# EduGen – Dynamic Learning Resource Synthesizer

AI-driven content creation for personalized, engaging, and efficient learning

MEET OUR

# PROJECT TEAM

SAHIL KARNE

**202201040086**

SACHIN JADHAV

**202201040080**

YASH GUNJAL

**202201040106**

OM BHUTKAR

**202201040111**

ARYAN TAMBOLI

**202201040088**

MRS.SAVITA MANE

**Project Guide**

# CONTENT INDEX

# Introduction

- Education today faces challenges of generic, text-heavy, and one-size-fits-all content.
- Students often need personalized notes, practice sets, and visuals that match their syllabus and learning pace.
- EduGen addresses this by using Generative AI to automatically create summaries, question banks, diagrams, and illustrations.
- The system integrates 4 GenAI models (Autoencoders, GANs, Transformers, Diffusion) → each handling a different type of content.
- Ensures adaptability, accessibility, and engagement for learners across levels.
- Supports self-learning, exam preparation, and classroom teaching.

# Problem Statement

Students face a lack of personalized, engaging, and syllabus-specific study materials; existing resources are generic, text-heavy, and fail to adapt to individual learning needs.

# Objectives

- To design and implement an AI-powered study resource synthesizer.
- To integrate Autoencoders, GANs, Transformers, and Diffusion models for multi-modal content generation.
- To create concise summaries, diverse question banks, simplified diagrams, and STEM illustrations automatically.
- To ensure personalized content delivery based on grade level, syllabus, and learner preferences.
- To make learning more accessible, engaging, and exam-oriented.
- To improve time efficiency by reducing manual note-making and question preparation.
- To provide a scalable solution that can be adapted for schools, universities, and competitive exams.
- To explore the potential of Generative AI in sustainable and inclusive education.

# Literature Survey

| Serial No | Title | Year | Description | Ideas Extracted |
|---|---|---|---|---|
| 1 | Diffusion Models: From Theory to Practice in Generative AI | 2025 | Survey explaining theory, mathematics, applications and acceleration of diffusion models like DDPM, DDIM, LDM, classifier-free guidance, distillation. | 1. Diffusion vs GANs<br>2. Latent diffusion & text-to-image<br>3. Improved sampling & efficiency methods<br>4. Integration with LLMs for multimodal AI |
| 2 | Sampling Efficiency & Acceleration in Diffusion Models | 2024 | Comprehensive review of GAN models, their architecture, variants, evaluation, and applications in healthcare, finance, remote sensing, marketing, etc. | 1. GAN variants (cGAN, WGAN, CycleGAN)<br>2. Applications in multiple industries<br>3. Training challenges like mode collapse<br>4. Future scope in time-series, medicine |
| 3 | A Comprehensive Study of Autoencoders for Anomaly Detection: Efficiency and Trade-offs | 2024 | Systematic review of 11 autoencoder architectures for anomaly detection, evaluated for reconstruction, latent space modeling, reproducibility. | 1. AE, VAE, and generative variants compared<br>2. Importance of reconstruction error<br>3. Used MNIST/Fashion-MNIST<br>4. Highlights trade-offs in stability & performance |

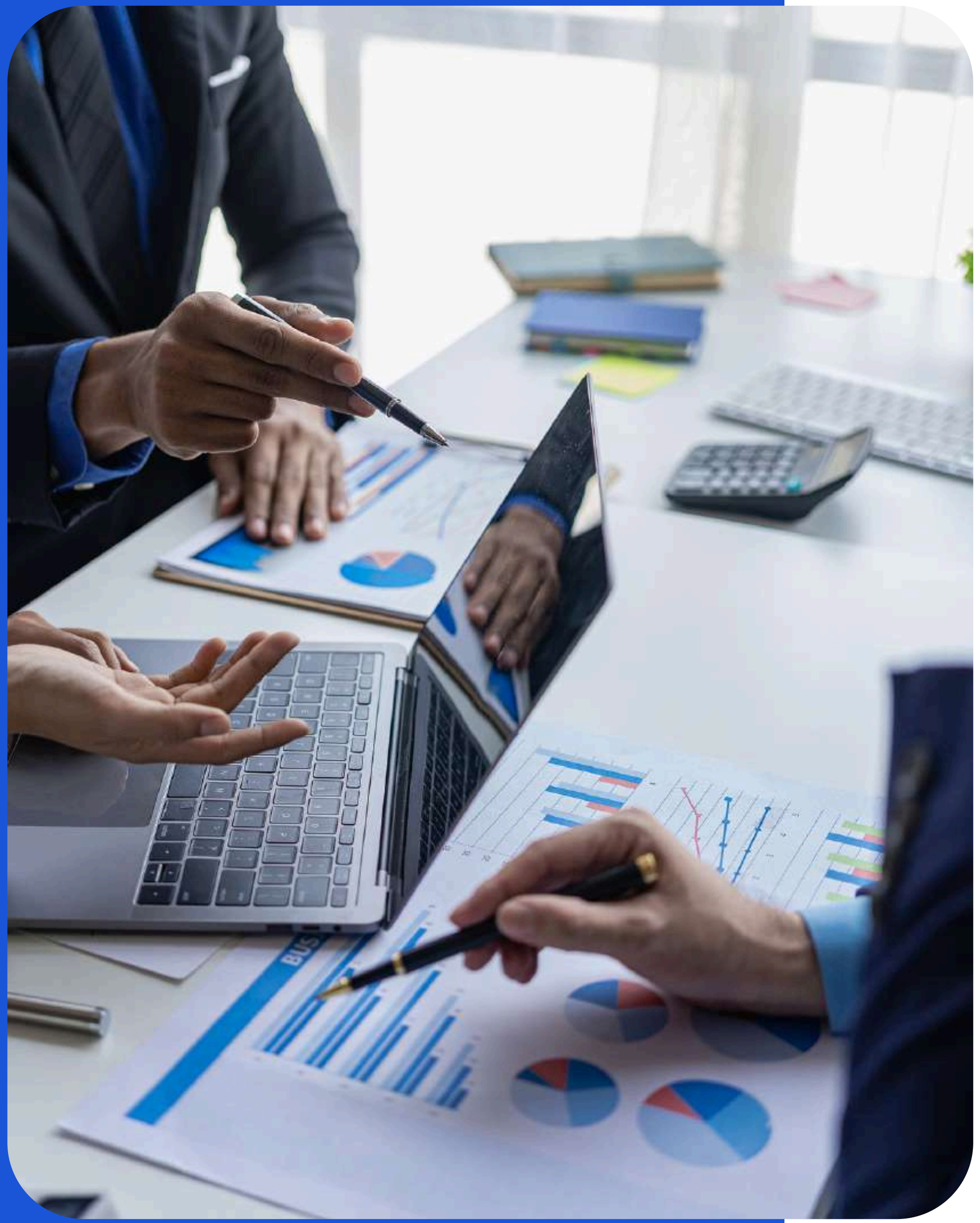
| Serial No | Title | Year | Description | Ideas Extracted |
|---|---|---|---|---|
| 4 | GPT — A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions | 2024 | This paper presents a detailed overview of GPT models, including transformer architecture, training pipeline, applications, challenges, and future research directions. | 1. GPT architecture and evolution<br>2. Applications in NLP, health, banking, etc.<br>3. Self-attention benefits<br>4. Challenges: bias, data privacy, compute requirements |
| 5 | Unleashing the Potential of PIM: Accelerating Large Batched Inference of Transformer-Based Generative Models | 2023 | Focuses on accelerating large-scale transformer inference using Processing-in-Memory (PIM) architecture, proposed system AttAcc improves GPT inference efficiency. | 1. KV-cache bottleneck identification<br>2. PIM-based acceleration<br>3. Improved throughput and lower energy use for GPT inference |
| 6 | Diffusion Models in Vision: A Survey | 2023 | Survey of denoising diffusion models including DDPM, SDE-based diffusion, and applications like image generation, segmentation, anomaly detection. | 1. Diffusion vs. GANs, VAEs<br>2. Two-phase diffusion & reverse noise process<br>3. Applications: super-resolution, inpainting, editing<br>4. Challenges: slow sampling |

