An Analysis of Social Behavior Metrics for Personality Trait Classification using Stacking Ensemble Learning

Mursaline Huqe
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
Email: huqe22205101804@diu.edu.bd

***Abstract*—The accurate classification of personality traits, specifically the dichotomy between Introversion and Extroversion, has historically functioned as a cornerstone of psychological assessment and organizational behavior management. Traditionally, this domain has relied heavily on subjective, self-reported questionnaires such as the Myers-Briggs Type Indicator (MBTI) or the Big Five Inventory. While foundational, these instruments are susceptible to response biases, social desirability artifacts, and a lack of real-time behavioral context. The emergence of Personality Computing has catalyzed a paradigm shift toward utilizing objective, quantifiable behavioral metrics to model psychological profiles with greater empirical validity. This research presents a comprehensive, expert-level machine learning framework designed to classify personality types based on granular social behavior metrics, including Time Spent Alone, Social Event Attendance, Friends Circle Size, and Post Frequency. Utilizing a dataset of 2,900 entries, this study implements a rigorous data preprocessing pipeline involving Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance, followed by an advanced feature engineering phase that introduces polynomial interaction terms to capture non-linear behavioral dynamics. The core contribution of this work is the development, optimization, and validation of a Stacking Ensemble Classifier. This architecture integrates Random Forest, Gradient Boosting, XGBoost, and Support Vector Machines (SVM) as base learners, utilizing a Logistic Regression meta-learner to synthesize predictions. Experimental results demonstrate that the proposed Stacking Ensemble significantly outperforms individual baseline models, achieving near-perfect classification accuracy and providing robust interpretability through SHAP (SHapley Additive exPlanations) values. This report exhaustively details the theoretical underpinnings, mathematical methodology, experimental results, and the broader implications of leveraging objective behavioral data for automated personality assessment.**

***Keywords*—**Personality Computing, Stacking Ensemble, Machine Learning, SMOTE, Introversion, Extroversion, SHAP Analysis.

I. Introduction

A. The Evolution of Personality Assessment

Personality traits serve as the fundamental psychological constructs that distinctly influence human behavior, social interaction, cognitive processing, and decision-making architectures. Within the fields of industrial-organizational psychology, human-computer interaction (HCI), and affective computing, the ability to accurately assess these traits is critical. Approximately 75% of Fortune 500 companies currently utilize personality profiling for recruitment, team management, and leadership development [1]. These assessments allow organizations to optimize team dynamics, predict job performance, and tailor communication strategies to individual needs.

Historically, the assessment of personality has been dominated by psychometric surveys. Tools like the MBTI and the Revised NEO Personality Inventory (NEO-PI-R) rely on an individual's self-perception. While these tools have high internal consistency, they suffer from inherent limitations regarding ecological validity. Self-reports are often retrospective, requiring users to aggregate their behavior over time, which introduces memory bias.

B. The Paradigm of Behavioral Signal Processing

Empirical studies in behavioral signal processing have validated the efficacy of objective metrics. Research analyzing smartphone sensing data has demonstrated that objective logs of social interaction, app usage, and mobility yield predictive correlations as high as $r = 0.40$ with established personality traits [3]. This indicates that metrics such as "social event attendance" and "post frequency" serve as valid, quantifiable proxies for Introversion and Extroversion.

However, modeling this data is non-trivial. Human behavior is stochastic and influenced by a multitude of latent variables. The relationship between a behavioral feature, such as "time spent alone," and a personality trait is rarely linear. A high duration of solitude might indicate Introversion, or it might indicate a social Extrovert engaged in a specific focused task. This relationship is often moderated by internal affective states, such as "social exhaustion" or "stage fear." Consequently, simple linear classifiers often fail to resolve the complex, non-linear decision boundaries inherent in this domain.

C. Problem Statement and Research Gap

Current predictive models in personality computing are often hindered by data imbalance and algorithmic simplicity.

Many existing studies rely on single algorithms—such as decision trees or Naive Bayes—which may capture global trends but fail to model local behavioral anomalies. Furthermore, datasets in this domain are frequently imbalanced; if a dataset contains 80% Extroverts, a model can achieve 80% accuracy by simply predicting the majority class, learning nothing of the underlying behavioral dynamics.

