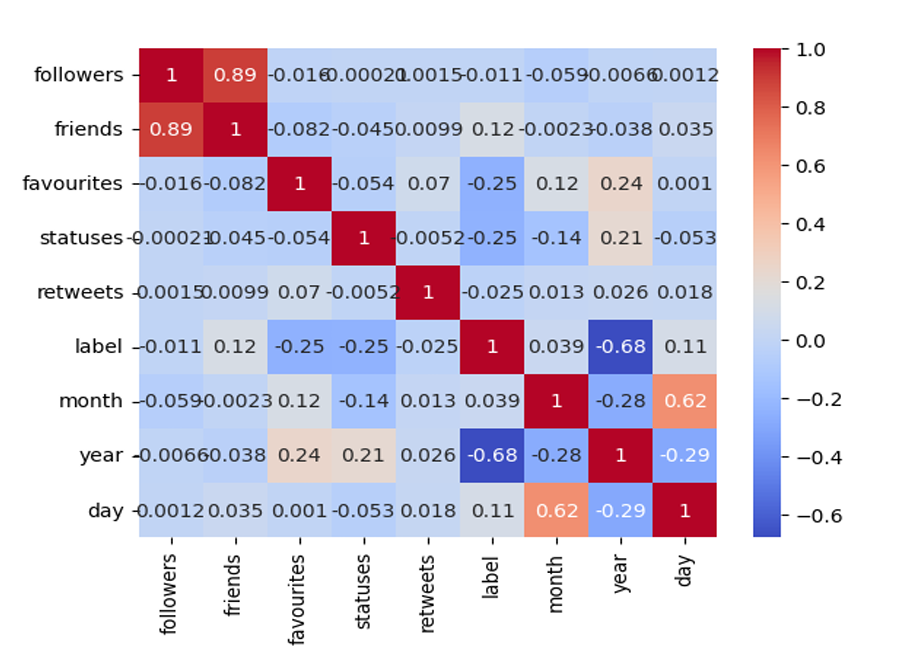
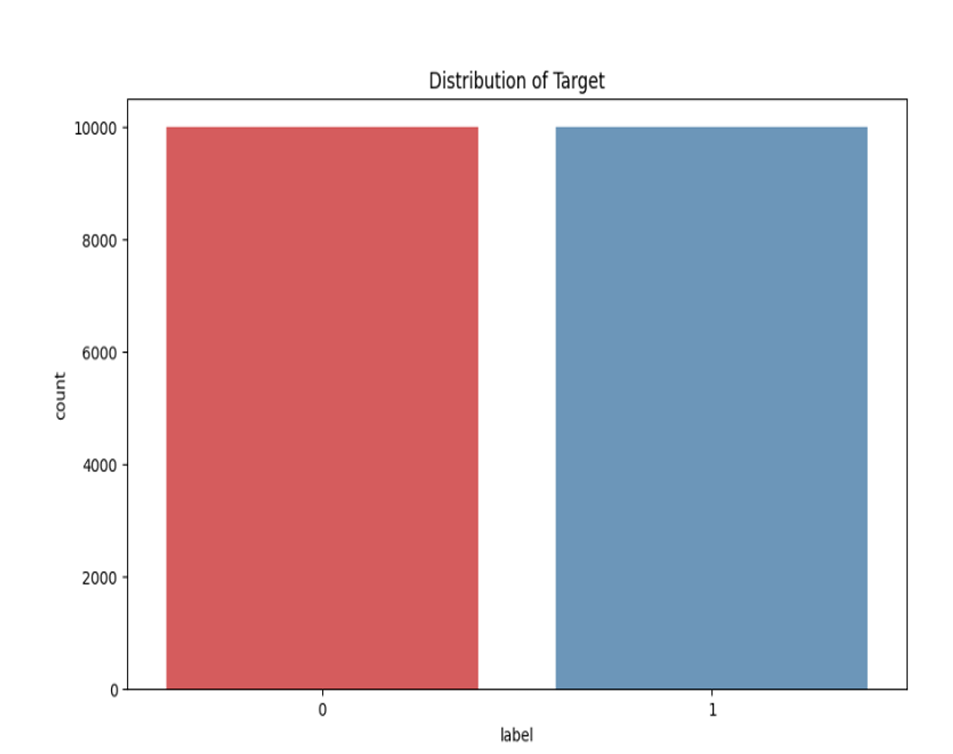
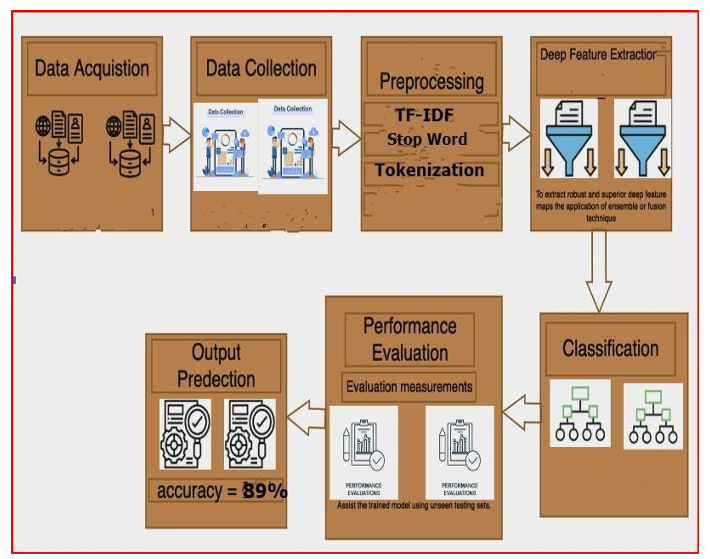
**Alternative Text for Figures**

