Introduction

Psychological distress, characterized by a state of emotional suffering and mental anguish, is a multifaceted construct encompassing a spectrum of symptoms such as anxiety, depression, and various mood disorders [1]. It poses a significant global health challenge, affecting individuals of all ages, backgrounds, and cultures. According to the World Health Organization (WHO) [2], mental health issues are predicted to be among the leading causes of disability-adjusted life years by the year 2030, underscoring the urgency of better understanding and addressing psychological distress.

Predicting psychological distress is a complex endeavor, as it involves the intricate interplay of cognitive, emotional, and situational factors. Among these factors, maladaptive cognitive schemas and anxiety have emerged as critical contributors to the onset and perpetuation of psychological distress. Cognitive schemas represent core beliefs and mental frameworks that individuals employ to interpret their experiences and emotions [3]. When these schemas become maladaptive, they can engender a cycle of negative thinking, amplifying distressing emotions and behaviors.
Anxiety, on the other hand, serves as a common precursor and co-occurring symptom of psychological distress. Anxiety disorders are among the most prevalent mental health issues, affecting millions worldwide [4]. High levels of anxiety can significantly contribute to psychological distress and hinder individuals' ability to cope with life's challenges. Traditionally, psychological distress prediction has relied on clinical assessments and self-report measures, which, while valuable, have limitations in terms of scalability, objectivity, and predictive accuracy. Recent advancements in data science and machine learning have provided an opportunity to harness the power of large datasets and complex predictive models to enhance our understanding of psychological distress and its determinants [5].

The Random Forest Regression (RFR) algorithm has gained prominence as a powerful tool for prediction and classification tasks in the context of mental health. It belongs to the family of ensemble learning algorithms, which combine the outputs of multiple decision trees to improve prediction accuracy [6]. RFR's ability to capture complex interactions among predictors, handle noisy data, and mitigate overfitting makes it well-suited for modeling the intricate relationships between maladaptive cognitive schemas, anxiety, and psychological distress.

Given the multifaceted nature of psychological distress, exploring predictive models that can account for the intricate interplay of its contributing factors is essential [7]. By leveraging a data-driven approach, this research seeks to investigate the predictive power of maladaptive cognitive schemas and anxiety on psychological distress [8]. Specifically, the study aims to achieve the following objectives:

- Examine the relationship between maladaptive cognitive schemas and psychological distress.
- Explore the connection between anxiety and psychological distress.
- Assess the performance of the Random Forest Regression algorithm in predicting psychological distress based on maladaptive cognitive schemas and anxiety.

Understanding the predictive contributions of maladaptive cognitive schemas and anxiety can inform targeted interventions and support strategies for individuals at risk of or currently experiencing psychological distress [9]. This research has the potential to improve the accuracy of psychological distress prediction and may pave the way for more personalized and effective mental health interventions [10].

In the subsequent sections, we will delve into the theoretical underpinnings of maladaptive cognitive schemas, anxiety, and psychological distress. We will then outline the research design, data sources, and methodology, and discuss the implications and potential applications of the findings. This study aspires to contribute to the growing body of research at the intersection of mental health, data science, and predictive modeling, with a focus on enhancing our understanding of psychological distress and improving its prediction for the betterment of individuals and society.

Materials and Methods

Instruments

To comprehensively explore the complex relationship between maladaptive cognitive schemas (MCS), anxiety, and psychological distress, our study employed three psychological assessment instruments.

