CS6120 : Natural Language Processing
Project Proposal

Title : Lyrics generator

Group Number : 15

Group Members :

- **Krishna Venkatesh [002774999]**
- **Laasya Anantha Prasad [002822242]**
- **Shikha Tiwari [002677267]**

Proposal :

We were inspired by Ed Sheeran's music and lyrics, and have created a series of lyrical pieces that capture the emotion, energy, and storytelling of his music, using the power of machine learning and language modeling.

Each piece in the exhibition reflects characteristics from Ed Sheeran's songs. The lyrics of each of his songs have been analyzed and learned by our models to generate and create unique new lyrics, imitating Ed Sheeran himself.

Using N-Gram language models, RNN-LSTM, and a state-of-the-art GPT-2 Transformer, we have analyzed Ed Sheeran's lyrics and music to generate new, original song lyrics that capture the essence of his music. Each piece in the exhibition features lyrics generated by the AI model, accompanied by visual art that reflects the mood and themes of the music.

One of the highlights of the exhibition is a piece inspired by Ed Sheeran's hit song "Shape of You." The AI model has generated a new set of lyrics that capture the romance and rhythm of the original song while adding a fresh, unexpected twist.

Overall, this exhibition is a testament to the power of technology to create new forms of art and creativity. By harnessing the power of AI, we have generated new song lyrics that pay homage to the music of Ed Sheeran while exploring and comparing the capabilities of different language models. We invite you to explore the exhibition and discover the exciting possibilities of AI-generated art.

Basic idea: Introduce the problem your group wants to solve.

Traditional language models, such as N-Gram, LSTM, and GPT-2 based Transformers, have shown significant prowess in understanding and generating human language by analyzing extensive training data. While these models are effective at generating text based on learned patterns, they also open new possibilities in creative fields, such as music and lyrics generation. However, generating high-quality and contextually relevant song lyrics that capture the unique style and emotional depth of a specific artist remains a challenging task.

In this project, we aim to generate lyrics that reflect the style and themes of Ed Sheeran's music. By using advanced language models, we seek to create new, original song lyrics that could seamlessly fit into Ed Sheeran's discography. Our goal is to analyze the characteristics of his lyrics and develop models that can generate text emulating his style while ensuring the lyrics are meaningful and artistically coherent

Our Approach:

We will utilize various language models, including N-Gram, LSTM, and GPT-2 based Transformers, to generate lyrics inspired by Ed Sheeran. These models will be trained on a comprehensive dataset of Ed Sheeran's songs to learn the patterns, themes, and structures commonly found in his lyrics. By comparing and contrasting the performance of these models, we aim to identify the strengths and limitations of each approach in generatmodel's performance improvements in real-world applications, providing insights into scalability and adaptability.

Algorithms:

Workflow of Lyrics Generation Using Language Models

1. **Dataset Preparation:** We start with a collection of Ed Sheeran's lyrics, which will be used for training our models.
2. **Pre-processing:** The lyrics will be pre-processed to clean the text, remove unnecessary characters, and standardize the format for model training.
3. **N-Gram Model:** This model will be used to capture the statistical properties of the lyrics, such as word frequency and n-gram sequences.
4. **RNN-LSTM Model:** An LSTM model will be employed to learn the sequential dependencies and contextual information within the lyrics.
5. **GPT-2 Based Transformer Model:** A state-of-the-art GPT-2 Transformer will be fine-tuned on the dataset to generate more sophisticated and contextually rich lyrics.
6. **Comparison and Evaluation:** The generated lyrics from each model will be compared to evaluate their quality, coherence, and adherence to Ed Sheeran's style.

Regularization Techniques:

1. **Laplace Smoothing**
2. **L1 and L2 Regularization**
3. **Data Augmentation**

Related Work:

Paper 1: Lyrics Generation Using LSTM and RNN
Authors: Mateusz Modrzejewski, Jakub Szachewicz & Przemysław Rokita

1. **Architectures Used:**
   - **LSTM Networks:** LSTM networks are a type of RNN designed to remember long-term dependencies and avoid issues like the vanishing gradient problem. They are particularly suited for sequence generation tasks like lyrics generation because they can maintain context over longer sequences.
   - **RNNs:** Traditional RNNs are used for comparison. They are simpler but tend to struggle with maintaining context over long sequences due to the vanishing gradient problem.

