International Journal of Social Science Studies
Vol. 13, No. 3; September 2025
ISSN 2324-8033 E-ISSN 2324-8041
Published by Redfame Publishing
URL: http://ijsss.redfame.com

Analysis of Disaster Scene Narratives Generated by Language Models to Provide Relevance with Individuals

Shogo Matsumoto[1], Hiromitsu Shimakawa[1], Fumiko Harada[1]

[1] Graduate School of Information Science and Engineering, Ritsumeikan University, Osaka, Japan

Correspondence: Shogo Matsumoto, Graduate School of Information Science and Engineering, Ritsumeikan University, Osaka Ibaraki Campus (OIC), 2-150 Iwakura-cho, Ibaraki City, Osaka 567-8570, Japan. E-mail: sritsuaka@gmail.com

Received: April 27, 2025 Accepted: June 6, 2025 Available online: July 14, 2025

doi:10.11114/ijsss.v13i3.7679 URL: https://doi.org/10.11114/ijsss.v13i3.7679

Abstract

Despite the fact that Japan is a country prone to natural disasters, the participation rate in evacuation drills, especially among young people, has remained flat in recent years. This is considered the fact that it is difficult to accept disasters that have occurred in other regions as their own. As a result of preliminary experiments, it became clear that viewers tended to be less likely to feel it as their own if they simply watched disaster videos.

The purpose of this study is to generate disaster site episodes that viewers feel as their own, and to clarify the characteristics of the episodes that are common to them. To avoid black boxing, we used a small AttentionSeq2Seq model. We selected tsunami and landslide disaster site episodes that viewers felt as their own, generated using ChatGPT based on a questionnaire, and trained the model to output sentences that were identical to the input sentences. After training, the test data was fed into the trained model and we analyzed the generated episodes and the attention texts based on the attention maps and examined the final layer of the encoder. As a result of the analysis, it was confirmed that the model captures the factors that viewers feel as their own for each disaster type. By including these elements in the training data, it was shown that even a small amount of data could generate disaster episodes that felt more familiar to the viewer.

Keywords: AttentionSeq2Seq, Tsunami Attention Texts, Landslide Attention Texts, episodes, ChatGPT

1. Introduction

Japan is a disaster-prone country and evacuation drills are one of the most effective ways to minimize disaster damage (Gwynne et al., 2020). However, in recent years, the number of participants, especially young people, in evacuation drills has been declining. One reason for this is that people do not view disasters occurring elsewhere as their own. It is thought that if people do not view disasters as their own, their motivation to participate in evacuation drills will decrease, leading to a decline in their participation in evacuation drills.

There are studies using Augmented Reality (AR) and Virtual Reality (VR) techniques to help evacuation training (Catal, Akbulut, Tunali, Ulug, & Ozturk, 2020)(Feng et al., 2020). However, these studies have not investigated whether the subjects had felt the presented disasters as their own or not. AR and VR utilize videos. We conducted a preliminary experiment using videos of tsunami disaster sites to determine whether people feel the disaster scene as if it were their own or not. The experiment revealed that while the subject could understand that the videos of a disaster scene is scary, she did not feel the disaster scene as their own. This suggests that modern people, who have become accustomed to videos of disasters on television and the Internet, tend to think of situations in disasters that occur elsewhere as someone else's problem. From preliminary experiments, we found that it was necessary to use episodes that viewers feel as their own.

In this study, we use a small language model, to analyze the patterns of effective episodes for each disaster site that viewers feel as their own, in order to convert disaster site episodes that occur elsewhere into disaster site episodes that occur in your own immediate area through the language model. In this study, we use a small language model to learn episodes and generate new episodes. By using a small language model, we can prevent the model from becoming a black box, visualize what information has been learned about disaster scene episodes, and improve the breadth of analysis.

In this study, we firstly extracted the episodes of tsunami and landslide disaster sites from those generated by the ChatGPT (Brown et al., 2020) through a questionnaire evaluating whether the viewers have felt as their own or not. We trained a small language model based on AttentionSeq2Seq with the extracted episodes to generate new episodes. We analyzed the attention maps and the final layer of the encoder to obtain the characteristics of the texts which the viewers felt as their own, for each of the disaster.

From the experimental results, it was found that the tsunami disaster site episodes are characterized by the texts describing the destruction of towns as elements that viewers feel as their own. It was found that the landslide disaster site episodes are characterized by the texts expressing the fear of affected residents as elements that viewers feel as their own.

This paper is organized as follows. Section 2 describes the AttentionSeq2Seq technique. Section 3 shows the preliminary experimental result. Section 4 proposes the methods to analyze the characteristics of the texts viewers feel as their own. The experimental result is shown in Section 5. Section 6 discusses the experimental results. Section 7 concludes this paper.

2. AttentionSeq2Seq

AttentionSeq2Seq (Li et al., 2022) is one of a deep learning model generating a sequential output from a sequential input, including texts. Various studies using AttentionSeq2Seq have been being conducted (Zhang, Deng, & Zhang, 2021). In this study, we aim to analyzing disaster site episodes that viewers feel as their own using AttentionSeq2Seq and generate new disaster site episodes.

AttentionSeq2Seq consists of the components that incorporate the Attention mechanism (Bahdanau, Cho, & Bengio , 2014) into Seq2Seq (Sutskever, Vinyals, & Le,2014). Seq2Seq consists of two components: an encoder and a decoder. The encoder and decoder have the Long Short-Term Memory (LSTM) (Staudemeyer & Morris, 2019) structure and the LSTM cells are connected in chronological order.

LSTM overcomes the vanishing gradient problem and exploding gradient problem in traditional Recurrent Neural Networks (RNNs). The basic unit of LSTM is a block structure called a memory cell. A memory cell has three types of gates: Forget Gate, Input Gate, and Output Gate, and one memory called Cell State. The Forget Gate determines how much of the information stored in the Cell State from the past should be forgotten. The Input Gate determines new information to add to the Cell State from the current input and previous hidden state. The Input also calculates new candidate Cell States. Output Gate determines how much Cell State is output when generating the next hidden state. The Cell State has a structure that stores or forgets important information for a long time and updates it using the output of Forget Gate and Input Gate. The hidden state is calculated using the Output Gate and updated Cell State.

The attention mechanism is a mechanism in which each time step i in the decoder output sequence dynamically accesses all hidden states output $h_1, h_2 \dots h_n (i = 1,2, \dots n)$ by all encoders and calculates their importance. With this mechanism, AttentionSeq2Seq can pay attention to highly relevant parts of the input sequence. By visualizing the importance, it is possible to observe the degree of attention (Hao, Lee, & Zhao, 2019).

3. Preliminary Experiment

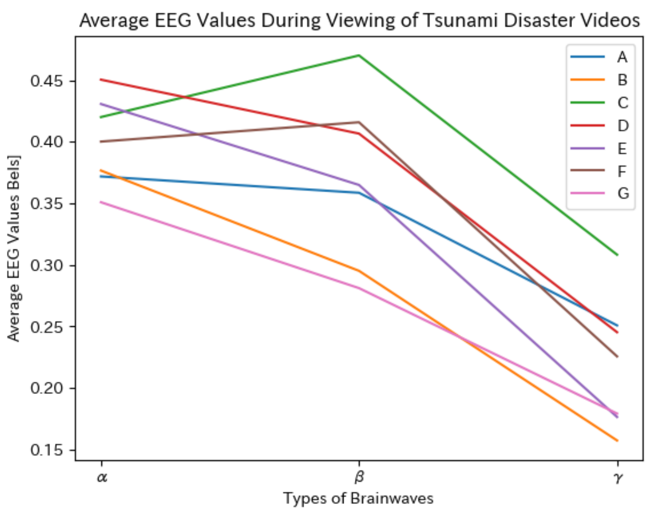
Studies on simulations of evacuation training for disasters using AR and VR have been conducted in the past. Simulated evacuation drills using videos such as AR and VR improve disaster prevention knowledge (Feng et al., 2020). Moreover, by using videos, it is easier to understand evacuation drills (Catal, Akbulut, Tunali, Ulug, & Ozturk, 2020), which may be because the video allows the viewers imagine the actual evacuation site However, such studies have focused only on the content of evacuation drills and improvement knowledge about evacuation and have not focused on whether people perceive the video as their own or not. If people do not perceive it as their own, they may lack the motivation to participate in evacuation training.