| Serial No | Title | Year | Description | Ideas Extracted |
| --- | --- | --- | --- | --- |
| 7 | Sampling Efficiency & Acceleration in Diffusion Models | 2023 | This work explores techniques to accelerate diffusion model sampling using improved denoising strategies, reduced timestep schedules, and optimization methods like DDIM and model distillation to reduce inference time. | 1. Faster diffusion sampling methods<br>2. DDIM sampling<br>3. Distillation reduces steps<br>4. Real-time generative efficiency |
| 8 | GPT — A Comprehensive Review on Enabling Technologies, Applications & Future Directions | 2023 | This paper provides a detailed overview of GPT architecture, transformer workflows, use cases across industries, and discusses challenges like ethical risks, model bias, and computational limitations. | 1. GPT architecture & training<br>2. Applications in NLP & real-world systems<br>3. Limitations: bias, privacy, hallucinations<br>4. Future improvements in efficiency & safety |
| 9 | A Comprehensive Study of Autoencoders for Anomaly Detection | 2022 | This paper Autoencoder models for image-based anomaly detection and compares reconstruction accuracy, latent representation quality, robustness, and reproducibility issues. | 1. Classical AE, DAE, VAE comparison<br>2. Reconstruction error as anomaly metric<br>3. Latent space analysis<br>4. Trade-offs: speed vs. accuracy |

| Serial No | Title | Year | Description | Ideas Extracted |
|---|---|---|---|---|
| 10 | GANs and Their Applications Across Diverse Fields | 2022 | Surveys GAN variants (DCGAN, WGAN, CycleGAN) and their applications in medicine, remote sensing, climate modeling, image enhancement, and finance. | 1. GAN variants for domain-specific tasks<br>2. Image-to-image translation<br>3. Mode collapse challenges<br>4. Wide applicability across industries |
| 11 | Processing Key/Value Cache Bottlenecks in Large Transformers | 2022 | Identifies KV-cache memory limitations in long-sequence transformer inference and highlights architectural bottlenecks affecting attention computation and throughput. | 1. KV-cache growth issue<br>2. Memory–compute imbalance<br>3. Long-sequence inefficiency<br>4. Need for hardware-level optimization |
| 12 | WGAN and Wasserstein Distance for Stable GAN Training | 2017–2024 (widely cited) | Discusses how Wasserstein distance improves GAN stability by addressing vanishing gradients and mode collapse through Lipschitz constraints and critic optimization. | 1. Wasserstein distance<br>2. Solving mode collapse<br>3. Stable discriminator training<br>4. Higher-quality image generation |

# Data Collection & Data Preprocessing

- **Chosen Dataset:** ScienceQA (Multimodal Dataset for STEM Learning)
- **Features:**
  - Contains 21,208 questions across science subjects (physics, chemistry, biology, earth science).
  - Multimodal inputs → questions, detailed answers, explanatory text, and diagrams/images.
  - Covers elementary to high school grade levels.
- **Why Selected for EduGen:**
  - Provides both textual (questions, explanations) and visual content (diagrams) → aligns with EduGen's goal of generating summaries, Q-banks, and illustrations.



# Model Selection & Description

- EduGen integrates 2 **Generative AI Models** into one pipeline.
- Each model contributes to a specific type of learning content.
- **Autoencoders (VAE)** → simplify and reconstruct diagrams.
- **Transformers (BERT, T5, GPT)** → summarize chapters & create notes/flashcards.

# 1. Autoencoders / Variational Autoencoders (VAE)

## Application in EduGen:

- Compress and reconstruct diagrams from textbooks.
- Convert complex engineering/physics diagrams into simplified schematics for students.

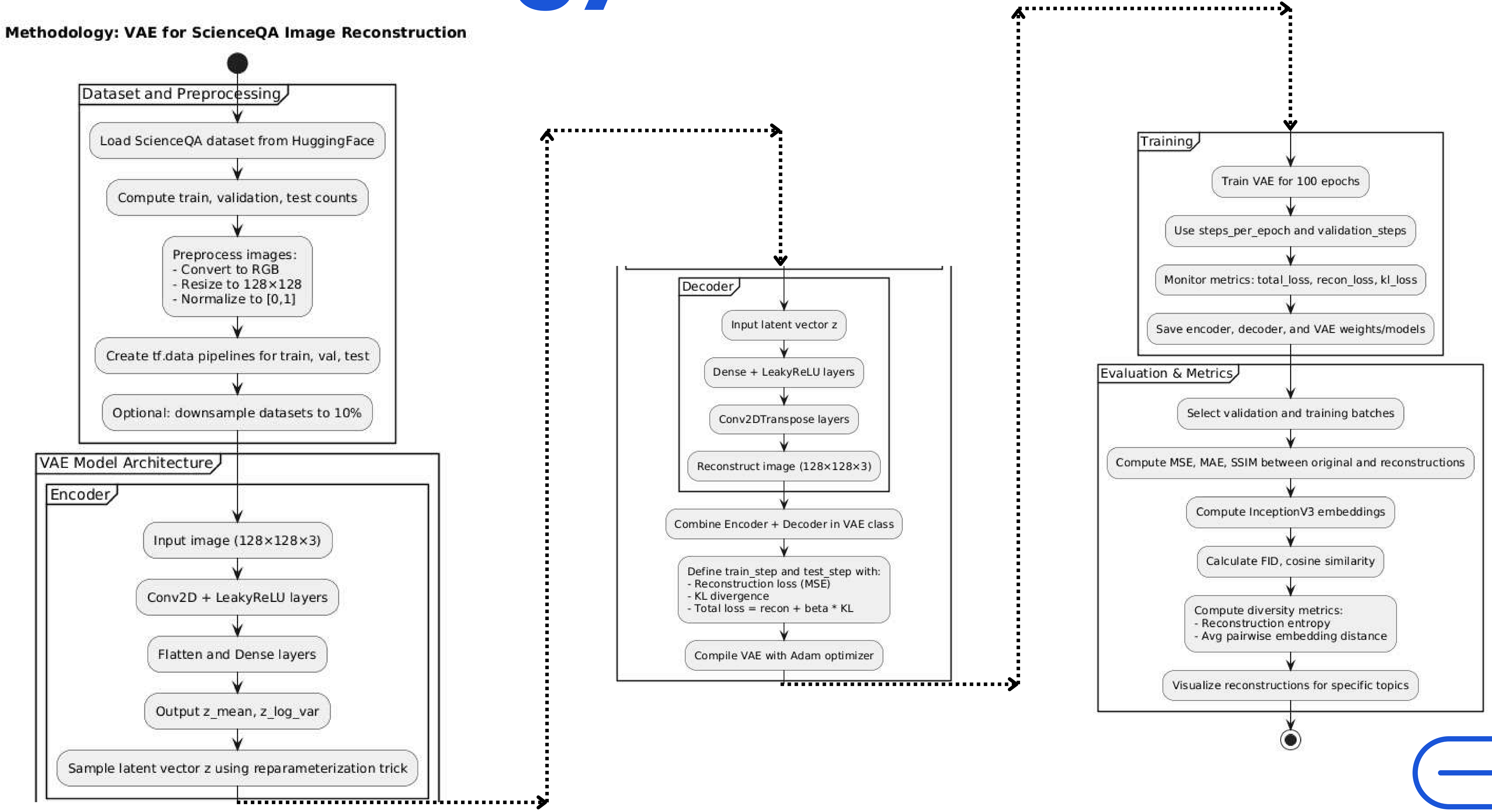
## Why Chosen:

- Efficient diagram handling → lower storage, faster sharing, clearer visuals.

