SentiCore

An emotionally intelligent AI that listens, learns, and responds with real empathy

Guide Name
Ms. Philo Sumi
Assistant Professor
Department of AIDS

Team 11 (Batch 1) :
Marc George (MUT22CA043)
Merlin Sarah Jiju (MUT22CA045)
Timon K. John (MUT22CA068)

Contents

Problem Statement

Problem

Traditional emotion-aware AI systems merely detect human affect from text, audio, or vision, but they do not internally simulate or evolve emotional states over time. They lack temporal memory, contextual empathy modeling, and generate only reactive outputs, resulting in one-dimensional, non-adaptive interactions. This limitation prevents such systems from maintaining a coherent emotional "persona" that can learn from past interactions and adjust its responses in future conversations.

Solution

SentiCore, An emotionally intelligent AI that listens, learns, and responds with real empathy

Objective

- The objective of this project is to build an emotionally intelligent AI system that doesn't just recognize emotions ,it actually feels and evolves its emotional state over time, like a human would.
- The system integrates brain-inspired models called Spiking Neural Networks (SNNs), processes speech tone, facial expressions, and text sentiment to understand how a person feels.
- It then uses that understanding to generate emotionally appropriate, empathetic responses — not just one-time replies, but ones that take into account the past conversation and adjust over time.
- To create an emotionally intelligent agent capable of supporting applications such as digital companions, mental health assistants, and adaptive tutors.

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
F. Ye, H. Gao, L. Yang, M. Li, J. Wu, and D. Liu, "EESCN: A novel spiking neural network method for EEG-based emotion recognition", (2024)	The paper Proposes EESCN, a hybrid CNN-SNN model that converts EEG data into neuromorphic 2D frames for emotion classification. Tested on DEAP and SEED-IV datasets.	EEG to neuromorphic frames, CNN + Spiking Neurons, NeuroSpiking framework, Parameter optimization .	High accuracy, 3× faster inference, captures global and local features, low memory use.	Limited to two datasets, Lacks testing on unseen subjects, No deep comparison with newer LLMs.
Rebecca Mobbs, Dimitrios Makris, Vasileios Argyriou :"Emotion Recognition and Generation: A Comprehensive Review of Face, Speech, and Text Modalities",(2025)	Survey on emotion recognition and generation across face, speech, and text; covers preprocessing, datasets, metrics, and future trends.	CNNs, LSTM, BiLSTM, Transformers, BERT, GANs, Diffusion, CLIP, Attention, PromptVC, Style Encoders, LLMs (GPT-4, LLaMA), PPLM.	Covers recognition and generation; multimodal; state-of-the-art models; includes evaluation and ethics	No standard benchmarks; high compute demand; misuse risk; emotion control unresolved.

Table: Literature Survey

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
Mohammed Kayed, Rebeca P. Díaz-Redondo, Alhassan Mabrouk, "Deep Learning-based Sentiment Classification: A Comparative Survey," (2021)	Comparative review of deep learning models for sentiment classification across text, image, and multimodal data; categorizes models (DNN, CNN, RNN, attention-based) and discusses performance, datasets, and trends.	DNN, CNN, RNN, LSTM, GRU, Attention, Transformers (BERT, RoBERTa), Multimodal Fusion, Hybrid Models.	Broad coverage of architectures; cross-domain comparison; benchmark datasets; highlights trends and applications.	High computational cost; limited interpretability; depends on large labeled data; weak on sarcasm, negation, domain shifts.
"Yuling Luo, Qiang Fu, Juntao Xie, Yunbai Qin, Guopei Wu, Junxiu Liu, Frank Jiang, Yi Cao, Xuemei Ding, "EEG-Based Emotion Classification Using Spiking Neural Networks," (2020)	Proposes SNN-based framework for EEG emotion classification using DEAP dataset; encodes EEG into spike trains and applies convolutional SNN for valence-arousal classification.	EEG processing, DEAP dataset, SNNs, Poisson encoding, Convolutional SNN, STDP learning.	Energy-efficient; biologically plausible; robust to noisy EEG; suitable for low-power edge devices.	Complex preprocessing; immature SNN training; sensitive to parameters; limited scalability for high-resolution data.