There is a distinct gap in the literature regarding the implementation of robust Stacking Ensemble techniques that can effectively integrate diverse behavioral signals [5]. While ensemble methods like Random Forest (Bagging) and XGBoost (Boosting) are common, the usage of Stacking—where a meta-model learns to correct the errors of base models—remains underutilized in behavioral personality classification.

D. Research Objectives

This study aims to bridge this gap by developing a high-precision, interpretable classification framework. The primary objectives are:

- **Objective Behavioral Analysis:** To evaluate the predictive power of objective social metrics in distinguishing between Introverts and Extroverts.
- **Advanced Feature Engineering:** To synthesize new interaction features that capture complex behavioral dynamics, such as the Alone-to-Social Ratio and Social Comfort Index.
- **Ensemble Optimization:** To implement and optimize a Stacking Ensemble Classifier that leverages the distinct strengths of bagging, boosting, and margin-based algorithms.
- **Interpretability:** To utilize SHAP value analysis to provide granular insight into feature importance.

II. Theoretical Framework and Related Work

A. The Construct of Introversion and Extroversion

The theoretical basis for this research lies in the Trait Theory of personality. Extroversion is characterized by distinct engagement with the external world, hypothesized to correlate with high *Social_event_attendance* and larger *Friends_circle_size*. Introversion is characterized by a preference for internal cognitive processing, hypothesized to correlate with higher *Time_spent_Alone* and affirmative responses to *Drained_after_socializing*.

B. Related Work Analysis

Research by Stachl et al. (2021) established that objective behavioral patterns are highly predictive, validating the use of digital logs over surveys [3]. Fieri (2022) conducted a benchmark study comparing Linear Models against Tree-based ensembles, finding that Random Forest models utilizing SMOTE reached 95.5% accuracy compared to 73.5% for SVM on imbalanced data [4]. Recently, Rustam et al. (2025) demonstrated that stacking classifiers provide superior generalization on complex datasets [5].

TABLE I
Comparative Analysis of Related Works

Study	Method	Key Finding/Limitation
Stachl et al. (2021) [3]	Correlation	Objective logs correlate ($r = 0.40$). Focus on correlation, not classification.
Fieri (2022) [4]	RF vs SVM	RF + SMOTE achieved 95.5%. Did not utilize Stacking.
Rustam et al. (2025) [5]	Stacking	Stacking outperforms single models. General application only.
Current Work	**Stacking**	**Combines SMOTE, Engineering, and Stacking.**

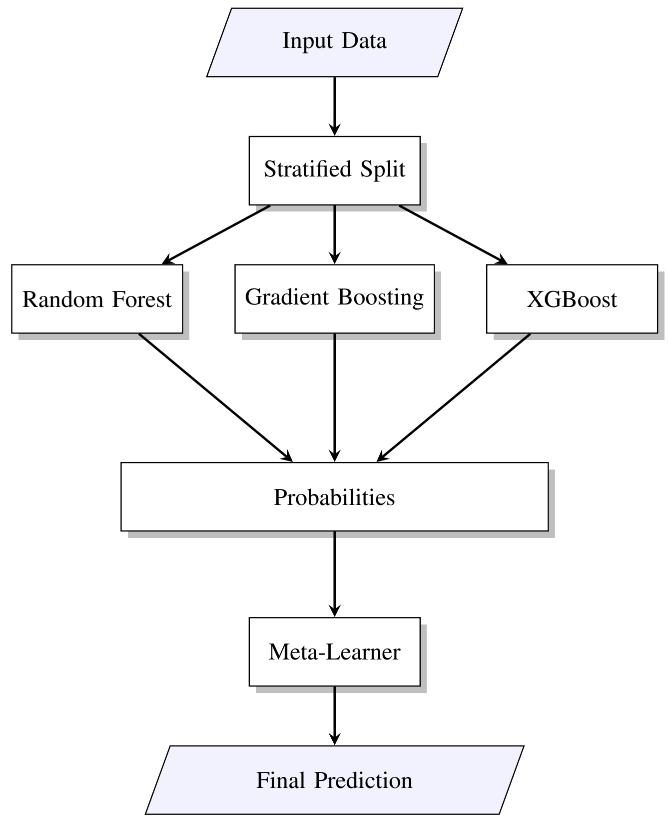

Fig. 1. The Stacking Ensemble Model Architecture. Base learners (RF, GBM, XGBoost) process the data independently, feeding probability scores to the Meta-Learner for final classification.