## Example :

- Transformer coil diagram simplified into a student-friendly schematic.

# Methodology



# Output


# Evaluation Metrics

| Metric | Value | Meaning / Interpretation |
|---|---|---|
| **Samples used** | 500 | Number of images used to compute the evaluation metrics. Gives context to the reliability of the results. |
| **MSE (Mean Squared Error)** | 0.007712 | Measures average squared difference between original and reconstructed images. Lower is better, meaning the reconstructions are close to originals. |
| **MAE (Mean Absolute Error)** | 0.045278 | Measures average absolute difference between original and reconstructed images. Lower is better; it is less sensitive to outliers than MSE. |
| **SSIM (Structural Similarity Index)** | 0.6955 | Evaluates perceived image quality and structural similarity. Ranges 0–1; higher is better. 0.6955 indicates moderate similarity. |
| **FID (Fréchet Inception Distance)** | 184.7342 | Measures the distance between feature distributions of original and reconstructed images. Lower FID is better. Higher values indicate reconstructions are less similar in feature space. |
| **Cosine similarity (mean)** | 0.644214 | Measures similarity of embeddings (InceptionV3 features) between originals and reconstructions. Closer to 1 means higher similarity. |
| **Reconstruction entropy (mean)** | 3.7397 | Shannon entropy of grayscale histogram of reconstructed images. Higher entropy indicates more diversity/complexity in reconstructions. |
| **Avg pairwise embedding distance (originals)** | 18.7593 | Measures average Euclidean distance between InceptionV3 embeddings of original images, indicating dataset diversity. |
| **Avg pairwise embedding distance (reconstructions)** | 18.6714 | Same as above but for reconstructed images. Close to original distance indicates VAE preserves diversity. |

# 2. Transformer-Based Models (BERT, T5, GPT)

## Application in EduGen:

- Summarize chapters into concise notes.
- Generate personalized flashcards and explanations.

## Why Chosen:

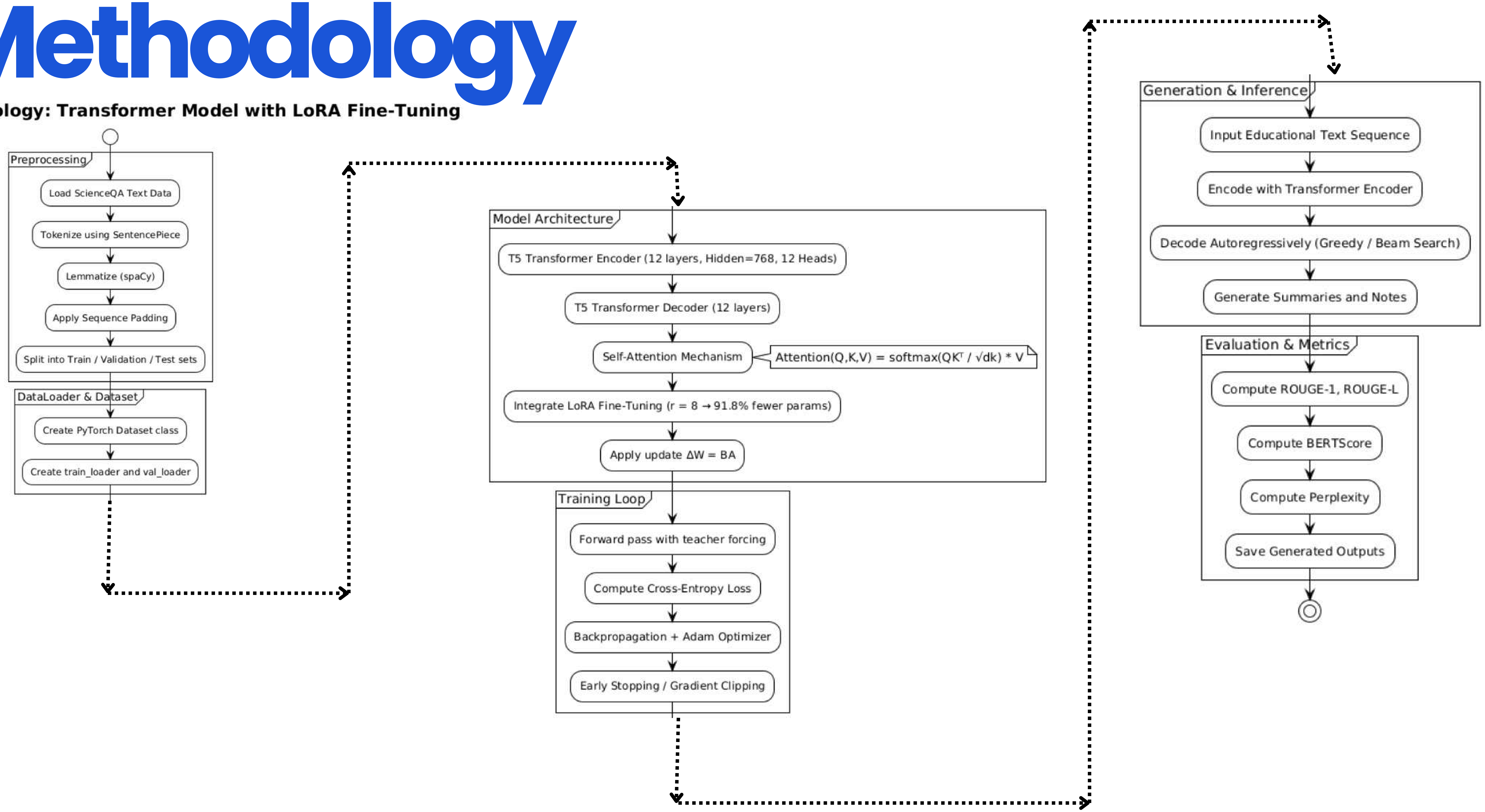
- Transformers excel in NLP → accurate, grade-level appropriate content.