Table: Literature Survey

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
A. Sam , R. Boostani , S. Hashempour , M. Taghavi, and S. Sanei, "Depression Identification Using EEG Signals via a Hybrid of LSTM and Spiking Neural Networks," (2023)	Proposes LSTM–SNN hybrid model for EEG-based depression detection using MODMA dataset; EEG encoded to spikes via Poisson encoding; LSTM extracts temporal features, SNN handles energy-efficient classification.	EEG processing, MODMA dataset, LSTM, SNN, Poisson encoding, Convolutional SNN.	Combines LSTM's temporal learning with SNN's efficiency; higher accuracy for depression detection; suitable for wearable/portable use.	High training complexity; difficult LSTM–SNN synchronization; less robust to noise and subject variability.
R. K. Chunduri and D. G. Perera, "Neuromorphic Sentiment Analysis Using Spiking Neural Networks," (2023)	Sentiment analysis via ANN-to-SNN conversion on SpiNNaker hardware, achieving 100% accuracy on IMDB with reduced energy use.	SNNs, ANN-to-SNN conversion, SpiNNaker hardware, IMDB dataset	Energy efficient, Real-time, Outperforms ANN, Novel approach	Limited dataset, Complex SNN training, Small-scale testing, Hardware-specific.

Table: Literature Survey

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
G. Udahemuka, K. Djouani, and A. M. Kurien, "Multimodal Emotion Recognition Using Visual, Vocal and Physiological Signals: A Review," (2024)	Reviews multimodal emotion recognition combining visual, vocal, and physiological signals, highlighting deep learning and fusion methods.	CNN, LSTM, 3D-CNN, EEG, ECG, EMG Visual and speech signals, Fusion techniques	High accuracy,captures micro-expressions,useful in HCI, healthcare, security	Limited datasets, noise sensitivity,cultural/context bias,poor generalization
J. Seekings, P. Chandarana, M. Ardakani, M. R. Mohammadi, and R. Zand, "Towards Efficient Deployment of Hybrid SNNs on Neuromorphic and Edge AI Hardware," (2024)	Proposes hybrid SNN-ANN models trained with PyTorch and Lava, deployed on Intel Loihi (SNN) and Jetson Nano (ANN). Shows reduced energy and latency with few spiking layers while maintaining accuracy.	Hybrid SNN-ANN,PyTorch and Lava,Intel Loihi,NVIDIA Jetson Nano,DVS Gesture dataset.	Lower energy use, Unified training, Real hardware deployment, Balances accuracy and efficiency.	Communication cost not measured, Hardware-dependent, Accuracy drops with more spiking layers, Limited dataset.

Table: Literature Survey

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
A. Bittar and P. N. Garner, "Surrogate Gradient Spiking Neural Networks as Encoders for Large Vocabulary Continuous Speech Recognition," (2023)	Tests surrogate gradient SNNs as replacements for LSTMs in speech recognition. On TIMIT and LibriSpeech, SNNs maintain near-LSTM accuracy with fewer parameters and avoid exploding gradients.	Surrogate gradient SNNs, LIF neuron model, PyTorch, SpeechBrain, TIMIT dataset, LibriSpeech dataset.	Energy-efficient, Robust to exploding gradients, Comparable to LSTM on LVCSR, Fewer parameters.	Slight accuracy drop vs. LSTM, Needs hybrid models for best performance, Not state-of-the-art, High training cost.
W. Fang et al., "SpikingJelly: An Open-Source Machine Learning Infrastructure Platform for Spike-Based Intelligence," (2023)	Presents SpikingJelly, a PyTorch-based framework for SNNs with APIs, dataset support, CUDA acceleration, and neuromorphic hardware deployment; achieves up to $11\times$ speedup.	SpikingJelly framework, PyTorch backend, CUDA acceleration, Neuromorphic datasets (N-MNIST, DVS Gesture, CIFAR10-DVS), Intel Loihi, Lynxi chips.	High simulation efficiency, Full-stack support, Easy PyTorch integration, Widely adopted in SNN research.	Requires PyTorch knowledge, Still maturing vs. ANN frameworks, High GPU demand, Limited hardware support.