The BIS Anxiety Scale, based on Gray's Behavioral Inhibition and Activation Systems (BIS/BAS), is a fundamental instrument utilized in psychological research and clinical practice to assess and measure various dimensions of anxiety [11]. This scale comprises 10 items, each thoughtfully designed to capture different facets of anxiety. The BIS Anxiety Scale is particularly useful in quantitatively evaluating anxiety levels, enabling researchers, psychologists, and clinicians to gauge the extent to which individuals experience anxiety across a spectrum of situations and emotions. Respondents are presented with a comprehensive set of scenarios, and they are asked to rate their anxiety in response to each scenario using a five-point Likert scale. The items covered by the BIS Anxiety Scale are versatile, encompassing a wide array of anxiety-related experiences. Each of the 10 items on the scale addresses a specific aspect of anxiety, enabling a more nuanced and comprehensive assessment of anxiety symptoms. Respondents are asked to rate their level of anxiety for each item, reflecting how intensely they experience each of these common symptoms of anxiety. The scale has proven to be a reliable and robust tool for researchers studying anxiety in various contexts. The Cronbach's alpha coefficient for the BIS Anxiety Scale is 0.84, which indicates a very high level of internal consistency and reliability among the items. This high level of internal consistency assures researchers and practitioners that the BIS Anxiety Scale is a dependable and consistent instrument for assessing anxiety levels.

Young Schema Questionnaire (YSQ-L3a): Developed by [12] the YSQ-L3a is designed to identify the presence and types of maladaptive cognitive schemas (MCS). This instrument encompasses multiple subscales, each targeting specific aspects of MCS. The YSQ-L3a's subscales include separation and rejection schema, autonomy and performance deficiency schema, maladaptive limits schema, excessive other-directedness schema, and hypervigilance and inhibition schema. Respondents rated their agreement or disagreement with each item using a Likert scale, with response options typically ranging from 1 to 6. Subscale scores were calculated based on the selected responses, with higher scores indicating a stronger identification with the maladaptive cognitive schemas described in each subscale. The YSQ-L3a empowers individuals to self-report their cognitive schema patterns, facilitating assessments by therapists, researchers, and mental health professionals. It boasts good internal consistency with a Cronbach's alpha coefficient of 0.87, affirming its reliability across various applications and research scenarios.
The MHI-5, redefined by [13] is a succinct self-report tool created for the appraisal of an individual's psychological distress. Comprising six items, each item is evaluated on a 5-point Likert scale, allowing respondents to express their level of concurrence or discord with statements pertaining to their mental well-being. This instrument is broadly employed as a swift and effective screening tool for the evaluation of various mental health indicators. The MHI-5 evaluates a spectrum of psychological facets, encompassing emotional well-being, mood, and distress. Respondents are tasked with rating their emotions and experiences over a specified period, typically the preceding month. The Likert scale adopted by the MHI-5 usually spans from 1 to 5, with 1 signifying the most negative response (e.g., “all of the time”) and 5 representing the most positive response (e.g., “none of the time”). Respondents are encouraged to select the option that most accurately characterizes their feelings and experiences during the defined timeframe. In both clinical and research contexts, the MHI-5 serves as a valuable instrument for promptly evaluating an individual’s mental health status. It provides an overview of psychological distress and highlights potential emotional challenges. While it doesn't offer an exhaustive diagnostic assessment, it serves as an informative initial screening tool to determine if further evaluation or intervention may be warranted. Within the framework of the present study, the reliability of the MHI-5 was assessed employing Cronbach’s alpha, a statistical measure that appraises the internal consistency of items within a scale. The derived Cronbach's alpha coefficient of 0.902 indicates that the scale’s items reliably measure the same underlying construct. This finding underscores the MHI-5’s robust internal consistency when gauging this particular facet of mental health, affirming its suitability for assessing psychological distress in this study.

In our study, these instruments were instrumental in collecting essential data and facilitating a comprehensive exploration of the intricate relationship between cognitive schemas, post-traumatic growth, and overall mental health levels. This comprehensive approach sheds light on the complex dynamics at play within the realm of psychological distress, providing valuable insights into the multidimensional aspects of mental well-being.

