

Academic Adjustment: State Versus Trait

1. Research Question / Problem Definition

Given measurements of change in an individual's reported satisfaction with their academic study over time, is it possible to identify patterns which relate to broader personal wellbeing? This question was investigated through the use of multiple predictive approaches, ranging from initial clustering analysis to determine validity of categories, and further complex regression analysis to determine predictive ability of specific psychological measures on students academic performance, wellbeing, and satisfaction.

Specifically, clustering analysis was used to track $k = 4$ clusters after analysis determined significant variance between 4 unique categories of students based on changes in students self reported satisfaction with their studies over the 14 day period. This measured change in satisfaction resulted in categories delineated by Stable High Performing, Stable Low Performing, Declining Performance, and Improving Performance metrics. The clusters were then individually analyzed with linear regressions targeting specific and hypothetically predictive psychometric self report data to verify category validity and predictive strength. Results determined a significant difference in predictive ability between stable states (Stable Low, Stable High), and changing states (Declining, Improving).

After initial analysis determined a significant difference between the predictive ability of patterns of change vs stability, supervised approaches were employed to determine what psychometric factors may play a role in determining how a student is divided between reported satisfaction and wellbeing categories. Regression analysis included 3 linear models with

regularization (Ridge, Lasso, Elastic Net), followed by a gradient boosting tree based model (XGBoost). Results provided insight into 2 specific categories of variable, related to whether the variable captures a State (immediate momentary feeling) or a Trait (Stable personality characteristics). Furthermore, the State category was divided into two types of momentary state, of either Affective, momentary emotional experiences, or Meta-Cognitive, momentary perceptions of study. Models ultimately provided genuine insight into causes and characteristics which drive differences in academic satisfaction and related metrics between the categories of student performance, and further outlined possible reasons for variance in prediction ability between State and Trait variables.

This analysis truthfully produced actual insight into emotional regulation trends and patterns between categories of student performance, such that there may exist many practical real world applications of actual interest. The ability to predict academic success trajectory at a confident level raises options for early detection of possible academic stress. Furthermore, models showed accuracy in predicting psychological measure variables which also provide context for group clustering, and served to mutually verify validity of approaches, thereby providing a framework for predicting academic success without the need for a sufficient sample for clustering analysis.

2. Data and Features

The dataset that was chosen for this analysis was taken from a research publication concerning student academic adjustment at an undergraduate level. Thus, the data tracks each of 321 individual students, measured across 14 days in a single university semester. Data collection

was performed using ESM with multiple daily conducted surveys covering 145 variables of measure.

Features were constructed for clustering analysis by dividing the 14 days for each individual into two categories of the first 6 days, and the last 8, and then computing a linear regression on the reported satisfaction levels over the entries within both categories respectively. The slopes of these regressions were then aggregated and tracked by individuals, such that the overall trend of individuals reported satisfaction with studies across the 14 day period provides the mechanism for categorizing trajectories in the 4 clusters. Furthermore, for analysis purposes, individuals were categorized as “At Risk,” (Stable Low + Declining) or “Thriving,” (Stable High + Improving), such that a confusion matrix could be employed to track prediction accuracy. Next, regression models were trained on the Thriving/At Risk binary based on the original clusters for prediction of a target variable, to allow for extra comparison to supervised model approaches later on and extra verification of cluster validity.

Features created for supervised approaches utilized PCA for dimension reduction to reduce the original 22 features to 18, while retaining 95% of the variance. Further, K-Means clustering analysis was conducted on components which yielded $k = 3$ centers for clustering, which represented three distinct student profiles for supervised regression analysis. Furthermore, a custom engineered feature was created to combine the original features before PCA, the PCA features, and hierarchical cluster membership metrics.

Data cleaning for both processes was handled similarly, due to the dataset remaining the same and thus requiring the same pre-processing steps. This consisted of initial cleaning by removing rows that do not include the relevant satisfaction data, and further groups these by individual, as each row corresponds to one day of measurements, and each individual has 14

rows respectively. Due to the dataset having a large amount of random NA entries, with at least 1 NA value per row, the need to piecewise address missing data within the pipeline prevented total exclusion of all data; a problem I ran into when attempting to pre-process all NA values prior to data analysis. The solution that was employed was to remove any NA values at any stage within the processing pipeline that depends on an existing value, rather than attempt to remove all NA at the start of the process, allowing for measurements on variable combinations otherwise uncaptured by the cleaned dataset.

