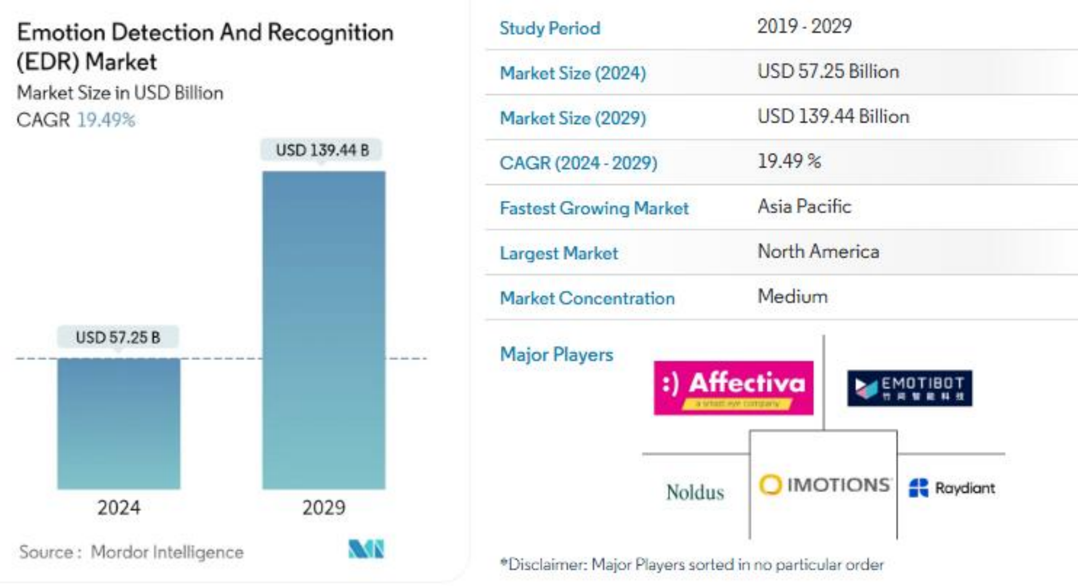
Ufa University Science and Technology, Russia

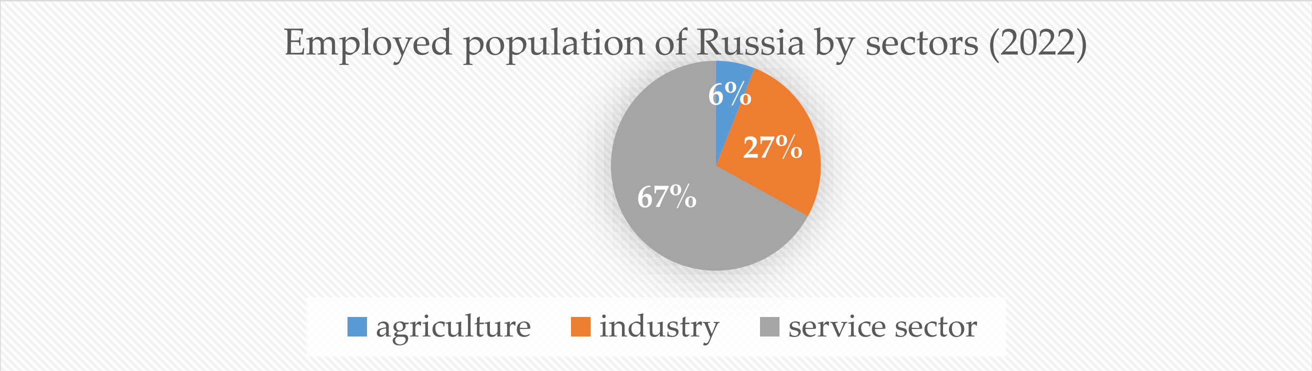
Application of AI technologies to assess the quality of healthcare services based on multidimensional client emotion recognition

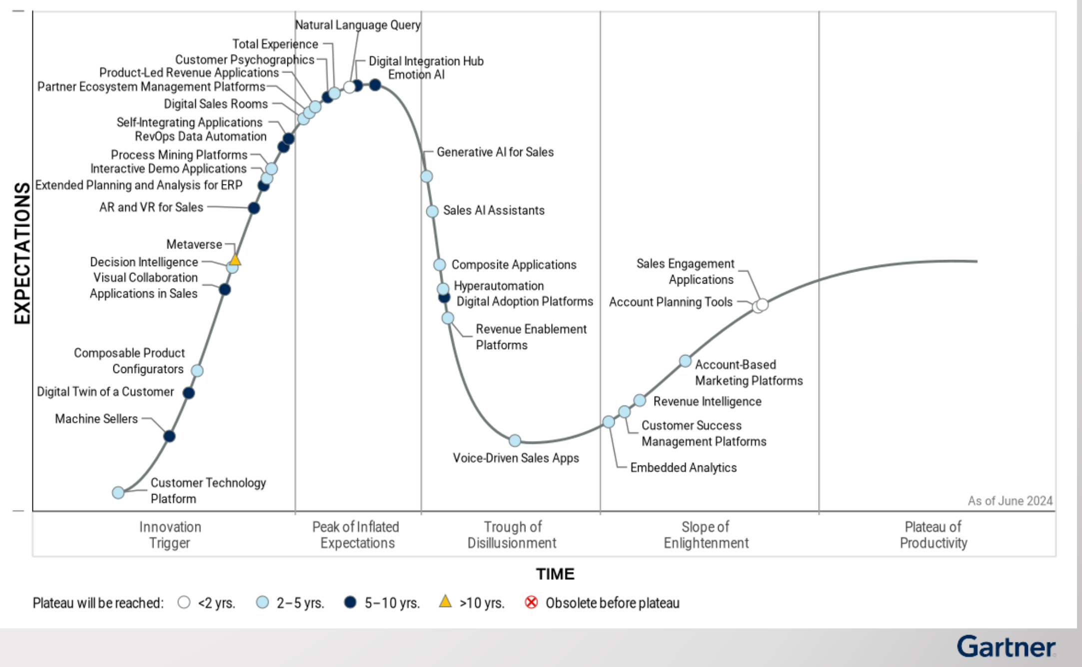
Institute of Computer Science,
Mathematics and Robotics
Associate Professor
Dr. Diana Bogdanova

Relevance

- One of the factors of quality of services provided in healthcare is their direct connection with patient satisfaction, which affects their stability.
- The quality of medical services provided has a direct impact on such indicators as profitability and competitiveness of medical companies.
- Emotions are an excellent reflection of a person's internal feelings about what is happening around them. In particular, emotions can reflect personal perception of the quality of medical services provided, expressed in the form of patient satisfaction.
- According to market estimates, the emotion analysis systems market is projected to grow to over $139 billion by 2029.
- On the Gartner curve, emotion recognition is currently at the peak of inflated expectations, with a projected transition to the plateau of productivity in 5–10 years (Gartner, 2024).