Thus, we conducted a preliminary experiment to verify the effect of the videos on disaster sites. In the preliminary experiment of this study, one female subject in her 20s was asked to watch seven videos of the tsunami disaster site. Tsunami is a disaster that causes a lot of damage in Japan. To increase fear while watching videos, we added three types of external influences (Choi, Bang, Heo, & Park, 2015) (Komatsu, Kawai, & Sakaguchi, 2018); Takahashi, 2020). We examined whether the subject who felt more fear perceive the content of the videos as ones that she feels as her own. For each video, the brain wave data and the Electrodermal activity (EDA) values were collected with and without external influences (normal state). The three types of external influences are: (1) water applied to the feet, (2) wind applied to the feet, and (3) sound changed. When the sound is changed, only the tsunami sounds included in the tsunami video are extracted as much as possible. By applying mathematical processing to the audio data, we extracted audio only in a specific frequency range. We conducted an experiment by replacing it with the audio from the original video. The seven videos of the tsunami disaster site were collected from YouTube. Each video was labeled as A (vientorio, 2011), B (Yomiuri Shinbun onrain dōga, 2011), C (ANNnewsCH, 2020), D (WankoDon, 2011), E (FNN311, 2013), F (TBS NEWS DIG Powered by JNN, 2021), and G (Wēzā Nyūsu Sutadi, 2023), respectively. All the videos were on a tsunami approaching and destroying the town. The video D included an interview with the cameraman who shot the video of the tsunami disaster site. In the preliminary experiment, in order to reduce the detention time of the subject as much as possible, we conducted the experiment using the external influence (1) on of the videos A, B, and C, the external influence (2) on the videos D and E, and the external influence (3) on the videos F and G, respectively. We denote the video added the external influence (1), (2), and (3) by adding "_water", "_wind", and "_sound" after the video's name, respectively.

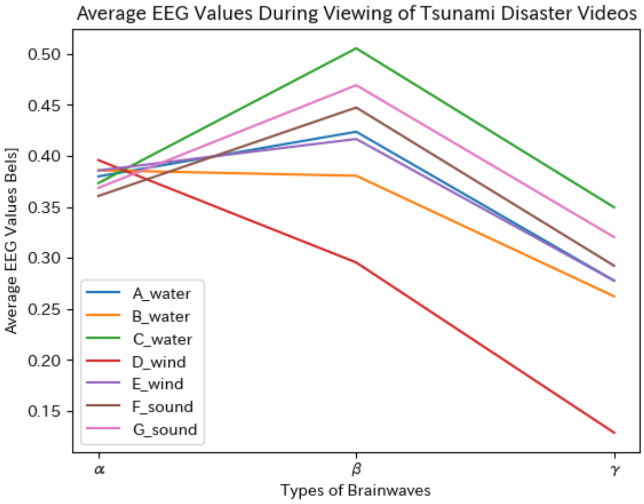
In each video, data are acquired by the Electroencephalography (EEG) devices Muse2 (InteraXon, n.d.)(Saia, Carta, Fenu, & Pompianu, 2022) and E4wristband (Empatica Inc, n.d.) at any time. Muse2 acquires three types of EEG from two frontal locations and two temporal locations. The EEG data to be acquired are α waves, β waves, and γ waves. α waves tend to increase when people are relaxed and decrease when they are stressed or nervous (Clayton, Yeung, & Cohen Kadosh, 2018). β waves, also known as stress waves, tend to increase when we are stressed, such as when we are afraid, when we make complex calculations, when we do something, we don't like, or when we are frustrated (Herrmann, Strüber, Helfrich, & Engel, 2016). γ waves are involved in the processing and integration of visual information, and tend to increase in concentrated states (Hughes, 2008)(Schneider et al., 2018). In the data obtained from the four locations, the average value of the four locations was obtained, and the average value for the number of data was obtained and compared. E4wristband acquires EDA. EDA refers to the electrical properties of the skin in response to the secretion of sweat on human skin (Benedek & Kaernbach, 2010). Changes in human emotions affect EDA (Dawson, Schell, & Filion, 2007). It has been found that drastic change in EDA implies change of emotions (Seguchi, 2022)(Uehara, Shimakawa, & Harada, 2024).

After the preliminary experiment, we asked whether she felt the content of the video as her own or not, and her opinions on the videos that left an impression on her . Opinions could be given freely. we analyzed the changes in EDA and brain waves when watching the videos and investigated whether the subject felt the contents of tsunami disaster videos as her own, taking into account the opinions after the preliminary experiments.

Figure 3.1 shows the results of the average values of each EEG during the viewing of the videos, both in the state without the external influences (normal state) and after adding external influences.



(a) Average Value of Each EEG under Normal State



(b) Average Value of Each EEG under External Influences

Figure3.1 Average EEG Values for Each EEG When Viewing Seven Types of Tsunami Disaster Site Video

From Fig. 3.1(a), it can be seen that the α wave is the lowest in the video G. β and γ waves are low in G similar to the video B. It is also clear that the β and γ waves are highest in the video C. From Fig. 3.1(b), it can be seen that the value of the α wave is decreasing, and the β and γ waves tend to increase.

For the EDA results, there were occasional small changes in all cases, but almost no significant changes. Figure 3.2 shows EDA for the video C under normal state.

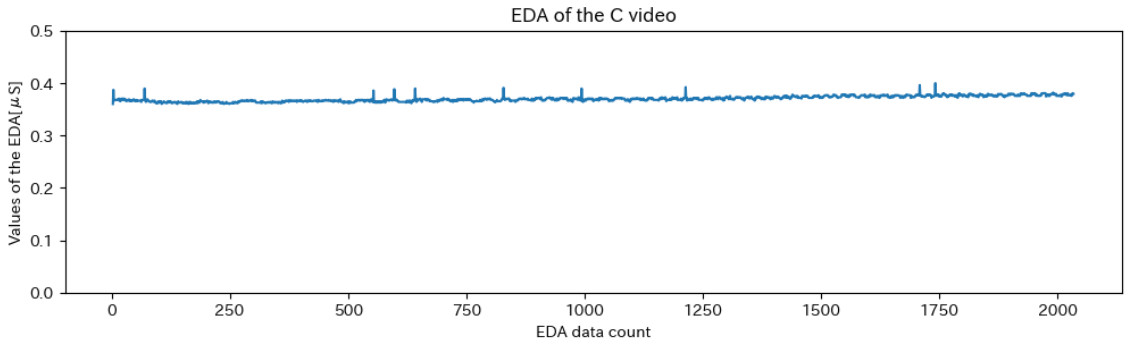


Figure 3.2. EDA Value for C Video under Normal State

From Figure 3.2, there are an occasional changes, but it is about 0.03 and relatively small. The changes are also temporary.