The supervised regression methods similarly piecewise excluded NA entries, however, this model created a matrix for individuals related to their positive/negative affects, and exhaustion. Furthermore, rough data imputation was attempted for NA values with overlapping median values after the feature matrix was created.

3. Cluster Models

For clustering, the K-Means method was used with $k = 4$ centers, with 4 determined as the optimal number through elbow and silhouette plot analysis. Next, data was combined across the Affect variables, and the mean and variance of the respective target variables was standardized to the timescale of measurement. Likewise, this process was completed for the first half predictors, as well as the second half predictors. After aggregating the students into the 4 respective clusters, visuals were created to show both individual satisfaction trajectories within and between the clusters (Figure 1), and the average trajectories by cluster (Figure 2). The Stable High Performing cluster maintained 67 of the total 193 students retained after data cleaning. The Declining and Improving Performance clusters retained 30 students total each, and the Stable Low Performing category captured 66 of the total students. Next, confusion matrix results

showed confidence in accurate predictive abilities for categorizing students as At Risk or Thriving at a rate of 68% (Figure 3). Furthermore, regression analysis served to test predictive ability of clusters on psychological metrics, and showed inversely related results for predicting levels of reported happiness and exhaustion between clusters which are contextually consistent with what would be expected from the actual psychology of the respective clusters.

Furthermore, Cohen's d effect sizes for the respective clusters bring forth the first clear evidence of the pattern difference between State and Trait measures, showing large effect sizes for Stable High and Stable Low clusters, and significantly reduced effect sizes for the improving and declining clusters.

4. Regression Models

For regularized model approaches, Ridge, Lasso, and Elastic Net methods were used. These models attempted to predict the 8 previously introduced variables across the two main categories of State and Trait variables, of Positive Affect (happy composite), Negative Affect (sad composite), and Exhaustion (self report level) comprising the State Affective group. Further, Study Demands (perceived study workload) and Study Satisfaction (satisfaction with studying) variables comprised the State Meta Cognitive Group, and finally Cognitive Reappraisal (adaptive emotion regulation), Expressive Suppression (maladaptive emotion regulation), and Rumination (repeat negative thinking) made up the Trait group. The tree based XGBoost model was designed to only target the State Affect variables for specific interest in capturing non-linear context of this category. All model and target variable results are displayed simply within Figure 5.

Due to the modular nature of the feature engineering and classification, the multiple categories and model approaches could be interchanged and measured to determine best

intersections of predictive ability and model accuracy. This resulted in 24 total tested models, across 3 linear methods of 2 feature sets, and 1 tree method of 2 feature sets, and compared against 3 State Affect targets. The Ridge, Lasso, and Elastic Net methods utilized a nested cross validation structure with a 10 fold outer split across the full data, with each fold used as a test set once, and then a 5 fold inner split specifically concerning the training split of the data to prevent data leakage, and hyperparameter tuning occurred within the 5 inner training folds.

5. Results

Clustering

The clustering analysis approach yielded an intriguing result concerning the Stable vs Changing behavior patterns and the respective influence these groups of variables have on prediction accuracy. Results show strong prediction capability for Stable Low and Stable High clusters for predicting self reported levels of both happiness, and exhaustion, and produces inverse results for these metrics. Furthermore, the Improving and Declining clusters showed low predictive ability (Figure 4). Ultimately, it is hypothesized that this relationship shows that an individual with a stable trajectory provides predictive evidence towards their underlying characteristics through their self-report metrics over time, while changing trajectories signal variance in self report data and thus result in lower predictive abilities.

Finally, the regression model trained to predict the Thriving vs At Risk students based on the clustering analysis provided 68% accuracy, with no significant bias towards false positive/negative, and further maintained .37 kappa denoting greater than chance prediction. With a resulting p-value of 0.005, this regression model is able to genuinely identify to a

relatively confident degree whether a student will maintain positive versus negative academic outlooks over time (Figure 3).