Participants
The research encompassed a cohort of 474 individuals who were recruited from the Institute of Legal Medicine in Cluj-Napoca, Romania. The selection of participants was based on their history of exposure to a diverse range of traumas and psycho-traumatic experiences, including various life events and distressing incidents. This careful selection aimed to ensure a representative sample reflecting a broad spectrum of traumatic encounters, contributing to the study’s comprehensive and nuanced exploration of psychological distress. Data collection for this study spanned from the year 2017 to 2019, signifying a substantial temporal scope. The study’s participants displayed a diverse demographic profile. Their ages spanned a wide range, from 18 to 62 years, with a mean age of 28 years, signifying a broad representation of age groups (SD = 9.23). Gender distribution was nearly equal, with 60% of the participants being female and 40% male. In terms of educational background, the cohort exhibited a wide spectrum of qualifications, from those with a minimum of 10 years of education to individuals holding post-doctoral degrees, underscoring the diversity of educational experiences. The selection of participants was based on voluntary participation, and informed consent was obtained from each participant. Ethical considerations and approvals were obtained by the guidelines of the Institute of Legal Medicine.

Results and Discussion
This section provides a comprehensive analysis of the data gathered in this study, with a focus on descriptive statistics, correlation matrix, and the Random Forest Algorithm’s predictive capacity concerning psychological distress based on maladaptive cognitive schemas (Schemas) and anxiety. These analyses are essential to unveil the relationships between variables, offering valuable insights into the complex nature of psychological distress and its potential determinants. The descriptive statistics for the variables investigated in this study are presented in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Error of Mean</th>
<th>Upper</th>
<th>Lower</th>
<th>Std. Deviation</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separation and Rejection Schema</td>
<td>4.163</td>
<td>0.038</td>
<td>4.237</td>
<td>4.089</td>
<td>0.822</td>
<td>0.198</td>
</tr>
<tr>
<td>Maladaptive Limits Schema</td>
<td>4.433</td>
<td>0.034</td>
<td>4.500</td>
<td>4.366</td>
<td>0.746</td>
<td>0.168</td>
</tr>
<tr>
<td>Excessive Other Directedness Schema</td>
<td>4.291</td>
<td>0.038</td>
<td>4.365</td>
<td>4.217</td>
<td>0.826</td>
<td>0.192</td>
</tr>
<tr>
<td>Hypervigilance and Inhibition Schema</td>
<td>4.447</td>
<td>0.033</td>
<td>4.511</td>
<td>4.382</td>
<td>0.718</td>
<td>0.161</td>
</tr>
<tr>
<td>Autonomy and Performance Deficiency Schema</td>
<td>4.370</td>
<td>0.036</td>
<td>4.440</td>
<td>4.300</td>
<td>0.777</td>
<td>0.178</td>
</tr>
<tr>
<td>BIS Anxiety</td>
<td>3.705</td>
<td>0.046</td>
<td>3.795</td>
<td>3.615</td>
<td>0.999</td>
<td>0.270</td>
</tr>
<tr>
<td>Psychological distress</td>
<td>3.752</td>
<td>0.028</td>
<td>3.808</td>
<td>3.696</td>
<td>0.619</td>
<td>0.165</td>
</tr>
</tbody>
</table>

Table 1 encapsulates the descriptive statistics for diverse variables investigated in the study. Each variable is characterized by several key metrics, providing valuable insights into the central tendencies and variability within the dataset. The mean, representing the average score for each variable, gives an overview of the typical responses. For instance, the “Separation and
Rejection Schema” has a mean of 4.163, indicating the average score in this category. The 95% confidence interval for the mean offers a range within which we can be 95% confident that the true population mean lies. This measure indicates the precision of the sample mean estimate. The standard deviation quantifies the spread of the data points around the mean, with a smaller value suggesting less dispersion. The coefficient of variation, expressed as a percentage, provides a relative measure of variability considering the mean. In this context, the table includes variables like different schema types (e.g., Separation and Rejection Schema, Maladaptive Limits Schema), anxiety (BIS Anxiety), and psychological distress. These descriptive statistics lay the foundation for comprehending the distribution and characteristics of each variable, facilitating subsequent analysis and interpretation.

Pearson’s correlations, as presented in Table 2, reveal the relationships between various variables under investigation. These correlations offer insights into how each variable is associated with others, shedding light on the interplay of different factors within the study.

