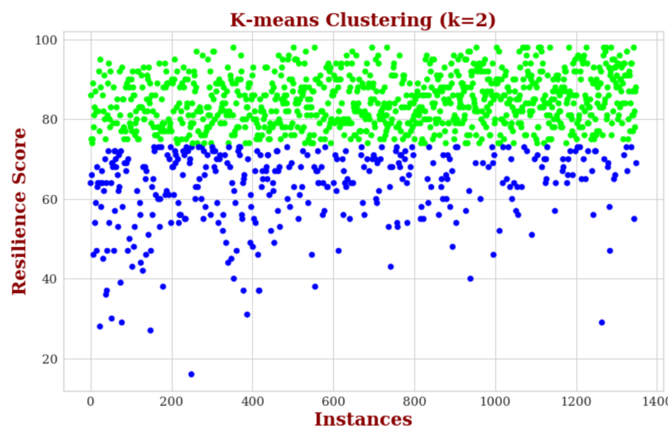


Fig. 2. Cluster Arrangement when k=2

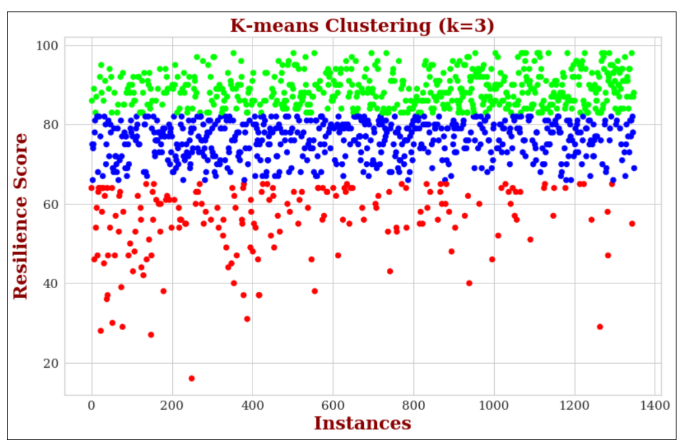


Fig. 3. Cluster Arrangement when k=3

of correlation in order to see which of the measures will give us the most accurate correlation estimate. The correlation measures that we used are Pearson's Correlation Coefficient, Spearman rank order correlation and Brownian Correlation. The pearson correlation is a measure of correlation which gives a value between -1 and 1. Usually, the Pearson coefficient is achieved via a Least-Squares fit and a value of 1 expresses a perfect positive relationship, -1 a perfect negative relationship, and 0 indicates the absence of a relationship between features.

$$\rho = \frac{\text{cov}(X,Y)}{\sigma_x \sigma_y}$$

The spearman rank order correlation measures the strength and directionality of the relationship between the two factors. Spearman's rho can be interpreted as a rank-based version of Pearson's correlation coefficient, which can be utilized for variables that are not normal-distributed and have a non-linear relationship.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

where, d = the pairwise distances of the ranks of the variables. n = the number of samples.

The Brownian covariance is motivated by the generalization of the notion of covariance to stochastic methods. The square of the covariance of random variables X and Y can be formulated in the following way:

$$cov(X,Y)^2 = E[(X - E(X))(X' - E(X'))(Y - E(Y))(Y' - E(Y'))]$$

In addition to that, a heat map was constructed to visualize the correlation among all features. It is a graphical representation of data that uses a system of color-coding to represent different values. The features having various correlation values had various colours representing the correlations.

E. Feature Importance

A feature selection was performed in order to know the most significant features for predicting resilience. If the most important features can be found, more detailed information with respect to those features can be collected to improve classifier performance in the future. For selecting the features, random forest feature selection method was used. In this method, the important features are selected based on decrease in gini impurity on each node. The node where the gini impurity decrease is maximized is considered to be the most important feature. Therefore, the feature with highest impurity was considered as the most important feature.

