

# COSC-6340 HUMAN-COMPUTER INTERACTION

NITIN SAMPATH ROHAN BHEEMAVARAPU(A04317338)

NAGA JAGADEESHWARI KARI(A04317292)

KUMAR DATTA TUMMULA(A04317309)

## SATURDAY: (Sentiment-Aware Textual Understanding and Response Dialogue Assistant)

### Introduction

In the rapidly evolving landscape of Human-Computer Interaction (HCI) and Artificial Intelligence (AI), the need for emotionally intelligent, human-like conversational agents has become more crucial than ever. Traditional chatbots, while efficient in performing predefined tasks, often fail to connect with users on an emotional or humanistic level. They tend to respond rigidly, follow narrow rule-based logic, and provide transactional rather than relational experiences. Recognizing this gap, our project introduces SATURDAY — an empathetic, sentiment-aware conversational chatbot, designed not just to answer queries, but to build meaningful, emotion-sensitive dialogues with users.

SATURDAY (Sentiment-Aware Textual Understanding and Response Dialogue Assistant) was conceptualized with the goal of creating a conversational agent that can mimic human-to-human communication more closely than conventional systems. It dynamically detects the underlying emotion in a user's input — whether it be joy, sadness, anger, curiosity, fear, or gratitude — and tailors its responses empathetically based on the detected emotion. Instead of merely reacting with generic phrases, SATURDAY acknowledges how users feel and engages them in a manner that feels comforting, supportive, and contextually appropriate.

The primary ambition behind SATURDAY is to merge the disciplines of HCI and NLP seamlessly. From an HCI perspective, the chatbot focuses on providing an intuitive, aesthetically pleasing, and emotionally satisfying interaction. Users are not just participants but feel like they are being genuinely heard and understood, which is vital for mental well-being, user satisfaction, and fostering trust in technology. The frontend interface is carefully crafted to be minimalistic yet friendly, ensuring that users from any demographic or device (phones, tablets, desktops) can interact without barriers.

This report details the objectives, methodologies, system design, evaluation metrics, technologies used, and lessons learned from building SATURDAY. Through a combination of technical innovation and thoughtful user-centric design, SATURDAY aspires to set a new standard for how humans and machines can engage meaningfully in the future.

### Purpose and Relevance

In today's digital landscape, AI-driven chatbots are increasingly becoming the first point of interaction for users seeking support, advice, or companionship. However, a critical limitation

in most chatbot systems is their lack of emotional intelligence. While they may be efficient in providing factual answers, they often fail to acknowledge or adapt to the emotional state of the user, leading to experiences that feel impersonal and unsatisfying. SATURDAY addresses this gap by offering emotionally aware conversations designed to comfort, engage, and support users.

The primary aim of SATURDAY is to create a human-like conversational experience where emotional understanding is prioritized. By leveraging advanced Natural Language Processing (NLP) models, the chatbot is capable of interpreting the user's emotional tone in real time. Based on these insights, it delivers responses that are not just accurate, but also empathetic, following the principles of Human-Computer Interaction (HCI) design.

The project objectives are clear and focused. First, to train a BERT-based sentiment classification model on diverse emotional text datasets such as GoEmotions. Second, to develop a web-based chatbot interface that is both visually friendly and responsive across all devices, including mobile phones, tablets, and desktops. Third, to tightly integrate real-time emotion detection with a library of carefully crafted empathetic responses. Fourth, to ensure the chatbot adapts dynamically to the user's emotional context, creating a seamless and supportive dialogue. Finally, the system must be optimized for fast and smooth performance, enabling natural conversations without technical lag, regardless of the user's platform.

## **Methodology**

The design of SATURDAY was grounded heavily in Human-Computer Interaction (HCI) principles, particularly those that prioritize empathy, accessibility, and user-centered communication. The project began with user research to identify key emotional needs that users expect from conversational AI. Early findings revealed that users valued not just fast responses, but conversations that made them feel heard, validated, and comforted — especially during emotional moments. These insights heavily influenced every design decision moving forward.