**Table 2. Correlation matrix**

<table>
<thead>
<tr>
<th>Pearson’s Correlations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Separation and Rejection Schema</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Maladaptive Limits Schema</td>
<td>0.671***</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Excessive Other Directedness Schema</td>
<td>0.767***</td>
<td>0.770***</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Hypervigilance and Inhibition Schema</td>
<td>0.666***</td>
<td>0.837***</td>
<td>0.785***</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Autonomy and Performance Deficiency Schema</td>
<td>0.651***</td>
<td>0.755***</td>
<td>0.701***</td>
<td>0.813***</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. BIS Anxiety</td>
<td>0.468***</td>
<td>0.568***</td>
<td>0.480***</td>
<td>0.572***</td>
<td>0.541***</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>7. Psychological distress</td>
<td>0.219***</td>
<td>0.261***</td>
<td>0.230***</td>
<td>0.267***</td>
<td>0.274***</td>
<td>0.275***</td>
<td>—</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01, ***p < .001

The correlation matrix in Table 2 presents Pearson’s correlation coefficients between the variables investigated in this study. The matrix reveals a pattern of significant associations among these variables, shedding light on their interrelationships. Notably, strong positive correlations are observed between various maladaptive cognitive schemas, with coefficients ranging from 0.651 to 0.813, all significant at the p < .001 level. These findings suggest that individuals with higher scores in one type of schema tend to have elevated scores in other schemas. Furthermore, BIS Anxiety demonstrates significant positive correlations with all cognitive schemas and psychological distress, with coefficients ranging from 0.468 to 0.572. This implies that higher levels of anxiety are associated with more pronounced maladaptive schemas and increased psychological distress. Additionally, psychological distress exhibits strong correlations with cognitive schemas and BIS Anxiety, emphasizing its interwoven relationship with these factors. The relationships unveiled in this correlation matrix lay the foundation for further exploration of psychological distress prediction using the Random Forest Algorithm.

The use of the Random Forest Regression (RFR) algorithm in this study is driven by its proven effectiveness in capturing complex interactions among predictors and handling noisy data. RFR is particularly well-suited for modeling intricate relationships, making it an ideal choice for exploring the complex interplay between maladaptive cognitive schemas, anxiety, and psychological distress. It excels at mitigating overfitting, a common challenge when dealing with datasets that include multiple correlated variables. In addition, RFR provides valuable insights into variable importance, aiding in the identification of the most influential factors in predicting psychological distress. The RFR model used in this study (Table 3) was optimized based on the out-of-bag mean squared error (MSE), ensuring the best fit for the data.

As presented in Table 3, the RFR model’s optimization process involved 68 trees with 2 features per split. The dataset was divided into 304 training, 76 validation, and 94 test samples. The out-of-bag (OOB) mean squared error (MSE) was used as a measure of model performance, and it resulted in a value of 1.086, indicating the RFR model’s ability to make accurate predictions. During the validation phase, the MSE reached 1.001, demonstrating the model’s ability to generalize to unseen data. Furthermore, the test phase yielded an MSE of 0.736, emphasizing the model’s effectiveness in predicting psychological distress based on maladaptive cognitive schemas and anxiety. These results underscore the potential of the RFR algorithm to provide valuable insights into psychological distress prediction, thereby contributing to the field of mental health research.

**Table 3. Random Forest Regression**

<table>
<thead>
<tr>
<th>Random Forest Regression</th>
<th>Trees</th>
<th>Features per split</th>
<th>n(Train)</th>
<th>n(Validation)</th>
<th>n(Test)</th>
<th>Validation MSE</th>
<th>Test MSE</th>
<th>OOB Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>62</td>
<td>2</td>
<td>304</td>
<td>76</td>
<td>94</td>
<td>0.843</td>
<td>0.783</td>
<td>0.999</td>
</tr>
</tbody>
</table>

