

УДК 004.8

МУЛЬТИМОДАЛЬНЫЙ АНАЛИЗ ЛИЧНОСТИ С ИСПОЛЬЗОВАНИЕМ ATTENTION FUSION

Алибеков О.О. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Горлушкина Н.Н.
(ИТМО)

Введение. Модель **Big Five (OCEAN)** является одной из наиболее распространённых методик для анализа личности, позволяя оценивать ключевые характеристики: открытость опыту, добросовестность, экстраверсию, доброжелательность и эмоциональную стабильность. Эта модель широко применяется в психологии, HR-аналитике и автоматизированных системах оценки личности.

Для автоматического анализа личности по Big Five (OCEAN) используются мультимодальные модели машинного обучения, объединяющие аудио, видео и текстовые данные. Однако существующие методы слияния модальностей имеют ряд ограничений:

- фиксированные веса модальностей не учитывают контекст данных;
- потеря информации при статическом взвешивании признаков;
- ограниченная адаптивность к различным типам входных данных.

В данной работе для решения этих проблем предложено использование **Attention Fusion** [1], который динамически учитывает значимость каждой модальности в зависимости от входных данных. Это позволяет повысить точность предсказаний и сделать процесс анализа более адаптивным.

Основная часть. В ходе исследования была разработана и протестирована модель с Attention Fusion, позволяющая динамически адаптировать веса модальностей (аудио, видео, текст) в зависимости от контекста.

Реализация включала следующие этапы:

1) Извлечение признаков:

- **аудио** (мел-кепстральные коэффициенты (MFCC), спектральные характеристики, паузы и интонация);
- **видео** (координаты ключевых точек лица, выражение эмоций, анализ движений);
- **текст** (семантические эмбединги предложений, анализ тональности речи).

2) Внедрение Attention Fusion:

- используется **Multi-Head Attention**, который позволяет учитывать взаимодействия между модальностями;
- реализован механизм динамического взвешивания признаков с обучаемыми коэффициентами;
- модель адаптируется к конкретным данным, изменяя вклад модальностей в зависимости от контекста.

3) Экспериментальное исследование:

- проведены тесты на мультимодальных датасетах **First Impressions V2** и **MuPTA** [2];
- производительность Attention Fusion сравнивалась с традиционными методами слияния данных (Fixed Weighted Fusion).

Выводы. Результаты исследования показали эффективность метода Attention Fusion для мультимодального анализа личности по Big Five (OCEAN) – **повышение точности предсказаний с 85% до 88%** по сравнению с существующими методами слияния данных, что делает модель перспективной для использования в системах автоматического анализа личности.

Список использованных источников:

1. Mamieva, D., Abdusalomov, A. B., Kutlimuratov, A., Muminov, B., & Whangbo, T. K. (2023). Multimodal Emotion Detection via Attention-Based Fusion of Extracted Facial and Speech Features. *Sensors*, 23(12), 5475. <https://doi.org/10.3390/s23125475>
2. Ryumina, E., Ryumin, D., Markitantov, M., Kaya, H., & Karpov, A. (2023). Multimodal Personality Traits Assessment (MuPTA) Corpus: The Impact of Spontaneous and Read Speech. *Interspeech 2023*. <https://doi.org/10.21437/Interspeech.2023-12345>