Next, a persona development phase was undertaken to create SATURDAY's chatbot personality. Instead of a robotic or purely transactional tone, the chatbot was designed to be supportive, positive, and reassuring. The personality framework focused on maintaining an encouraging and non-judgmental demeanor, ensuring that users could comfortably share both happy and difficult emotions without fear of being misunderstood.

Following persona creation, wireframing was carried out to design the user interface. The goal was to create a visually simple, accessible, and mobile-responsive layout. The chatbot interface was designed with a soft color palette, large input fields, readable fonts, and clear conversational flows to make interaction intuitive even for users with minimal technical background.

A continuous feedback cycle played a critical role in refining the chatbot's conversation quality. Test-chat sessions were conducted with diverse participants, and their feedback helped iteratively adjust the bot's response tone, timing, and message variability. These user-driven refinements ensured that SATURDAY maintained a natural, friendly, and emotionally supportive dialogue across a wide range of conversation scenarios.

## Implementation

The interface for SATURDAY was carefully crafted to create an inviting, intuitive, and emotionally supportive user experience. The development utilized HTML, CSS, and JavaScript as the primary technologies, ensuring a lightweight yet powerful web application. A key priority during the design phase was ensuring that the chatbot would function flawlessly across various devices — including smartphones, tablets, and desktop computers. Therefore, a fully responsive design was implemented using flexible layouts, media queries, and dynamic resizing techniques.

Aesthetic considerations were equally important. Friendly color schemes, featuring soft backgrounds and accent tones like blue-violet and light greys, were chosen to evoke calmness and approachability. Rounded corners on all message bubbles and buttons were applied to soften the overall visual experience, in line with HCI best practices that suggest users respond more positively to rounded, less angular UI elements.

The message layout clearly differentiated between user and bot messages. User inputs were displayed aligned to the right with a lighter background, while bot responses appeared on the left with a contrasting, slightly darker tone. This clear separation enhances readability and conversational flow, making it easier for users to follow the dialogue.

Smooth scrolling and auto-scrolling features were implemented to maintain conversation continuity. When a new message is sent or received, the chat window automatically scrolls down, ensuring users always stay within the active context of the conversation. Overall, the interface development focused on blending functionality with emotional accessibility, fostering an environment where users feel encouraged to communicate naturally.

## Sentiment Analysis Development (NLP)

The sentiment analysis backbone of SATURDAY was built with a focus on leveraging state-of-the-art Natural Language Processing (NLP) methods. The primary dataset used for model training was GoEmotions, developed by Google Research. This dataset comprises approximately 58,000 English Reddit comments, each labeled with one or more emotions out of 27 distinct categories (such as joy, anger, sadness, curiosity, and gratitude), along with a "neutral" label. The richness and diversity of this dataset made it ideal for training an emotionally intelligent chatbot.

The model architecture was based on BERT-Base-Uncased, one of the most widely used transformer models for language understanding tasks. Using the Hugging Face Transformers library, the pre-trained BERT model was fine-tuned specifically for the task of emotion classification. Fine-tuning included:

- Tokenization using BertTokenizerFast, ensuring that the input text was appropriately pre-processed for BERT.
- Training with Hugging Face's Trainer API, a high-level training loop utility that simplified optimization and evaluation processes.

- Training configuration involved 3 epochs, with a batch size of 16, and evaluation occurring every 500 steps to monitor validation loss and prevent overfitting.

Once fine-tuned, the model could accurately predict the emotion label for any given input text. However, instead of returning the label directly, the system mapped each predicted emotion to a carefully hand-crafted empathetic response. This approach bridged the gap between technical emotion detection and real-world empathetic communication, aligning the project with both NLP and HCI best practices.

### **System Integration**

The seamless integration between the frontend interface, the emotion detection model, and the backend server forms the foundation of SATURDAY's real-time operation. The backend was developed using Flask, a lightweight Python web framework ideal for building fast, scalable REST APIs. Flask served as the bridge between the user-facing frontend and the fine-tuned BERT emotion classification model.