Purpose of the study

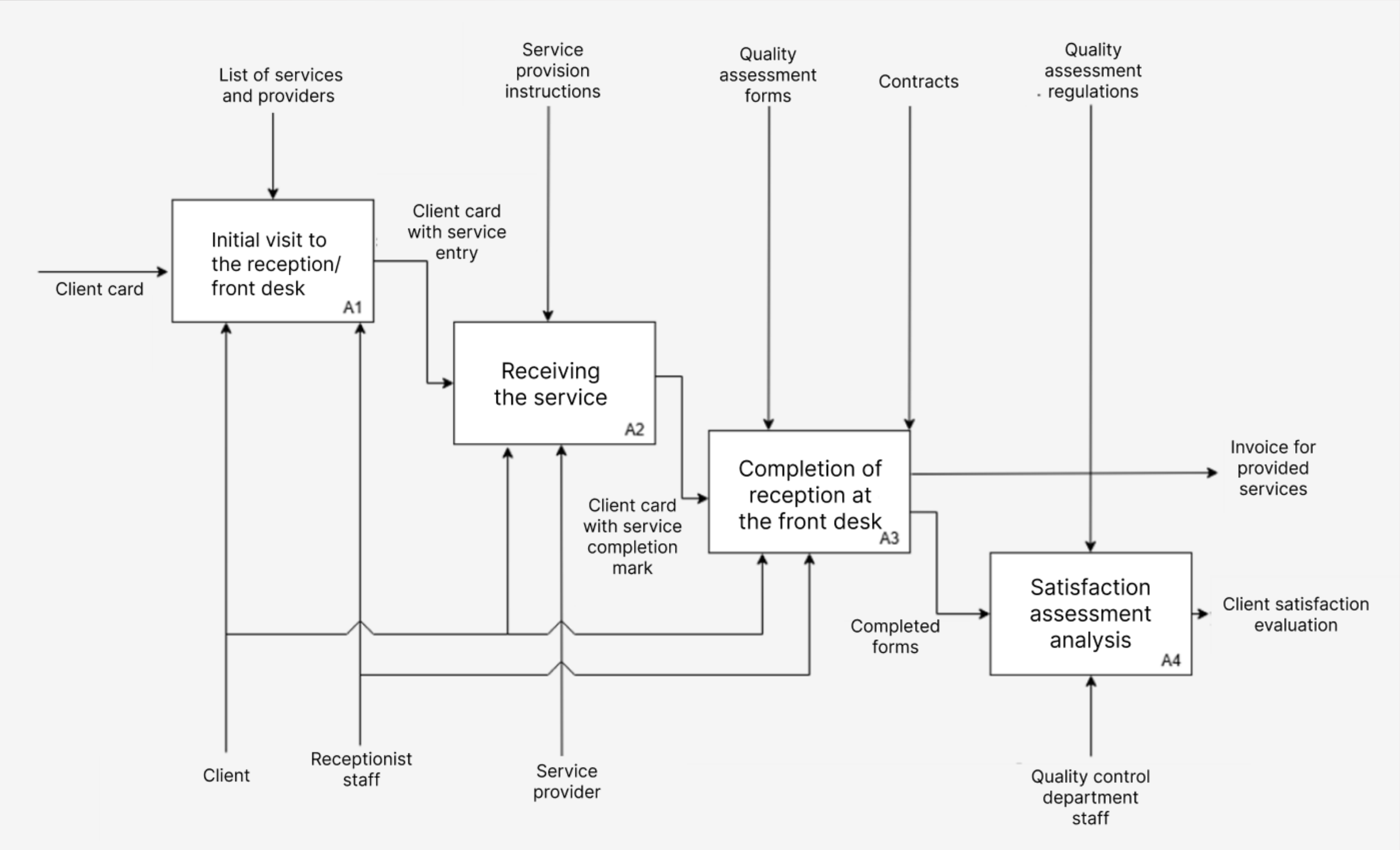
Development of a decision support system for managing the quality of medical care for patients based on their emotional satisfaction, which will automate the process of assessing patient satisfaction and integrate it into medical care quality management processes.

Research objectives

1. Conduct an analysis of current approaches to assessing the quality of medical care and patient satisfaction, as well as an analysis of facial emotion recognition systems. Based on the analysis results, develop a concept for integrating facial emotion recognition into the patient satisfaction assessment process.
2. Develop a model of a facial emotion recognition system based on the analysis conducted and adapt it to the concept of the patient satisfaction assessment process.
3. To develop a model and method for assessing patient satisfaction based on the developed integration concept, allowing for a control impact on the quality of medical care, and an algorithm for supporting decision-making based on the resulting model.
4. To develop a software implementation of a decision support system for managing the quality of medical care based on the emotional satisfaction of patients.

Quality Assessment Process (1/2)

