GENERATING MUSIC FROM LITERATURE USING TOPIC EXTRACTION AND SENTIMENT ANALYSIS

[1] Aafia Nausheen, [2], Abhilasha Yadav, [3] Ashley Thomas, [4] Gayathri M
Under the guidance of Mrs. Varsha Nilugal, Assistant professor,
Department of Information Science and Engineering, The Oxford College of Engineering,
Bangalore, Karnataka, India.

Abstract: This study presents Tambr, a new software for translating literature into audio using Topic Extraction and Sentiment Analysis for the way in which their timbre relates to the meaning and sentiment of the topics conveyed in the story. The durations, intervals and pitches of the output audio are generated using sentiment analysis which corresponds to sentiment of the text. Natural language processing algorithms are used for topic extraction in the text. Sentiment analysis is used to vary the intensity of the audio. The sounds generated by the system are very characteristics of ambiance music, as the emphasis was placed on selecting musical timbres that match with the themes of the text.

Index Terms **- Topic Extraction, Sentiment Analysis, Natural Language Processing, Text Summary, Tambr.**

I. INTRODUCTION

Musical timbre is one of the most defining characteristics of how a piece of music sounds. Timbre refers to the character and quality of a sound, as opposed to the pitch or loudness. A saxophone and an electric guitar, for example, could play the same notes, but would still be distinguishable as different instruments; that difference is in their timbre. Because of the large number of features that define the timbre of a sound, be it from anacoustic musical instrument or a computer generated synthesizer, academic discussion of timbre is often limited.

TransProse is an existing software that automatically generates musical pieces from text. TransProse uses known relations between elements of music such as tempo and scale, and the emotions they evoke.

Further, it uses a novel mechanism to determine sequences of notes that capture the emotional activity in text. Transpose focuses on novels and generate music that captures the change in the distribution of emotion words.

Although, TransProse has several advantages, there are certain challenges that it faces such as generating sequences of notes, given the infinite possibilities of pitch, duration, and order of the notes. Computational approaches to analyzing timbre are still in their early stages and are often limited. It also lacks intentional harmony and discord between the melodies. It is unable to evaluate the impact of textual features such as the length of the novel and the style of writing on the generated music.

Tambr is a new software that we are proposing for translating literature into sound using multiple synthesized voices selected for the way in which their timbre relates to the meaning and sentiment of the topics conveyed in the story. It uses sentiment analysis to generate the pitches, durations, and intervals of the output audio in a way corresponding to the sentiment of the novel. It also uses natural language processing algorithms to extract the topics in the novel. It varies the intensity of notes based on sentiment analysis.It takes musical timbre into account when selecting voices.

There are several applications of the proposed software which include helping People with learning disabilities. Some people have difficulty reading large amounts of text due to dyslexia and other learning disabilities. Translating text into audio on the basis of sentiment helps. Supporting people who have literacy difficulties. Some people have basic literary levels. By offering them an option to hear the text instead of reading it, they can get valuable information in a way that is more comfortable for them. Aiding people who speak the language but do not read it - Many people who come to a new country learn to speak and understand the native language effectively, but may still have difficulty reading in a second language. This allows them to take in the information in a way they are more comfortable with, making your content easier to comprehend and retain. Assisting people who multitask. A busy life often means that people do not have time to do all the reading they would like to do online. Having a chance to listen to the content instead of reading it allows them to do something else at the same time. It benefits people with visual impairment**.** Text to speech can be a very useful tool for the mild or moderately visually impaired. Even for people with the visual capability to read, the process can often cause too much strain to be of any use or enjoyment. With text to speech, people with visual impairment can take in all manner of content in comfort instead of strain. Guiding people with different learning styles**.** Some people are auditory learners, some are visual learners, and some are kinesthetic learners – most learn best through a combination of the three. This system helps people to read with better understanding of the pronunciation of words. These are some of the main applications of the proposed system. It is very helpful for millions of people in their day-to-day lives. This system is useful for children as well as adults. There are several other applications but only a few have been listed in this article.