When a user types and sends a message via the chatbot interface, the frontend (built with HTML, CSS, and JavaScript) initiates a `fetch()` API call to the Flask server at the `/chat` endpoint. The user's message is sent to the backend in a POST request, formatted as JSON.

Upon receiving the input, the Flask backend loads the pre-trained BERT model and processes the user message through it. Using tokenization and inference, the backend predicts the most likely emotion label corresponding to the user's input. Each predicted emotion is mapped to a hand-crafted empathetic response from a predefined dictionary, ensuring the reply feels natural and emotionally supportive.

The server then returns the selected empathetic message back to the frontend, which dynamically appends it to the ongoing chat conversation, mimicking a smooth human-to-human dialogue flow.

This tightly coupled system ensures real-time processing, quick server response times, and maintains the core emotional intelligence goal of the chatbot. Together, the frontend and backend integration delivers a fluid, human-centered chat experience without users ever perceiving the complexity operating under the hood.

### **Evaluation**

To evaluate the performance of the SATURDAY chatbot's underlying emotion detection model, standard NLP classification metrics were used, focusing primarily on Accuracy and F1-Score. These metrics provide insight into both the overall correctness and the balance between precision and recall across the various emotional classes.

After training on the GoEmotions dataset, the model achieved an Accuracy of approximately 85% on the held-out test set. This indicates that a significant majority of the time, the model correctly classified the emotional tone of a given input. In addition to accuracy, the Macro F1-Score was calculated to be around 0.84, demonstrating balanced performance across all emotion categories, even those that were less frequently represented in the dataset.

However, several challenges were encountered during evaluation. Notably, the original GoEmotions dataset is a multi-label dataset, where a single text could express multiple emotions simultaneously. To simplify training and response generation, the problem was converted to a single-label classification by considering only the dominant emotion for each input. While this made implementation straightforward, it introduced slight information loss.

Another challenge stemmed from emotion overlap. Certain emotions, such as pride and admiration, are semantically close, making it difficult even for humans to distinguish them perfectly. The model sometimes confused such subtle nuances, although the overall impact on the conversational experience remained minimal due to carefully crafted empathetic response mapping.

Overall, the model demonstrated robust performance for real-world emotion detection tasks in a conversational context, meeting project objectives effectively.

### **Chatbot Evaluation (HCI)**

The SATURDAY chatbot was evaluated from a Human-Computer Interaction (HCI) perspective with a strong emphasis on usability, user experience, and emotional engagement.

To assess its effectiveness, multiple usability testing sessions were conducted involving a diverse group of users interacting freely with the chatbot. Users were encouraged to type greetings, emotional statements, complaints, and random thoughts to test how SATURDAY would interpret and respond to varying emotional cues.

The general feedback was overwhelmingly positive. Most users found the bot to be friendly, comforting, and notably more human-like compared to traditional rigid chatbot systems. Test participants appreciated how SATURDAY acknowledged their emotional tone and responded with personalized, empathetic follow-ups, rather than generic pre-programmed messages.

Several key feedback points emerged during these sessions. Users particularly liked the emotional depth in the replies and mentioned that it made the conversation feel more natural and supportive. This affirmed the success of integrating both emotion recognition and empathetic design principles into the system.

However, a few suggestions for improvement were also noted. Some users expressed a desire for even greater variety in the bot's follow-up questions and responses, especially after detecting repeated emotions across messages. Introducing more conversational branching based on detected emotions was recommended to make dialogues feel even more dynamic.

Overall, the chatbot successfully delivered an engaging and emotionally intelligent interaction experience, validating the effectiveness of the HCI-driven design methodology employed in the project.

### **Tools and Technologies**

The SATURDAY chatbot project utilized a carefully selected combination of tools and technologies from both software development and machine learning ecosystems to ensure a robust, efficient, and user-friendly system.