"AS-IS" patient Satisfaction Assessment Process Model

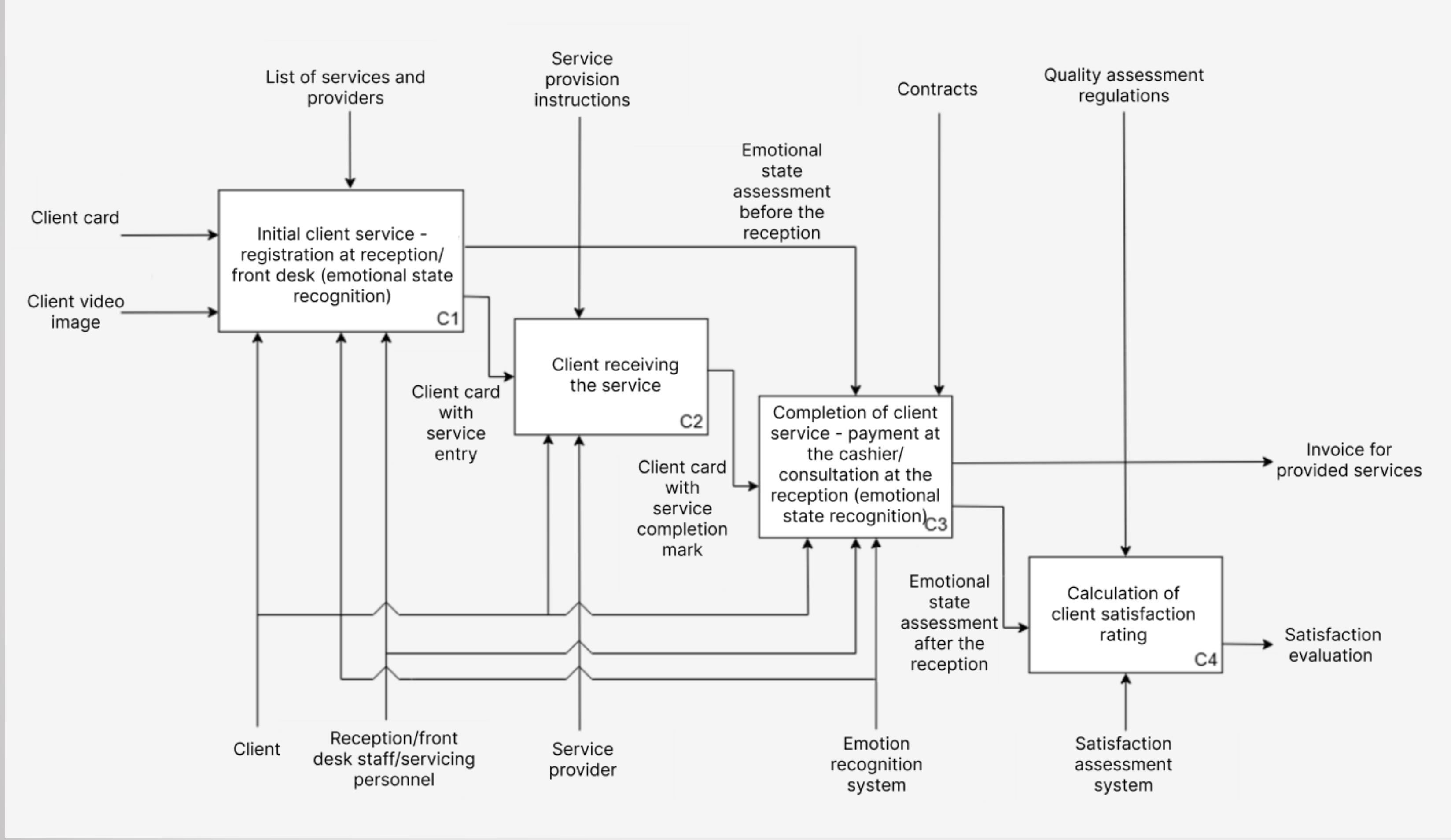


Disadvantages of this process:

- The assessment uses outdated methodologies developed 20–40 years ago, which rely on classical information collection methods such as surveys and feedback forms.
- Requires active patient participation in all quality assessment variations (leaving reviews, filling out questionnaires, complaint books, etc.).
- Manual processing of results is required (feedback, questionnaires).
- patient reluctance to spend time on quality assessment procedures
- patients may be afraid to express their true opinion about service quality due to personal contact.
- Medical companies are forced to offer incentives (discounts, coupons, certificates, etc.) to encourage patient participation in quality assessment.

Quality Assessment Process (2/2)

"AS-TO-BE" patient Satisfaction Assessment Process Model
"



Key Points:

- The patient's emotional state is captured at their first visit (registration/consultation) and again after receiving the service.
- This allows comparison of the patient's emotional state before and after the medical service, enabling conclusions about patient satisfaction based on the change in emotional state.
- Any need for additional actions from the patient is eliminated.
- For the patient, this process occurs exactly as if no satisfaction assessment were being conducted.
- Manual data processing is not required; the system automatically assigns the evaluation.

=> This creates the need to develop such an automation system.

Concepts of Emotion Typology

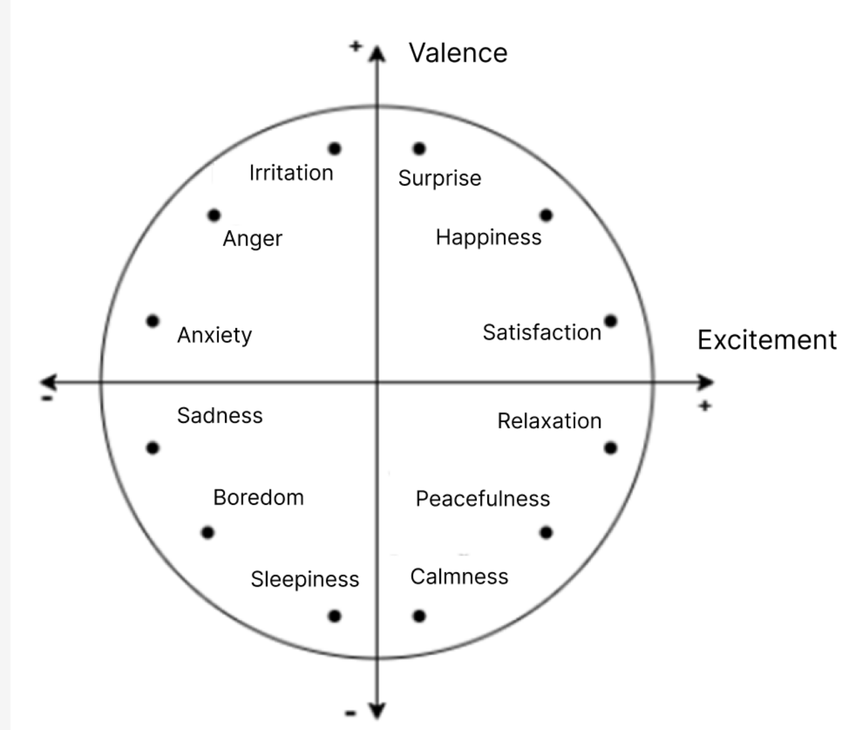
Basic Emotions Concept:

- A set of basic emotions is identified. For example, Paul Ekman's six basic emotions: joy, anger, disgust, fear, sadness, and surprise.
- Complex emotions are formed as combinations of basic ones
- There is no consensus on the number and composition of basic emotions.
- The most popular concept. The largest number of training sets.

Multidimensional Emotion Concept:

- Certain emotional bases (primary characteristics of emotions) are identified in a multidimensional space.
- Emotions are represented as points within this space.
- There is no agreement on the number of bases.
- Less common due to complexity and high cost of creating training datasets.
- Better suited for facial emotion recognition

We use the multidimensional theory of emotion typology, which defines two characteristics: arousal and valence. This model is based on James Russell's two-dimensional circumplex model of emotions.



Russell's Circumplex Model

Emotion Recognition System Model (1/4)

