Predicting depression history from a short reward/aversion task with behavioral economic features

L. Stefanopoulos[1,5], S. Lavlani[1], BW Kim[2], N. Vike[2], S. Bari[2], E. Azcona[1], S. Woodward[3], M. Block[4], N. Maglaveras[5], A. K. Katsaggelos[1] and H. Breiter[2,6]

[1] Northwestern University/Department of Electrical and Computer Engineering, Evanston, IL, USA
[2] University of Cincinnati /Departments of Computer Science & Biomedical Engineering, Cincinnati, OH, USA
[3] Northwestern University / Feinberg School of Medicine, Chicago, IL, USA
[4] Northwestern University/Integrated Marketing Communications, Evanston, IL, USA
[5]Aristotle University of Thessaloniki/Department of Medicine, Thessaloniki, Greece
[6]Massachusetts General Hospital/Department of Psychiatry, Boston, MA, USA

Abstract—**This paper presents a novel example of depression prediction, merging cognitive science with data-driven machine learning. Behavioral economic features were engineered from a short picture rating task. Relative Preference Theory was applied to rating data for quantifying the degree to which participants liked, disliked, or were neutral to several types of pictures; thus, behavioral economic variables including loss aversion, risk aversion, and 13 others that are amenable to psychological interpretation were mined. These variables were features of a logistic regression predictive model that targeted depression in a population-based sample (N=281) with high test accuracy and no overfitting. Per our review of the literature, we cannot identify other papers that explicitly use behavioral economic features to predict depression with machine learning.**

Keywords— **Depression prediction, Behavioral Economic Features, Machine Learning**

I. INTRODUCTION

The use of simple behavioral features for the prediction of mental health conditions, such as depression, may provide faster diagnosis and patient monitoring. Currently, no framework exists to assess the probability of depression with a short behavioral task, without use of clinical symptoms. Tools for the detection and characterization of depression are needed for individual diagnostic tests, continuous depression assessment/monitoring, risk stratification, personalized therapeutic intervention decisions. Psychiatry and behavioral science have been traditionally dependent on constellations of clinical symptoms for diagnosing depression [1].

Our method utilizes a task where participants rate a set of affective pictures in conjunction with Relative Preference Theory (RPT) analysis [2,3,4] to mine information regarding behavior. In particular, the resulting RPT features provide information about one's judgment, reflecting behavioral economic variables, which are then used by machine learning (ML) to predict the historical occurrence of self-reported depression.

The paper is structured as follows: section "Materials & Methods" is dedicated to data acquisition, feature engineering, and the methodology applied for this research. Sections "Results" and "Conclusions" detail the outcomes and conclusions, respectively.

II. MATERIALS AND METHODS

A. Data

The data used for this study were drawn from the Emotion and Behavior Study, an online study of American adult consumers, conducted by Research Results, Inc. (Boston, MA). Specifically, the rating experiment and brief mental health survey were included in the subsequently described analyses. A total of 501 participants gave consent for their anonymized response data to be used and released to Northwestern University. Participants were asked to self-report how often they had experienced depression on a Likert scale from 1 to 7 (distributions are shown in Fig.1). Participants were also asked to complete a rating task which used 48 pictures from the International Affective Picture Set (IAPS) [5]. Participants rated pictures belonging to one of six categories (food, aggressive animals, cute/nice animals, disasters, sports and nature scenes of mountains or beaches) on a 7-point Likert-like scale from -3 to +3. Each picture category consisted of eight images with similar valence and intensity scoring per the IAPS database. Out of the total 501 participants, the picture rating task was successfully completed by 281 participants. For these participants, all 15 RPT indices were calculated as described in [6], and five demographic variables were collected (Age, Sex, Education level, Education years, and Dominant hand). These 20 variables constituted the final dataset as listed in Table 1. Of the 281 participants, 168 identified as female (59.79%) and 113 identified as male (40.21%). The distributions of all levels for all the demographic variables are represented graphically in Fig. 2