Table: Literature Survey

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
Ma, H., Zhang, B., Xu, B., Wang, J., Lin, H., Sun, X. "Empathy Level Alignment via Reinforcement Learning for Empathetic Response Generation," (2024)	Introduces EmpRL, an RL-based framework for empathetic dialogue generation. Uses pre-trained T5 as generator, fine-tuned with PPO and an empathy reward function (emotional reaction, interpretation, exploration) to align generated and target empathy levels.	Reinforcement Learning (PPO), T5 model, Pre-trained empathy identifiers, Empathy reward function.	Enhances response quality, Aligns empathy levels with human responses, Covers affective and cognitive empathy dimensions.	Requires pre-trained empathy identifiers, High computational cost, Dependent on reward function quality.
Yamazaki, K., Vo-Ho, V.-K., Bulsara, D., Le, N., "Spiking Neural Networks and Their Applications: A Review" (2022)	Comprehensive survey of SNNs: neuron/synapse models, training (surrogates, STDP, conversion), frameworks, and applications in vision/robotics with neuromorphic hardware perspectives.	HH, LIF, Izhikevich, AdEx; STDP/R-STDP, SpikeProp, SuperSpike, SLAYER; ANN-to-SNN; encoding (rate/temporal); toolkits (Nengo, Brian2, SNN Toolbox); Loihi, SpiNNaker.	Broad, structured coverage; links biology to ML; practical guidance on training and deployment; highlights energy efficiency and temporal processing.	Training remains hard (non-differentiability, latency-timestep trade-offs); limited large-scale benchmarks vs. ANNs; hardware/software fragmentation.

Table: Literature Survey

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
Martinez-Navarro, J.A. et al., "Comparison of Neural Networks for Emotion Detection" (2023)	Compares SNN, CNN, and MLP for audio emotion recognition on EmoDB, SAVEE, RAVDESS; CNN achieves highest accuracy, MLP moderate, SNN lowest due to lack of encoder and training complexity.	Mel-spectrogram features; CNN with conv/pooling/BN and dense layers; MLP with 8 hidden layers; SNN with Izhikevich neurons and DE/PSO/COSA optimization; datasets: EmoDB, SAVEE, RAVDESS.	Clear benchmarking across models; shows CNN feature-extraction strength; highlights SNN hardware potential and compactness.	SNN accuracy/training time limitations without encoder; small, imbalanced datasets; high data need for CNN; limited SNN tooling.
Alzhrani, W., Doborjeh, M., Doborjeh, Z., Kasabov, N., "Emotion Recognition and Understanding Using EEG Data in a Brain-Inspired SNN Architecture" (2021)	Uses NeuCube BI-SNN with STDP + deSNN to classify EEG-based emotions; achieves 94.83 percent on four emotions and 83.5 percent positive vs. negative, outperforming MLP/RBF/MLR; provides brain-region interpretability.	NeuCube 3D SNN reservoir; TBR spike encoding; SW connectivity; STDP unsupervised learning; deSNN supervised classifier; DREAMER EEG (14 channels, 23 subjects).	High accuracy with EEG; interpretable neural connectivity and regional activation patterns; end-to-end spatio-temporal modelling.	Requires parameter grid search; dataset size limited; specialized framework dependency.

Table: Literature Survey

Existing System and Literature Survey

Title	Summary	Technology Used	Advantages	Disadvantages
Wang, J., "A Review of Spiking Neural Networks" (2022)	Intro review on neuron models, coding, learning, and neuromorphic platforms; contrasts bio-mimicry vs. compute cost; advocates mixed coding and unsupervised learning.	Models (HH, LIF, Izhikevich); coding (rate, temporal, population, AER); learning (Hebb, STDP, SpikeProp, ReSuMe, ANN-to-SNN); platforms (Neurogrid, DYNAPS, TrueNorth, BrainScaleS, SpiNNaker, Loihi).	Clear comparison of models/coding; practical platform overview; emphasizes energy-efficient, event-driven computing.	High-level; limited experimental results; notes BP weight-transport issue and STDP limits for deep multilayer training.