Supervised Regression

Since my design attempted to cast a wide net to analyze as many models and combinations as possible in an attempt to determine the most optimal combinations of analysis, there were many results to parse through. Ultimately, the results of all of the linear methods tested showed high predictive ability for State Affect variables, at $R^2 = 0.6$ for positive and negative affect, and $R^2 = 0.62$ and 0.61 for Study Demands and Study Satisfaction respectively. (Figure 6). Exhaustion variance was captured at $R^2 = 0.49$ by the XGBoost, representing the most valuable improvement over iterations of any of the methods, which is hypothesized to be due to this variable capturing a large amount of non-linear interactions, and thus allowing the tree method to perform stronger than the linear methods. Ultimately, the results show that the Stable State variables did provide the most prediction accuracy above that of the Trait variables (Figure 6), with the meta-cognitive variables providing the most robust predictive capability, hypothesized due to the context proximity to the predictor variables. Furthermore, it was noted that emotional regulation strategies provided limited predictive value, possibly supporting the understanding that stable personality characteristics are difficult to change, even with emotional regulation strategies, and thus result in greater predictive ability due to resistance towards intentional emotion regulation.

6. Interpretation / Takeaways

After conducting supervised and unsupervised analysis of the data by means of a broad range of modelling methodology, a general resulting trend was quickly identified concerning variation in predictive ability as a result of relatively stable vs unstable individual behavior patterns. Clustering categories related to stable emotional characteristics, even those that are stable but low performing, provide a framework with greater relevant predictive information than that of the data from Improving or Declining students. However, ultimately, the most at risk group of the 4 clusters was that of the Stable Low Performers, who showcased both 0.68 SD lower happiness on average, as well as 0.54 SD higher reported exhaustion averages. As this category represents 34% of the data population, this group is a particularly good starting point for any applicable intervention strategies. Conversely, the Stable High Performing students showcased both 0.65 SD greater happiness and 0.55 SD lower exhaustion, denoting a particularly emotionally strong cluster.

Another important result from this analysis concerned the State Affect variables, and further the Meta Cognitive subcategory for State variables which showcased by far the most accurate and compelling predictive abilities. This result further supports the understanding that the greatest predictive accuracy is derived from analysing self report metrics of perceived emotions over time, and has limited influence from intentional emotional regulation strategies within the Trait variables category, hypothesized to be due to underlying personality characteristics that are resilient to intentional momentary effects when viewed in aggregate over time.

While the results captured within this analysis provide meaningful interpretations of the students' emotional states and long term regulation abilities, it must be addressed that the sample size was limited to a few hundred students at a single university and therefore the results may

lack broad generalizability beyond the study. Furthermore, the lack of objective metrics like GPA which could show more direct causal links to the data is also limiting. These limitations can be addressed effectively with further steps in this research, such as introducing a long timeframe of study, including measurements that track not just self-reported perceptions of emotion and satisfaction, but also objective measures like GPA and enrollment status over time. Furthermore, conducting the study across more universities and settings would extend generalizability of the results.

References

Neubauer, A. B., & Schmiedek, F. (2024). Approaching academic adjustment on multiple time scales. *Zeitschrift für Erziehungswissenschaft*, 27, 147–168.

<https://doi.org/10.1007/s11618-023-01182-8>

Anthropic. (2024). Claude 3.5 Sonnet (Version December 2024) [Large language model].

<https://claude.ai>

Figures

Figure 1:



Figure 2:

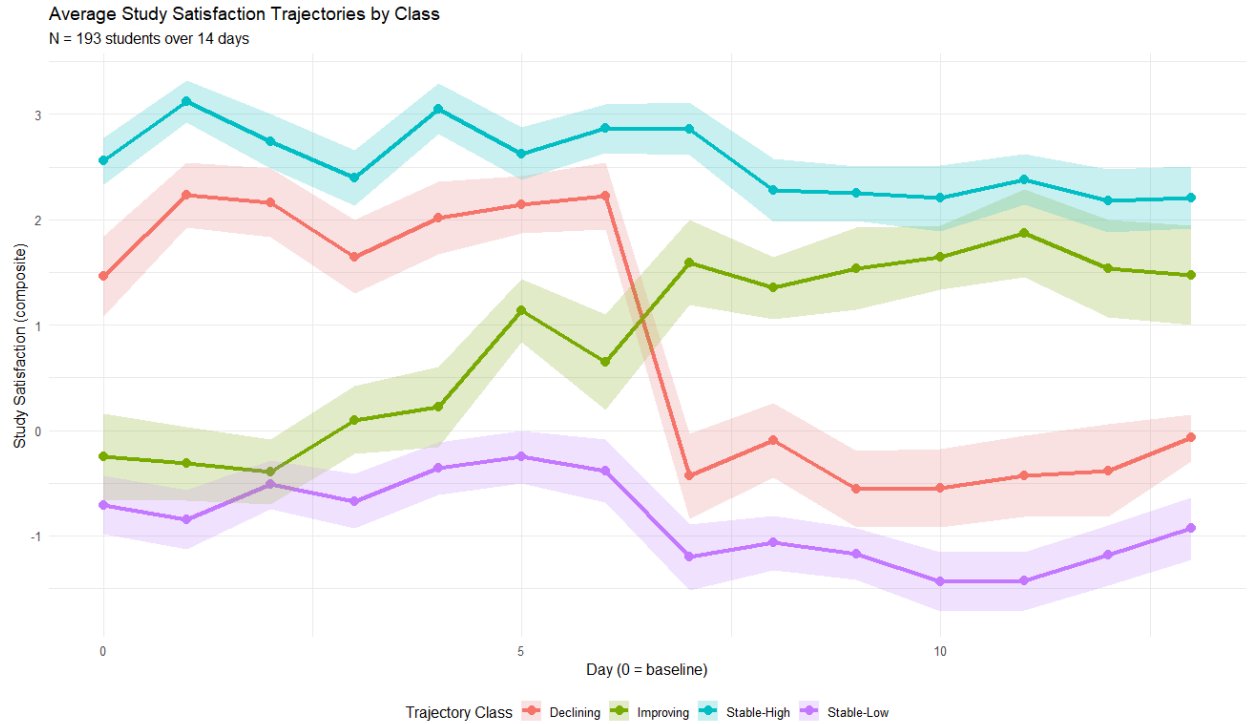


Figure 3:

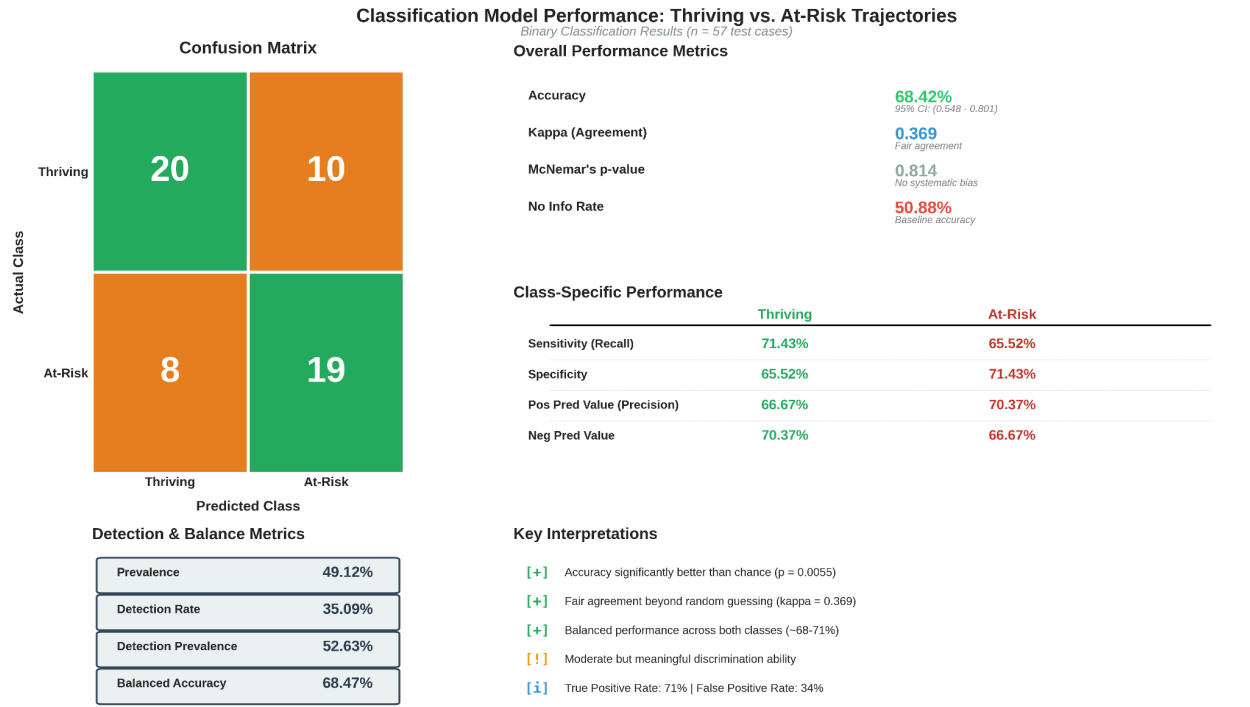