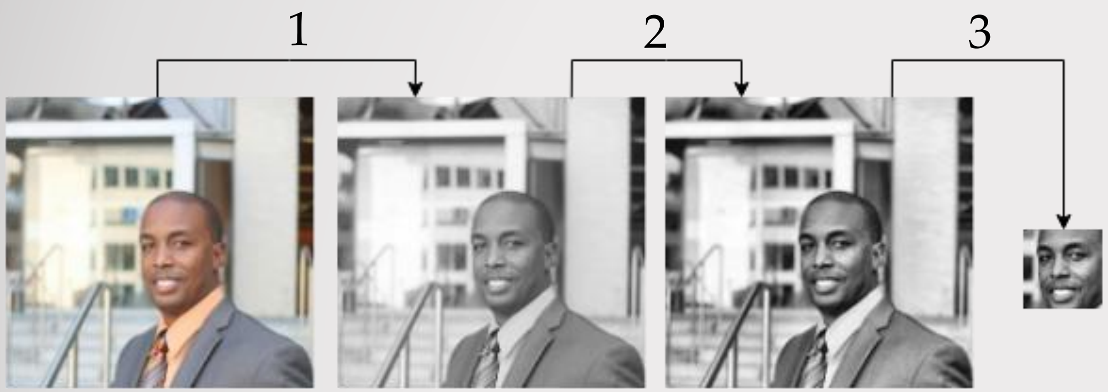
Mathematical formulation of the problem

- The input data of the task are images containing the patient's face: $K_i = \{K_i^{t_1}, K_i^{t_2}, \dots, K_i^{t_l}\}$. The parameter K_i ,within the mathematical model, represents an array of parameters K_i^t that define a specific facial emotion of the i-th client at time t.
- To build the model, it is necessary to find a mapping F such that, given an input image K_i^t containing the person's face, it returns the multidimensional emotional characteristics: $F(K_i^t) \to \{A_i^t, V_i^t\}$. ,where A_i^t is the arousal characteristic and V_i^t is the valence characteristic at time t for the i-th patient.
- Before features can be extracted from an image K_i^t, it must first be processed to enhance feature quality and remove unnecessary information. This requires constructing a transformation $P(K_i^t) \to K'^t_i$, hat converts the original image K_i^t into K'^t_i.
- To extract features from the processed image K'^t_i it is necessary to define a transformation $E(K'^t_i) \to \{x_j\}_1^n$, which would return an array of features x_j ,where j is the feature index from 1 to n.
- Furthermore, to improve the model's quality, the feature quality must be enhanced by applying a frontalization algorithm and a dimensionality reduction method. This requires finding a mapping $T(\{x_j\}_1^n) \to \{x'_j\}_1^k$, where k is the new feature dimension.
- Finally, it is necessary to construct a function R that can map the features x_j' to values of multidimensional characteristics $R\left(\{x'_j\}_1^k\right) \to \{A_i^t, V_i^t\}$

Proposed Approach to Emotion Recognition (2/4)

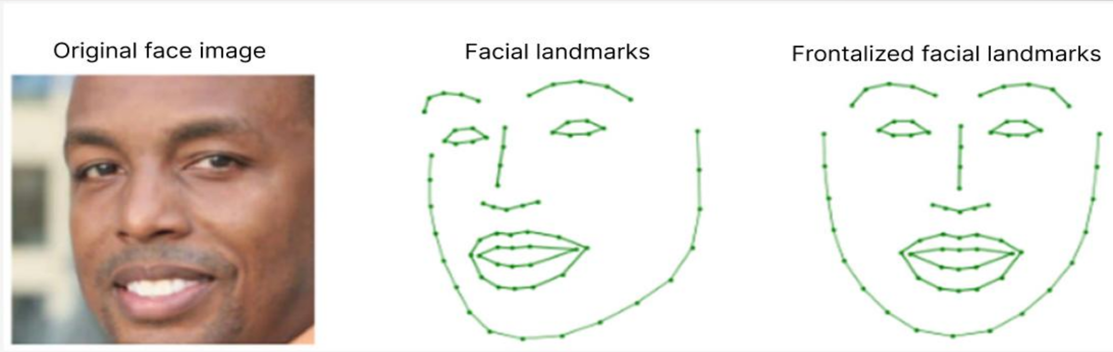
Description of Image Preprocessing and Feature Extraction Stages

Image Preprocessing Stages

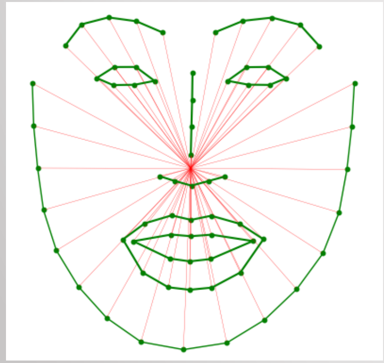


1. Conversion of the image to grayscale.
2. Image illumination normalization using Contrast-Limited Adaptive Histogram Equalization
3. Face coordinate detection using Haar features and cropping the image based on the detected coordinates

Feature Extraction Stage



- Extraction of facial landmark coordinates – 68 points.
- Application of a frontalization algorithm to align facial landmark coordinates to a standard frontal position.



- Based on the obtained frontalized 68 landmarks, the lengths from each point to the center of gravity, as well as the angles formed by these lines and the horizon line, are calculated. Thus, for each point, we have 4 real-valued features: x-coordinate, y-coordinate, distance to the center of gravity, and the angle formed with the horizon line. The original x and y coordinates are excluded from consideration. This results in a total of 136 features.
- The resulting 136 features are then subjected to PCA (Principal Component Analysis) while preserving 99% of the variance. Finally, 25 features remain.

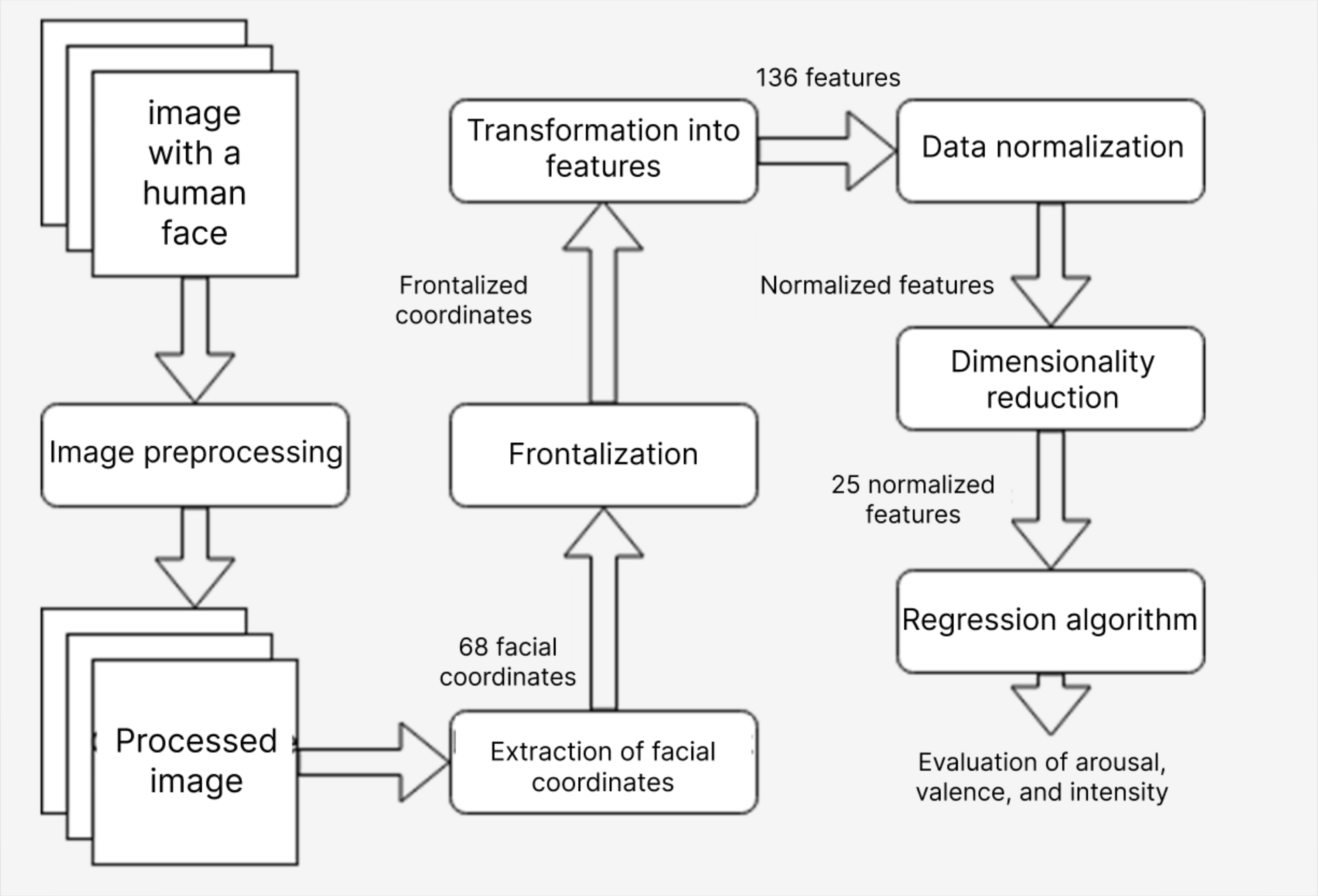
Proposed Approach to Emotion Recognition (3/4)