Table: Literature Survey

Proposed System

- The proposed system converts inputs like tone of voice and the emotional meaning of text into synchronized spike patterns using a neuromorphic-inspired encoding approach and feature extraction.
- A spiking neural core then integrates these spike patterns over time, simulating a continuously evolving emotional state.
- An affective appraisal module compares incoming patterns with learned emotional examples to decide whether to reflect the user's feelings or respond with supportive guidance.
- Finally, the current state of the spiking network, along with the appraisal result, guides a language model to generate responses that are both meaningful and emotionally appropriate.
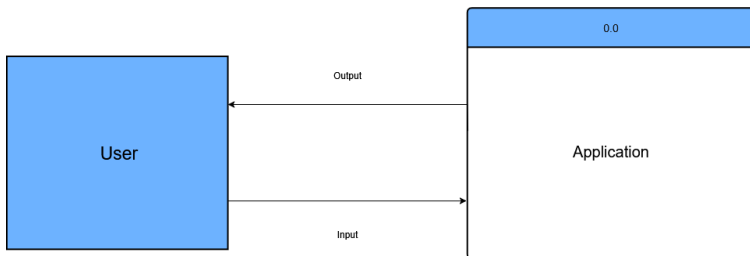
Proposed System

A. Data Flow Diagrams

Fig. 1: Level 0 Data Flow Diagram

Proposed System
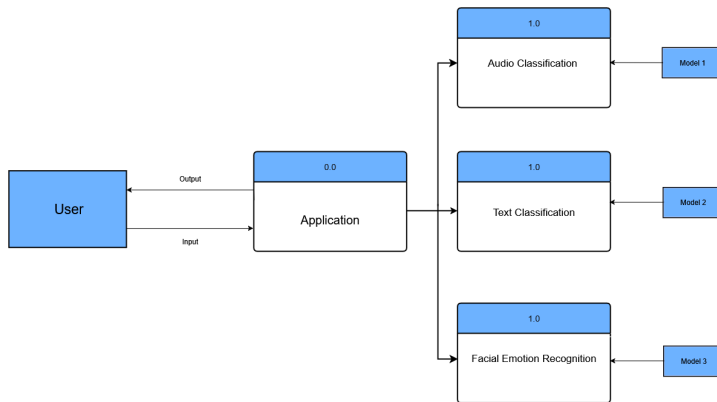
A. Data Flow Diagrams

Fig. 1: Level 1 Data Flow Diagram

Modules

- **Data Preprocessing Module**
  - Cleans and normalizes raw text data and converts audio signals into MFCC (Mel-Frequency Cepstral Coefficients).
- **Spike Generation Module**
  - Transforms preprocessed text and audio features into spike trains for SNN input.
- **Audio Classification Module**
  - Classifies audio-derived spike patterns into emotion categories using SNNs and detects emotion states.
- **Text Classification Module**
  - Processes cleaned text data through spiking/ANN hybrid layers and classifies emotional sentiment expressed in textual input.
- **User Interface (UI) Module**
  - Web-based interface (React/Node.js) enabling real-time user interaction with the system.

Advantages of the Proposed System

- **Multimodal Input**
  - Includes speech tone and text sentiment into the model for richer emotion understanding.
- **Neuromorphic Emotion Reasoning**
  - Helps the model understand, feel, and reason about human emotions using brain-like spikes and signals that mimic how we recognize emotions.
- **Evolving Emotional Memory**
  - Simulates a continuously changing mood that builds on past interactions for more coherent replies.
- **Adaptive Empathy**
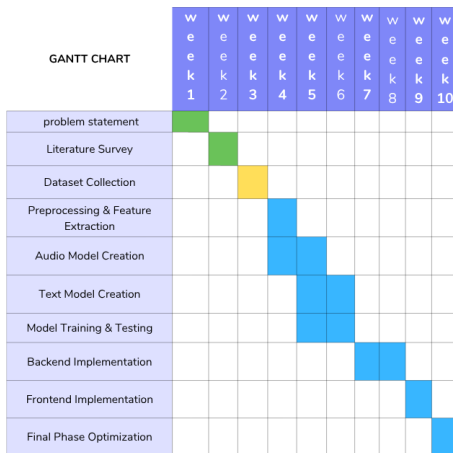 - Learns when to mirror your feelings or offer comfort by adjusting its behavior based on feedback.
- **Emotion-Driven Responses**
  - Uses its internal mood state to guide the chatbot so responses feel consistent and empathetic.

Software Requirements

- **Python** — Core language for SNN modeling, data processing, and API services.
- **PyTorch** — Deep learning framework powering surrogate-gradient training of spiking networks.
- **SpikingJelly/SNNTorch** — High-performance library for building and training LIF-based SNN layers on GPU.
- **openSMILE** — Toolkit for extracting audio emotion features from speech.
- **Hugging Face Transformers** — Pretrained GPT-style models for generative, tone-conditioned text output.
- **LangChain** — Orchestrates multimodal pipelines, chaining SNN state vectors into LLM prompts.
- **BindsNET** — Reinforcement-learning library for shaping empathy behaviors via reward feedback.