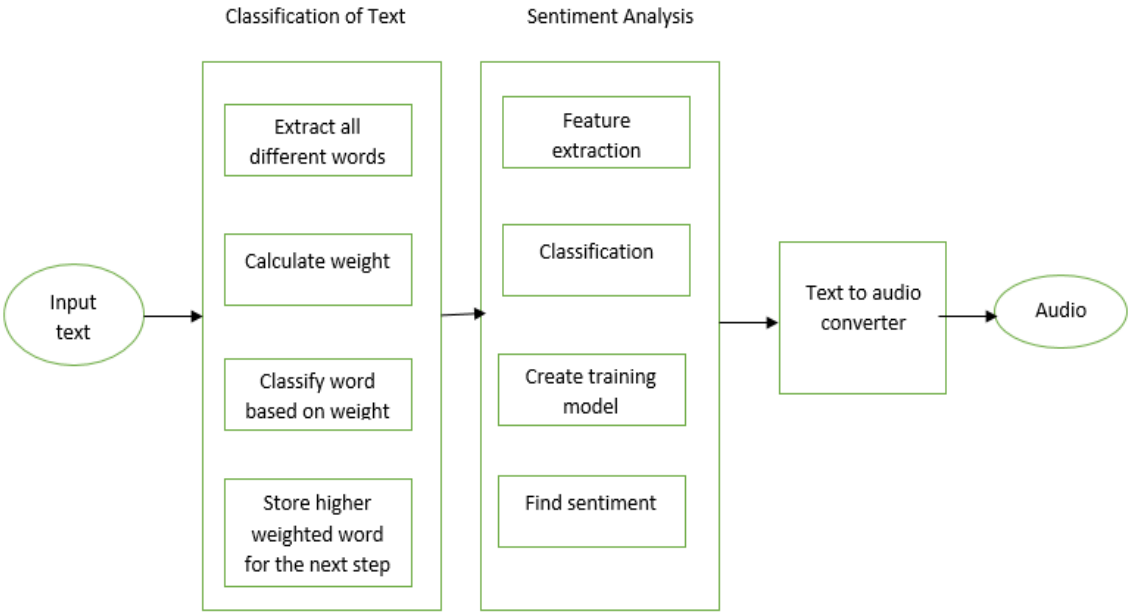

Figure 1: Proposed System Architecture

1.1 Text Summarization

Text summarization is the problem of creating a short, accurate, and fluent summary of a longer text document. Automatic text summarization methods are greatly needed to address the ever-growing amount of text data available online to both better help discover relevant information and to consume relevant information faster. There is an enormous amount of textual material, and it is only growing every single day. Think of the internet, comprised of web pages, news articles, status updates, blogs and so much more. The data is unstructured and the best that we can do to navigate it is to use search and skim the results. There is a great need to reduce much of this text data to shorter, focused summaries that capture the salient details, both so we can navigate it more effectively as well as check whether the larger documents contain the information that we are looking for. Automatic text summarization, or just text summarization, is the process of creating a short and coherent version of a longer document. There are two main approaches to summarizing text documents. They are extractive methods and abstractive methods. They are explained further in detail.

1.1.1 Extractive Methods

Extractive text summarization involves the selection of phrases and sentences from the source document to make up the new summary. Techniques involve ranking the relevance of phrases in order to choose only those most relevant to the meaning of the source.

1.1.2 Abstractive Methods

Abstractive text summarization involves generating entirely new phrases and sentences to capture the meaning of the source document. This is a more challenging approach but is also the approach ultimately used by humans. Classical methods operate by selecting and compressing content from the source document.

1.2 Sentiment Analysis

Natural language processing is only half the battle though. Human communication isn't just words and their explicit meanings. Human communication is nuanced and complex. You can tell based on the way a friend asks you a question whether they're bored, angry, or curious. Sentiment is like a combination of tone of voice, word choice, and writing style all rolled into one. As shown in Fig.2, the input to natural language processing will be a simple stream of Unicode characters (typically UTF-8). Basic processing will be required to convert this character stream into a sequence of lexical items (words, phrases, and syntactic markers) which can then be used to better understand the content. The basics include: Structure extraction – identifying fields and blocks of content processing and understanding. Tokens can be words, numbers, identifiers or punctuation (depending on the use case). Lemmatization / Stemming which reduces word variations to simpler forms that may help increase the coverage of NLP utilities. Lemmatization is strongly preferred to stemming if available. Search Technologies has lemmatization for English and our partner, Basis Technologies, has lemmatization for 60 languages**.** Decompounding for some languages (typically Germanic, Scandinavian, and Cyrillic languages), compound words will need to be split into smaller parts to allow for accurate NLP. Entity extraction for identifying and extracting entities (people, places, companies, etc.) is a necessary step to simplify downstream processing. There are several different methods: regex extraction, dictionary extraction, complex pattern-based extraction and