Alternatives to the Face-Landmark-Based Approach

Approach Based on Ekman's FACS	Neural Network–Based Approach	Facial Landmark-Based Approach
• FACS encodes all possible facial muscle movements → a guideline for emotion recognition. • Strong scientific foundation behind this theory. • No need for a large training dataset • Used in most commercial applications.	• The most accurate emotion classification systems • Robust to various face tilt angles and rotations. • May not require an additional classifier or feature extractor (a neural network can both extract features and perform classification).	• A relatively new approach using computer vision. • A wide variety of annotation options (5 points, 68 points, etc.). • Flexibility in processing point coordinates. • Good accuracy results when preprocessing and classifier are properly selected.
• Complex implementation of algorithms (requires training to identify facial action units) • Additional equipment is often desirable (e.g., a depth sensor) • Rigidity of the approach (strict adherence to instructions).	• Dependence on neural networks, and consequently inheritance of their issues (large number of parameters, uncertainty in architecture and resource demands, “black box” nature) • Requirement for a large training dataset • High demands on computational resources and time.	• Need to find an appropriate method for processing coordinates. • Possibility of generating an excessively large number of features. • Inability to extract features under extreme face rotations.

Proposed Approach to Emotion Recognition (4/4)

General System Model



Experiment Results (1/2)

Description of experiments

Computing resources:
The program was written in Python using the libraries dlib, OpenCV, and sklearn. Training and experiments were conducted on a computing machine with an 8-core processor clocked at 3.6 GHz (GPU computations were not used).

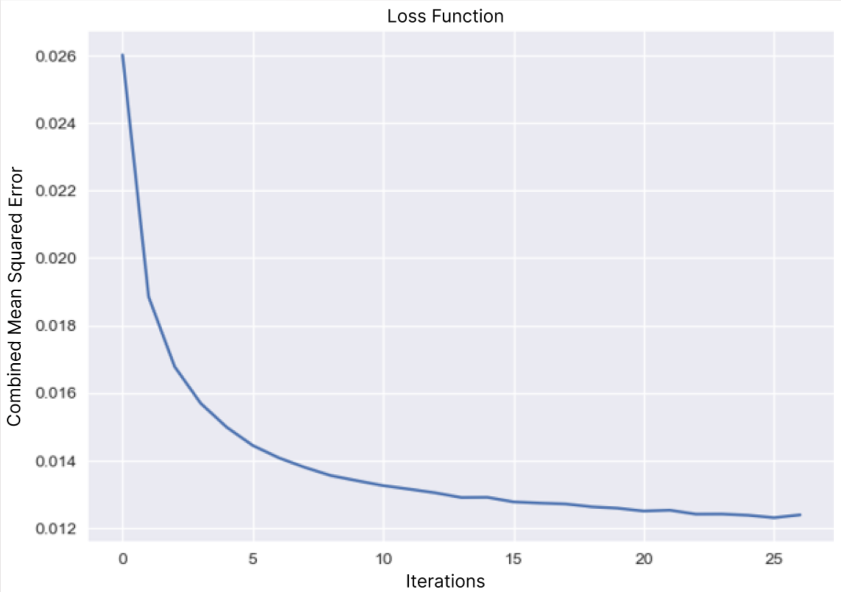
Input data:
For training and testing, a combination of three discrete datasets was used: RaFD, KDEF, and WSEFEP, which were transformed into a multidimensional dataset using the MorphSet framework. The final dataset contains more than 150,000 images from 712 subjects.

Validation methodology:
For validation of accuracy results, the K-fold cross-validation method was used with k = 5 (the dataset is divided into K blocks, training occurs on k-1 blocks, and validation is performed on the k-th block; this process is repeated k times).

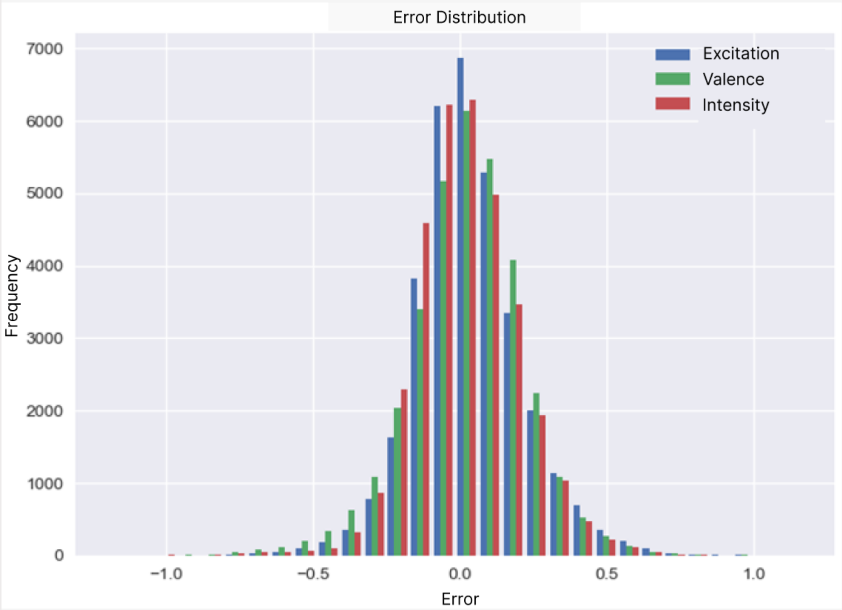
Algorithm parameters:
100 neurons in the hidden layer, activation function – logistic, alpha parameter – 0.0001, initial learning rate – 0.0028, optimizer – adaptive moment estimation (Adam). Optimal parameters were found using grid search.

Experiment Results (2/2)

- The multilayer perceptron achieved the best performance across all measurements.
- The perceptron stops after 25 iterations and reaches a combined mean squared error of 0.0123.
- On the validation subset, the error increases to a combined value of 0.032.
- The mean squared error is relatively low compared to the standard deviation, which indicates good accuracy of the built model.
- The histogram of the error distribution on the validation set shows that most errors are concentrated around 0, and the graph has a symmetric distribution.
- The classification speed for one object is 0.01 seconds.