Regarding the opinions after the preliminary experiment, she said, “I understood that the content of the videos themselves were scary, but I did not feel the video contents as my own.” Moreover, the scariest video that left an impression on her was the video A. Another thing that impressed her was the video D, which was commented as “stimulated my imagination and increased fear.” The video A is a video of the destruction of a city inhabited by residents. Along the way, the voices of the victims are included, and the contents are about the destruction of their own homes. The video D contains mainly an interview with a victim of the tsunami disaster.

From the results of Figure 3.1, the α wave decreased at a rate of 4/7 compared to the case under normal conditions. This indicates that people are less relaxed when they are in the external influences. For the β waves, the rate increased by 6/7. This indicates that people are more likely to feel tension, stress, and fear when they are exposed to the external influences. For the γ waves, the rate increased by 6/7. This indicates that visual attention and conscious processes have become active due to the external influences. These are thought to represent the characteristics of feeling fear. From the EDA results, it can be seen that the subject had almost no emotional changes even if she was affected by the external influences. In addition, from the opinions of the subject in the preliminary experiment, it was found that she did not feel the content of the videos of the tsunami disaster site as their own. From these results, it can be considered that just watching videos of the tsunami disaster site does not make the viewers feel their contents like their own. On the other hand, the video that left a strong impression on the subject, as it stirred their imagination and increased her fear, is the video D. This showed a woman who was about to be swallowed by the tsunami and thought she could not be saved but managed to save her life in a story-like manner. According to the above results, it can be considered that having the subject read episodes about the disaster site and encouraging her to imagine it would make her feel more personally connected to the content of the disaster, rather than just showing her the video.

4. Pattern Analysis of Involvement-inspiring Disaster Site Episodes Using AttentionSeq2Seq

4.1 Involvement-inspiring Disaster Site Episodes and Small Language Model

From the preliminary experiment, it was found that while the subject understands that a disaster is dangerous and evokes fear by watching disaster-related videos, she did not feel the content of the disaster videos as her own. In modern times, with the spread of SNS and the Internet society, it is thought that we have become accustomed to seeing images of disasters. In this study, based on disaster episodes that occurred elsewhere, we used a small language model to generate disaster site episodes occurring in the subject's own local area, making the content of the episodes feel as their own. The natural phenomena caused by disasters and the nature of disasters differ between other regions and their own regions. To generate disaster episodes in your own region, it is required to convert episodes of other regions to suit your own region. Thus, we analyze textual episodes of different natural disasters and extract their characteristics. In this study, we use a small language model, AttentionSeq2Seq, to investigate attention to analyze the episodes.

AttentionSeq2Seq is an encoder-decoder model that includes an attention structure. Attention stores information about what words are important in each episode for the two types of disaster episodes to be learned. Additionally, by visualizing the information stored in the attention, it becomes possible to analyze the two types of disaster site episodes elements that viewers feel as their own. In recent years, AI models (Zhang et al., 2022) that perform tasks to generate sentences have a large number of parameters and are capable of complex language learning. However, it requires a large amount of training data, and the cost of training is enormous. In addition, due to the complexity of the structure, the contents of the model at the time of training are black boxes, which makes it difficult to discover the elements necessary for optimal learning. In this study, we aim to train an AttentionSeq2Seq model by the episodes of two different types of disasters, white box the model, and clarify the elements necessary for learning. The overview of the proposed method is shown in Figure 4.1.

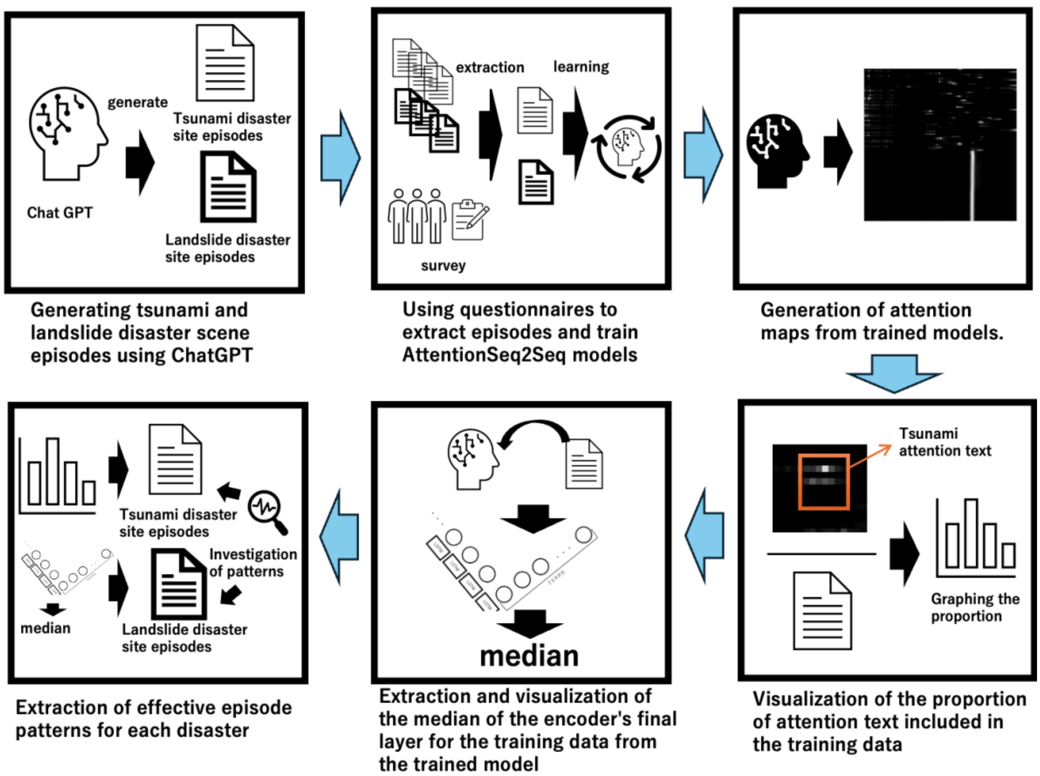


Figure 4.1 Overview of the Proposed Method

In this study, we first use ChatGPT to generate disaster site episodes related to tsunamis and landslides. From the generated episodes we extract the episodes that the viewers feel as their own through a questionnaire. These extracted episodes is utilized as the training data of an AttentionSeq2Seq model. An attention map from the trained AttentionSeq2Seq is generated. For output of the model, we extract texts with large attention among the test data(hereinafter referred to as attention text). Use a graph to visualize what percentage of the total training data text the attention text contains. Additionally, the training data is input to the trained AttentionSeq2Seq model, and the median of the hidden state vector, which is the final layer output of the encoder obtained at that time, is extracted and visualized. Finally, based on the information visualized from the attention text and the visualized information of the median of the encoder's final layer output, we analyze the pattern of effective episodes per disaster.

4.2 Disaster Site Episode Generation Using ChatGPT

In this study, we use ChatGPT to generate disaster site episodes related to tsunamis and landslides to secure training data in a format suitable for small language models. Japan is a disaster-prone country, and although it has done a lot of disaster prevention especially against earthquakes, it has a past of carrying out large debts due to tsunami disasters after earthquakes. It is considered that the damage could have been suppressed if evacuation drills for tsunamis had been conducted on a regular basis. In this study, we will generate tsunami disaster site episodes as symbols of disasters that occurred in other regions of the subject and caused great damage to Japan. It also generates disaster site episodes of landslides as disasters that occur in the subject's immediate area. If episodes for each disaster are generated under different conditions, the nature of the data may be biased and affect the training results of the model. In this study, we unified the prompt content given to ChatGPT in episode generation for each type of disaster. The prompts used to generate the tsunami disaster site episode and the landslide disaster site episode were designed based on the policies in Table 4.1.