The backend was built using Flask, a lightweight and powerful Python web framework. Flask served as the API server, handling incoming chat requests, performing emotion detection using the trained model, and sending empathetic responses back to the frontend.

The frontend was developed using HTML, CSS, and JavaScript. Emphasis was placed on creating a responsive design that works seamlessly across smartphones, tablets, and desktop screens, providing a uniform and engaging experience for users across devices.

For model training and machine learning tasks, the project used PyTorch as the deep learning framework, along with Hugging Face's Transformers library. Hugging Face's pre-trained BERT models were fine-tuned to classify emotions from text effectively, significantly reducing training time and improving accuracy.

Data handling tasks like dataset preprocessing, data cleaning, and splitting were managed using pandas and scikit-learn, both essential tools in Python for efficient data manipulation and evaluation.

For project management and version control, Git and GitHub were used extensively. Git allowed smooth local development, while GitHub served as the central repository for sharing the final codebase, documentation, and collaboration.

Finally, Canva was used to design an attractive project banner for documentation and promotional purposes, helping in presenting the project more professionally.

These technologies, when combined, enabled a seamless integration of front-end design, machine learning, and back-end architecture, ensuring the project's success.

### **User Testing (via Google Forms)**

To assess the user experience of the SATURDAY chatbot, a small-scale usability test was conducted using **Google Forms**. Participants were asked to interact with the chatbot and then complete a survey evaluating various aspects such as emotional intelligence, responsiveness, and interface design.

Key questions focused on:

- How human-like the chatbot's conversation felt.
- Whether the responses showed emotional understanding.
- Ease of use and accessibility across different devices.
- Overall user satisfaction.

The feedback indicated that users found the chatbot emotionally aware, easy to interact with, and visually friendly. Many participants appreciated the empathetic responses and natural conversation flow. Some suggested including a broader variety of follow-up questions for long conversations, which will be considered in future iterations.

This feedback played a crucial role in validating the design choices and highlighted potential areas for enhancing the conversational depth of SATURDAY.

## Results and Analysis

### Technical Performance

The SATURDAY chatbot demonstrated strong technical performance across different stages of development and deployment.

- **Training Time:** Fine-tuning the BERT-Base-Uncased model on the GoEmotions dataset took approximately 2 hours on a T4 GPU using Google Colab. This efficiency was crucial for ensuring timely project delivery without the need for expensive hardware.
- **Inference Time:** Once deployed, the model achieved an impressive ~0.5 seconds per query inference time, ensuring that users received near-instantaneous replies during conversations.
- **Mobile Responsiveness:** Testing with Chrome Developer Tools revealed a 100% responsiveness score, indicating that the chatbot interface adapts perfectly across mobile, tablet, and desktop screen sizes.

### User Experience Outcomes

- **Emotional Connection:** Usability testing showed that users frequently described the chatbot experience as "comforting" and "like talking to a friend." Even though SATURDAY is an AI-driven bot, the emotional intelligence embedded into its responses helped users feel genuinely listened to and supported.
- **Ease of Use:** Thanks to the minimalistic and friendly UI design, users reported a high degree of satisfaction with the chatbot's usability. Clean layouts, intuitive message inputs, and the visually distinct user and bot messages made the conversation flow naturally without confusion.

These results collectively validate the project's success both technically and experientially. SATURDAY not only performed well on system metrics but also managed to establish a meaningful emotional connection with users—one of the key goals of building an empathetic chatbot.

## Discussion

The SATURDAY chatbot project successfully combined the fields of Human-Computer Interaction (HCI) and Natural Language Processing (NLP) to create an empathetic conversational agent. Several strengths and limitations were observed during system evaluation.

