**DTI Project – SentiBot(Your Virtual Friend)**

**Milestone1**

**Write Your Name**

Aditya Raj

**Write your BU Email Address**

[E22CSE0649@bennett.edu.in](mailto:E22CSE0649@bennett.edu.in)

**Write your Mobile Number**

7004230852

**Select your team's batch**

22

**Your team number [check LMS for the same]. If your team is not listed then write N/A**

N/A

**Team member names [separated by comma]**

Aditya Raj, Shivam Garg, Sanjana Nayak

**Team member enrolment# [separated by comma]**

e22cseu0649, e22cseu1506, e22cseu0644

**Bennett email ID of Team Members [separated by comma]**

[e22cseu0649@bennett.edu.in](mailto:e22cseu0649@bennett.edu.in), [e22cseu1506@bennett.edu.in](mailto:e22cseu1506@bennett.edu.in), [e22cseu0644@bennett.edu.in](mailto:e22cseu0644@bennett.edu.in)

**Project Title:** SentiBot

**Describe Your Project:**

Our sentiment analysis webapp introduces a novel way for users to engage with the platform using voice and video interactions. When users visit the app, they can communicate by sharing their thoughts and feelings through spoken words, which the app will efficiently convert to text using advanced speech-to-text technology. Simultaneously, the app processes video data to enhance the overall understanding of the user's expressions.

Our model, trained to analyze text and video data, provides users with instant insights into their expressed sentiments such as Happy, Sad, Neutral, Angry, etc. This real-time sentiment analysis adds a new dimension to user interaction.

Looking ahead in future, we plan to integrate generative AI to create a more adaptive and personalized user experience. This future enhancement will enable the app to give responses based on the user's expressed sentiments, offering a unique and empathetic interaction.

**Why Team Has Chosen This Project?**

Our team is passionate about using technology to address contemporary societal needs. In today's fast-paced and technologically advanced environment, we observed a gap in open conversations about personal feelings. Our sentiment analysis web app goal is to fill this gap by giving the users a virtual space to express their thoughts and feelings. We believe that interacting with our app could offer users a sense of relief, providing a non-judgmental platform to share their emotions.

The project aligns seamlessly with our team's interests and skills. As individuals with a specialization in AI-ML, we are enthusiastic about exploring model training and implementing various algorithms. The integration of generative AI into our future aligns with our ambition to create a technologically advanced and adaptive platform. Furthermore, our collective skills, which include proficiency in AI-ML and front-end development, position us well to bring this project to life on a web app.

**What Problems it solve?**

Our sentiment analysis web app addresses the contemporary need for open conversations about personal feelings in a technologically advanced environment. Recognizing the existing gap in expressive outlets, the webapp serves as a space for the users to share their thoughts and feelings freely, providing a non-judgmental platform for emotional expression. The project fulfills the need for a supportive and technologically adaptive solution, offering users a sense of relief through seamless interaction with AI-ML models.

**Who will be the potential clients/customers/users/startups/beneficiaries of the project. Describe appropriately. Why you think they are your target audience?** Be specific so that the prospective target for your project is clear.

The sentiment analysis web app is designed to cater to a diverse user base, including individuals in search of an emotional outlet, tech enthusiasts intrigued by AI-ML innovations, mental health advocates promoting open expression, online community platforms looking for enhanced engagement tools, and startups in the AI-ML space seeking unique applications. Our target audience comprises those who value both emotional well-being and technological innovation, seeking a supportive and technologically advanced platform for expressive interactions.

**What will be the innovation in this project?**

In this project, we are incorporating spoken words (via speech-to-text) and using video data (facial expressions, etc.) for a more comprehensive understanding of sentiment. It makes the communication more effectively. If you speak something it will predict if you are happy, sad, angry, shock. it captures the nuances of your voice, from the tremor of excitement to the sigh of frustration. This allows for a deeper understanding of your sentiment, uncovering hidden emotions that might otherwise remain veiled.

Our tool also focuses on Facial Recognition. It will provide the user with instant insights into their expressed emotions in real time is novel and valuable for self-awareness and communication improvement. Additionally, we are also in plan to include a video response of the user sentiment after the user gave their input.

We will also be including generative AI in future which allows for a truly personalized dialogue between you and the tool. It will act as a system that not only understands your emotions, but also responds in a way that resonates with your unique needs. This personalized interaction could offer guidance, support, or simply a listening ear, fostering a deeper level of connection and understanding.

**How it will be different from similar existing solutions.**

Our project based on sentiment analysis is different from other existing solutions because of its different and unique combination of capabilities and other features. For example, if we compare our project with similar existing solutions such as **Google’s Dialog flow, IBM Watson Tone Analyzer, Microsoft Azure Text Analytics**, our project offers some different advantages.