F. Classification

To classify the resilience level, three machine learning algorithms were used, K-Nearest Neighbor, Random Forest and Multilayer Perceptron.

KNN is a lazy learning method that was used to classify this problem. It is considered to perform optimally as it only considers the k nearest points around it instead of computing for the whole dataset. Based on the k neighbors, the algorithm determines class label by taking the most frequent value among the neighbours. The neighbors are determined by calculating distance of all the points from test data points. The k data points having minimum distance from the points are considered as k neighbors. In order to calculate distance, various distance measures can be used. In our case, we used euclidean distance to calculate the distance between instances. In K-NN, K is a hyper-parameter that can be altered to find the optimal value of k.

Random forests are an ensemble learning method for classification, regression. Multiple decision trees are formed at the time of training and a bootstrap sample technique is used for outputting the class that is the mode of the classes of the individual trees for classification purposes. A set of decision trees are constructed from a randomly selected subset of the training set. Then it aggregates the majorities from all the decision trees to decide the final class of the test instance. In random forest, the measure the quality of a split can be calculated using gini impurity or entropy. In our study, we exploited gini index for the purpose of split. Also, we considered $\sqrt{M}$ (M=number of features) number of features when looking for the best split. In addition to that, choosing the number of decision trees that should be regarded to make a voting decision is also very important. If the number of trees is very small, the algorithm will not use enough data to build a proper model, on the other hand, if the number of trees is very high, same features will be used several times increasing the correlation between trees. Therefore in random forest, the number of trees can be considered as a hyper-parameter for the model.

Multilayer perceptron (MLP) is a type of artificial neural network. It is composed of more than one perceptron. They are composed of one or more layers of neurons. The three major components are input layer, hidden layers and output layer. Data is inserted into the input layer and after being interpreted at the hidden layers, predictions are made at the output layer or visible layer. Here, the number of hidden layers can be multiple, it depends on the type of problem being solved. For our model, we have used only one hidden layer. While predicting the binary classification problem, the number of output layer neurons was 1, however, while predicting the multiclass problem where there were 3 classes, the number of output neurons were considered as 3.

For K-Nearest neighbor and Random Forest, the hyperparameters were tuned using grid search. The best hyperparameter that gives maximum accuracy was found out. Then that configuration was utilized to measure the performance on the test dataset as well. The grid search was validated using 5-fold cross validation.

G. Regression

Two regression techniques were applied to predict the continuous value, resilience score of the dataset. The techniques were ridge regression and random forest regression. The features were all scaled using min-max feature scaling in order to build the model. The class variable was taken in two different ways, once it was scaled using min-max feature scaling and another time the values were divided by 100 as scaling caused the ratio of the values to be altered when brought within a range of 0 to 1. The models were validated using 5 fold cross validation. A value was predicted and the error of the predicted value from the original value was calculated. To observe regression performance, three error measures were used, mean absolute error, mean squared error and root mean squared error.

Ridge Regression is a technique for analyzing multiple regression data having multicollinearity issue. Multicollinearity occurs when a linear relationship exists between the independent variables. When multicollinearity takes place, least squares estimates are unbiased, but their variances are large so they may be far from the true value. Ridge regression adds a degree of bias to the regression estimates and reduces the standard errors. Therefore, a more reliable estimate of regression can be expected due to the bias variance trade-off. For our model, the regularization strength, that states the level of variance reduction alpha was used as 1.

Random Forest algorithm, as mentioned in classification section, can also be used for regression problems. The idea is pretty much similar to the classification problem, except here regression trees are used to build the forest. Therefore, they are an ensemble of different regression trees and are used for nonlinear multiple regression. Each leaf contains a distribution for the continuous output variables and the final output is decided based on the mean of all the regression tree outputs. For our problem, number of trees were considered as 1000 and maximum number of features in each tree was $\log_2 n$.