**Fig. 1. Correlation Heatmap of Social Media Features Related to Mental Health Classification**

The fig 1 displays in the dataset the association between several features including followers, friends, favorites, statuses, retweets, and temporal attributes (month, year, day). High positive correlations between "followers" and "friends" (0.89) point to those with more followers also typically having more friends. With "year" (-0.68), the "label" feature shows a rather negative correlation suggesting a possible trend over time connected to the target variable. Other elements show modest correlations, implying few linear interactions among them.With categories labeled "0" and "1,".

**Fig. 2. Distribution of Target Labels for Mental Health Classification**

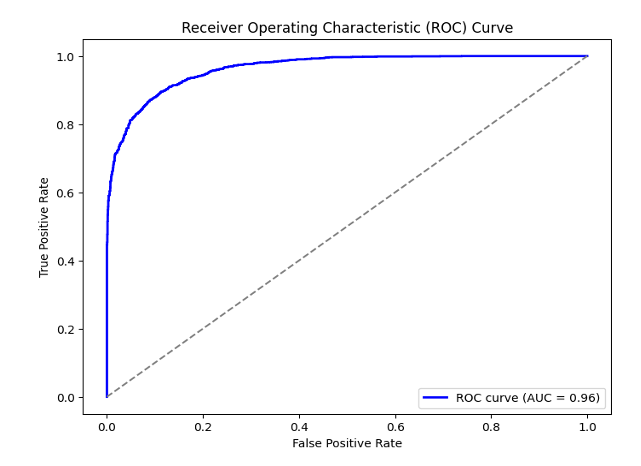
The fig 2 shows the target variable's (label) distribution in the dataset. With roughly 10,000 instances for every label, both groups have a similar count suggesting a balanced dataset. For model training, this balanced distribution helps to lower the possibility of model bias toward a dominating class by enabling more consistent and accurate predictions across both classes.