On first, the existing solutions are mainly text-based sentiment analysis, but our web app integrates both voice and video interactions. This multi-feature approach will allow the users to express themselves more comprehensively, capturing the emotion and facial expressions that get missed in text-only analysis. Also, it will provide the users sentiment in real time with minimum delay.

Additionally, in future our integration of generative AI gives us an edge than others by offering a personalized experience. The software will interact with user based on their sentiment and the talk they are having with software. By this, the user will feel deeper connections and understanding.

**What makes it challenging enough to be chosen to be done in four months (Jan-May)**

The complexity and timeline for our project are very big and challenging. The selection of models, finding datasets, learning domain knowledge (like different NLP techniques, web-app development, etc..) of our project, and consulting with appropriate specialists is very time-consuming.

Our project cannot be done within the scope of 4 months. As our tool will be targeted to the audience having some minor mental health issues like depression, anxiety etc, which is very sensitive and moral concerning matter, so training a model that is ethically tested and moral supported, will require time more than 4 months.

**How you will measure the success/outcome/quality of your project?**

In the time being of 4 months, it is challenging to complete every aspect of the project within the given timeframe. However, the main focus will be on developing and integrating the text and video analysis components of the sentiment analysis web app which includes selecting and training models for text and video data processing, implementing speech-to-text technology for audio input, and analysing facial expressions and body language from video inputs. Additionally, the plan is to include a video response of the user sentiment after the user gave their input.

Our success will be measured by the accuracy and performance of these components, as well as user feedback and engagement with the web app.

While we may not be able to integrate generative AI within the 4-month timeframe, we will lay the groundwork for future enhancements by designing the system with adaptability in mind.

**Resources required for the project and the feasibility of their availability as per the plans?**

To execute our project successfully, we will require various resources. In terms of software, we will utilize Python for machine learning , TensorFlow and Py Torch for model training, JavaScript, HTML, CSS etc for web application development, and OpenCV for video processing. Adequate computing power with GPU support will be essential for efficient model training. While accessing larger and more diverse datasets will be pivotal for enhancing model generalization. The challenge lies in acquiring suitable datasets for different aspects of the project, including text and image-based data. Moreover, ensuring the availability of high-quality microphones and cameras is crucial for capturing clear audio and video inputs during user interactions.

**If you are a team of 2 or more then clearly define responsibilities and how each member should be assessed differently**? It should be by Name and Roll No of each Member

Shivam Garg E22CSEU1506 - Responsible for training generative model that will reply to user based on the processed data of their sentiment(based on face emotion and the user input). Also responsible for creating backend of the web application.

Aditya Raj E22CSEU0649 – Responsible for finding and processing the datasets then separating and extracting text from audio and training model for sentiment analysis of text. Also responsible for converting output text to video.

Sanjana Nayak E22CSEU0644 – Responsible for extracting image from video and training model for emotion recognition. Also responsible for creating frontend of the web application.

**Short and long term planning with detailed and logical steps and timelines.**

Short term planning includes sentiment analysis which includes speech to text recognition and facial recognition. We are also planning to take the dataset from the specialist so that it is more convenient for the user. Speech to text recognition will capture the nuances of your voice, from the tremor of excitement to the sigh of frustration. This allows for a deeper understanding of your sentiment, uncovering hidden emotions that might otherwise remain veiled. Facial Recognition It will provide the user with instant insights into their expressed emotions in real time is novel and valuable for self-awareness and communication improvement. Short term planning includes creating a website with these ideas.

Long term planning will include the generative AI plan and text to video which will be a little complex. Generative AI will provide empathetic and supportive responses tailored to individual emotions and situations. Text to video will take the text of the user develop AI-powered avatars that can display emotions in a natural and engaging way, mirroring the user's sentiment.

**Evidence/literature/research/survey etc. in support of the hypothesis or idea to be successful. (Local/National/International Context) \*** What makes you believe that it will be successful. If there are some validations, supporting theory or literature available, then please cite it here. You can also give some references in this answer that have been used to build up your proposal.

***Speech to text-***

[1] Jose D V, Alfateh Mustafa, Sharan R, "A Novel Model for Speech to Text Conversion," International Refereed Journal of Engineering and Science (IRJES), vol 3, no. 1, 2014.

[(PDF) Speech to text conversion and summarization for effective understanding and documentation (researchgate.net)](https://www.researchgate.net/publication/342147736_Speech_to_text_conversion_and_summarization_for_effective_understanding_and_documentation)

[2] Y. H. Ghadage and S. D. Shelke, "Speech to text conversion for multilingual languages," 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, pp. 0236-0240, 2016.