IV. Justification

During data set pre-processing, age binning was done to categorize the age values. It is believed that the resilience of a person varies within a scope of certain years, rather than every passing year. Therefore age was grouped to find the resilience among age groups rather than with particular ages. Also, one hot encoding method was used to convert the categorical data as it gives a better interpretation of the categorical data rather than other numerical representations of it, say label-encoding. In addition to that, the values were scaled with z score normalization technique as applying this makes the data more centralized towards 0. It also takes into account the variance of the dataset. Therefore, representing the data more meticulously. As mentioned earlier, there was a huge variation in the number of instances when the class variables were divided into six labels. This induced an issue of class imbalance in the dataset. The problem with class imbalance is that, during classification, the majority classes dominates the minority class. This might lead to wrong classification of the dataset. Therefore,SMOTE technique was used to deal with this issue, especially because, it creates new datapoints from the dataset instead of taking an existing point randomly.

Clustering was performed on the resilience score in order to define our own resilience level. This was done because the resilience level according to the previous study was not verified, and therefore there was a chance of discrepancy in the ranges they used. Therefore, we performed clustering on the dataset to see how many groups we can get from the provided resilience scores. Clustering detects the similarities among groups of data and divides them accordingly. Thus, defining the resilience groups this way appeared to be a more accurate way of finding a pattern in the resilience score. This procedure can also increase the model performance as data is said to be more organized into similar groups.

Feature correlation is an important part of classification. Two correlation factors should be taken into account, the correlation between class and features, and the correlation between features with other features. The first one should be maximized and the later one should be minimized as much as possible. In order to figure out these factors, correlation measures were taken. The pearson correlation was used to measure any existing linear correlation in the data. The spearman correlation measure was used to know the presence of any monotonic relationship among the data. Finally, the brownian correlation measures the degree of any kind of relationship between two vectors, feature and class. Therefore, the goal was to determine correlation from various aspects of the data. A heat map was also constructed to give us an idea about the feature to feature correlation.

For classification, K-Nearest Neighbours, Random Forest and Multilayer Perceptron was used. K nearest neighbour was used, as it is a lazy learning method, it computes the model locally for each test instance, thus constructing a less complex model. The computational time is therefore, much lower, and the classification also give good results mostly. Random forest was used as it is an ensemble classifier and encompasses the concept of voting. Therefore, it avoids any type of overfitting that might happen during classification and give more accurate result as it considers vote of majority trees. This was also one of the reasons why random forest was used as a regressor for our model. Finally, neural network model, multilayer perceptron was used because it is considered one of the most efficient models for performing classification. They are very flexible and can be used generally to learn a mapping from inputs to outputs. For regression, another technique was used which is the ridge regression. We chose ridge regression instead of any other models like linear regression is because ridge regression is a more precise variation of linear regression which also takes into account factors like regularization strength for referring to the bias variance trade-off factor to construct a more accurate model.

V. Results and Discussion

A. Feature Importance

Table II demonstrates feature importance scored generated using random forest classifier and feature correlation with resilience score. It is evident that the most important features appear to be more correlated with resilience. Trait anxiety has the highest importance score among all the features as well as it is the highly correlated (negatively correlated) with resilience score. Also, the least important features appear to be the least correlated features as well. As we can see, the four least important features (group, sex, school location and school status) had a very low correlation with resilience score as well which explains the reason for being less important in classifying resilience.

Feature	Feature Importance	Correlation with Resilience
Trait_anxiety	0.1068	**-0.6055**
psychological	0.0952	**0.6298**
academic_self_perception	0.0769	0.4865
BDI	0.0750	-0.5710
State_Anxiety	0.0686	-0.4394
learning	0.0635	0.3377
atmosphere	0.0632	0.4454
teachers	0.0618	0.2818
social_self_perception	0.0571	0.4398
environment	0.0553	0.2964
physical_health	0.0532	0.4359
social_relationships	0.0480	0.3508
Age	0.0468	0.0767
Medical_QoL	0.0370	0.2851
Overall_QoL	0.0326	0.2836
Group	0.0199	-0.0267
Sex	0.0135	-0.0442
School_location	0.0132	0.0883
School_legal_status	0.0116	0.0798

TABLE I: Feature Importance and Correlation with Resilience (Sorted in descending order according to Importance Score)

B. Feature Correlation

Correlations of the topmost important features with resilience are shown in Figures 4 and 5. Figure 4 demonstrates the Pearson and Spearman correlation.