Project Planning

GANTT CHART	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
problem statement	█									
Literature Survey		█								
Dataset Collection			█							
Preprocessing & Feature Extraction				█						
Audio Model Creation				█	█					
Text Model Creation					█	█				
Model Training & Testing					█	█				
Backend Implementation							█	█		
Frontend Implementation									█	
Final Phase Optimization										█

Task Allocation

- **Marc George** – Working on the **Audio Emotion Recognition pipeline**, including MFCC feature extraction, spike generation, and SNN-based classification; also contributes to the literature survey.
- **Merlin Sarah Jiju** – Handling **Data Analysis & Preprocessing**, lead **Frontend** (React/Node.js), and prepare **documentation** alongside literature survey contributions.
- **Timon K. John** – Developing the **Text Emotion Recognition module** (text preprocessing + hybrid classification) and implement the **Backend Services** for system integration.

Conclusion

- SentiCore introduces a neuromorphic, spike-based core that includes speech and text into a single emotional stream.
- Its SNN model yields a self-evolving "mood," enabling more coherent, persona-driven dialogue.
- An RL-trained appraisal module dynamically chooses between empathetic mirroring and supportive guidance.
- Conditioning a GPT-style generator on the SNN's state produces responses that are both contextually relevant and emotionally resonant.
- SentiCore offers an explainable, adaptive framework for digital companions, tutors, and mental-health assistants.

References

- F. Ye, H. Gao, L. Yang, M. Li, J. Wu, and D. Liu,"EESCN: A novel spiking neural network method for EEG-based emotion recognition" ,2024.

- Rebecca Mobbs, Dimitrios Makris, Vasileios Argyriou,"Emotion Recognition and Generation: A Comprehensive Review of Face, Speech, and Text Modalities," 2025.

- Mohammed Kayed, Rebeca P. Díaz-Redondo, Alhassan Mabrouk,"Deep Learning-based Sentiment Classification: A Comparative Survey" , 2021.

- Yuling Luo, Qiang Fu, Juntao Xie, Yunbai Qin, Guopei Wu, Junxiu Liu, Frank Jiang, Yi Cao, Xuemei Ding,"EEG-Based Emotion Classification Using Spiking Neural Networks", 2020.

- A. Sam , R. Boostani , S. Hashempour , M. Taghavi, and S. Sanei,"Depression Identification Using EEG Signals via a Hybrid of LSTM andSpiking Neural Networks", 2023.

References (contd.)

- R. K. Chunduri and D. G. Perera, "Neuromorphic Sentiment Analysis Using Spiking Neural Networks," 2023.
- G. Udahemuka, K. Djouani, and A. M. Kurien, "Multimodal Emotion Recognition Using Visual, Vocal and Physiological Signals: A Review," 2024.
- J. Seekings, P. Chandarana, M. Ardakani, M. R. Mohammadi, and R. Zand, "Towards Efficient Deployment of Hybrid SNNs on Neuromorphic and Edge AI Hardware," 2024.
- A. Bittar and P. N. Garner, "Surrogate Gradient Spiking Neural Networks as Encoders for Large Vocabulary Continuous Speech Recognition," 2023.
- W. Fang et al., "SpikingJelly: An Open-Source Machine Learning Infrastructure Platform for Spike-Based Intelligence," 2023.

References (contd.)

- H. Ma, B. Zhang, B. Xu, J. Wang, H. Lin, and X. Sun, "Empathy Level Alignment via Reinforcement Learning for Empathetic Response Generation," 2024.
- K. Yamazaki, V.-K. Vo-Ho, D. Bulsara, and N. Le, "Spiking Neural Networks and Their Applications: A Review," 2022.
- J. A. Martinez-Navarro et al., "Comparison of Neural Networks for Emotion Detection," 2023.
- W. Alzhrani, M. Doborjeh, Z. Doborjeh, and N. Kasabov, "Emotion Recognition and Understanding Using EEG Data in a Brain-Inspired SNN Architecture," 2021.
- J. Wang, "A Review of Spiking Neural Networks," 2022.

Thank You!