Название метода	Mean Squared Error			R^2
	Excitation	Valence	Intensity	
Linear Regression	0.053	0.064	0.046	0.41
Elastic Net	0.045	0.053	0.038	0.50
Gaussian Mixture Regression (GMR)	0.047	0.053	0.040	0.49
Random Forest	0.033	0.043	0.030	0.63
Multilayer Perceptron	0.030	0.039	0.028	0.65



Patient Satisfaction Rating (1/5)

Transformation of Valence and Excitation Scores

- To find $OУ_{ij}$ - the emotional satisfaction rating of procedure j for client i, it is necessary to use the proposed model of the patient satisfaction evaluation process and utilize the following obtained scores:$\{A_{ij}^{t}\}$, $\{A'^{t}_{ij}\}$, $\{V_{ij}^{t}\}$, $\{V'^{t}_{ij}\}$,
- Here A_{ij}^{t} – the excitation score of client i for procedure j at time t before receiving the service, A'^{t}_{ij} after receiving the service. Similarly, V_{ij}^{t} is the valence score before receiving the service, and V'^{t}_{ij} is the score after.
- The parameter t will vary for different clients, as each client may spend different amounts of time in the registration area before and after the procedure.
- To mitigate this problem and reduce the size of the data with minimal loss of key emotions, it is proposed to use a sliding window method.
- Within each window, the mean value of the data falling within the window will be computed. Then, applying this algorithm to the original time series, we obtain new data arrays: $\{A^{w}{}^{n}_{ij}\}$, $\{A'^{w}{}^{n}_{ij}\}$, $\{V^{w}{}^{n}_{ij}\}$, $\{V'^{w}{}^{n}_{ij}\}$, where $A^{w}{}^{n}_{ij}$ – the excitation score before receiving the service for the n-th window, a $A'^{w}{}^{n}_{ij}$ is the excitation score after receiving the service. Similarly, $V^{w}{}^{n}_{ij}$ is the valence score before the procedure, and $V'^{w}{}^{n}_{ij}$ is the score after the procedure.
- These resulting arrays will still have different lengths, and the number of elements will be too large for fuzzy logic algorithms to process efficiently. Therefore, we will gather their statistical characteristics, specifically compute their mean values. Finally, we obtain four characteristics:
- $A_{ij}^{E}, A'^{E}_{ij}, V_{ij}^{E}, V'^{E}_{ij}$ – the mean values of the excitation and valence scores before and after the procedure.

Patient Satisfaction Rating (2/5)

Fuzzy System Design

Input values:

- A_{ij}^{E} – the mean value of the excitation window before receiving
- A'^{E}_{ij} – the mean value of the excitation window after receiving
- V_{ij}^{E} – the mean value of the valence window before receiving
- V'^{E}_{ij} – the mean value of the valence window after receiving

Each of these values forms a corresponding linguistic variable. For the valence scores, a basic fuzzy set is introduced, consisting of 5 fuzzy variables with a reasoning range from -1 to 1, and triangular membership functions:

- Strong Negative. (a = -1, b = -1, c = -0.6)
- Moderate Negative. (a = -0.75, b = -0.5, c = -0.15)
- Neutral. (a = -0.2, b = 0, c = 0.2)
- Moderate Positive. (a = 0.15, b = 0.5, c = 0.75)
- Strong Positive. (a = 0.6, b = 1, c = 1)

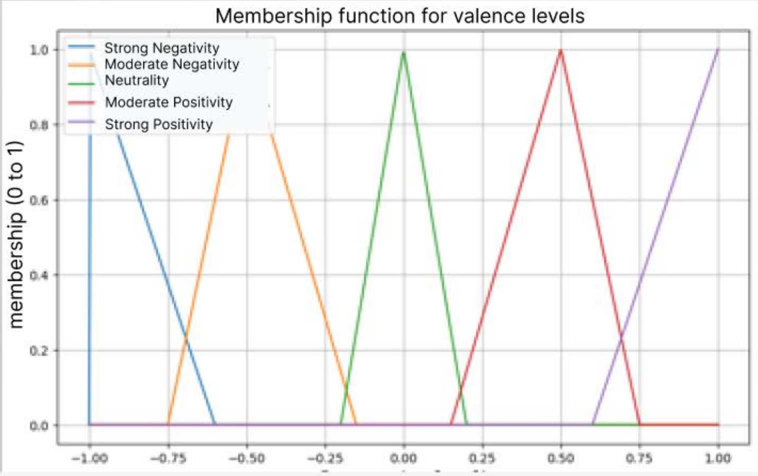
For excitation scores, we introduce a basic fuzzy set consisting of 3 fuzzy variables with a reasoning range from -1 to 1 and triangular membership functions:

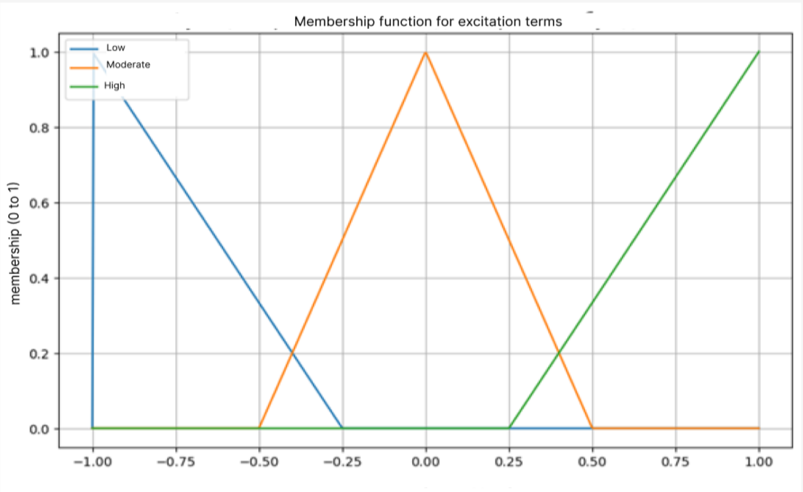
- High (a = -1, b = -1, c = -0.25)
- Moderate (a = -0.5, b = 0, c = 0.5)
- Low (a = 0.25, b = 1, c = 1)