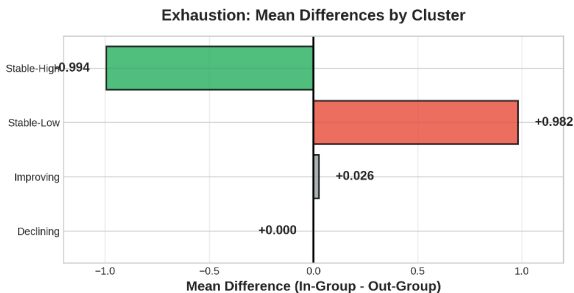
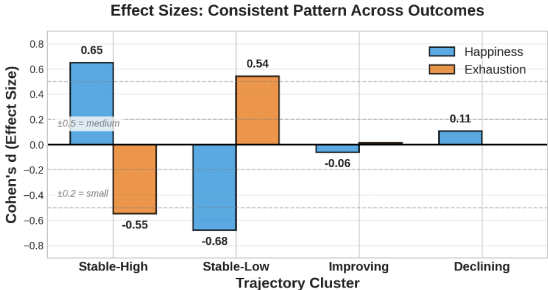
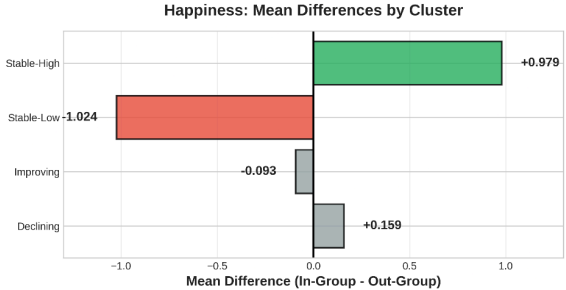