Table 4.1. Prompt Design Policy

Generate episodes that include the characteristics of each disaster
Prompt chatGPT to include the characteristics of each tsunami and landslide disaster and have it generate a disaster site episode. If the generated episodes are exactly the same, we prompt it to generate a new episode including each disaster-likeness.
Number of characters
When generating an episode, give a prompt so that the number of characters is about 100 characters.

Regarding the number of characters policy in Table 4.1, if we let ChatGPT generate many short texts, the content will be almost the same. It is necessary to generate as long a text as possible to avoid overshadowing the content. However, in this study, we train on a small language model. For the above reasons, in this study, the length of the data used for training was set at about 100 characters.

4.3 Questionnaire on The Generated Disaster Site Episodes

We conduct a questionnaire to extract text that viewers feel as their own from the disaster site episodes generated using ChatGPT. The questionnaire items and evaluations are shown in Table 4.2.

Table 4.2. Questionnaire Items and Five-Point Scale Evaluation

Questionnaire item	
Read each episode to see if the content feels realistic	
Evaluation	
1	I can't feel it at all
2	I don't really feel it
3	I can't say either way
4	Feels so-so
5	I feel it very muc

In the questionnaire in Table 4.2, the number assigned to this five-point scale is the evaluation score as it is. We use the questionnaire results to calculate the average score from the scores of the participants for each disaster site episode. We then extract episodes with an average score of 3.5 or higher and use some of them as training data to train the Attention

Seq2Seq model.

4.4 Generation of Attention Maps

In this study, we generate attention maps using the trained AttentionSeq2Seq. The attention map is a visualization of how the attention mechanism in AttentionSeq2Seq assigns weights to the input data. The trained AttentionSeq2Seq receives the text of a new disaster site episode as input data and generates a disaster scene episode from the obtained input data. The attention map is a heatmap visualization of the attention scores for each text, showing which text in the input data is important for each character in the generated data sentence. In this study, we use an attention map to improve interpretability by analyzing which part of the input the trained AttentionSeq2Seq focuses on for the data generated based on the input data. An example of an attention map is shown in Figure 4.2.

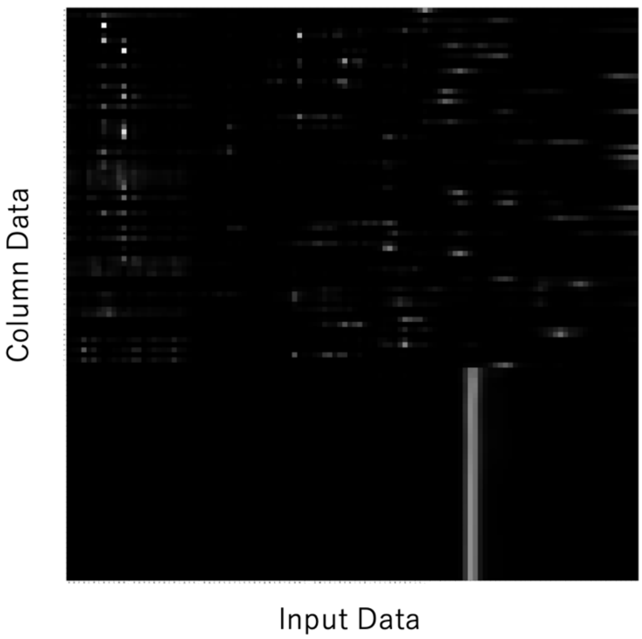


Figure 4.2. Example of Attention Map

Figure 4.2 shows an example of an attention map when the trained AttentionSeq2Seq is used for text prediction using tsunami test data as input data. The horizontal axis represents the input data, and the vertical axis represents the target column data. In this study, Input Data is inverted and input to AttentionSeq2Seq for the purpose of improving learning. Therefore, when inputting test data to the trained model, Input Data is also inverted. The Input Data and Column Data in Figure 4.2 are shown in Table 4.3.

Table 4.3. Input and Column Data from Figure 4.2

Input Data	津波が港に押し寄せた瞬間、停泊中の船が波に押し流され、陸地に打ち上げられた船が建物を次々に破壊し、瓦礫と共に再び海へと戻っていく光景が広がった。
Onput Data	波が港に押し寄せた瞬間、停泊中の船が波に押し流され、陸地に打ち上げられた船が建物を次々に破壊し、瓦礫と共に再び海へと戻っていく光景が広がった。

In the example of Figure 4.2, there is a dark white color vertically in the upper left corner of the figure, and occasionally a region with a slight gray color next to it, and attention is paid to the text“瞬間”(It means“a moment”) in the input data. By generating an attention map in this way, it is possible to check which part the trained model focuses on when generating a new output from a new input.

4.5 Extracting Attention Text

The text in the input data corresponding to the white area on the attention map is defined as the attention text. We give the trained AttentionSeq2Seq test data as input. From the resulting attention map, extract the attention text for each tsunami and landslide. Some attention maps may have relatively similar white areas. In this study, three attention texts are extracted from one attention map. For the three attention texts, the attention map is visually checked, three locations are selected in order from the whitest part, and the corresponding text is extracted. By extracting multiple attention keywords, the range of analysis of what the trained AtentionSeq2Seq is learning expands. The percentage of the extracted attention texts included in the total number of training data is calculated and visualized. In addition, we graph the proportion of attention texts for each type of disaster in the number of training data.

4.6 Extraction of Hidden State Vectors and Pattern Analysis of the Final Layer of the Encoder

In this study, we will check what kind of information is stored in the final layer of the encoder of the trained model. The AttentionSeq2Seq model is a model of an encoder/decoder structure that includes an attention structure. The AttentionSeq2Seq model has a structure in which the hidden state vector of the final layer of the encoder is used as input to the decoder. It is considered that the hidden state vector of the final layer of the encoder retains the most features about tsunami-like and landslide-like properties in the encoder of the trained model. We firstly input training data for each disaster site episode of tsunami and landslide into the trained model. The hidden state vector of the encoder's final layer for each disaster site type obtained from the input is extracted and visualized to see if there are any differences between the disasters. Moreover we extract the median of the output from the encoder's final layer. Based on the hidden state vector of the final layer contained in these encoders, its median, and the visualization information of the attention text, we analyze the tsunami-like and landslide-like patterns included in each disaster site episode learned by AttentionSeq2Seq.

5. Experiments and Results

5.1 Small Language Model

In this study, we use AttentionSeq2Seq as a small language model to learn two types of disaster site episodes and generate new disaster site episodes. Figure 5.1 shows the overall image of AttentionSeq2Seq, which is used in this study.