statistical extraction. Phrase extraction which extracts sequences of tokens (phrases) that have a strong meaning which is independent of the words when treated separately. These sequences should be treated as a single unit when doing NLP. For example, "Big Data" has a strong meaning which is independent of the words "big" and "data" when used separately. All companies have these sorts of phrases which are in common usage throughout the organization and are better treated as a unit rather than separately. Techniques to extract phrases include: part of speech tagging, statistical phrase extraction and hybrid. Based on tagging, Identify and mark sentence, phrase, and paragraph boundaries. These markers are important when doing entity extraction and NLP since they serve as useful breaks within which analysis occurs**.** Language identification will detect the human language for the entire document and for each paragraph or sentence. Language detectors are critical to determine what linguistic algorithms and dictionaries to apply to the text**.** Tokenization in order to divide up character streams into tokens which can be used further.

1.3 Decide on Macro versus Micro Understanding

Before you begin, you should decide what level of content understanding is required:

1.3.1 Macro Understanding provides a general understanding of the document as a whole. It is typically performed with statistical techniques which are used for: clustering, categorization, similarity, topic analysis, word clouds, and summarization.

1.3.2 Micro Understanding extracts understanding from individual phrases or sentences. It is typically performed with NLP a technique which are used for: extracting facts, entities (see above), entity relationships, actions, and metadata fields.

1.4 Macro Understanding

Once you have decided to embark on your NLP project, if you need a more holistic understanding of the document this is a "macro understanding." This is useful for: classifying / categorizing / organizing records, clustering records, extracting topics, general sentiment analysis, record similarity, including finding similarities between different types of records (for example, job descriptions to résumés / CVs), keyword / key phrase extraction, duplicate and near-duplicate detection, summarization / key sentence extraction and semantic search.