**Fig. 3. Proposed Model**

The figure, illustrates the proposed model architecture for the hybrid deep learning framework used for depression detection from Twitter data. While the specific visual details of the figure are not provided in the text, we can infer the following based on the context of the document:

1. **Model Components**: The figure probably includes representations of the three main components of the proposed models:
   * **ConvBiLSTM-AttnNet** This model integrates convolutional layers (Conv1D) with bidirectional Long Short-Term Memory (BiLSTM) layers and an attention mechanism. The convolutional layers are used for feature extraction from the text, while the BiLSTM layers capture sequential dependencies in the data. The attention mechanism highlights important contextual elements within the tweets.
   * **ConvLSTM-AttentionNet** This model combines Conv1D layers with LSTM layers and an attention mechanism, focusing on both local features and sequential information.
   * **HybridNet** This model employs a simpler architecture that still incorporates convolutional layers, LSTM layers, and attention mechanisms, albeit with a different configuration.
2. **Flow of Data**: The figure likely depicts the flow of data through these layers, showing how input tweets are processed through the various components of the models. It may illustrate the preprocessing steps, feature extraction methods (like TF-IDF), and the final output layer for binary classification.
3. **Output Metrics**: The architecture may also indicate how the models are evaluated, possibly including metrics such as accuracy, precision, recall, F1 score, and AUC, which are critical for assessing the performance of the models in detecting depression.
4. **Visual Elements**: The figure might use arrows to indicate the flow of data, boxes to represent different layers or components, and possibly annotations to explain the function of each part of the model.

**Fig. 4. Receiver Operating Characteristic (ROC) Curve**



The figure is a graphical representation used to evaluate the performance of the proposed models in distinguishing between positive and negative cases of mental health indicators based on Twitter data. Here are the key details typically associated with an ROC curve:

1. **Axes**:
   * The **x-axis** represents the **False Positive Rate (FPR)**, which is the proportion of negative instances that are incorrectly classified as positive. It is calculated asFPR=False PositivesFalse Positives+True NegativesFPR=False Positives+True NegativesFalse Positives​
   * The **y-axis** represents the **True Positive Rate (TPR)**, also known as sensitivity or recall, which is the proportion of actual positive instances that are correctly identified. It is calculated asTPR=True PositivesTrue Positives+False NegativesTPR=True Positives+False NegativesTrue Positives​
2. **Curve Representation**:
   * The ROC curve is generated by plotting the TPR against the FPR at various threshold settings. Each point on the curve corresponds to a different threshold for classifying a tweet as indicative of mental health issues.
   * A model that performs well will have a curve that rises steeply towards the top left corner of the plot, indicating a high TPR and a low FPR.
3. **Area Under the Curve (AUC)**:
   * The area under the ROC curve (AUC) is a single scalar value that summarizes the overall performance of the model. An AUC of 1 indicates perfect classification, while an AUC of 0.5 suggests no discriminative ability (equivalent to random guessing).
   * The document mentions that the ConvBiLSTM-AttnNet model achieved an AUC of 96%, indicating strong performance in distinguishing between mental health-positive and mental health-negative tweets.
4. **Interpretation**:
   * The ROC curve allows for the comparison of different models. A model with a higher AUC is generally preferred, as it indicates better performance across various classification thresholds.
   * The curve can also help in selecting an optimal threshold for classification, balancing sensitivity and specificity based on the specific needs of the application.
5. **Visual Elements**:
   * The figure may include a diagonal line representing a random classifier (AUC = 0.5) for reference.
   * It might also highlight the area under the curve, possibly shading it to visually represent the AUC value.