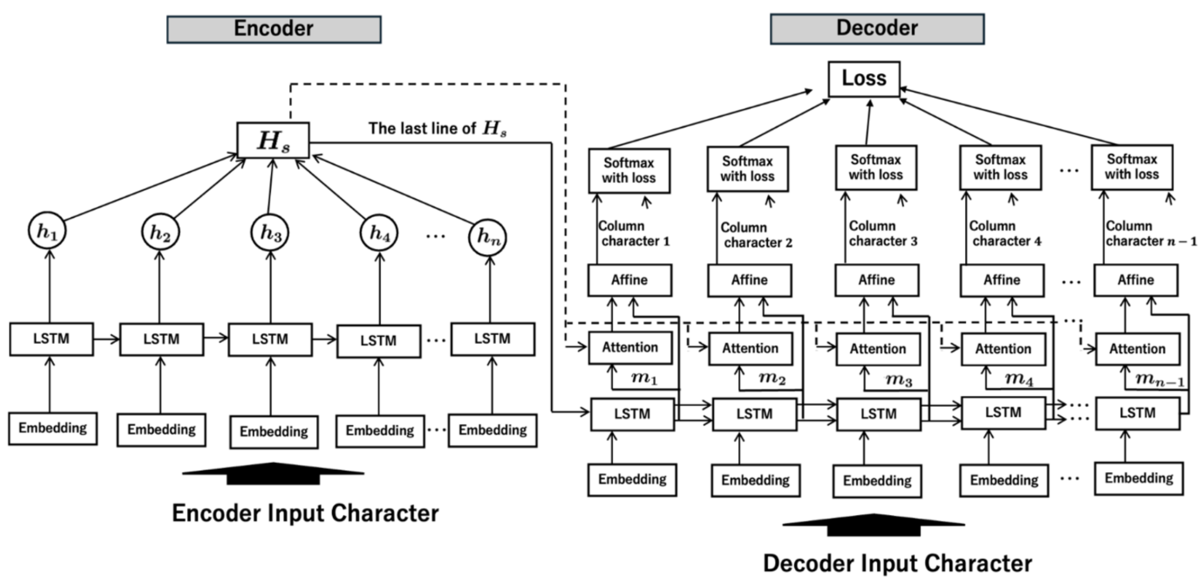


Figure 5.1. Overview of AttentionSeq2Seq

In Figure 5.1, AttentionSeq2Seq has three main structures: an encoder, an attention structure, and a decoder structure. In Figure 5.1, each cell of Embedding vectorizes a numerically-coded Encoder Input Character. The number of dimensions of the vector from Embedding is 16. The LSTM at the beginning of the encoder takes a vectorized number as input and outputs a hidden state h_1. The second and subsequent LSTMs receive the corresponding embedding number and the output of the previous LSTM, and output the hidden state $h_i(i = 2,3,4 \dots n)$. In this study, n is 114. $\boldsymbol{H_s}$ is the summary of the hidden states. The number of dimensions of $\boldsymbol{H_s}$ is (114,256). When the decoder is trained, it receives embeddings of numeric Input Characters, just like the encoder. Secondly, the LSTM at the beginning of the decoder side receives the output of the Embedding and the final layer of $\boldsymbol{H_s}$ as input, and outputs m_1. The second and subsequent LSTMs receive the output of the previous LSTM and output $m_i(i = 2,3,4 \dots n-1)$. Attention receives the output of LSTM and $\boldsymbol{H_s}$, and outputs a context vector. Affine takes the output of the LSTM and the context vector and linearly transforms it. Softmax with loss uses the softmax function to calculate the errors of the column character and outputs the Loss that summarizes them. Using the error backpropagation method, the parameters are updated, and the loss is learned to be low. Table 5.1 shows the hyperparameters of AttentionSeq2Seq in training disaster site episodes.

Table 5.1. The Hyperparameters During AttentionSeq2Seq Training

vocab size	114
wordvec size	16
hidden size	256
batch size	128
max epoch	350
max grad	5

In Table 5.1, vocab size represents the length of the training data string. Wordvec size represents the dimensionality of each character in the string input to the model when it is embedded. Hidden size represents the dimension of the hidden vector output by LSTM. Batch size represents the number of samples of the model to be processed in one iteration during model training. Max epoch represents the maximum number of times the model scans the entire learning model. In this study, since the amplitude of the loss value decreased around 350, 350 times was set as the max epoch. Max grad represents the maximum value of the gradient during model training. Setting max grad prevents gradient explosion during model training and improves the stability of model training. In this study, the length of the string input to the encoder at the time of training is 114 . Since the maximum length of the string of training data obtained through the questionnaire was 114, the data was filled with blanks so that it was 114 characters. The correct label is 113 characters. This is because the correct answer label is the character string after the second character of the string input to the encoder. In this study, we set the length of all the training data to 114 characters. However, since most of the data ends with blank spaces, we modified the correct answer label to start from the second character of the training data onward to reduce the learning of blank spaces. Using the above structure, we trained two types of disaster scene episodes using AttentionSeq2seq and generated new disaster scene episodes.

5.2 Questionnaire Results

We conducted a questionnaire about 100 texts generated using ChatGPT on tsunami disaster site episodes and landslide disaster site episodes. Nine male subjects, including undergraduate and graduate students, were subjected. The five-point-scaled evaluations by the nine subjects were averaged for each disaster site episode. Examples of the episodes with an average evaluation score of 3.5 or higher are shown in Table 5.2 .

Table 5.2. Examples of Disaster Site Episodes with an Average Score of 3.5 or More

Types of disasters	Disaster Site Episode
Tsunami Disaster Site Episode Example 1	津波が港を飲み込み、船舶が波に持ち上げられて陸地に投げ出され、次々と建物に衝突し、街全体が瓦礫と水に覆われる壮絶な光景が広がった。
Tsunami Disaster Site Episode Example 2	津波が町を飲み込み、建物が次々と崩壊し、車や家具が波に押し流され、瓦礫が遠くまで広がって町全体が壊滅していった。
Landslide Disaster Site Episode Example 1	土砂崩れの影響で道路が厚い土砂に覆われ、周囲には助けを求める声が微かに響き、住民たちはその状況を不安な目で見守っていた。
Landslide Disaster Site Episode Example 2	崩れ落ちた山から流れ出た重い土砂が、湿った泥と混ざり合いながら激しく流れ込み、無情にも住宅を押しつぶしていく様子は、住民たちに恐怖を植え付けた。

From Table 5.2, the tsunami disaster scene episode is relatively similar in Example 1 and Example 2. Moreover, the flow of episodes is relatively similar. In the tsunami disaster scene episode, the tsunami first hits the town, and the story is about destroying the town, including the expression of buildings and objects being swept away by the waves. On the other hand, the flow of the episodes at the site of the landslide disaster is relatively similar in Example 1 and Example 2. The first half of the episode is about landslide occurring, and the second half shows the emotional content of the fear of the residents affected by the landslide. When comparing the tsunami disaster scene episode and the landslide disaster scene episode, the flow of the episode was different. In this study, we used these two types of disaster scene episodes to train a small language model, AttentionSeq2Seq.

5.3 Training Results of AttentionSeq2Seq

Table 5.3 shows the size of the training data of the disaster site episode trained by AttentionSeq2seq in this study.

Table 5.3. Data Size of Disaster Site Episodes

Data	Data size
Data of All Disaster Site Episodes	900
Data of Tsunami Disaster Site Episodes	449
Data of Landslide Disaster Site Episodes	451

Figure 5.2 shows the loss graph during training with respect to the correct labels, using training data of disaster site episodes with an average evaluation score of 3.5 or higher.