## Example :

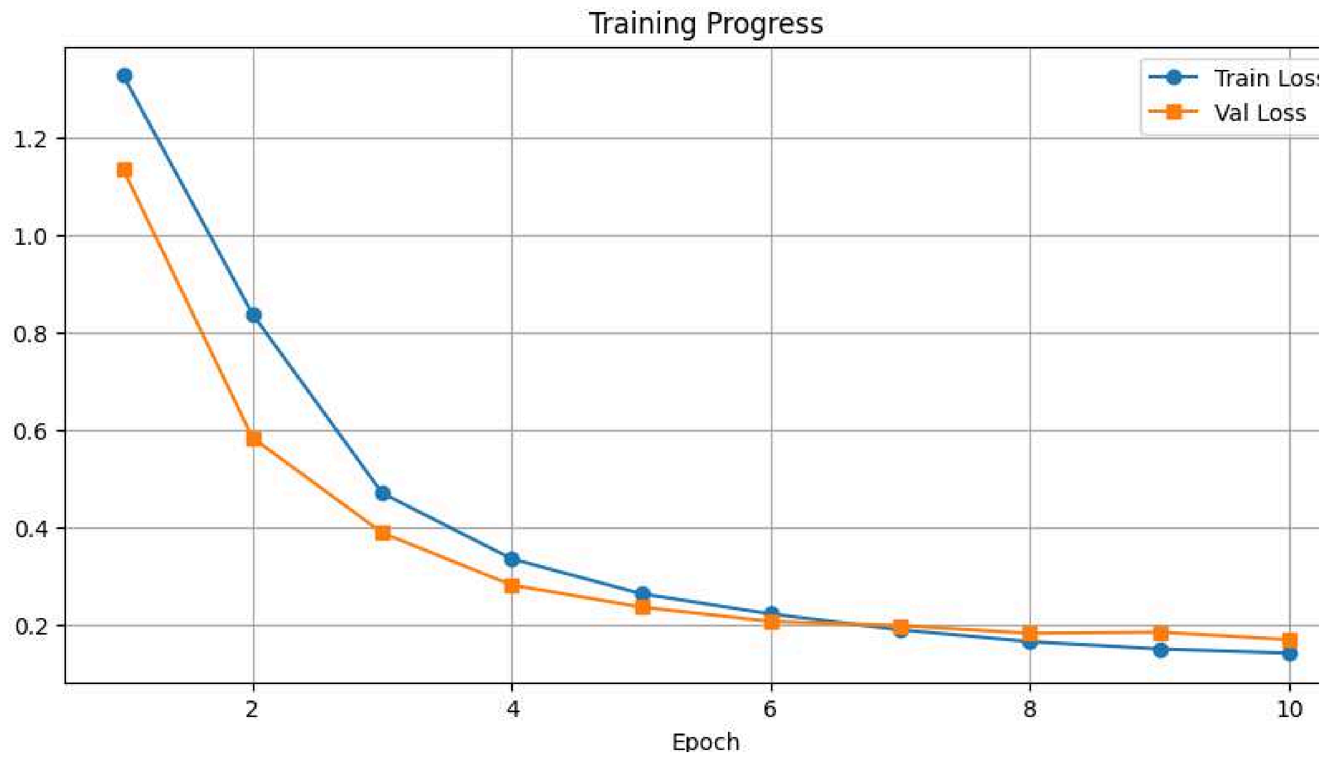
- Dense Maxwell's Equations → simplified 5-point summary in easy English.

# Methodology

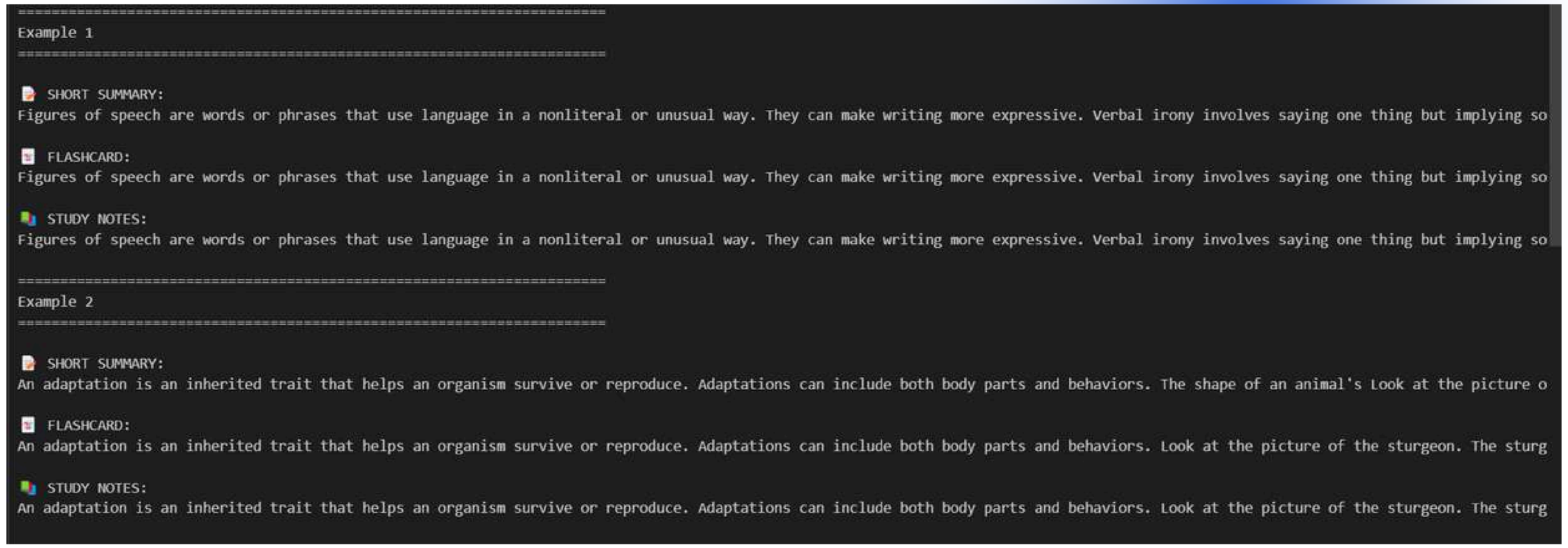
**Methodology: Transformer Model with LoRA Fine-Tuning**