2. **Dataset and Preprocessing:**
   - The dataset typically includes a large corpus of song lyrics which are preprocessed to remove noise, tokenize words, and structure the data into sequences that the models can learn from.
3. **Training and Evaluation:**
   - The models are trained on the preprocessed dataset using techniques like backpropagation through time (BPTT).
   - Evaluation metrics include BLEU (Bilingual Evaluation Understudy), ROUGE (Recall-Oriented Understudy for Gisting Evaluation), and human judgment to assess the coherence, creativity, and stylistic accuracy of the generated lyrics.
4. **Challenges and Solutions:**
   - **Maintaining Coherence:** Ensuring that the generated lyrics make sense and follow a logical sequence.
   - **Stylistic Consistency:** The generated lyrics should reflect the style of the training data, capturing the unique features of the original songs.
   - **Overfitting:** Regularization techniques like dropout and early stopping are used to prevent the models from overfitting to the training data.

Findings:

- LSTM networks generally outperform traditional RNNs in generating coherent and contextually relevant lyrics.
- The generated lyrics are evaluated to be closer to human-like writing, with better handling of long-term dependencies.

Paper 2: Deep Learning in Musical Lyric Generation: An LSTM-Based Approach
Authors: Harrison Gill, Daniel Lee, and Nick Marwel
Published in: Journal of Artificial Intelligence Research (JAIR 2021)
Objective:

- To explore the capability of LSTM networks in generating genre-specific lyrics that maintain stylistic and contextual coherence.

2. **Methodology:**
   - **LSTM Network:** The primary architecture used for generating lyrics. LSTM networks are chosen for their ability to handle long-term dependencies and sequence data effectively.
   - **Dataset:** The training data consists of lyrics from various genres, which are preprocessed and tokenized. Each song's lyrics are used to train the LSTM model to learn the patterns and structures typical of the genre.
   - **Training:** The model is trained using sequences of lyrics, optimizing it to predict the next word in a sequence given the previous words.

3. **Evaluation:**
   - The generated lyrics are evaluated using linguistic metrics to assess how closely they match the patterns and styles of the training data.
   - Metrics such as average line length, word variation, and thematic consistency are used to evaluate the quality of the generated lyrics.
   - Human evaluators also assess the quality, coherence, and creativity of the generated lyrics.

4. **Findings:**
   - The LSTM model successfully generates lyrics that capture the stylistic elements of the target genres.
   - The generated lyrics exhibit average line lengths and word variations similar to the original training data.
   - Thematic consistency is generally maintained, with the generated lyrics fitting well within the specified genre.

Conclusion:

- The study demonstrates that LSTM networks are effective in generating genre-specific lyrics that are coherent and stylistically consistent with the training data.
- Future work could focus on improving the diversity and creativity of the generated lyrics by incorporating additional context and fine-tuning the model.

Assessment Methodology:

Generation Performance:

- **BLEU (Bilingual Evaluation Understudy):** Measures the similarity between the generated text and a reference text.
- **ROUGE (Recall-Oriented Understudy for Gisting Evaluation):** Assesses the overlap of n-grams between the generated text and reference text.
- **PERPLEXITY :** Considers precision, recall, and synonymy matching for evaluating generated text.

Cross-Validation Strategy: Use k-fold cross-validation to ensure that the model generalizes well to unseen data. This involves splitting the data into k subsets and training the model k times, each time using a different subset as the validation set and the remaining subsets as the training set. The results are then averaged to obtain a robust estimate of the model's performance.

Ablation Settings: Ablation studies will be conducted to evaluate the impact of various factors on the performance of the lyrics generation models. The following settings outline the different dimensions, pre-processing steps, and algorithm complexities that will be altered systematically to understand their effect on the model's performance.