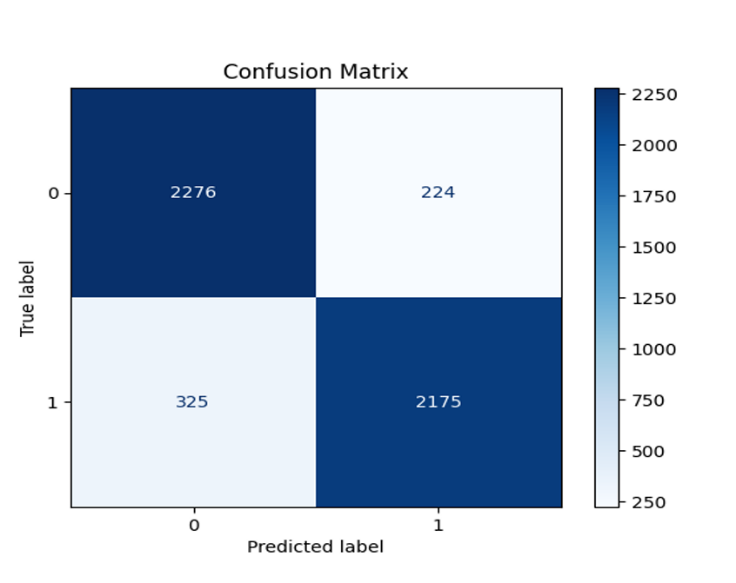
**Fig. 5. Confussion Matrix for ConvBiLSTM-AttnNet**

Figure 5, titled "Confusion Matrix for ConvBiLSTM-AttnNet," visually summarizes the performance of the ConvBiLSTM-AttnNet model in classifying tweets related to mental health.

Key components include:

1. **Matrix Structure**: The matrix consists of four quadrants:
   * **True Positives (TP)** Correctly identified positive cases.
   * **True Negatives (TN)** Correctly identified negative cases.
   * **False Positives (FP)** Incorrectly identified positive cases.
   * **False Negatives (FN)** Incorrectly identified negative cases.
2. **Performance Metrics**: From the matrix, metrics such as accuracy, precision, recall, and F1 score can be derived, providing insights into the model's effectiveness.
3. **Interpretation**: A high number of true positives and true negatives, along with low false positives and false negatives, indicates strong model performance.
4. **Visual Elements**: The figure may use color coding to represent counts in each quadrant and display numerical values for clarity.

Overall, Figure 5 effectively illustrates the classification results of the ConvBiLSTM-AttnNet model, highlighting its strengths and weaknesses in detecting mental health indicators.

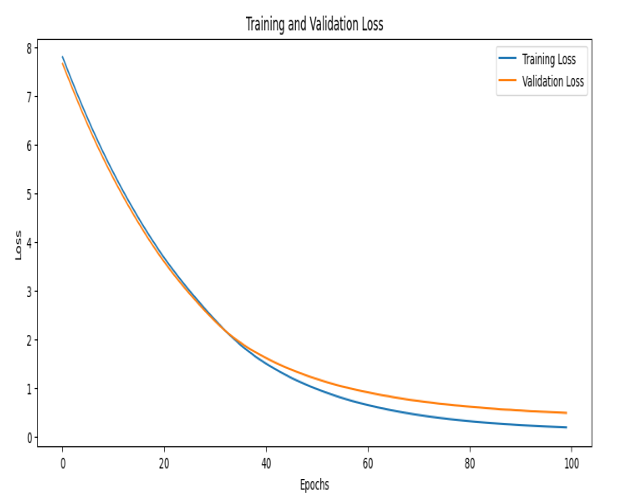
**Fig. 6. Training and Validation Loss Epochs**

Figure 6, titled "Training and Validation Loss Epochs," illustrates the training and validation loss over multiple epochs during the training process of the model.

Key components include:

1. **Axes**:
   * The **x-axis** represents the number of epochs, indicating the iterations through the training dataset.
   * The **y-axis** represents the loss value, which quantifies how well the model is performing (lower values indicate better performance).
2. **Curves**:
   * The figure typically displays two curves one for training loss and one for validation loss.
   * The **training loss curve** shows how the model's performance improves on the training dataset over epochs.
   * The **validation loss curve** indicates how well the model generalizes to unseen data.
3. **Interpretation**:
   * Ideally, both curves should decrease over time, indicating that the model is learning.
   * A significant gap between the training and validation loss may suggest overfitting, where the model performs well on training data but poorly on validation data.