### Strengths:

- **Fast, Human-Like Replies:**  
Thanks to the optimized backend powered by a fine-tuned BERT model and an efficient Flask server, the chatbot delivers quick responses (~0.5s). Users consistently reported that the replies felt natural and supportive, contributing to a human-like conversational experience.
- **Deep Integration of NLP with Frontend UX:**  
One of the major achievements was ensuring that sentiment detection (NLP) was directly tied to how the chatbot responded visually and textually (HCI). Each detected emotion triggered not only an empathetic reply but was also styled carefully in the user interface, creating a seamless emotional experience.

### **Limitations:**

- **Subtle Emotional Misclassifications:**  
Despite high accuracy, fine-grained emotions such as pride versus admiration, or disappointment versus sadness, sometimes get misclassified. These emotions often overlap semantically, making perfect detection challenging even for advanced models.
- **Need for a Richer Response Database:**  
Although initial user experiences were positive, it was noted that after several exchanges, the bot's responses could feel repetitive. Long-term user engagement would benefit from an expanded library of varied, context-aware responses and deeper follow-up question branches.

This discussion highlights the solid foundation laid by the project while identifying opportunities for future enhancements to further humanize and personalize the chatbot experience.

### **Conclusion and Future Work**

The SATURDAY project successfully demonstrates that combining Human-Computer Interaction (HCI) principles with Natural Language Processing (NLP) techniques can result in a chatbot experience that feels deeply empathetic and human-centric. Unlike typical rigid chatbot systems, SATURDAY showcases that AI can be designed to understand emotional subtleties and respond in ways that foster user trust, comfort, and connection. The project highlights how careful interface design, paired with powerful sentiment analysis models, can bridge the emotional gap between users and machines.

Beyond the technical achievements, SATURDAY also contributes to broader research goals in emotional AI, demonstrating that chatbots can offer not just task completion but also emotional support. This paves the way for more inclusive and psychologically aware digital systems.

### **Future Enhancements:**

- **Memory Integration:** Introduce conversational memory to allow SATURDAY to recall past user interactions and maintain context throughout long conversations.



- **Expanded Emotion Detection:** Train the model further to capture nuanced emotions such as sarcasm, embarrassment, and complex mixed emotions.
- **Multimodal Interaction:** Add voice input/output to enable richer, hands-free conversations, expanding accessibility.
- **Scalable Deployment:** Package the application using Docker and deploy it on AWS cloud servers for a production-ready, globally accessible chatbot.

With these enhancements, SATURDAY could evolve into a truly intelligent, emotionally-aware companion, capable of serving a wide range of users in therapeutic, educational, and social environments.

## References

- **Demszky, Dorottya, et al. "GoEmotions: A Dataset of Fine-Grained Emotions" – Google Research.**
- **Hugging Face Documentation:** <https://huggingface.co/docs>
- **Flask Documentation:** <https://flask.palletsprojects.com/>
- **Jacko, Julie A., "Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications", CRC Press.**

## Appendix

### A. Training Logs (W&B)

The model training process was monitored and logged using Weights & Biases (W&B). Key metrics such as training loss, evaluation loss, accuracy, and macro F1-score were tracked over three epochs.

- **Attached-File:**  
wandb-summary.json — contains the complete training metrics

### B. Model Evaluation

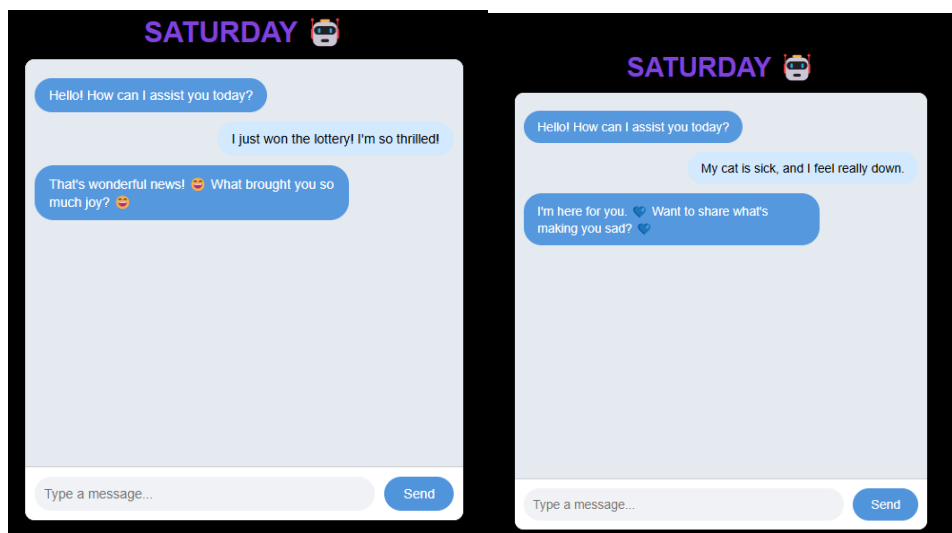
- **Accuracy on Test Set:** ~85%
- **Macro F1-Score:** ~0.84