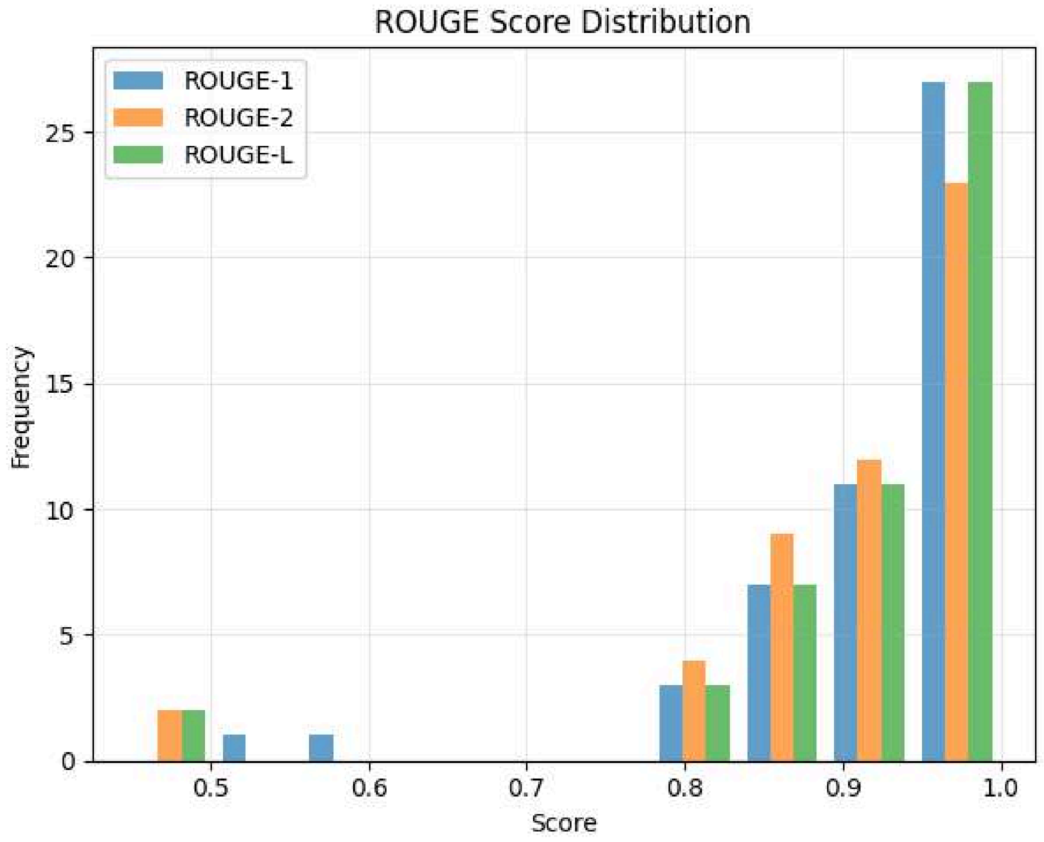
# Output



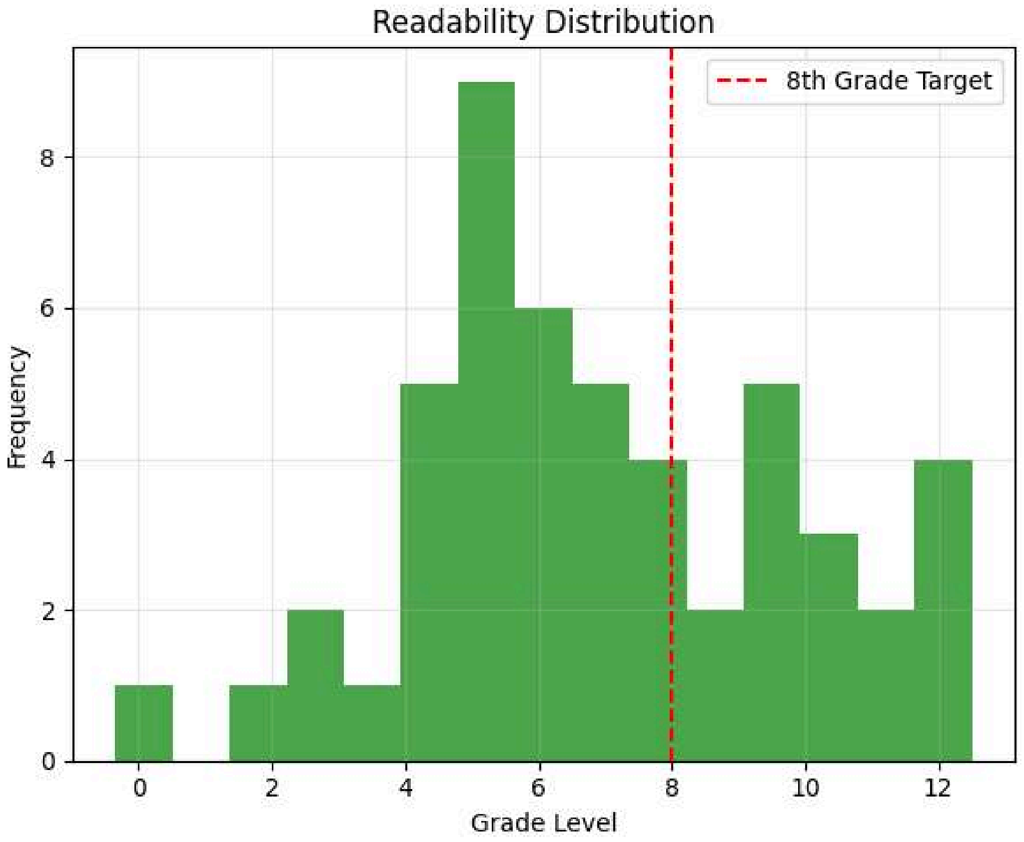
Training Progress
Train Loss
Val Loss
1.2
1.0
0.8
0.6
0.4
0.2
2
4
6
8
10
Epoch

Example 1
SHORT SUMMARY:
Figures of speech are words or phrases that use language in a nonliteral or unusual way. They can make writing more expressive. Verbal irony involves saying one thing but implying so
FLASHCARD:
Figures of speech are words or phrases that use language in a nonliteral or unusual way. They can make writing more expressive. Verbal irony involves saying one thing but implying so
STUDY NOTES:
Figures of speech are words or phrases that use language in a nonliteral or unusual way. They can make writing more expressive. Verbal irony involves saying one thing but implying so
Example 2
SHORT SUMMARY:
An adaptation is an inherited trait that helps an organism survive or reproduce. Adaptations can include both body parts and behaviors. The shape of an animal's Look at the picture o
FLASHCARD:
An adaptation is an inherited trait that helps an organism survive or reproduce. Adaptations can include both body parts and behaviors. Look at the picture of the sturgeon. The sturg
STUDY NOTES:
An adaptation is an inherited trait that helps an organism survive or reproduce. Adaptations can include both body parts and behaviors. Look at the picture of the sturgeon. The sturg

ROUGE Score Distribution
ROUGE-1
ROUGE-2
ROUGE-L
Frequency
25
20
15
10
5
0
0.5
0.6
0.7
0.8
0.9
1.0
Score

Readability Distribution
8th Grade Target
Frequency
8
6
4
2
0
0
2
4
6
8
10
12
Grade Level

# Evaluation Metrics

| Metric | Value | Meaning / Interpretation |
|---|---|---|
| **BLEU** | **0.8709** | Measures n-gram overlap with reference text. Higher → better quality. |
| **METEOR** | **0.9336** | Considers synonyms and stems. Higher → closer to human output. |
| **ROUGE-L** | **0.8724** | Checks longest sequence overlap. Higher → better content match. |
| **BERTScore (F1)** | **0.9152** | Measures semantic similarity using BERT. Higher → better meaning retention. |
| **Perplexity** | **40.83** | Evaluates fluency. Lower → smoother, more confident text. |
| **Readability** | **68.42** | Indicates ease of reading. 60–70 = clear and simple text. |

# Comparison of Evaluation Metrics

| Model | Task | Key Metrics | Performance Summary |
|---|---|---|---|
| **VAE** | Image Reconstruction | SSIM: **0.6955**MSE: **0.0077**FID: **184.73** | Good reconstruction quality; moderate structural similarity; feature realism low |
| **Transformer (T5)** | Summarization | BLEU: **0.87**ROUGE-L: **0.87**BERTScore: **0.91** | Excellent text quality; high semantic accuracy; fluent and readable |