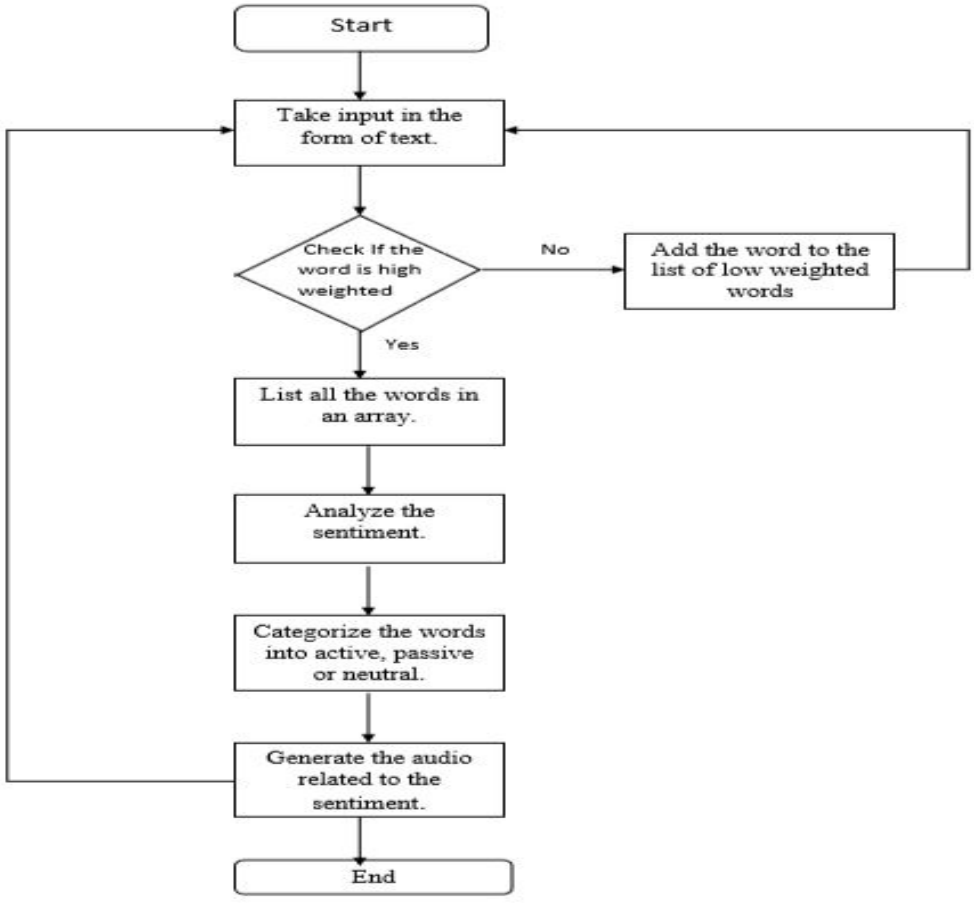

Figure 2: Flow Chart

II. ALGORITHM

1. Take the full CONTENT and split it into PARAGRAPHS.
2. Split each paragraph into SENTENCES.
3. Compare every sentence with every other. This is done by Counting the number of common words and then Normalize this by dividing by average number of words per sentence
4. These intermediate scores/values are stored in an INTERSECTION matrix
5. Create the key-value dictionary

    - Key: Sentence
    - Value: Sum of intersection values with this sentence

6. From every paragraph, extract the sentences with the highest score.
7. Sort the selected sentences in order of appearance in the original text to preserve content and meaning.

III. CONCLUSION

The software will transform each paragraph of the text providing a weight of how relevant each word is to a given paragraph based on the number of other paragraphs in which it appears. It generates a list of topics, where each topic is a set of, at most, ten terms that define the topic. It interprets the sequence of sentiment scores for each sentence as a series of signals generated by the sentiment analysis module corresponding a sensible way with the plot structure that a reader can verify in the text. This system will generate audio that is related to the characteristic ambiance of the text. It provides intentional harmony and discord between the audio.

IV. ACKNOWLEDGEMENT

We would like to specially thank Mrs.VarshaNilugal, Assistant Professor, Department of Information Science, The Oxford College of Engineering, for her advice, contribution and encouragement throughout this project. We would also like to thank Dr. D Jayaramiah, Professor and Head of Information Science Department, The Oxford College of Engineering, for his constant support and encouragement.

REFERENCES

[1] K. Siedenburg, I. Fujinaga, and S. McAdams, *"A comparison of approaches to Timbre descriptors in music information retrieval and music psychology,"* J. New Music Research, vol. 45, no. 1, 2016.
[2] H. Davis and Saif M. Mohammad, *"Generating music from literature,"* in Proc. EACL Workshop on Computational Linguistics for Literature, 2014.
[3] P.N. Vassilakis, *"Representing sound energy, phase, and interference using three-dimensional signals"* presented at the 153nd meeting of the Acoustical Society of America, Salt Lake City, UT, 2007
[4] Eck, D., and Lapalme, J. 2008. *"Learning musical structure directly from sequences of music"*, UniversityofMontreal, Department of Computer Science, CP 6128.
[5] H. L. F. Helmholtz, *"On the Sensations of Tone as a Physiological Basis for the Theory of Music"*, 2nd ed., A.J. Ellis, Trans. New York: Dover Publications, Inc., 1885 (1954).
[6] G. von Békésy, *"Experiments in Hearing. New York: Acoustical Society of America Press"*, 1960 (1989).
[7] E. Terhardt, *"On the perception of periodic sound fluctuations (roughness)"*,Acustica, vol. 30(4), pp. 201-213, 1974.
[8] R. Plomp, *"The ear as a frequency analyzer"*, J. Acoust. Soc. Am., vol. 36(9), pp. 1628-1636, 1964.
[9] E. Zwicker, *"Subdivision of the audible frequency into critical bands,"* J. Acoust. Soc. Am., vol. 33(2), pp. 248-249, 1961.
[10] M. Campbell and C. Greated, *"The Musician's Guide to Acoustics"*, 2nd ed. New York: Oxford University Press, 1994.