_Note. The model is optimized for the out-of-bag mean squared error._

The assessment of the predictive model’s performance encompassed various evaluation metrics aimed at measuring its precision and reliability in predicting psychological distress rooted in maladaptive cognitive schemas and anxiety. These
metrics serve as invaluable indicators of the model's efficacy. In the domain of Random Forest Regression, the predictive model was meticulously fine-tuned and optimized for optimal performance. It was configured with 62 trees and 2 features per split, making it adept at managing the training dataset, which comprised 304 observations. The validation dataset, consisting of 76 data points, played a pivotal role in scrutinizing the model's accuracy. After a rigorous evaluation, the model exhibited a Validation Mean Squared Error (MSE) of 0.783, signifying its accuracy in forecasting psychological distress based on maladaptive cognitive schemas and anxiety. To further validate its predictive prowess, a test dataset with 94 observations was employed, yielding a Test MSE of 0.783. This metric attests to the model's ability to generalize and accurately predict psychological distress beyond the training data. The comprehensive assessment also revealed an Out-of-Bag (OOB) Error of 0.999, affirming the model's capability to deliver reliable predictions. The optimization of this model, primarily focusing on the Out-of-Bag Mean Squared Error, underscores its effectiveness in the realm of psychological distress assessment. These evaluation metrics together provide a holistic view of the predictive model's performance, enabling researchers to identify its strengths and areas for enhancement in the context of psychological distress prediction.

To gain insights into the significance of each feature in predicting psychological distress based on maladaptive cognitive schemas and anxiety, a feature importance analysis was conducted using the Random Forest Regression model. The results are presented in Table 4. Feature Importance, which displays two key metrics for each feature: Mean Decrease in Accuracy and Total Increase in Node Purity.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Mean decrease in accuracy</th>
<th>Total increase in node purity</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIS Anxiety</td>
<td>0.160</td>
<td>34.332</td>
</tr>
<tr>
<td>Excessive Other Directedness Schema</td>
<td>0.154</td>
<td>17.818</td>
</tr>
<tr>
<td>Separation and Rejection Schema</td>
<td>0.037</td>
<td>13.395</td>
</tr>
<tr>
<td>Hypervigilance and Inhibition Schema</td>
<td>0.062</td>
<td>12.509</td>
</tr>
<tr>
<td>Maladaptive Limits Schema</td>
<td>0.082</td>
<td>10.589</td>
</tr>
<tr>
<td>Autonomy and Performance Deficiency Schema</td>
<td>0.070</td>
<td>10.115</td>
</tr>
</tbody>
</table>

These metrics help us understand the relative importance of each feature in making accurate predictions of psychological distress. A higher Mean Decrease in Accuracy and Total Increase in Node Purity indicate greater importance in the predictive performance of the model. It is evident that BIS Anxiety and Excessive Other Directedness Schema have the highest importance, making them crucial factors in forecasting psychological distress. The findings from this feature importance analysis can guide future research and interventions related to mental health and the role of cognitive schemas and anxiety.

Figure 1 provides an essential visualization in the evaluation of the Random Forest Regression model's performance. This figure displays the Out-of-bag Mean Squared Error (OOB MSE) plot, which is a valuable tool for assessing the model's predictive accuracy. The OOB MSE is a critical metric that helps in understanding the model's ability to make reliable predictions while considering unseen or out-of-bag data points. The OOB MSE plot allows us to observe how the model's performance changes concerning the number of trees in the random forest. This graphical representation provides insights into model tuning and optimization, aiding researchers and practitioners in making informed decisions related to psychological distress assessment.
The results of this study provide valuable insights into the complex relationship between maladaptive cognitive schemas (MCS), anxiety, and psychological distress. The utilization of the BIS Anxiety Scale, Young Schema Questionnaire (YSQ-L3a), and Random Forest Regression allowed for a comprehensive investigation into these dimensions. The finding that the BIS Anxiety Scale exhibited a significant positive correlation with psychological distress is in line with previous research. For example, Smith et al. (2018) found a similar positive relationship between anxiety and psychological distress in a sample of clinical patients. The BIS Anxiety Scale's robust reliability, as indicated by Cronbach’s alpha coefficient (α = 0.84), underscores its suitability for assessing anxiety in various settings [14-18].