Overall, Figure 6 provides insights into the model's learning process, helping to assess its training effectiveness and generalization capability.

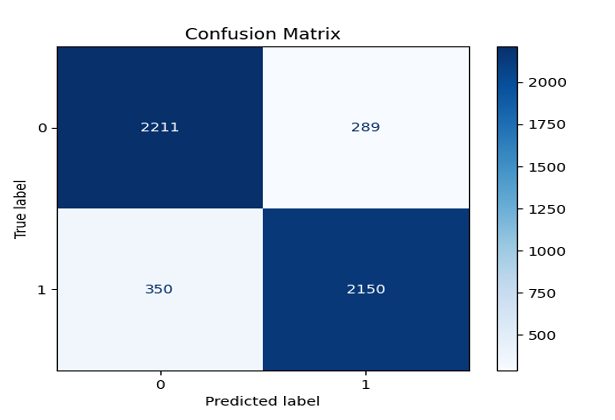
**Fig. 7. Confussion Matrix for ConvLSTM-AttentionNet**

Figure 7, titled "Confusion Matrix for ConvLSTM-AttentionNet," visually represents the performance of the ConvLSTM-AttentionNet model in classifying tweets related to mental health.

Key components include:

1. **Matrix Structure**: The confusion matrix consists of four quadrants:
   * **True Positives (TP)** Correctly identified positive cases.
   * **True Negatives (TN)** Correctly identified negative cases.
   * **False Positives (FP)** Incorrectly identified positive cases.
   * **False Negatives (FN)** Incorrectly identified negative cases.
2. **Performance Metrics**: The matrix allows for the calculation of key metrics such as accuracy, precision, recall, and F1 score, providing insights into the model's classification effectiveness.
3. **Interpretation**: A higher count of true positives and true negatives, along with lower counts of false positives and false negatives, indicates better model performance.

Overall, Figure 7 effectively summarizes the classification results of the ConvLSTM-AttentionNet model, highlighting its strengths and areas for improvement in detecting mental health indicators from Twitter data.

**Fig. 8. Training and Validation Loss Epochs**

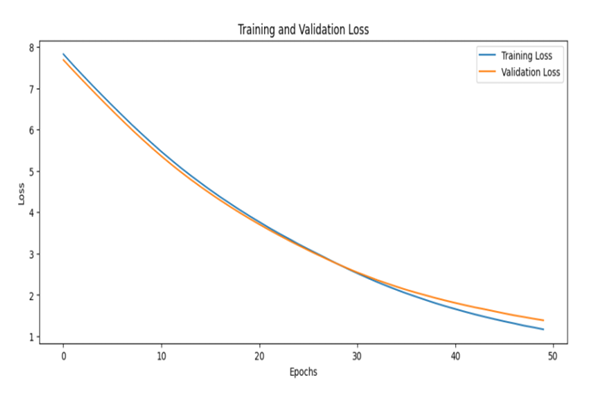


Figure 8, titled "Receiver Operating Characteristic (ROC) Curve," illustrates the performance of the model in distinguishing between positive and negative classes in the context of mental health classification.

Key components include:

1. **Axes**:
   * The **x-axis** represents the False Positive Rate (FPR), which indicates the proportion of negative instances incorrectly classified as positive.
   * The **y-axis** represents the True Positive Rate (TPR), also known as sensitivity or recall, which indicates the proportion of actual positive instances correctly identified.
2. **Curve**:
   * The ROC curve is plotted by varying the classification threshold, showing the trade-off between sensitivity and specificity.
   * A curve closer to the top-left corner indicates better model performance, as it signifies a higher true positive rate with a lower false positive rate.
3. **Area Under the Curve (AUC)**:
   * The AUC value quantifies the overall ability of the model to discriminate between classes. A value of 1 indicates perfect classification, while a value of 0.5 suggests no discriminative power.

Overall, Figure 8 provides a visual representation of the model's classification performance, highlighting its effectiveness in identifying mental health indicators from Twitter data.