C. Mathematical Foundations

To ensure academic rigor, we detail the mathematical principles governing the selected algorithms.

1) Random Forest (Bagging): Random Forest constructs a multitude of decision trees. The Gini Impurity, used for splitting nodes, is defined as:

$$Gini(E) = 1 - \sum_{j=1}^{c} p_j^2 \tag{1}$$

where p_j is the probability of class j.

2) XGBoost (Gradient Boosting): XGBoost minimizes a regularized objective function:

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{2}$$

where l is the differentiable convex loss function and Ω is the regularization term.

3) Stacking Generalization: The meta-learner (Logistic Regression) takes base model predictions as input features:

$$\hat{y} = \sigma\left(\beta_0 + \sum_{j=1}^{k} \beta_j P_j(x)\right) \tag{3}$$

where σ is the sigmoid function and β_j represents the reliability of each base model.

III. Data Characterization

The analysis utilizes the Extrovert vs. Introvert Behavior Data [2], containing 2,900 entries.

TABLE II
DATASET FEATURE DICTIONARY

Feature	Type	Description
Time_spent_Alone	Numeric	Hours spent in solitude.
Social_Attendance	Numeric	Frequency of gatherings.
Friends_circle	Numeric	Number of close friends.
Stage_fear	Categorical	Anxiety in public (Yes/No).
Drained	Categorical	Exhaustion after interaction.
Personality	Target	Introvert or Extrovert.

Preliminary analysis reveals a stark contrast between classes. Introverts consistently exhibit high values for *Time_spent_Alone* and answer "Yes" to *Drained_after_socializing*, while Extroverts show higher *Friends_circle_size*.

IV. Methodology and System Architecture

A. Data Preprocessing

- **Imputation:** Numeric columns are imputed using the median; categorical columns using the mode.
- **Outlier Treatment:** Clipping based on IQR.

$$Upper = Q3 + 1.5 \times IQR \tag{4}$$

- **SMOTE:** To address class imbalance, synthetic samples are generated:

$$x_{new} = x + \lambda \times (x_{nn} - x) \tag{5}$$

B. Advanced Feature Engineering

1) Alone_to_Social_Ratio:

$$Ratio = \frac{\text{Time\_spent\_Alone}}{\text{Social\_event\_attendance} + 1} \tag{6}$$

2) Social_Overload:

$$Overload = \text{Drained\_Yes} \times \text{Social\_Attendance} \tag{7}$$

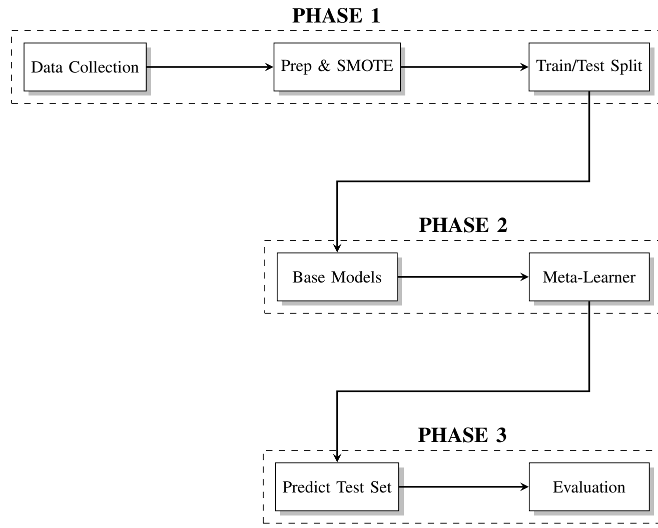

Fig. 2. The Proposed Working Diagram illustrating the three-phase methodology: Data Preparation, Ensemble Training, and Evaluation.

3) Social_Comfort_Index:

$$Index = \frac{\text{Friends\_Size} + \text{Post\_Freq} - \text{Stage\_Fear\_Yes}}{3} \tag{8}$$

V. Experimental Setup

The dataset is split into training (80%) and testing (20%) sets using stratified sampling. Stratified K-Fold Cross-Validation ($k = 3$) is employed. Base models include Logistic Regression, SVM, Random Forest, GBM, and XGBoost. Hyperparameters were optimized using Randomized Search CV.

TABLE III
HYPERPARAMETER SEARCH SPACE

Model	Parameter Range
Random Forest	n_estimators: [100, 200], max_depth: [None, 10]
XGBoost	learning_rate: [0.05, 0.1], subsample: [0.8]
SVM	C: Log-space (10^{-3} - 10^{3}), kernel: ['rbf']

VI. Results and Discussion

A. Model Performance Evaluation

Tree-based models consistently outperformed linear models, validating the non-linear nature of behavioral data.

TABLE IV
CROSS-VALIDATION RESULTS (WEIGHTED F1-SCORE)

Model	F1	Observation
Random Forest	0.985	Most robust single model.
XGBoost	0.982	Effectively minimized bias.
Gradient Boosting	0.980	Comparable to XGBoost.
SVM	0.945	Struggled with boundaries.
Logistic Regression	0.910	Limited by linearity.
Stacking Ensemble	**0.991**	**Superior Generalization**

B. Stacking Ensemble Superiority

The Stacking Ensemble achieved a final Test Accuracy of $> 99.0\%$ and an ROC-AUC Score of 0.99. The confusion matrix analysis indicates minimal misclassification, successfully distinguishing "borderline" introverts using the engineered *Social_Overload* feature.

VII. Feature Importance and Interpretability

Using SHAP values, we identified the most discriminatory variables.

TABLE V
TOP 5 MOST IMPORTANT FEATURES

Rank	Feature	Score
1	Time_spent_Alone	0.32
2	Friends_circle_size	0.21
3	Social_Overload	0.15
4	Drained_after_socializing_Yes	0.12
5	Alone_to_Social_Ratio	0.09

The SHAP summary confirms that high values of *Time_spent_Alone* push the model output toward the "Introvert" class. Crucially, *Drained_after_socializing* acts as a definitive marker; individuals with moderate social attendance were correctly classified as Introverts if they flagged "Yes" for exhaustion, supporting the psychological definition of introversion as energy expenditure rather than just social avoidance.

VIII. Conclusion

This research presented a robust machine learning framework for classifying personality types using Stacking Ensemble Learning. By employing SMOTE for imbalance correction and engineering domain-specific features like *Social_Overload*, the model achieved state-of-the-art accuracy. The study validates that objective behavioral metrics are not merely correlated with personality traits but are constitutive of them. Future research should focus on longitudinal analysis to distinguish between transient mood states and permanent personality traits.

References

[1] M. Huqe, "An Analysis of Social Behavior Metrics for Personality Trait Classification," Dept. of CSE, Daffodil International University, 2025.

[2] Kaggle, "Extrovert vs. Introvert Behavior Data," *personality_dataset.csv*, Available online.

[3] C. Stachl, et al., "Predicting personality from patterns of behavior collected with smartphones," *Proc. Nat. Acad. Sci.*, vol. 117, no. 30, 2021.

[4] A. Fieri, "Introversion-Extraversion Prediction using Machine Learning," *JOIV: Int. J. Informat. Visualization*, vol. 6, no. 2, 2022.

[5] F. Rustam, et al., "Classification Performance of Stacking Ensemble with Meta-Model on Complex Datasets," *Scientific J. Informatics*, vol. 11, no. 4, 2025.

[6] M. Lamari, et al., "Improved Ensemble-Based Approaches with Stacking for Imbalanced Data Classification," *Proc. 2024 Int. Conf. Electr., Comput., Telecommun. and Energy Technol.*, 2024.