



SIKSHA 'O' ANUSANDHAN
(Deemed to be University)
Faculty of Engineering & Technology (ITER)

Department of Computer Science and Engineering

Project Proposal Form

SENIOR DESIGN PROJECT-2025

SECTION: L

GROUP NO: L15

PROJECT TITLE: Rhythm Composer: Music Generator.

PROJECT ABSTRACT:

This project introduces an innovative approach to AI-powered music generation by leveraging cutting-edge Transformer models and PyTorch. By combining natural language processing with deep learning, the system follows a structured two-phase process to create original music compositions. Initially, it takes a simple text prompt from the user and employs a pre-trained GPT-Neo model to generate imaginative lyrics. These lyrics are then refined and formatted for better readability, ensuring coherence and structural clarity, which enhances their suitability for musical composition.

In the second phase, the polished lyrics are passed into a music generation model that produces an audio output, effectively crafting a musical piece. The entire workflow is integrated into a FastAPI-based web interface, enabling seamless user interaction. Users simply provide a text prompt and select a preferred duration, and the system generates a unique AI-composed track. Thanks to its modular architecture, this solution offers complete creative control while eliminating reliance on costly third-party APIs, making it a budget-friendly option for AI-driven music production.

Built upon robust open-source libraries and modern frameworks, the project prioritizes flexibility and scalability. Its ability to merge AI-generated lyric creation with automated music composition highlights the adaptability of deep learning in artistic applications. By harnessing advanced AI techniques, this system provides a valuable tool for musicians, artists, and enthusiasts interested in exploring AI-assisted music generation. Ultimately, it serves as a compelling example of how deep learning can revolutionize the arts and entertainment sector, pushing the boundaries of AI-driven creativity.

(1) SOFTWARE, HARDWARE OR METHODS/ALGORITHMS SPECIFICATIONS:

Software: Programming Language: Python

Web Framework: FastAPI for creating API endpoints and managing web requests.

Deep Learning & NLP: Hugging Face Transformers (using the GPT-Neo 1.3B model) for text generation and PyTorch for underlying model support.

Music Generation: Replicate's integration with the suno-ai/bark model for generating audio from text.

Templating & Environment: Jinja2 for HTML templating and dotenv for managing environment variables.

Hardware: Development Environment: A standard PC running Windows, Linux, or macOS with Python 3.10

Processing: While the system runs on a regular CPU, a GPU can significantly speed up processing for deep learning tasks.

Cloud Options: Services like Google Colab or Kaggle Cloud can be used for GPU-accelerated processing when needed.



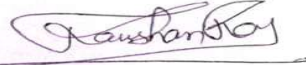

Methods/Algorithms:

Text Generation: Uses the GPT-Neo model to generate lyrics from a user-supplied prompt with adjustable parameters such as temperature and maximum length.

Audio Generation: The generated lyrics serve as input to a music generation model that creates audio using advanced deep learning techniques.

Modular Design: The system is structured to allow easy updates and integration of alternative models or algorithms in the future.

(2) **NAME, REG. NO AND SIGNATURE OF GROUP MEMBERS:**

NAME	RegNo	Signature
Saurav Kumar	2141011106	
A Durga Madhab Patro	2141016231	
Raushan Raj	2141011803	
Dipanshu Kumar Mahato	2141013040	

(3) **APPROVAL STATUS (To be filled in by the Section Coordinator of SDP):**



Project Supervisor

Section Coordinator, SDP

SDP Coordinator