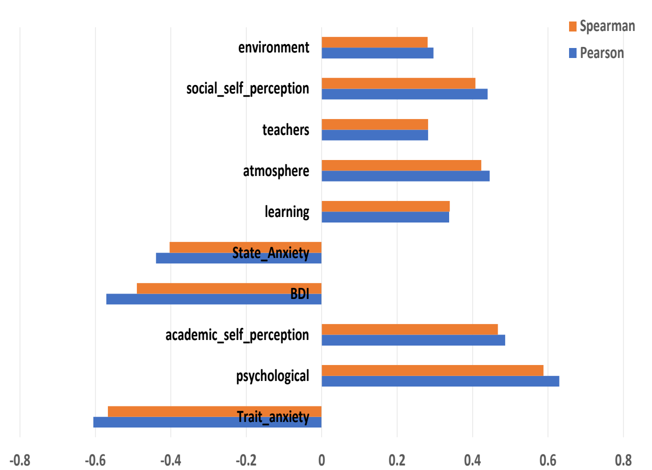


Fig. 4. Pearson and Spearman correlation comparison of top 10 important features

Both the correlation measure technique gave very similar results which are visible from the figure. Figure 5 is further demonstrating the Brownian correlation of the important features. From Figures 4 and 5 we can see that all three correlation measures gave very identical results as trait anxiety and psychological features had the highest correlation in all three measures of correlation.