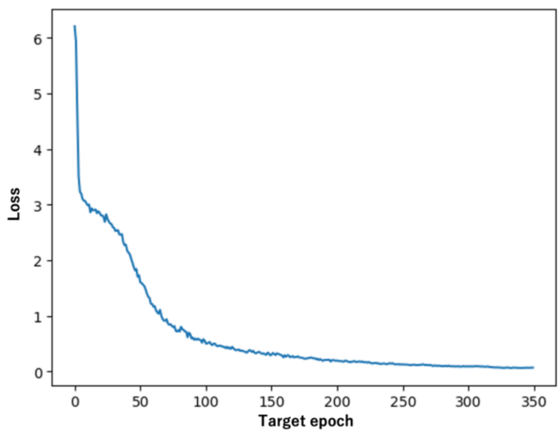


Figure 5.2 Graph of Training Loss Results

In the graph shown in Figure 5.2, the horizontal axis, labeled as "Target epoch," represents the count of the first batch for each epoch during training. The vertical axis represents the loss value for the first batch of each epoch. From Figure 5.2, it can be observed that starting slightly before a Target epoch of 350, the loss values level off, suggesting that they are converging.

5.4 Attention Maps in Generated Statements

The trained AttentionSeq2Seq was used to generate attention maps using 30 test data samples each from the tsunami disaster site episodes and the landslide disaster site episodes. Examples of the attention maps generated using the test data of each disaster site episode is shown in Figure 5.3.

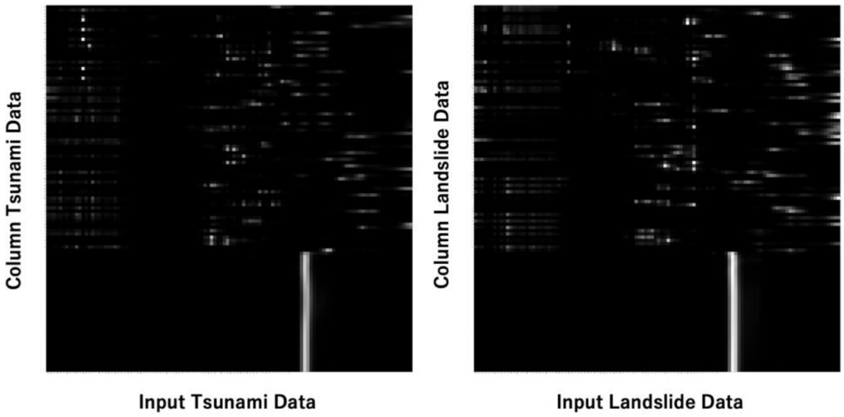


Figure 5.3. Example of Attention Map for Each Disaster Site Episode

The attention map on the left in Figure 5.3 was generated using the test data of the tsunami disaster site episode, and the attention map on the right was generated using the test data of the landslide disaster site episode. The white areas in the attention maps indicate where the trained model is particularly focusing when generating new disaster episodes using the test data. From Figure 5.3, it is considered that the locations of the white areas differ between the tsunami disaster site episode and the landslide disaster site episode. However, in all the attention maps, a white vertical bar-like area appears in the bottom right corner.

Table 5.4 shows the prediction data generated using the horizontal and vertical axes and test data of each of the attention maps in Figure 5.3.

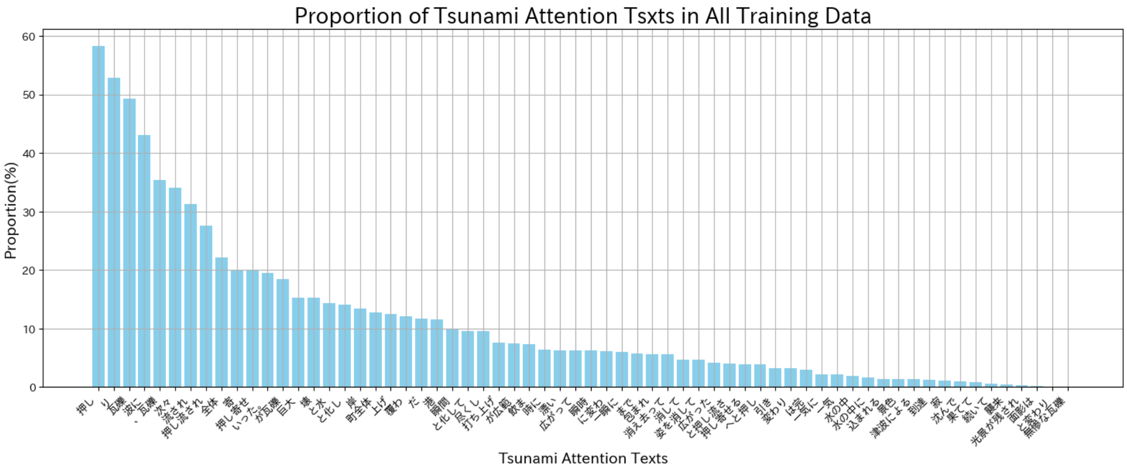
Table 5.4. Attention Map Axes and Generated Prediction Data

Data Name	Text
Input Tsunami Data	津波が海岸を越えて押し寄せると、家々が波に飲まれて崩れ、瓦礫となったものが海へと運ばれていく中、街の景色は一瞬で無残な瓦礫と水の広がりに変わってしまった。
Column Tsunami Data	波が海岸を越えて押し寄せると、家々が波に飲まれて崩れ、瓦礫となったものが海へと運ばれていく中、街の景色は一瞬で無残な瓦礫と水の広がりに変わってしまった。
Guess Tsunami Data	波が海岸を越えて押し寄せると、家々が波に飲まれ、次々に崩壊し、瓦礫が広範囲にわたって波に乗って海へと戻され、町全体が瓦礫と化して静かに消え去っていった。
Input Landslide Data	崩れた山から流れ出た土砂が激しい勢いで住宅を襲い、周囲には散乱した土の塊や泥が無造作に積み重なっており、住民たちはその恐ろしい光景を目の当たりにしていた。
Column Landslide Data	れた山から流れ出た土砂が激しい勢いで住宅を襲い、周囲には散乱した土の塊や泥が無造作に積み重なっており、住民たちはその恐ろしい光景を目の当たりにしていた。
Guess Landslide Data	れた山から流れ出た土砂が、激しい音を立てて流れ込み、住宅がその力に飲み込まれる様子は、まさに自然の力の恐ろしさを痛感させるものであり、住民たちは恐れを抱いていた。

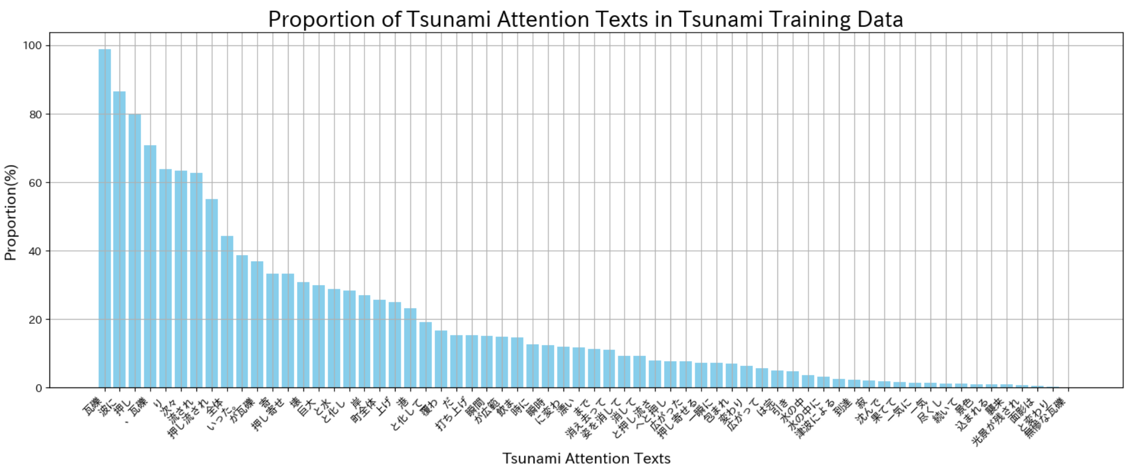
In the example in Table 5.4, “Guess Tsunami Data” begins in the same way as “Column Tsunami Data.” However, as the text progresses, the content of “Guess Tsunami Data” starts to differ significantly from “Column Tsunami Data.” Moreover, “Guess Landslide Data” begins similarly to “Column Landslide Data,” but the content in the middle section differs relatively, while the content in the latter half is more similar between the two.