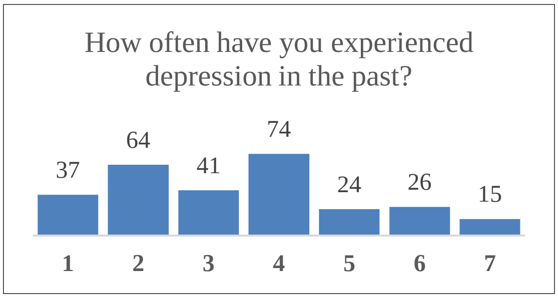

Fig. 1 Distributions by depression scale

Table 1 Features used in the analysis

Name
Risk aversion
Loss resilience
Loss aversion
Ante
Insurance
Peak positive risk
Peak negative risk
Reward tipping point
Aversion tipping point
Total reward risk
Total aversion risk
Reward-aversion tradeoff
Tradeoff range
Reward-aversion consistency
Consistency range
Age
Sex
Education level
Education years
Dominant Hand

B. Software

Feature engineering and extraction of the RPT features were performed using MATLAB.

For the ML analysis, the Python programming language via the Spyder integrated development environment was used. The libraries used for data preprocessing were Pandas [7] and NumPy [8]; in combination with the statsmodel library [9], the exploratory data analysis was performed. The Scikit-learn library [10] was used to perform predictive machine learning and validation.