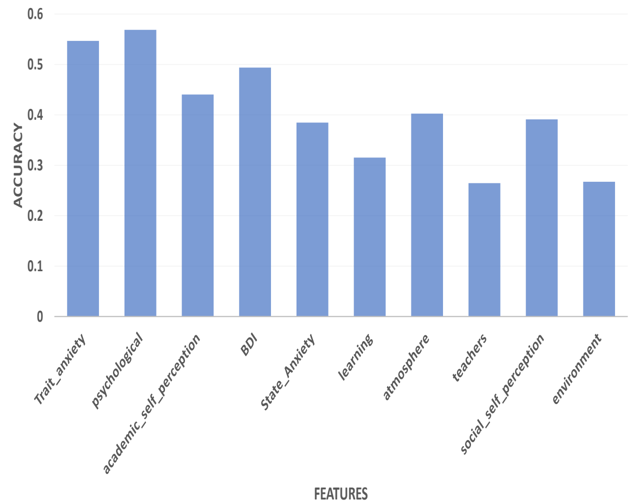


Fig. 5. Brownian correlation comparison of top 10 important features

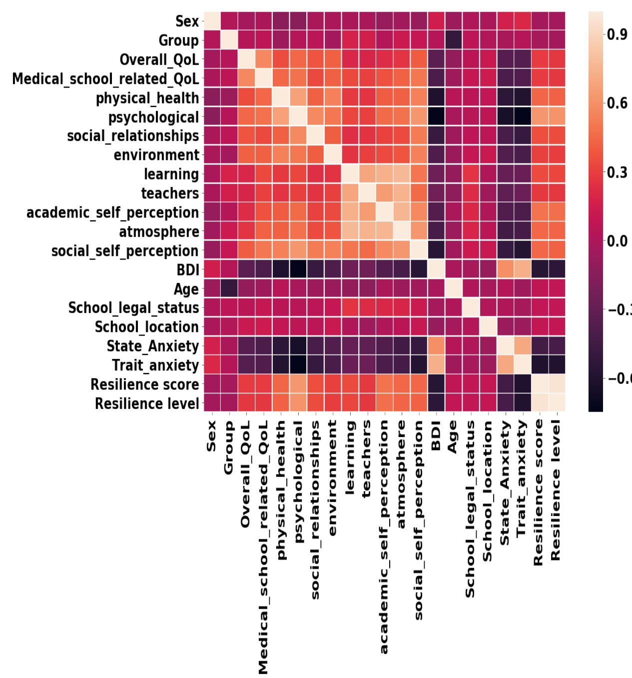


Fig. 6. Correlation heatmap among all features

Figure 6 represents the heatmap of correlations among all the features in the dataset. In this figure, the whitish cells represent a higher correlation whereas the blackish cells represent a lower correlation.

C. Classification of Resilience Levels

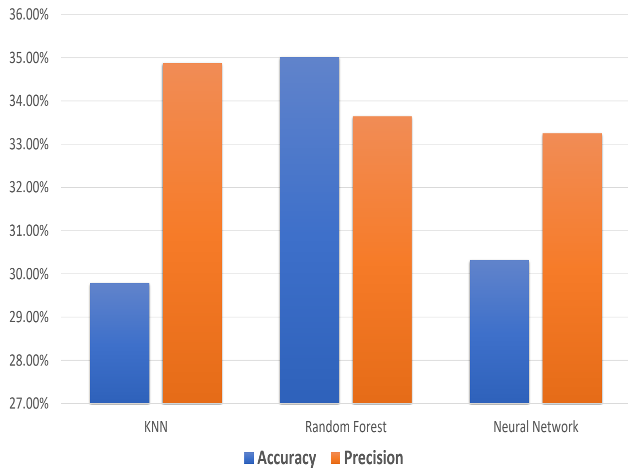


Fig. 7. Performance of classification of six resilience levels

Resilience level	Number of instances
Very High	435
High	348
Moderately High	203
Moderately Low	188
Low	89
Very Low	87

TABLE II: Class distribution for six levels of resilience

Figure 7 exhibits the classification performance after applying SMOTE and Stratified 10 fold cross-validation. Here, six levels of resilience were taken into consideration for classification. All three classifiers performed poorly and among the three Random Forest (35% accuracy) performed comparatively better than KNN and MLP in terms of accuracy and precision.

As we can understand from figure 6, our classification technique for class 6 did not perform much accurately. Hence, we thought of defining a different categorization of resilience. For that purpose, we applied K-means clustering only on the resilience score to find out for how many clusters the clustering fitness is the largest. This was represented in figure 1 where the silhouette score was the largest while the number of clusters was two. Therefore, in our further analysis, we converted the six classes resilience into two classes and then applied our classification algorithms. Furthermore, we also selected number of cluster 3 to generate 3 resilience levels and then applied the classification algorithms.

Figure 8 and 9 illustrates the result of hyper-parameter tuning for each of the methods. To choose the best hyper-parameters, we tested on the following list of parameters:

- For KNN: Adjusted the value of K from 1 to 30
- For Random Forest: Adjusted the number of trees from 2 to 120
- For Neural Network: One hidden layer was selected. We adjusted the number of nodes in the hidden layer from 1 to 20. The number of nodes in the output layer was 1 and 3 for binary and multi-class classification respectively.

The best parameters selected from the hyper-parameter tuning of binary classification and multi-class classification are presented in Table III. After learning the best possible value of these parameters from the training dataset, we have used these parameters for testing the test data. We have used the selected parameters to further apply the methods on the testing dataset. Figure 10 and 11 represents the performance of three classifiers using binary and multi-class classification scheme respectively. From Figure 10, we can see that, for binary classification of resilience, the Random Forest classifier performs better than both KNN and Neural Network with the accuracy of 84.75% for the dataset where age categorized and 83.63% where age numerical.