5.5 Rates of Attention Text

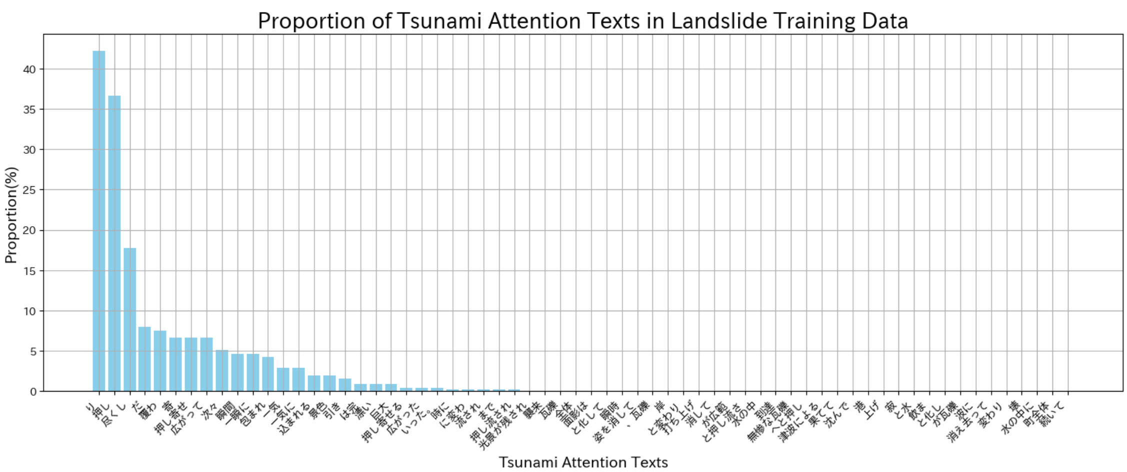
In this study, the words in the Input Data that particularly attract attention in relation to the Column Data in the attention map were defined as attention texts. Three attention texts were extracted from each attention map, and the proportions of how many times these attention texts appear in the total disaster scene episode training data, tsunami disaster scene episode training data, and landslide disaster episode training data were visualized as graphs. These graphs are shown in from Figure 5.4(a) to Figure 5.4(f).



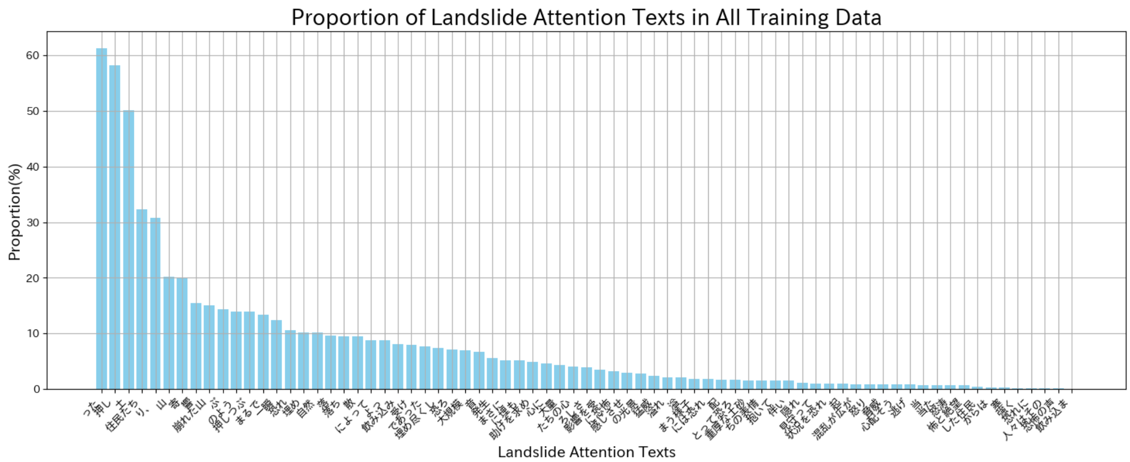
(a) Percentage of Tsunami Attention Texts Included in All Training Data



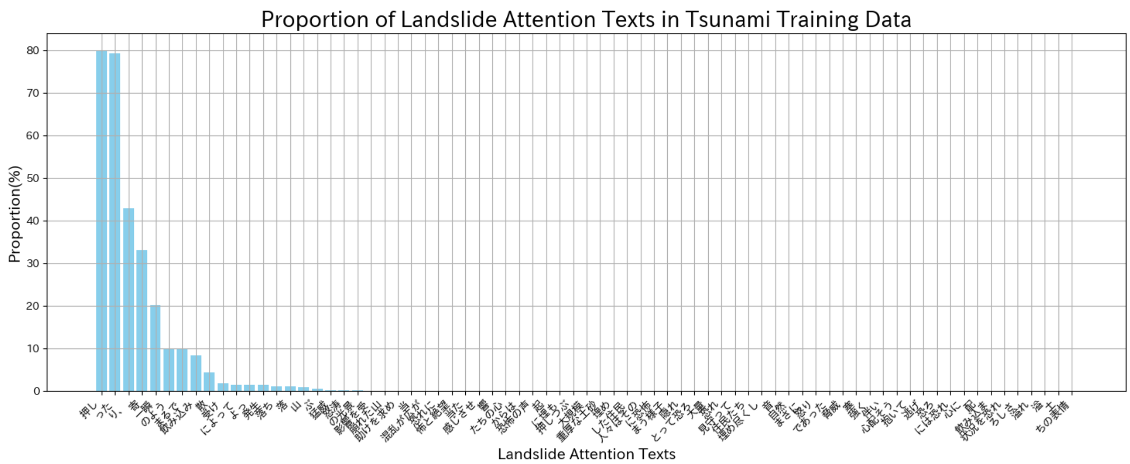
(b) Percentage of Tsunami Attention Texts Included in Tsunami Training Data



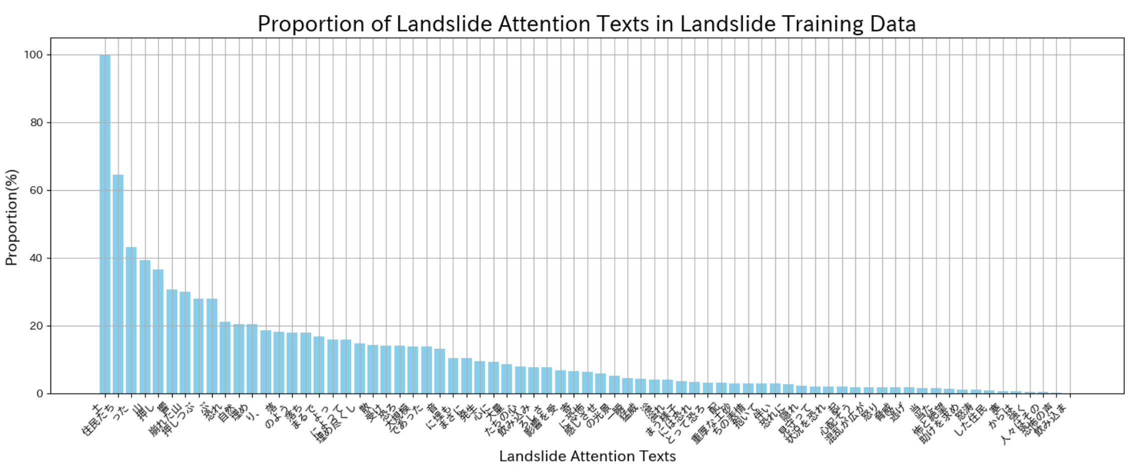
(c) Percentage of Tsunami Attention Text Included in Landslide Training Data