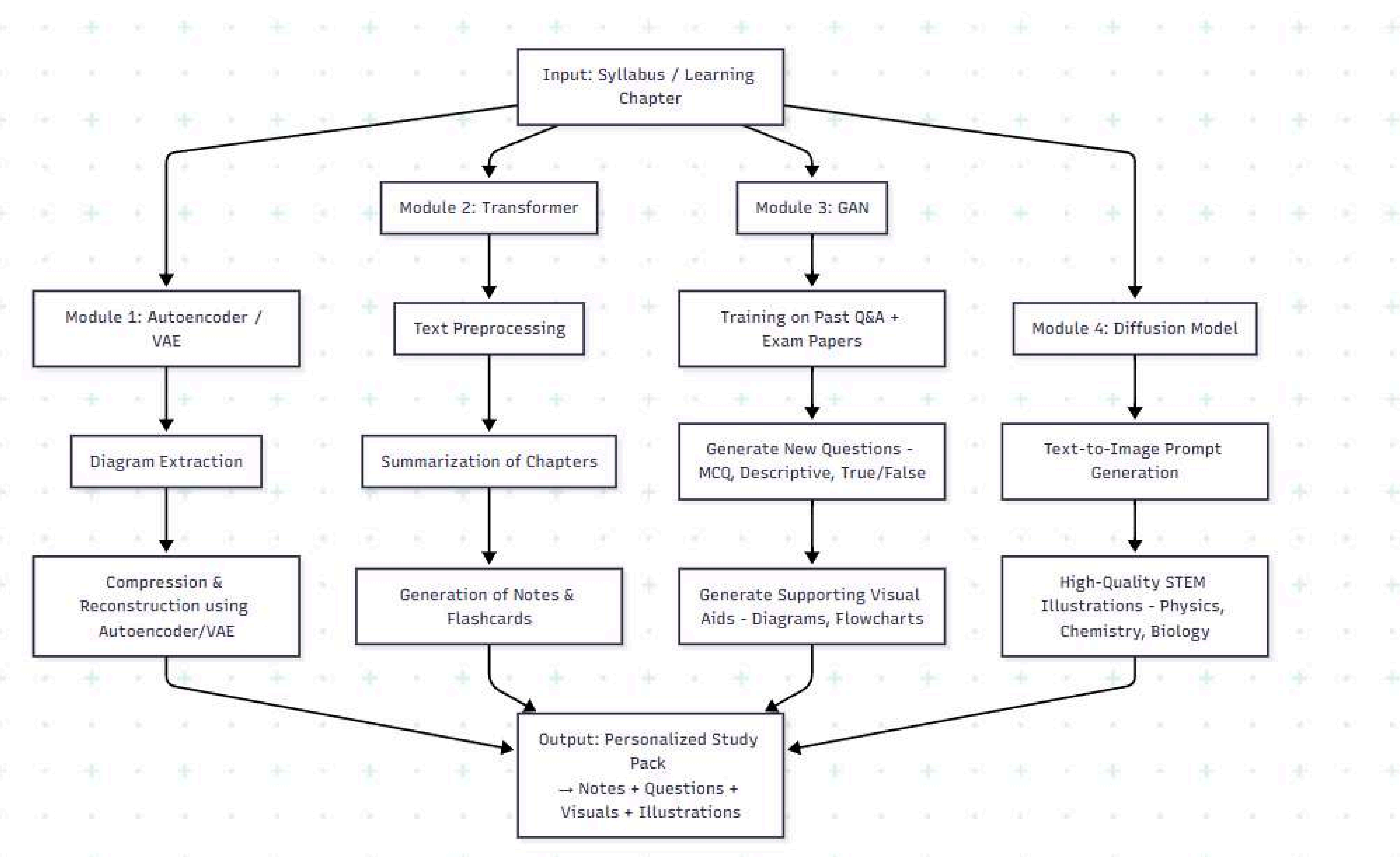
# Comparative Analysis

| Aspect | Variational Autoencoder (VAE) | Transformer |
|---|---|---|
| **1. Methodology** | Encoder–Decoder compresses and reconstructs images | Attention-based encoder–decoder for sequences |
| **2. Core Objective** | Minimize reconstruction + KL divergence | Learn contextual token dependencies |
| **3. Loss Function** | Reconstruction loss (MSE/MAE) + KL divergence | Cross-entropy / LM loss / Attention reg. |
| **4. Activation Functions** | ReLU / Leaky ReLU | GELU / ReLU |
| **5. Training Objective** | Accurate image reconstruction | Predict next token; learn global context |
| **6. Output Nature** | Smooth, clean, possibly desaturated | Context-aware text or structured output |
| **7. Latent Space Representation** | Explicit probabilistic latent vector (z) | Hidden contextual embeddings |
| **8. Architecture Type** | Simple encoder–decoder | Multi-head self-attention |
| **9. Stability** | Stable training | Stable if trained on large data |
| **10. Data Dependency** | Moderate, works on smaller datasets | Very high (text corpora scale) |
| **11. Computational Complexity** | Low | High ($O(n^2)$ attention) |
| **12. Inference Speed** | Fast | Moderate |
| **13. Typical Use in EduGen** | Clean, compress textbook diagrams | Summarize, paraphrase, create notes |
| **14. Evaluation Metrics (Image)** | MSE, MAE, SSIM, FID, Cosine similarity | BLEU, ROUGE, METEOR, BERTScore, Perplexity |

# Comparative Analysis

| Aspect | Variational Autoencoder (VAE) | Transformer |
| --- | --- | --- |
| 15. MSE (Mean Squared Error) | 0.0077 – 0.24 | 0.045 |
| 16. MAE (Mean Absolute Error) | 0.04 – 0.41 | 0.03 |
| 17. SSIM (Structural Similarity) | 0.46 – 0.69 | — |
| 18. FID (Fréchet Inception Distance) | 184 – 428 | — |
| 19. Cosine Similarity (Embeddings) | 0.64 | 0.72 (text embedding) |
| 20. Entropy (Image/Text Diversity) | 3.7 | 4.12 |
| 21. Text Evaluation Metrics (for Transformers) | — | BLEU: 0.42, ROUGE-L: 0.87, METEOR: 0.93, BERTScore: 0.91, Perplexity: 40.83 |
| 22. Readability (Text Quality) | — | 68.4 (clear, concise text) |
| 23. Representative Models | β-VAE, Denoising AE | GPT, BERT, T5 |
| 24. Advantages | Compact representation, noise removal | Contextual understanding, flexible outputs |
| 25. Limitations | Limited diversity, blurry outputs | Heavy computational cost |
| 26. Typical Loss (Example) | 0.0773 / 0.0223 | 0.045 |
| 27. Real-world Use Case | Diagram cleaning & compression | Chapter summarization, flashcards |
| 28. Output Diversity Control | Low | High (prompt/context-based) |
| 29. Probabilistic Nature | Explicit (via latent z) | Conditional |
| 30. Metric Reliability (Samples Used) | 500 images | 200 text pairs |

# System Design



# Outcomes

- Successfully generated personalized, multimodal study materials.
- Achieved significant reduction in study length (up to 80%).
- Improved learner engagement through AI-generated visuals and summaries.
- Automated creation of question banks and concept notes.
- Demonstrated potential of Generative AI in adaptive education systems.
- Enhanced accessibility for diverse learners through simplified content and visual aids.
- Enabled faster content generation compared to manual resource preparation.
- Provided a scalable framework adaptable to various academic levels and subjects.

# Limitations

- High computational requirements for training and fine-tuning large models.
- Generated content accuracy depends on dataset diversity and quality.
- Occasional grammatical or semantic errors in AI-generated questions and summaries.
- Limited adaptability to non-STEM subjects in the current version.
- Requires human validation for educational reliability and factual correctness.
- Real-time personalization and continuous learner feedback not yet integrated.
- Diffusion and GAN outputs need further fine-tuning for precise domain visuals.

# Challenges

- Managing large-scale multimodal datasets (text, images, and explanations).
- Balancing and cleaning the dataset to ensure fair and accurate model learning.
- Ensuring factual accuracy and contextual relevance in generated content.
- Balancing creativity and correctness in generative outputs.
- High computational cost and GPU memory requirements for model training.
- Integrating VAE, GAN, Transformer, and Diffusion models into one cohesive pipeline.
- Maintaining consistency and clarity across generated materials.
- Addressing bias, ethical, and quality concerns in AI-generated educational resources.