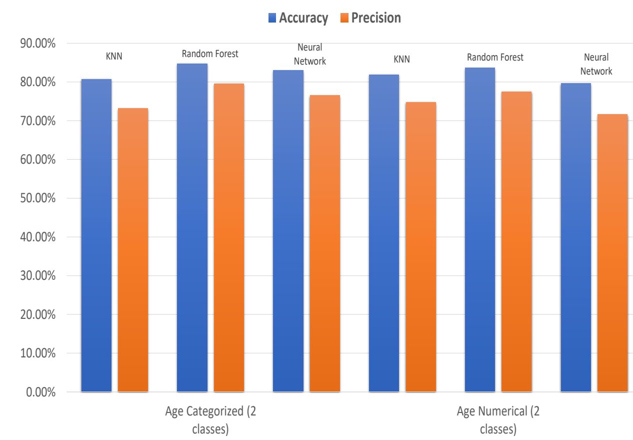


Fig. 10. Performance of binary classification on test data

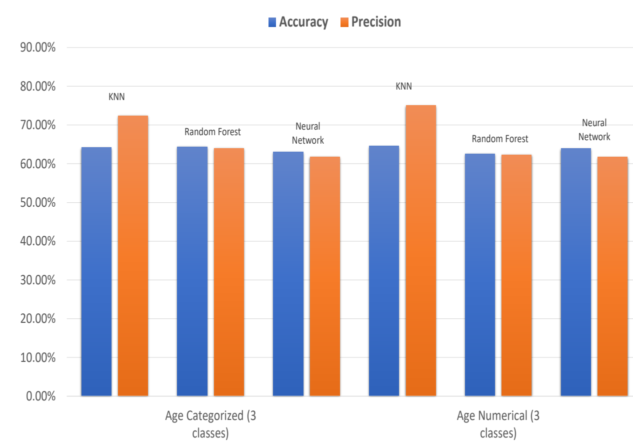
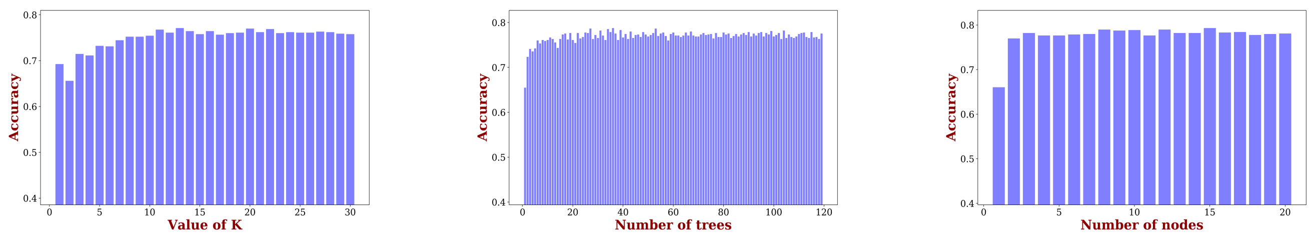


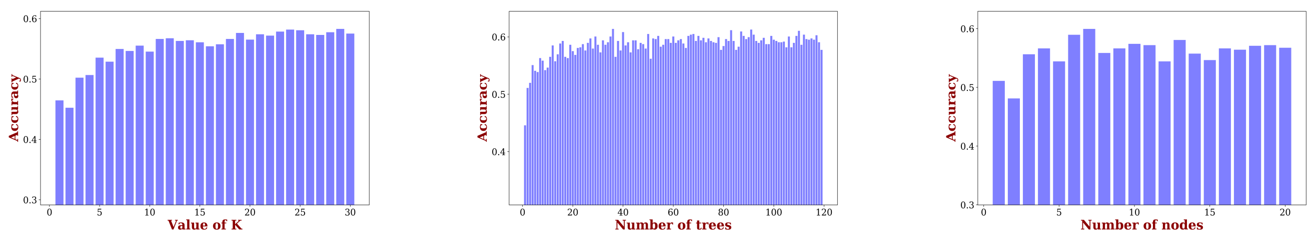
Fig. 11. Performance of 3 class classification on test data