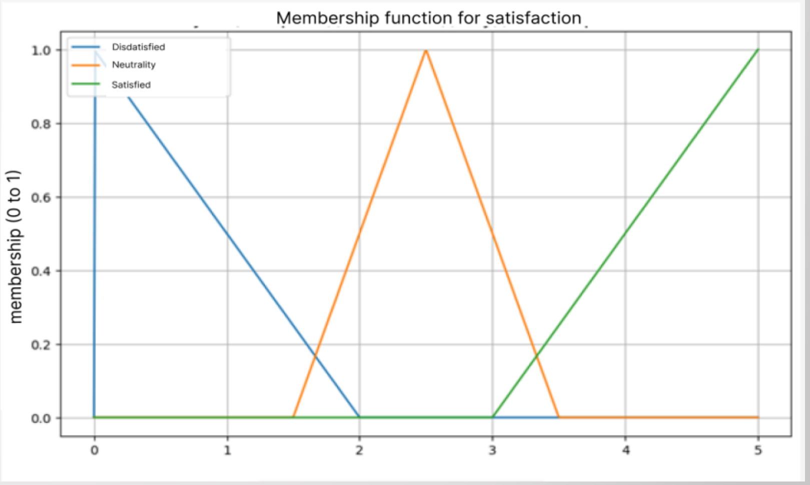
Based on the described linguistic variables, the patient satisfaction score needs to be calculated. To do this, we introduce a linguistic variable with the following set of terms (the reasoning range from 0 to 5):

- Dissatisfied (a = 0, b = 0, c = 2)
- Neutral (a = 1.5, b = 2.5, c = 3.5)
- Satisfied (a = 3, b = 5, c = 5)

The Mamdani algorithm is used for fuzzy inference. The defuzzification function is the center-of-gravity method.





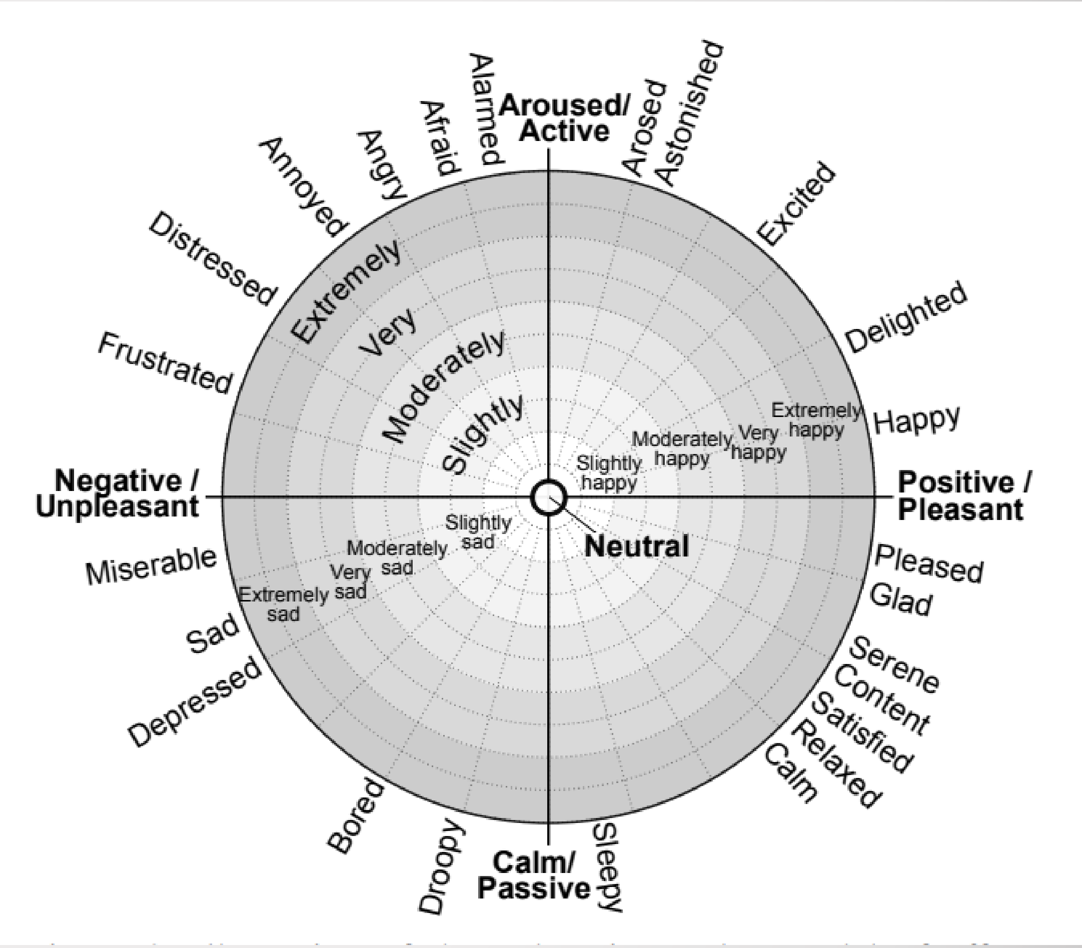


Patient Satisfaction Assessment (3/5)

Rule Construction Logic

Valence:

- Indicates the emotional quality (positive or negative).
- Improvement in valence (from negative to positive) leads to increased satisfaction.
- Deterioration in valence (from positive to negative) decreases satisfaction.
- Has higher priority. Improvement or worsening of valence is a key factor.



Excitement:

- Indicates the level of activity or excitement.
- Moderate values usually have a positive effect on satisfaction.
- Excitement levels that are too low or too high can be neutral or negative depending on the context. For example, very low or very high arousal levels may reduce satisfaction if valence does not compensate for these changes.

Patient satisfaction assessment (4/5)

Example of Inference Rules

Valence (Before)	Valence (After)	Excitement (before)	Excitement (after)	Satisfaction
Moderate Positivity	Strong Positivity	Moderate	Moderate	Satisfied
Improvement in valence and stable excitement indicate a positive emotional effect.				
Moderate Negativity	Neutrality	High	Moderate	Satisfied
Improvement in valence and balanced excitement indicate a positive emotional effect.				
Strongly Negative	Moderate Positivity	Low	High	Satisfied
Strong improvement in valence and activation of excitement indicate a positive emotional effect				
Neutrality	Neutrality	Moderate	Moderate	Neutrality
A complete lack of change in emotional state; no conclusion can be drawn about patient satisfaction.				
Strong Positivity	Moderate Positivity	Moderate	High	Neutrality
A slight deterioration in valence is compensated by an increase in excitement, the satisfaction result is unclear.				
Moderate Negativity	Strong Negativity	Moderate	Moderate	Dissatisfied
A strong deterioration in valence and unchanged excitement indicate a negative effect.				
Strong Negativity	Strong Negativity	Low	High	Dissatisfied
A strong deterioration in valence and unchanged excitement indicate a negative effect.				
Moderate Negativity	Strong Positivity	Moderate	High	Dissatisfied
Improvement in valence and excessive excitement indicate a negative effect.				

Patient Satisfaction Assessment (5/5)

Adaptation in Progress

- No training data.
- Initial rules are based on expert knowledge.

=> There is a need to adapt the model as it operates and accumulates a knowledge base.

- **In the system, patients can self-assess their satisfaction (providing feedback). However, this is optional.**
- **Patient ratings are entered into the knowledge base. For these patients, there is both the actual patient rating and the system-calculated rating.**

A genetic algorithm is used to supplement the system with new rules (for cases not covered by existing rules) and to refine existing rules.