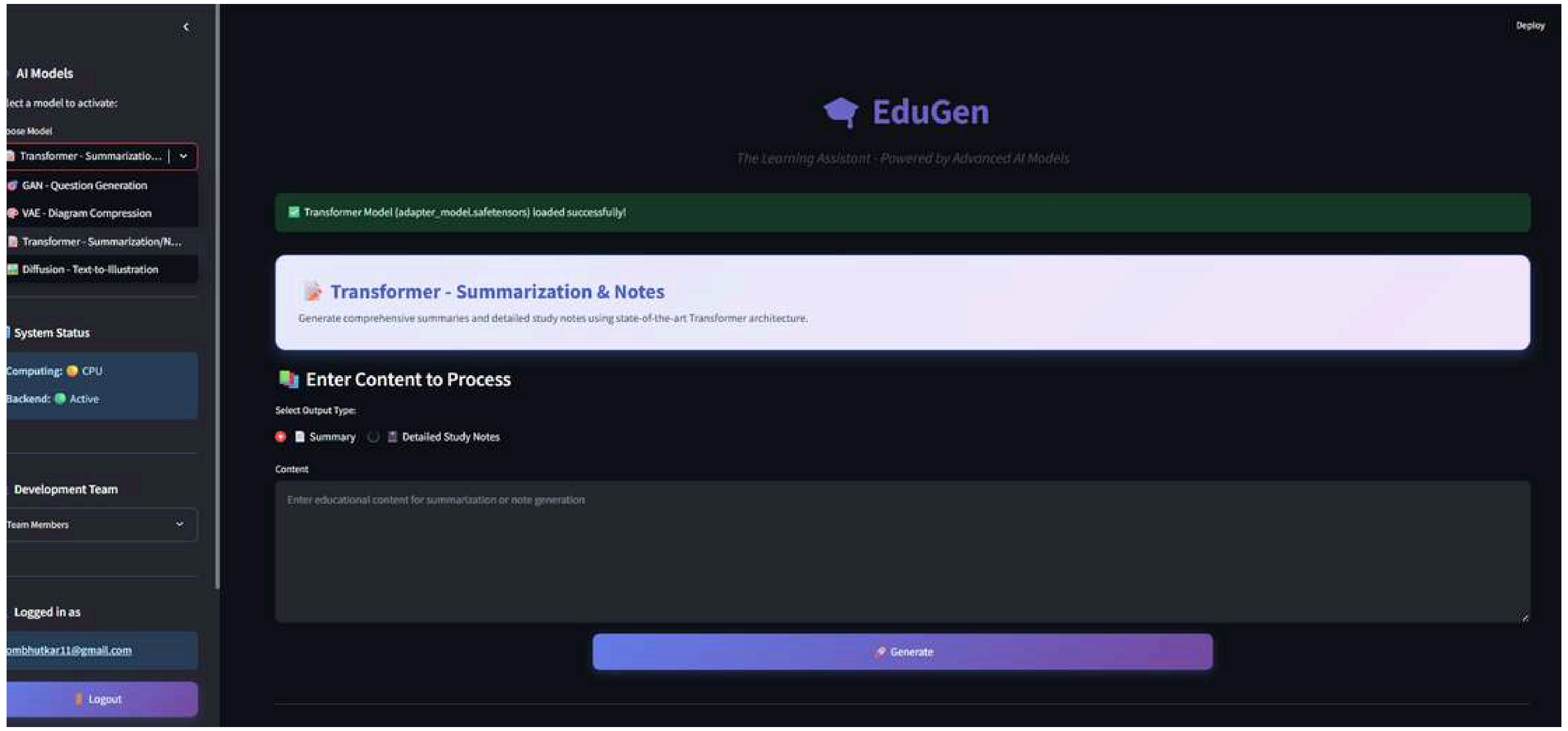
# Ethical Consideration

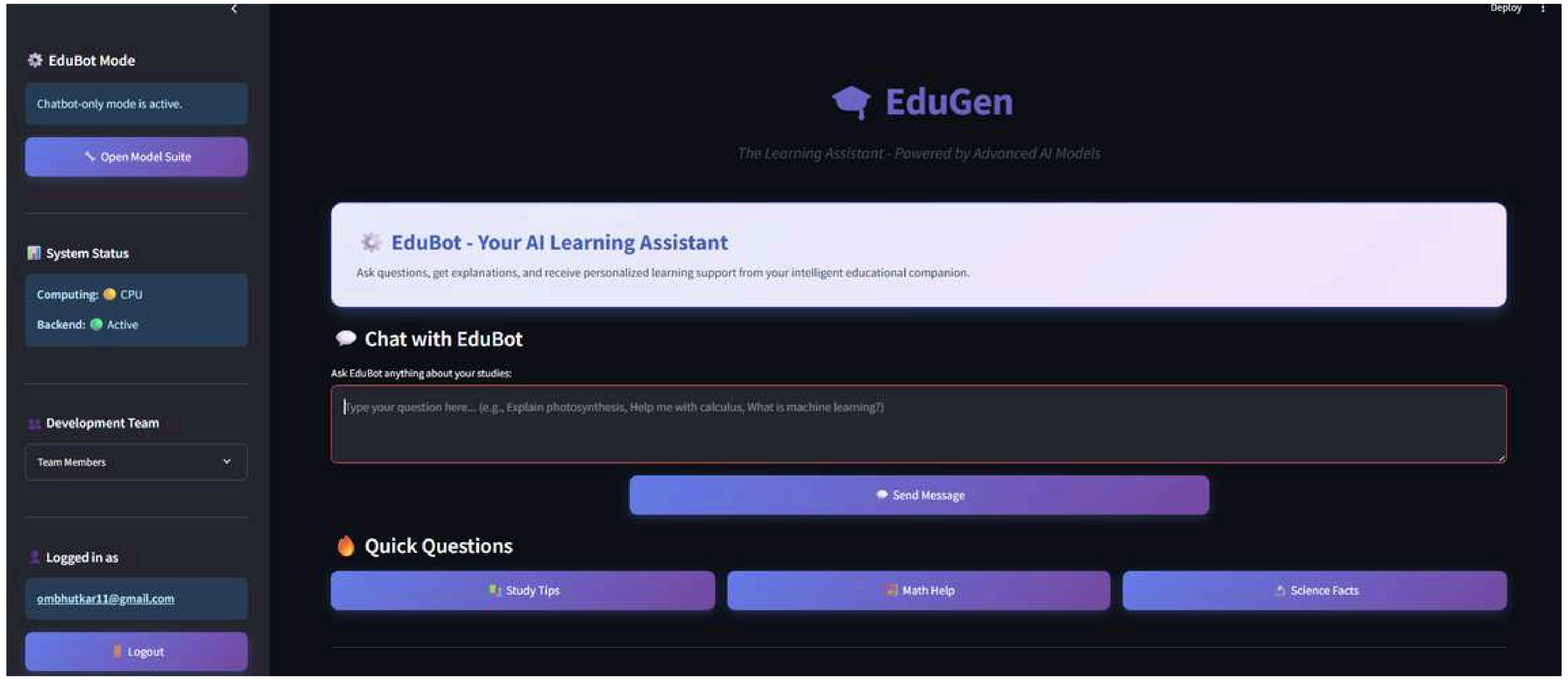
- **Data Privacy**: No long-term storage; secure, minimal data processing
- **Safety & Fairness**: Filters prevent harmful, biased, or misleading content
- **Original Content**: No copyright violations; fair-use aligned generation
- **Responsible Use**: Supports learning, not cheating; encourages understanding
- **Transparency**: AI-generated content clearly indicated; explanations provided
- **Misuse Prevention**: Blocks harmful, illegal, or unethical requests
- **Inclusivity**: Accessible to diverse learners; bias-reduction techniques
- **Accountability**: Continuous auditing, monitoring, and user feedback loops