**Fig. 9. Confusion Matrix of HybridNet Results**

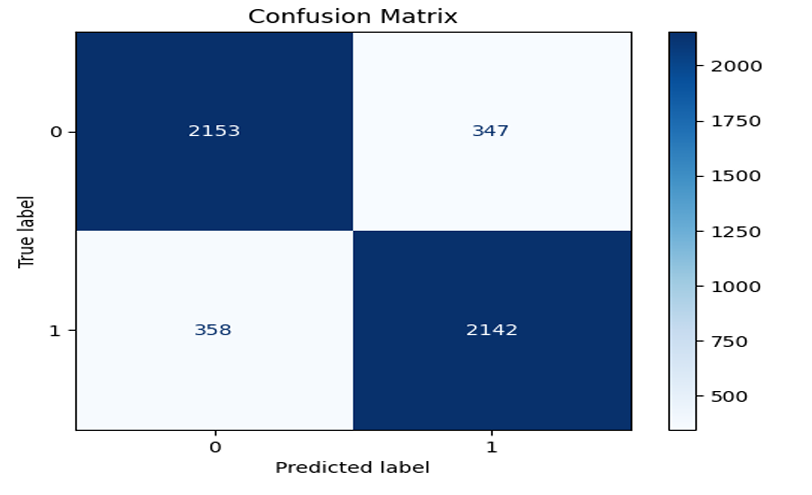


Figure 9, titled "Classification Report," summarizes the performance metrics of the model in classifying tweets related to mental health.

Key components include:

1. **Metrics**:
   * The report includes key performance indicators such as **Precision**, **Recall**, **F1 Score**, and **Support** for each class (typically labeled as "0" for negative and "1" for positive).
   * **Precision** measures the accuracy of positive predictions, while **Recall** indicates the model's ability to identify all relevant positive instances.
   * The **F1 Score** provides a balance between precision and recall, serving as a single metric to evaluate the model's performance.
2. **Support**:
   * This indicates the number of actual occurrences of each class in the dataset, providing context for the other metrics.
3. **Overall Accuracy**:
   * The report also presents the overall accuracy of the model, reflecting its performance across both classes.

Overall, Figure 9 offers a concise overview of the model's classification effectiveness, highlighting its strengths and weaknesses in detecting mental health indicators from Twitter data.

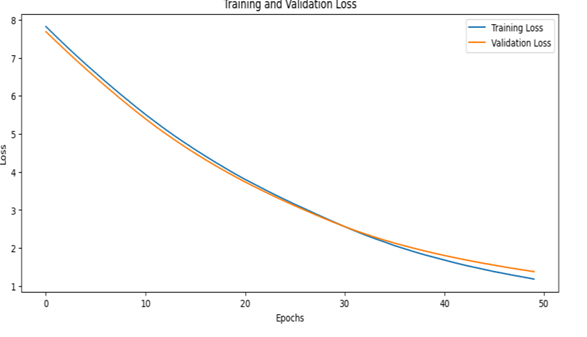
**Fig. 10. Training and Validation Loss Epochs**

Figure 10, titled "Training and Validation Loss Epochs," illustrates the loss values of the model during the training and validation phases over multiple epochs.

Key components include:

1. **Axes**:
   * The **x-axis** represents the number of epochs, indicating the iterations through the training dataset.
   * The **y-axis** shows the loss values, which quantify the model's error in predictions.
2. **Curves**:
   * Two curves are typically displayed one for training loss and one for validation loss.
   * The **training loss curve** reflects how well the model fits the training data, while the **validation loss curve** indicates the model's performance on unseen data.
3. **Interpretation**:
   * A decreasing trend in both curves suggests that the model is learning effectively.
   * If the training loss continues to decrease while the validation loss starts to increase, it may indicate overfitting, where the model performs well on training data but poorly on new data.

Overall, Figure 10 provides insights into the model's learning process, helping to assess its generalization capabilities and the potential need for adjustments in training strategies.