[Speech to text conversion for multilingual languages | Semantic Scholar](https://www.semanticscholar.org/paper/Speech-to-text-conversion-for-multilingual-Ghadage-Shelke/26ffe939459b01ba7d1962b9ae374319844e35a3)

[3] Umar Nasib Abdullah, Kabir Humayun, Ahmed Ruhan, Uddin Jia., "A Real Time Speech to Text Conversion Technique for Bengali Language," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), pp. 1-4, 2018.

[A Real Time Speech to Text Conversion Technique for Bengali Language | IEEE Conference Publication | IEEE Xplore](https://ieeexplore.ieee.org/abstract/document/8465680)

[4] L. Wan, "Extraction Algorithm of English Text Summarization for English Teaching," 2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Xiamen, China, 2018, pp. 307-310.

[(PDF) Speech to text conversion and summarization for effective understanding and documentation | International Journal of Electrical and Computer Engineering (IJECE) - Academia.edu](https://www.academia.edu/43708563/Speech_to_text_conversion_and_summarization_for_effective_understanding_and_documentation)

[5] Saiyed S., Sajja P. S., "Review on text summarization evaluation methods," Indian Journal of Computer Science and Engineering, vol. 8, no. 4, pp. 497, 2017.

[REVIEW ON TEXT SUMMARIZATION EVALUATION METHODS (ijcse.com)](https://www.ijcse.com/docs/INDJCSE17-08-04-021.pdf)

*Facial recognition-*

1] Takeo Kanade. Computer recognition of human faces, volume 47. Birkh¨auser Basel, 1977.

[Computer recognition of human faces | SpringerLink](https://link.springer.com/book/10.1007/978-3-0348-5737-6)

[2] Lawrence Sirovich and Michael Kirby. Low-dimensional procedure for the characterization of human faces. Josa a, 4(3):519–524, 1987.

[3] M. Turk and A. Pentland. Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1):71–86, Jan 1991.

[4] Dong chen He and Li Wang. Texture unit, texture spectrum, and texture analysis. IEEE Transactions on Geoscience and Remote Sensing, 28(4):509–512, Jul 1990.

[5] X. Wang, T. X. Han, and S. Yan. A hog-lbp human detector with partial occlusion handling. In 2009 IEEE 12th International Conference on Computer Vision, pages 32–39, Sept 2009.

*Generative AI-*

[1] J. Gui, Z. Sun, Y. Wen, D. Tao and J. Ye, "A review on generative adversarial networks: Algorithms theory and applications", IEEE Trans. Knowl. Data Eng., 2021.

[[2001.06937] A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications (arxiv.org)](https://arxiv.org/abs/2001.06937)

[2] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional GANs", Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., pp. 8798-8807, 2018.

[High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs | IEEE Conference Publication | IEEE Xplore](https://ieeexplore.ieee.org/document/8579015)

**Risk Analysis (What are the factors which pose risk of failure of your project and risk of not completing your project by deadline)**How you are taking care of these risks so that they can be avoided.

In speech to text analysis, Noise can be a factor if excess noise is coming from the background, then it won’t be able to give response to your voice. This can we corrected by using soundproof devices like ear dopes or headphones so that voice goes directly to the device.

In Facial recognition, it is necessary that the quality of the image is clear not blurry and high resolution. And should train the system so that it recognises different facial expression like smiling, crying. Also, error can be generated on side angles so to avoid it the system should be trained in such manner that it recognises the face at different angles or orientations, including side profiles or tilted heads.

In Generative AI, problem can be overfitting, models may be overfit to specific sentiment patterns present in the training data, resulting in poor generalization to unseen sentiments or contexts. It can be avoided by regularly validating the model's performance on diverse sentiment datasets to detect and mitigate overfitting. Employ techniques such as regularization, dropout, or early stopping during model training to prevent overfitting.

**Give Names and emails of Three Persons with whom you have discussed the details of the project and what was their reaction/suggestion. \* It is important to converse and take perspective on your proposal from others.** One person each from categories (Industry, Your parents, friends, teachers)

**List down the Ethics, Privacy, Moral and Legal issues related with the project \*** Even if there are no issues, then describe how it fulfils all the normsx

Ethics can be that the responses would be ethically moral and positive. User should be able to express themselves so that it gives positive responses while considering user’s feeling and privacy.

Privacy can be that user’s personal responses will be private, and the system won’t ask anything private like password or something. Personal information like the user’s name, number, his/her responses during the session would be totally secured.

Moral of our project is while aiming for empathy we must ensure that AI responses are authentic and don't exploit emotional vulnerabilities.