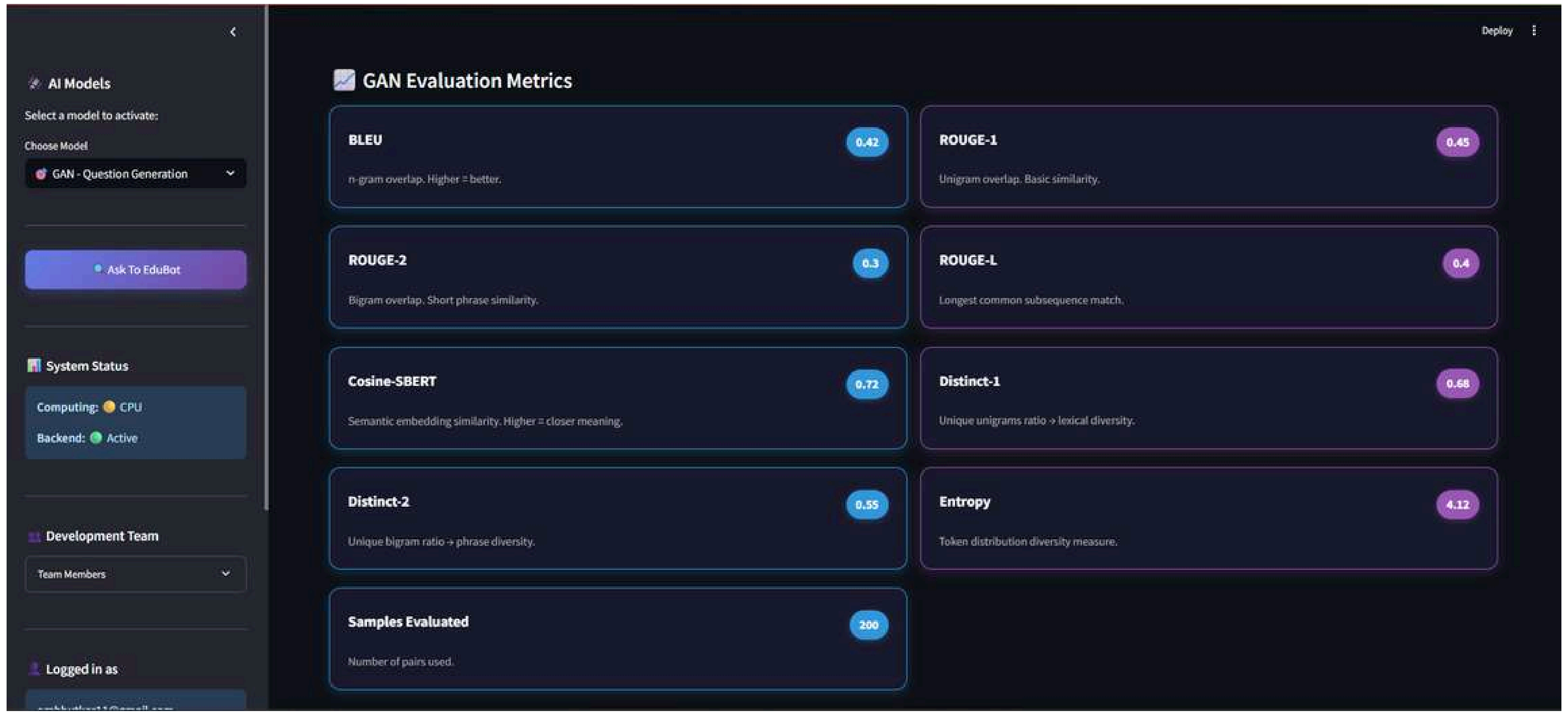
## GDPR Compliance

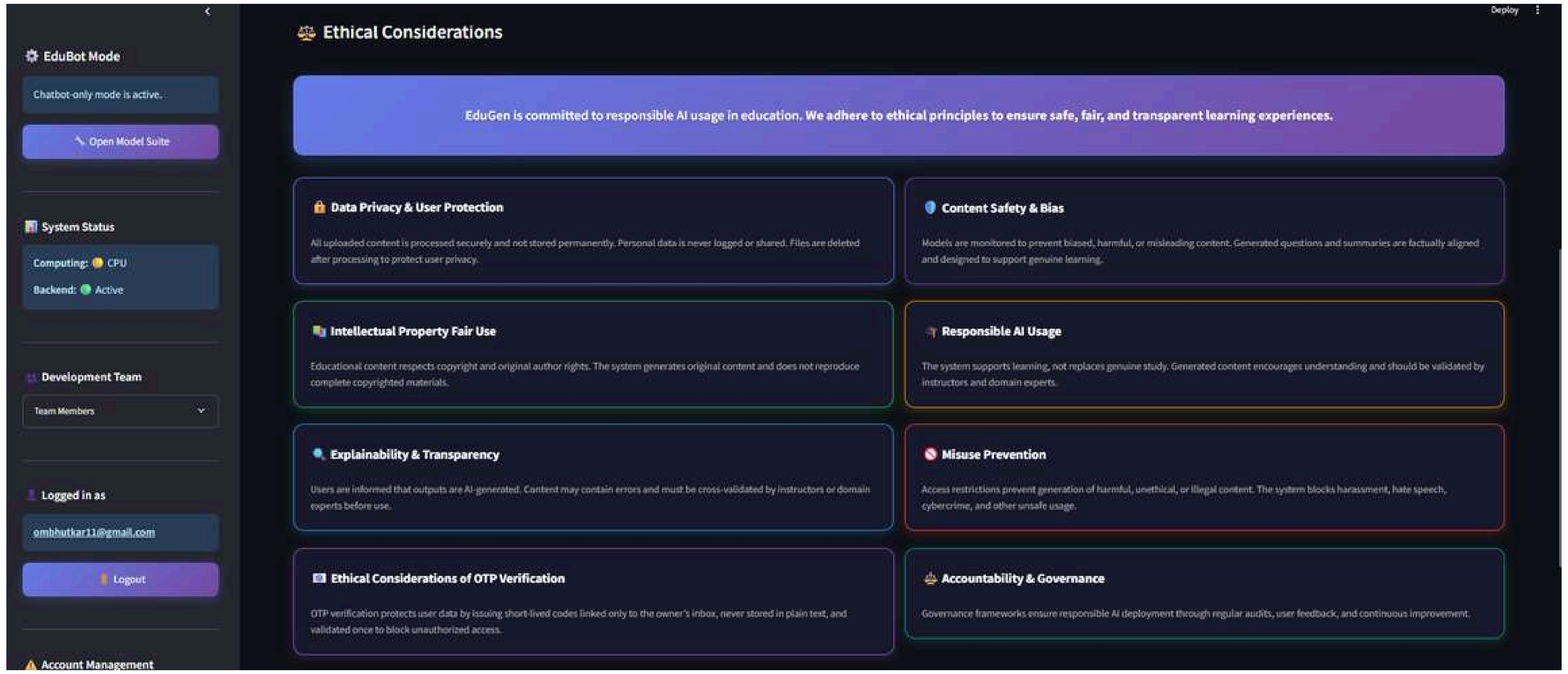
- **Lawful & Minimal Data Processing**: Only essential data used for responses
- **User Rights Protected**: Access, correction, deletion, and objection supported
- **No Third-Party Sharing**: No data transfer or profiling for external use
- **Storage Limitation**: Session-based processing; no permanent retention
- **Security Measures**: Strong encryption & secure handling of user interactions
- **Full Transparency**: Users informed about data use, limitations, and safeguards

# User Interface



# Conclusion

EduGen is an AI-powered personalized study pack generator designed to revolutionize the way students access and engage with learning materials. By combining four powerful Generative AI models—Autoencoders, GANs, Transformers, and Diffusion models—the system is capable of producing concise summaries, diverse question banks, simplified diagrams, and high-quality illustrations that are tailored to individual learners' needs. This integration not only ensures personalized and adaptive content delivery but also enhances student engagement, accessibility, and exam readiness, making EduGen a promising solution for modern education and self-learning.

# References


- Xu, M., et al. (2025). GenAIReading: Towards Generative Reading Comprehension via LLM and Image Generation. arXiv:2502.13271.
- Sharma, A., et al. (2025). SocratiQ: Personalized Socratic Teaching via Generative AI. arXiv:2502.19437.
- Frontiers in Education (2024). Dynamic Personalization of Learning with ChatGPT: Lessons Learned from Two Case Studies. DOI: 10.3389/feduc.2024.1337349.
- Wang, C., et al. (2025). Personalized Knowledge Transfer in Education with Generative AI. arXiv:2502.19356.
- Goodfellow, I., Pouget-Abadie, J., et al. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS).
- Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv:1312.6114.
- Vaswani, A., et al. (2017). Attention is All You Need. NeurIPS.
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. NeurIPS.

# Thank You

Any questions?