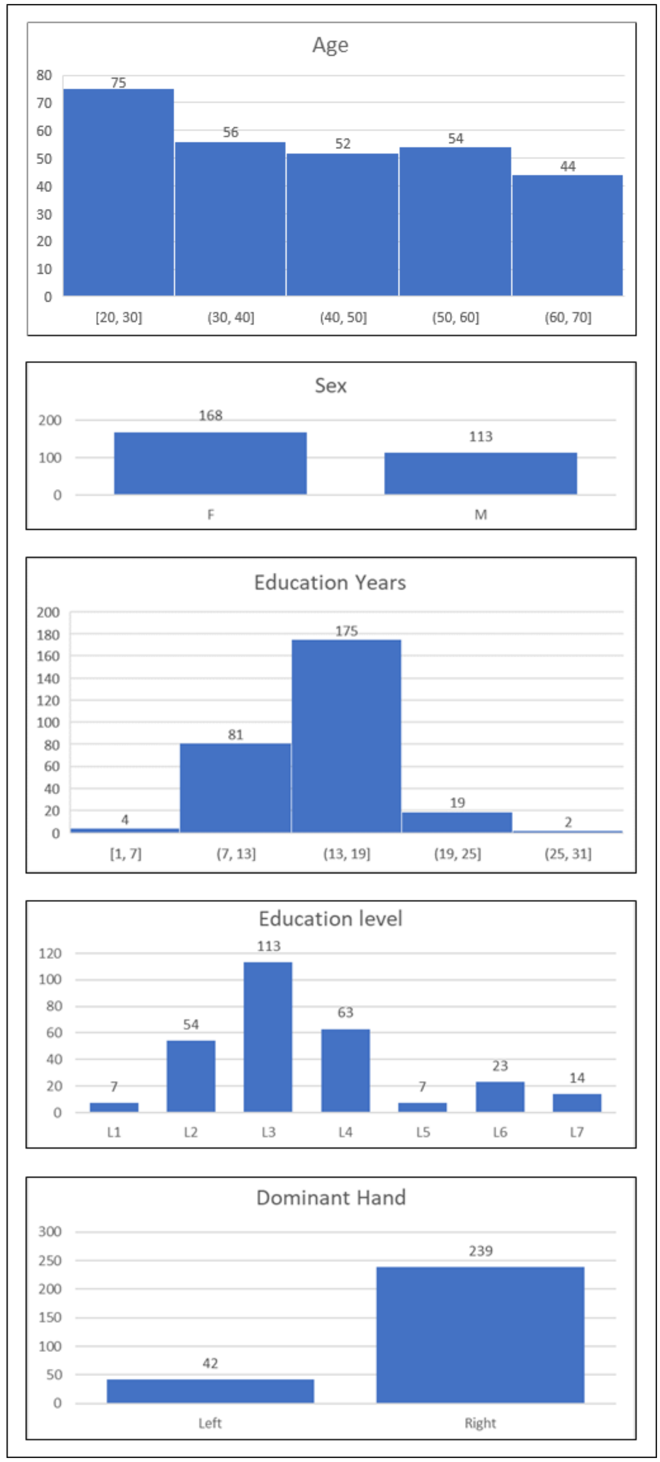

Fig. 2 Distributions of demographics

C. Feature Selection

The curse of dimensionality states that, as the number of features increases in a feature space, the volume of the state space increases exponentially, creating rapidly escalating analytic difficulties. One difficulty is that exploration of the feature space during model training becomes exponentially more difficult as features are added [11]. Therefore, feature selection is fundamental to identify a sparse set of RPT features that can be utilized to detect a history of depression. With this in mind, we aimed to understand the priority of the features for our classifier. For the ML modeling, we employed the Logistic Regression (LR) classifier and features were selected based on a backwards, stepwise approach that identified the best subset of the predictor variables. The best feature subset was the result of multiple iterations of evaluating the Variance Inflation Factor (VIF) [12], correlation, and significance of each feature. VIF measures the amount of multi-collinearity among features. A feature with a high VIF value (typically greater than 2.5 [13]) indicates that the associated independent variable is highly collinear with the other variables in the model. The correlation between features is explored with the calculation of the Pearson Correlation Coefficient, which evaluates the linear relationship between two continuous variables, and with the Spearman Correlation Coefficient, which evaluates the monotonic relationship between two continuous variables [14]. The significance (*p*-value) for each feature is used for determining the most significant feature subset and for deciding the "winner" between the correlated features. In each iteration, the cut-off thresholds for VIF, correlation, and significance were tightened, so that the final subset of features was strongly independent and highly significant.

III. Results

We trained multiple LR models to explore the predictive capabilities of our approach. All models presented below have been validated using 1000 times, 100 repeated, 10-fold cross-validation. First, all 7 classes of the depression scale were targeted individually; then, the neighbor classes were combined to create 6, 5, 4, and 3 class classification models; finally, binary models were created.

The feature selection procedure significantly distinguished two of the total 20 features (LR $p<0.05$) which also provided the best performance metrics across all examined models. These two features were the "reward tipping point" and the "tradeoff range", both belonging to the RPT feature set. The values for VIF, significance, and correlation can be found in Table 2.

Table 2 Features of highest importance

Feature	VIF	p-value	Pearson Coefficient	Spearman Coefficient
Reward tipping point	1.002	0.001	0.0044	0.13
Tradeoff range	1.000	0.001		

A. Multi-Class Classification

For the multi-class classification approach, we employed OneVsRest LR. All possible class combinations were investigated and the most noteworthy are shown in Table 3, where the pipe symbol "|" is used as an indicator for separating classes. The highest accuracy accrues when classes 2,3,4, and 5 were merged into one class, and classes 6 and 7 into another, thus training a 3-class classification LR model.

Table 3 Accuracies by model

Model	Combinations	Accuracy
7-class	1 | 2 | 3 | 4 | 5 | 6 |7	25.0%
6-class	1 | 2 | 3 | 4 | 5 | 6+7	25.1%
4-class	1 | 2+3 | 4+5 | 6+7	39.2%
3-class	1+2 | 3+4+5 | 6+7	50.0%
3-class	1 | 2+3+4 | 5+6+7	64.4%
3-class	**1 | 2+3+4+5 | 6+7**	**72.5%**

B. Binary Classification

To maximize the intensity of the examined condition, and for balancing the subject count for the remaining classes, it was decided - based on prior results – to keep classes 1, 6, and 7. These classes represent participants with lower (class 1) and higher (classes 6 & 7) reported history of depression. Furthermore, the target vector was transformed to binary by assigning class 1 to the "zero" value and classes 6 & 7 to the "one" value. The derived dataset included 78 participants, 37 (47%) in the "zero" class (54% males) and 41 (53%) in the "one" class (46% males).

With the 78-subject dataset, we performed Binary Logistic Regression using the two RPT features of highest importance (reward tipping point and tradeoff range) as predictors and the binary state of depression as the target. This binary model was trained with LR, Random Forest (RF) and Support Vector Machine (SVM) classifiers, and the results are summarized in Table 4, below. 1000 repeats of 100-Repeated-Stratified-10-fold cross validation were performed.