Figure 4:

Satisfaction Trajectories Predict Wellbeing Outcomes:
 Stable Patterns Matter, Changing Patterns Don't



Key Insights

Pattern Consistency
 Stable patterns are predictive, changing patterns are not

Stable-High

- Exhaustion: d = -0.55 (less exhausted)
- Happiness: d = +0.65 (more happy)

Stable-Low

- Exhaustion: d = +0.54 (more exhausted)
- Happiness: d = -0.68 (less happy)

Improving/Declining

- No effect on either outcome
- d = 0, R² = 0

Implication
 Trajectory stability matters more than direction

Figure 5:

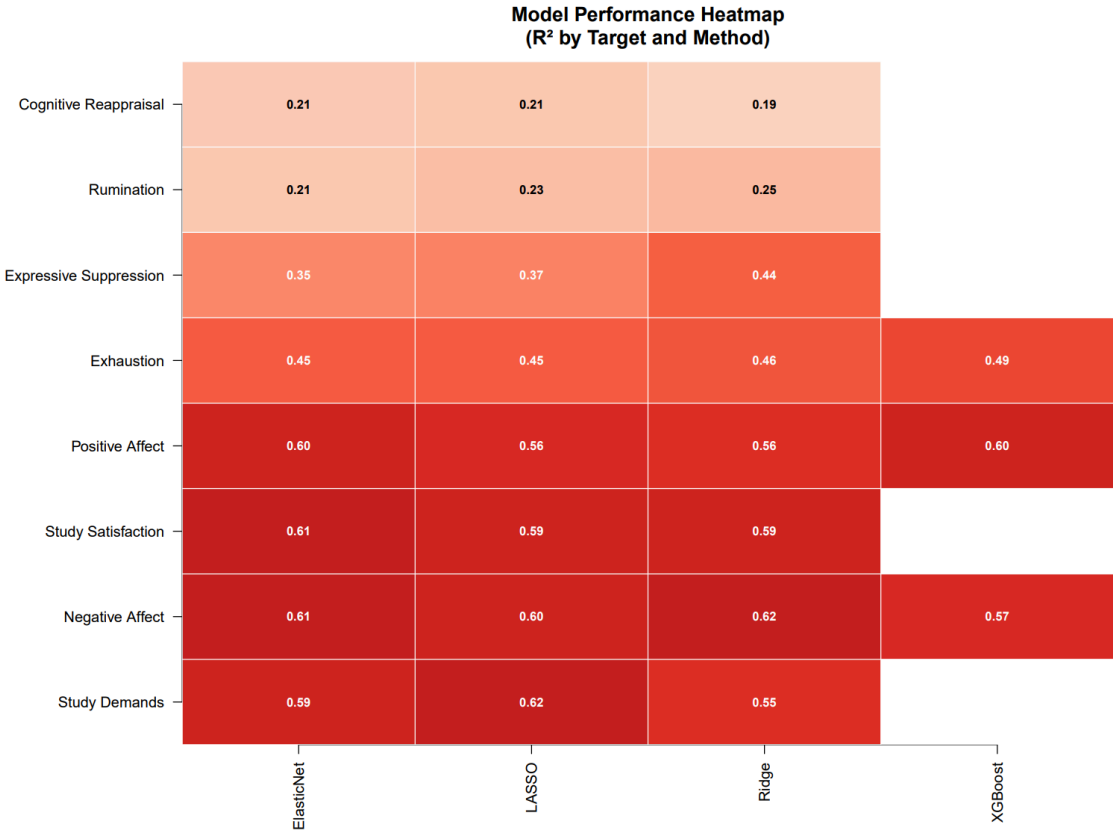


Figure 6:

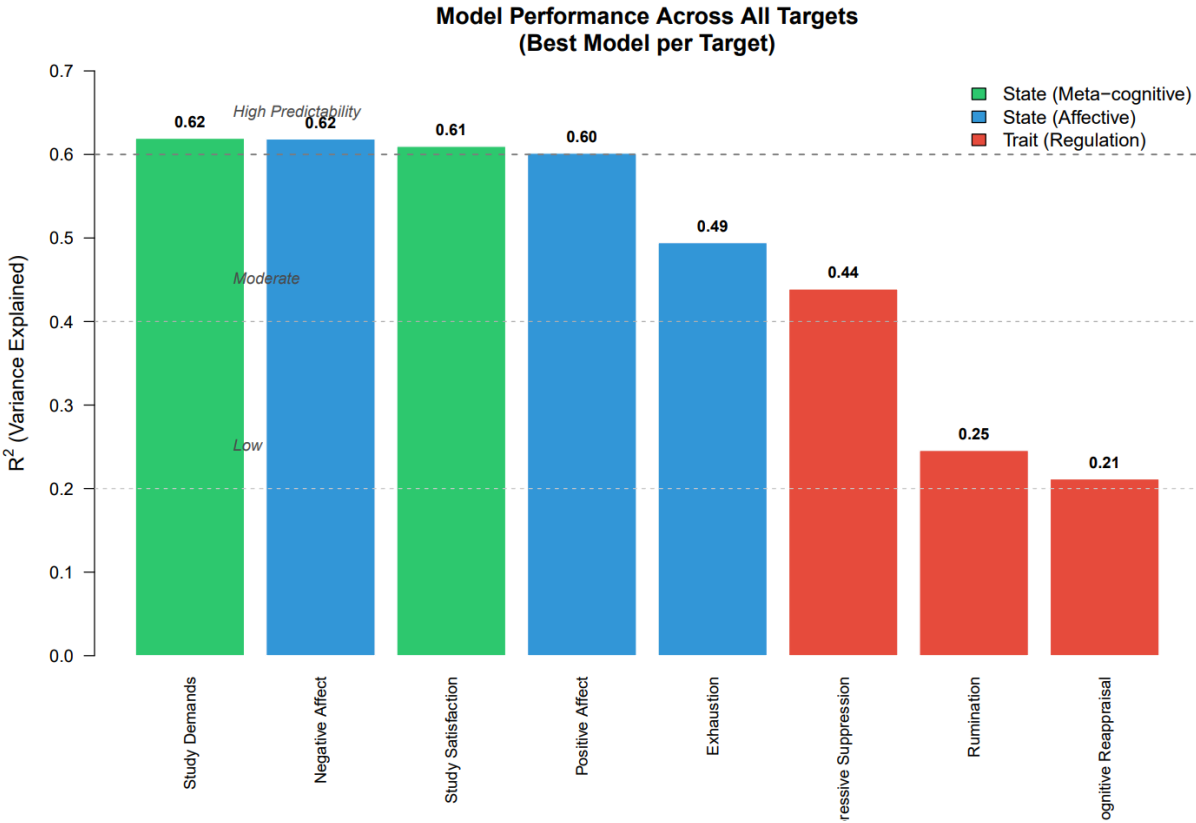


Figure 7:

