"Multimodal Machine Learning for Psychological Profiling: Predicting Psychological traits from Video, Audio, and Text"

---

1. Background and Rationale

Understanding human psychological states such as anxiety, depression, cognitive reflection, risk-taking, and personality traits is central to fields spanning psychology, behavioral economics, and affective computing. Traditionally, these constructs are assessed using self-report questionnaires, which are limited by subjective bias, time constraints, and situational variability.

Recent advances in **multimodal machine learning** and **deep learning** offer the possibility of inferring psychological traits automatically from naturalistic behavioral data, such as **speech, facial expressions, and linguistic patterns**. Video, audio, and text provide complementary cues: prosody and tone reveal affective states; facial dynamics reflect emotion and engagement; and language encodes cognitive style and self-referential tendencies.

The availability of a comprehensive dataset including **video, audio, and text recordings** alongside **validated psychometric scores** (e.g., anxiety, depression, CRT, risk-taking, financial literacy, and Big Five traits) provides a unique opportunity to develop data-driven models capable of predicting these constructs with high accuracy and interpretability.

2. Research Aim

The goal is to establish reliable, non-intrusive, and scalable computational methods for psychological assessment.

3. Research Objectives

1) **Feature Extraction and Data Representation**
    a) Extract interpretable behavioral features from each modality:
        i) **Video**: facial landmarks, micro-expressions, gaze patterns, and head movements using frameworks like OpenFace, MediaPipe, and DeepFace.
        ii) **Audio**: prosodic, spectral, and temporal features (e.g., pitch, jitter, formants) using libraries such as Librosa or OpenSMILE.
        iii) **Text**: semantic and psycholinguistic features via transformer-based embeddings (BERT, RoBERTa) and LIWC.
2) **Unimodal Modeling**
    a) Build baseline machine learning and deep learning models for each modality to predict Anxiety, depression levels, Cognitive reflection (CRT), risk-taking tendencies, Big Five personality traits, etc.
3) **Multimodal Fusion**

a) Design fusion architectures (e.g., late fusion, attention-based fusion, or transformer-based fusion models) to integrate heterogeneous data streams for improved psychological trait prediction.

4) **Model Interpretability and Validation**
   a) Use SHAP, LIME, and attention visualizations to interpret model decisions.
   b) Validate model performance using cross-validation and external datasets (if available).
5) **Ethical and Practical Implications**
   a) Address issues of privacy, fairness, and consent in the automated assessment of psychological states.

4. Methodology

- **Data**: Utilize the existing multimodal dataset collected from human subjects, including questionnaire responses and synchronized audio, video, and textual recordings.
- **Preprocessing**: Data cleaning, synchronization, and normalization across modalities.
- **Model Development**:
  - Implement and evaluate CNNs, RNNs, Transformers, and hybrid architectures.
  - Explore transfer learning from emotion recognition and personality datasets.
- **Evaluation Metrics**:
  - Regression and classification performance (e.g., MAE, R^2, F1-score).
  - Model interpretability and generalizability across demographics.

---

5. Required Candidate Skills

The ideal candidate will have:

- Strong **Python programming** skills and experience with data science libraries (NumPy, pandas, scikit-learn).
- Proficiency in **deep learning frameworks** (PyTorch, TensorFlow, Keras).
- Familiarity with **multimodal data processing** (video, audio, text).
- Knowledge of **machine learning for behavioral or affective computing**.
- Happy to perform **data collection.**
- Experirnced in **preprocessing, and experimental design**.
- Interest in **psychology, human behavior, or computational social science**.

---

6. Expected Outcomes

- A suite of machine learning models capable of inferring psychological traits from naturalistic multimodal data.
- A benchmark dataset linking psychometric scores with behavioral indicators.

- Academic publications in high ranking affective computing, computational psychology, and AI journals.

For any question, please contact Dr. Mohammad Tayarani

m.tayaraninajaran@herts.ac.uk

References

1. M. H. Tayarani- N., Anna Esposito and Alessandro Vinciarelli, “What an ”Ehm” Leaks About You: Mapping Fillers into Personality Traits with Quantum Evolutionary Feature Selection Algorithms”, IEEE Transactions on Affetive Computing, 2019.

2. M. -H. Tayarani-N. and Alessandro Vinciarelli, “Learning Classifier Performance in Ensemble of Classifiers for Personality Prediction Using Laughter”, IEEE Transactions on Affective Computing, 2025.

3. M. -H. Tayarani-N. and Shamim Ibne Shahid, “Detecting Anxiety via Machine Learning Algorithms: A Literature Review”, IEEE Transactions on Emerging Topics in Computational Intelligence, 2025.