Table 4 Binary classification results

Classifier	Accuracy	Precision	Recall	ROC AUC
RF	0.64	0.69	0.60	0.70
LR	0.72	0.75	0.73	0.78
SVM	0.73	0.78	0.70	0.77

IV. CONCLUSIONS

Mental health, and mood states in particular, are at the epicenter of modern life [15]. In recent years, depression has become the most dominant mental health problem in adults [16,17], and is the leading cause of suicide [18]. This study provides evidence that behavioral economic RPT features, derived from a simple and quick picture rating task, can successfully predict three depression states with increasing severity with 72.5% accuracy, and two states of "none" vs. "severe" depression with 73% accuracy. These findings suggest that RPT features may be useful to predict active depression, thereby objectively aiding the diagnosis, treatment, and monitoring of depression. Future work will examine larger cohorts and expand to other mental health conditions.

ACKNOWLEDGMENT

This work has been supported by the Office of Naval Research, Grant N000142112216.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

1. Richards, D. (2011). Prevalence and clinical course of depression: a review. Clinical psychology review, 31(7), 1117-1125.
2. Kim, B. W., Kennedy, D. N., Lehár, J., Lee, M. J., Blood, A. J., Lee, S., ... & Phenotype Genotype Project in Addiction and Mood Disorders (PGP). (2010). Recurrent, robust and scalable patterns underlie human approach and avoidance. PLoS One, 5(5), e10613.
3. Lee, S., Lee, M. J., Kim, B. W., Gilman, J. M., Kuster, J. K., Blood, A. J., ... & Breiter, H. C. (2015). The commonality of loss aversion across procedures and stimuli. PloS one, 10(9), e0135216.
4. Breiter, H. C., Block, M., Blood, A. J., Calder, B., Chamberlain, L., Lee, N., ... & Zhang, F. (2015). Redefining neuromarketing as an integrated science of influence. Frontiers in Human Neuroscience, 8, 1073.
5. Lang, P., & Bradley, M. M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. Handbook of emotion elicitation and assessment, 29, 70-73.
6. Azcona, E. A., Kim, B. W., Vike, N. L., Bari, S., Lalvani, S., Stefanopoulos, L., ... & Breiter, H. C. (2022). Discrete, recurrent, and scalable patterns in human judgement underlie affective picture ratings. arXiv preprint arXiv:2203.06448.
7. McKinney, W. (2010, June). Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference (Vol. 445, No. 1, pp. 51-56).
8. Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E.. (2020). Array programming with NumPy. Nature, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2
9. Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In 9th Python in Science Conference.
10. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.
11. Jain, A., & Zongker, D. (1997). Feature selection: Evaluation, application, and small sample performance. IEEE transactions on pattern analysis and machine intelligence, 19(2), 153-158.
12. Wasserman, W., Kutner, M. H., & Neter, J. (1983). Applied linear regression models. Homewood, IL: Richard D. Irwin. (pp. 391–393)
13. Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. Quality & quantity, 52(4), 1957-1976.
14. Zou, K. H., Tuncali, K., & Silverman, S. G. (2003). Correlation and simple linear regression. Radiology, 227(3), 617-628.
15. Marten, W. D., & Wilkerson, B. (2003). Stress, work and mental health: a global perspective. Acta neuropsychiatrica, 15(1), 44-53.
16. Centers for Disease Control and Prevention (CDC. (2010). Current depression among adults---United States, 2006 and 2008. MMWR. Morbidity and mortality weekly report, 59(38), 1229-1235.
17. Brody, D. J., Pratt, L. A., & Hughes, J. P. (2018). Prevalence of depression among adults aged 20 and over: United States, 2013-2016.
18. National Collaborating Centre for Mental Health (UK. (2010). Depression: the treatment and management of depression in adults (updated edition). British Psychological Society.

corresponding author:

Author: Leandros Stefanopoulos
Institute: Northwestern University
Street: 2145 Sheridan Rd.
City: Evanston, IL
Country: USA
Email: leandros@northwestern.edu