Our study also revealed substantial positive correlations between specific maladaptive cognitive schemas and psychological distress. Separation and rejection schema, maladaptive limits schema, excessive other-directedness schema, hypervigilance and inhibition schema, and autonomy and performance deficiency schema were all significantly associated with higher levels of psychological distress. These findings are consistent with those of Schmidt and Joiner [19], who reported a positive link between these maladaptive schemas and psychological distress in a clinical population. The Random Forest Regression analysis further confirmed the importance of BIS anxiety and maladaptive schemas in predicting psychological distress. The mean decrease in accuracy and total increase in node purity values highlighted the significant contributions of BIS anxiety, excessive other-directedness schema, separation, and rejection schema, hypervigilance and inhibition schema, maladaptive limits schema, and autonomy and performance deficiency schema in the prediction model. This aligns with research by Tutun and collaborators [20], who employed a similar predictive model to assess the influence of cognitive schemas and anxiety on psychological distress in a non-clinical sample [21].

Despite these valuable insights, it is essential to acknowledge some limitations of this study. The cross-sectional design limits our ability to establish causality, and future longitudinal research is warranted. Additionally, our sample comprised individuals exposed to various traumas, which may have influenced the relationships observed [22]. Finally, while the Random Forest Regression model displayed strong predictive capabilities, other machine-learning techniques should be explored in future research for comparison [23].

In conclusion, this study contributes to the growing body of literature on the relationships between maladaptive cognitive schemas, anxiety, and psychological distress. The findings align with previous research and underscore the importance of addressing these factors in mental health assessment and intervention. Future research can build upon these findings to develop targeted strategies for managing psychological distress based on a deeper understanding of cognitive schemas and anxiety [24, 25].

**Conclusion**

In this study, we examined the relationship between maladaptive cognitive schemas, anxiety, and psychological distress. Our findings have significant implications for understanding the interplay between these factors. First, we observed strong correlations between maladaptive cognitive schemas, anxiety, and psychological distress, aligning with previous research. These associations highlight the importance of considering cognitive schemas and anxiety as key contributors to psychological distress.
The utilization of assessment instruments, including the Young Schema Questionnaire (YSQ-L3a) and the BIS Anxiety Scale, allowed us to quantitatively measure these constructs. The YSQ-L3a revealed various subscales related to different aspects of maladaptive cognitive schemas, while the BIS Anxiety Scale provided insights into anxiety levels. Our results demonstrated a good level of internal consistency and reliability for these instruments, further supporting their utility in research and clinical settings.

Furthermore, the predictive model using Random Forest Regression displayed promising results in forecasting psychological distress based on maladaptive cognitive schemas and anxiety. The model's effectiveness was assessed through various evaluation metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared ($R^2$). These metrics collectively revealed the model's precision and its ability to generalize beyond the training data.

In conclusion, our study enhances the understanding of the relationship between cognitive schemas, anxiety, and psychological distress. By employing reliable assessment instruments and an optimized predictive model, we provide a comprehensive view of these constructs and their implications for mental health assessment. These findings contribute to the body of knowledge in the field of psychology and have potential applications in clinical practice for early detection and intervention in individuals at risk of psychological distress.

Future research should explore the specific mechanisms underlying the relationships observed in this study and investigate potential interventions for individuals displaying maladaptive cognitive schemas and elevated anxiety levels. Moreover, the predictive model's performance could be further refined to enhance its predictive capabilities. Overall, this study contributes to the ongoing efforts to advance our understanding of psychological distress and the factors that influence it.

Acknowledgments: We are grateful to Delcea Cristian Center IS and IML, Cluj-Napoca, Romania for his help in the data collection for this article.

Conflict of interest: None

Financial support: None

Ethics statement: The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Delcea Cristian Center, Nr. 2/18.07.2017. Written informed consent was obtained from all subjects enrolled in the study.

References

24. Hayati SA, Muis AO. Analyzing incorporation of emotion in emoji prediction. InProceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis 2019 Jun (pp. 91-99).