For multi-class classification, figure 11 represents that, each of the three methods performs significantly poorly compared to the results of binary classification. For each of the methods, the accuracy is mostly around 60% after using the best parameters selected using 5 fold cross-validation with grid search. KNN performs better than both Random Forest and Neural Network



(a) KNN (b) Random Forest (c) Neural Network

Fig. 8. Hyper-parameter tuning results for binary classification



(a) KNN (b) Random Forest (c) Neural Network

Fig. 9. Hyper-parameter tuning results for Multi-class (3 classes) classification

Classification Scheme	**KNN (Selected Value of K)**	**Random Forest (Selected Number of Trees)**	**Neural Network (Selected number of nodes in the hidden layer)**
Binary classification	13	37	15
Multi-class classification (3 classes)	29	37	7

TABLE III: Selected hyper-parameters after Grid Search with 5 Fold Cross-validation

both in terms of accuracy and precision as well. In terms of accuracy for KNN is slightly higher than the other two methods. However, KNN outperforms Random Forest and Neural Network in terms of precision value for both when age categorized (0.723592) and numerical (0.750057).

D. Regression

Figure 12 illustrates the comparative analysis of Ridge Regressor and Random Forest Regressor. It is evident that, in all three performance measures, Ridge Regressor was slightly more reliable than Random Forest Regressor. Also, we transformed the resilience score using two different methods, in one method it was scaled using min-max scaling and in another approach we normalized it by dividing the resilience score with 100 (max range of resilience score). The error rate significantly lower when the resilience score was normalized by dividing by 100.

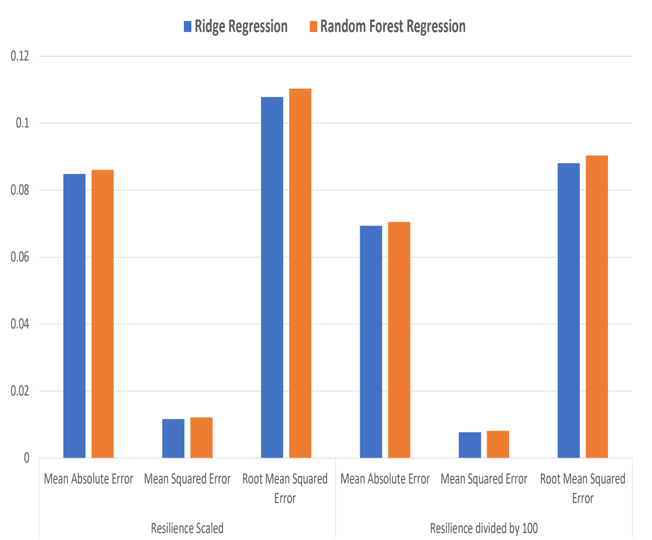


Fig. 12. Results of Ridge and Random Forest Regression after 5 Fold Cross-validation

VI. Conclusion and Future Work

Resilience is a significant factor for keeping a person mentally stable and strong against different kinds of adversarial situations. In this research, we have explored factors associated with medical students' mental health and applied Machine Learning techniques to classify the resilience of the students. Among different factors, we have found psychological factors and anxiety traits are highly correlated with resilience compared to other factors. We exploited K-Nearest Neighbour, Random Forest and Neural Network to classify the resilience of medical students. Among these three methods, the overall performance of Random Forest was more efficient than the other two in terms of classification. We also performed a regression analysis to predict resilience score where Ridge Regressor outperformed Random Forest Regressor.

There were some shortcomings in the project, for example, in the dataset the features did not have much covariance with the class label. This brings a question as to whether the right set of questions were asked in order to collect the data. Therefore, in our future work, a survey will be conducted with relevant set of questions in other medical schools as well as educational institutions to generate new dataset with which the experiments will be performed again. Additionally, determining the mental health of a patient is a crucial part of any treatment. Therefore, an application can be developed that will measure the resilience of a person. It can be used in emergency rooms of hospitals to measure resilience in addition to other initial medical examinations.

Acknowledgement

The authors had no conflict during any part of the research. Both authors were equally involved in this research. Authors would like to express their gratefulness to Dr. Thiago E. A. Oliveira for his valuable suggestions throughout the research.

References

[1] P. A. Atkinson, C. R. Martin, and J. Rankin. Resilience revisited. *Journal of Psychiatric and Mental Health Nursing*, 16(2):137–145, 2009.

[2] Patricia Tempski, Milton A. Martins, and Helena B M S Paro. Teaching and learning resilience: A new agenda in medical education. *Medical Education*, 46(4):345–346, 2012.

[3] Amanda Howe, Anna Smajdor, and Andrea Stöckl. Towards an understanding of resilience and its relevance to medical training. *Medical Education*, 46(4):349–356, 2012.

[4] Patricia Tempski, Itamar S. Santos, Fernanda B. Mayer, Sylvia C. Enns, Bruno Perotta, Helena B.M.S. Paro, Silmar Gannam, Munique Peleias, Vera Lucia Garcia, Sergio Baldassin, Katia B. Guimaraes, Nilson R. Silva, Emirene M.T. Navarro Da Cruz, Luis F. Tofoli, Paulo S.P. Silveira, and Milton A. Martins. Relationship among medical student resilience, educational environment and quality of life. *PLoS ONE*, 10(6):1–13, 2015.

[5] Canadian Mental Health Association. Mental health meter - cmha national, 2013.

[6] Sarath Burgis-kasthala, Nicholas Elmitt, Lillian Smyth, Malcolm Moore, Nicholas Elmitt, Lillian Smyth, Malcolm Moore, Nicholas Elmitt, Lillian Smyth, and Malcolm Moore. Predicting future performance in medical students . A longitudinal study examining the effects of resilience on low and higher performing students. *Medical Teacher*, 0(0):1–8, 2019.

[7] S U E Roff, Sean Mcaleeri, Ronald M Harden, Mona Al-qahtani, Ashraf Uddin, Horacio Deza, Guido Groenen, and Praorn Primparyon. Development and validation of the Dundee Ready Education Environment Measure (DREEM). 19(4):295–299, 1997.

[8] Z Metelko, S Szabo, Metabolic Diseases, Shuba Kumar, New Delhi, Van Heck, J De Vries, J Arroyo Sucre, M Kabanov, Yooth Bodharamik, Kitikorn Meesapya, United Kingdom, W Acuda, C D Jenkins, and J Oliver. Pergamon THE WORLD HEALTH ORGANIZATION QUALITY OF LIFE ASSESSMENT (WHOQOL): POSITION PAPER FROM THE WORLD HEALTH ORGANIZATION. 41(10), 1995.

[9] Gill Windle. What is resilience? a review and concept analysis. *Reviews in Clinical Gerontology*, 21:152 – 169, 05 2011.

[10] Fernanda Brenneisen Mayer, Itamar Souza Santos, Paulo SP Silveira, Maria Helena Itaqui Lopes, Alicia Regina Navarro Dias de Souza, Eugenio Paes Campos, Benedita Andrade Leal de Abreu, Itágores Hoffman II, Cleidilene Ramos Magalhães, Maria Cristina P Lima, et al. Factors associated to depression and anxiety in medical students: a multicenter study. *BMC medical education*, 16(1):282, 2016.

[11] Melissa Deziel, Dayo Olawo, Lisa Truchon, and Lukasz Golab. Analyzing the mental health of engineering students using classification and regression. In *Educational Data Mining 2013*, 2013.

[12] John Fleming and Robert J Ledogar. Resilience, an evolving concept: A review of literature relevant to aboriginal research. *Pimatisiwin*, 6(2):7, 2008.