- Each existing rule and membership function is encoded as chromosomes. The initial population consists of the existing set of rules.
- The algorithm's objective function minimizes the difference between predicted and actual patient satisfaction.
- The algorithm is launched when N (an adjustable parameter) new records are added to the knowledge base.
- The resulting new rules are compared with the existing ones for accuracy on a validation set. If the new model improves, it replaces the old one.

Recommendations

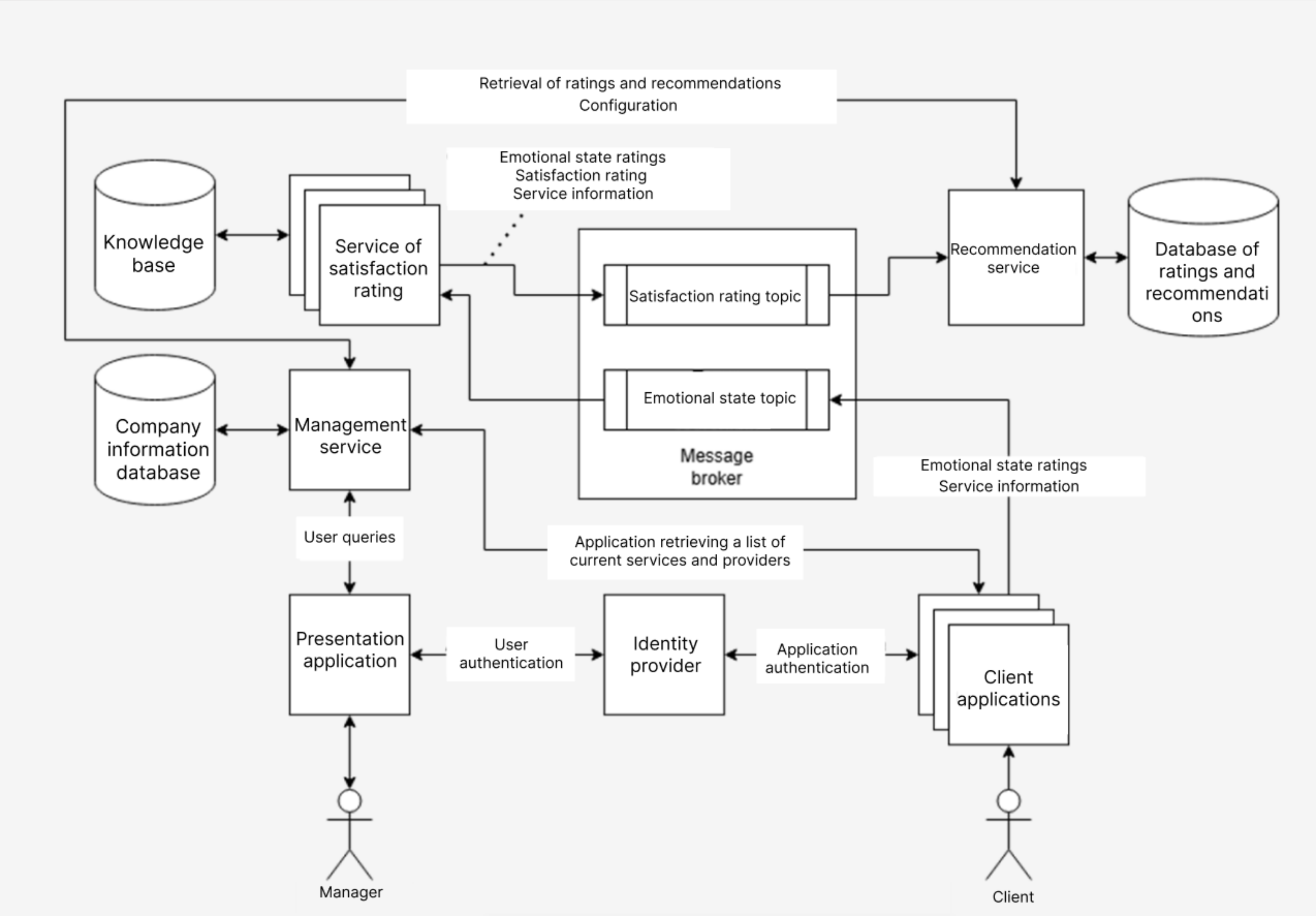
- Recommendations are based on accumulated patient satisfaction data, as well as their emotional characteristics.
- To minimize errors, recommendations are based on time periods consisting of N (configurable) days.
- Recommendations are implemented as notifications sent to decision makers.
- Each alert contains information about a negative change in indicators, as well as a description of possible causes and recommended actions.

Example notifications:

- <u>Notification:</u> Over the past N days, X% of patients were dissatisfied with the service. This exceeds the acceptable threshold of K%.
  Recommendation: Conduct an internal investigation to identify the causes of dissatisfaction.
- <u>Notification:</u> Over the past N days, X% of patients exhibited high arousal levels before the service, and Y% remained highly aroused after the service. This may indicate stress factors during the service process.
  Recommendation: Check the waiting area: eliminate irritating factors, analyze waiting times before service, and assess staff responsiveness.
- <u>Notification:</u> Over the past N days, the percentage of dissatisfied patients increased by X% compared to the previous period.
  Recommendation: Review changes in the system or staff performance over the past week.
- <u>Notification:</u> Over the past N days, Y% of patients remained satisfied with the service. This exceeds the expected threshold.
  Recommendation: Analyze successful cases to identify practices that can be implemented across general operations.
- <u>Notification:</u> Over the past N × K days, the level of dissatisfaction has remained at X% without significant changes.
  Recommendation: Conduct an in-depth analysis of the reasons for stagnation and consider involving external experts to evaluate processes.

Software Design

Architecture and Interaction Scheme



- System Modularity, facilitating maintenance and further development.
- Scalability through separation of system components and use of asynchronous inter-process communication (IPC).
- Flexibility and adaptability due to the loose coupling of system elements..
- Historical data storage enabled by using a message broker..

Results

- An analysis of the domains of service quality assessment and patient satisfaction evaluation was conducted, as well as an analysis of the field of emotion recognition and its potential integration with patient satisfaction assessment. During the analysis, problems and limitations of existing approaches and solutions were identified, highlighting the relevance of research and the development of a new approach to patient satisfaction evaluation. Based on the analysis, a new concept for assessing patient satisfaction was proposed.
- An emotion recognition system was developed that provides multidimensional characteristics of emotions (valence and arousal) based on human facial images. The system uses a computer vision and facial landmark-based approach, employing a small number of features, which increases its processing speed. Additionally, the system utilizes a frontalization algorithm, making it robust to various head rotations and tilts during recognition.
- A model and method for evaluating patient satisfaction based on the analysis of their emotional states were developed. The model operates using a fuzzy logic system and can adapt to new data over time through the use of a genetic algorithm.
- A decision support system was developed that leverages information on patient satisfaction as well as their emotional states, allowing the system to provide timely recommendations to decision-makers regarding service quality management processes.
- A prototype of the decision support system based on the proposed methods and models was implemented. It automates the patient satisfaction evaluation process, eliminating the need for direct patient involvement, thereby increasing the sample size of participating patients and reducing company expenses.

Thank you for your attention!