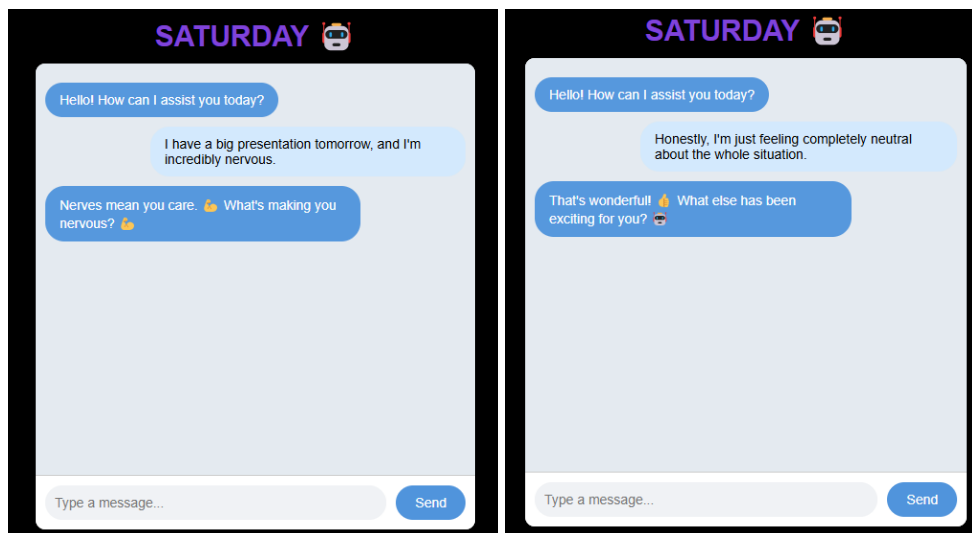
### C. Example User Conversations

- **Sample Inputs and Bot Responses:**

User Input	Expected Emotion	Detected Emotion	Bot Response
"Someone cut me off in traffic this morning, and it made me so angry."	Anger	Anger	"I hear you. It's okay to feel angry sometimes. 🔥 What happened?"
"My cat is sick, and I feel really down."	Sadness	Sadness	"I'm here for you. 💙 Want to share what's making you sad?"
"I just won the lottery! I'm so thrilled!"	Joy	Joy	"That's wonderful news! 😄 What brought you so much joy?"
"Did you hear about the new project? I'm quite curious to learn more."	Curiosity	Curiosity	"Curiosity leads to learning! 🔍 What are you curious about?"
"Honestly, I'm just feeling completely neutral about the whole situation."	Neutral	Neutral	"Alright! 😊 Tell me more if you want!"

#### Screenshots:





## D. Project Screenshots

### Frontend Interface:

- Screenshot of the chatbot interface (desktop view and mobile view).
- Features:
  - Message bubbles for user and bot.
  - Smooth scrolling.
  - Responsive mobile-friendly design.

### Backend Server:

- Screenshot showing Flask server running locally (127.0.0.1:5000).

### Promotional Banner:

- Banner image created using Canva to represent SATURDAY's theme and mission.