(d) Percentage of Landslide Attention Texts Included in All Training Data



(e) Percentage of Landslide Attention Texts Included in Tsunami Training Data



(f) Percentage of Landslide Attention Texts Included in Landslide Training Data

Figure 5.4. Percentage of Attention Texts Included in the Number of Training Data

The top five tsunami attention texts with the highest proportions in the entire training data, as shown in Figure 5.4, are introduced here. Among all the training data in Figure 5.4(a), the tsunami attention text with the highest proportion is “押し” (means “push”). This attention text ranks third in proportion in the tsunami training data in Figure 5.4(b) and second in proportion in the landslide training data in Figure 5.4(c). Additionally, this attention text is included in approximately 27% of the training data as “押し流され” (means “washed away”) and approximately 19% of the training data as “押し寄せ” (means “surge”) in Figure 5.4(a).

The second most frequent tsunami attention text in the entire training data in Figure 5.4(a) is “り.” This attention text appears in the word “広がり” (means “expansion”) within the episode of the test data, specifically as the character “り.” In Figure 5.4(b), which represents the tsunami training data, it ranks fifth in proportion. In Figure 5.4(c), representing the landslide training data, it has a higher proportion than “押し.”

The third most frequent tsunami attention text in the entire training data is “瓦礫” (means “rubble”). This attention text has a similar string content to “、瓦礫” (the word “瓦礫” with punctuation added), which ranks fifth in proportion among the entire training data, but “瓦礫” has approximately a 12% higher proportion. “瓦礫” has the highest proportion in the tsunami training data in Figure 5.4(b), while “、瓦礫” ranks fourth. However, these two attention texts were completely absent in the landslide training data in Figure 5.4(c).

The fourth most frequent attention text in the entire training data in Figure 5.4(a) is “波に” (includes the meaning of “wave”). This attention text has the second-highest proportion in the tsunami training data in Figure 5.4(b), accounting for over 80%. However, it was not included at all in the landslide training data in Figure 5.4(c).

Additionally, the top five landslide attention texts with the highest proportions in the entire training data in Figure 5.4 are introduced. The attention text with the highest proportion in the entire training data in Figure 5.4(d) is “った” (“った” is often used at the end of sentences). This attention text has the second-highest proportion in the tsunami training data in Figure 5.4(e), at approximately 79%. In the landslide training data in Figure 5.4(f), it ranks third, with a proportion of about 43%.

The second most frequent landslide attention text in Figure 5.4(d) is “押し.” This attention text has the highest proportion in the tsunami training data in Figure 5.4(e). Additionally, it ranks fifth in the landslide training data in Figure 5.4(f).

The third most frequent landslide attention text in Figure 5.4(d) is “土” (means “soil”). This attention text does not appear at all in the tsunami training data in Figure 5.4(e), but it has the highest proportion in the landslide training data in Figure 5.4(f), at 100%.

The fourth most frequent landslide attention text in Figure 5.4(d) is “住民たち” (means “residents” in plural form). This attention text does not present at all in the tsunami training data in Figure 5.4(e), but has the second-highest proportion in the landslide training data in Figure 5.4(f).

The fifth most frequent landslide attention text in Figure 5.4(d) is “り、” (includes the word “り” followed by punctuation). This attention text ranks third in the tsunami training data in Figure 5.4(e) and 13th in the landslide training data in Figure 5.4(f).

Table 5.5 shows the number of times and the attention texts for each disaster site episode appearing two or more times.

Table 5.5. Number of Attention Texts for Each Disaster

Tsunami Attention Text	Frequency of Tsunami Attention Texts Appearing Two or More Times
瞬間	7
寄	5
押し寄せ	5
巨大	4
全体	3
上げ	3
壊	2
町全体	2
り	2
押し	2
Landslide Attention Text	**Frequency of Landslide Attention Texts Appearing Two or More Times**
音	3
山	3

恐れ	3
溢れ	3
のよう	3
隠れ	2
強く	2
散	2
自然	2
住民たち	2
土	2
大量	2

From Table 5.5, the frequency ranges of occurrences of the tsunami attention text appearing two or more times is from 2 to 7 times. The range of occurrences of two or more occurrences of the landslide attention text is from two to three.

Regarding tsunami attention texts, the text with the highest frequency is "瞬間" (means "moment"), with a frequency of 7 occurrences. This attention text ranks 26th in terms of percentage across all training data in Figure 5.4(a), accounting for approximately 10%. Furthermore, comparing Figure 5.4(b) and Figure 5.4(c), the percentage for the landslide training data is relatively higher.

The second most frequent texts are "寄" (part of a word meaning "to approach") and "押し寄せ" (a textual expression meaning "to surge"), both occurring 5 times. These account for approximately 20% of all training data in Figure 5.4(a), with the same percentage. Additionally, comparing Figure 5.4(b) with Figure 5.4(c), the percentage for tsunami training data is relatively higher. Regarding "押し寄せ," a similar attention text, "押し," occurs 2 times. In Figure 5.4(a), Figure 5.4(b), and Figure 5.4(c), this attention text shows a relatively higher percentage than "押し寄せ."

The third most frequent attention text is "巨大" (means "huge"), occurring 4 times. It accounts for approximately 15% of all training data in Figure 5.4(a). Furthermore, about 30% of tsunami training data in Figure 5.4(b) includes this text, while about 1% of landslide training data in Figure 5.4(c) includes it.

The fourth most frequent attention texts are "全体" and "上げ," both occurring 3 times. In Figure 5.4(a), these occupy proportions mainly in the first half, with percentages of approximately 22% and 12%, respectively. Additionally, they maintain the same ranking in Figure 5.4(b) but are not present at all in the landslide training data of Figure 5.4(c).

Other attention texts with a frequency of 2 occurrences, besides "押し," include "壊" (means "to destroy"), "町全体" (means "the entire town"), and "り" (a Japanese word). In Figure 5.4(a), all of these occupy proportions mainly in the first half, with "り" standing out as having the second-highest proportion. Furthermore, among these, only "り" is included in the landslide training data of Figure 5.4(c), where it occupies the highest proportion.

From the above, it was found that, regarding the tsunami attention texts, the texts with high frequencies do not necessarily have a high percentage in the training data. Additionally, even for the attention texts with high percentages, their frequencies may not necessarily be high.

Moreover, an explanation is provided regarding landslide attention texts. The most frequently occurring texts are "音" (means "sound"), "山" (means "mountain"), "恐れ" (an expression representing "fear"), "溢れ" (part of a verb meaning "to overflow"), and "のよう" (an expression used as a suffix to describe something), each appearing 3 times. In Figures 5.4(d), 5.4(e), and 5.4(f), among these, the percentages for all training data place "音," "山," "恐れ," and "のよう" in the first half, while "溢れ" is positioned in the latter half. Additionally, for tsunami training data, "のよう" and "山" have percentages higher than 0%.

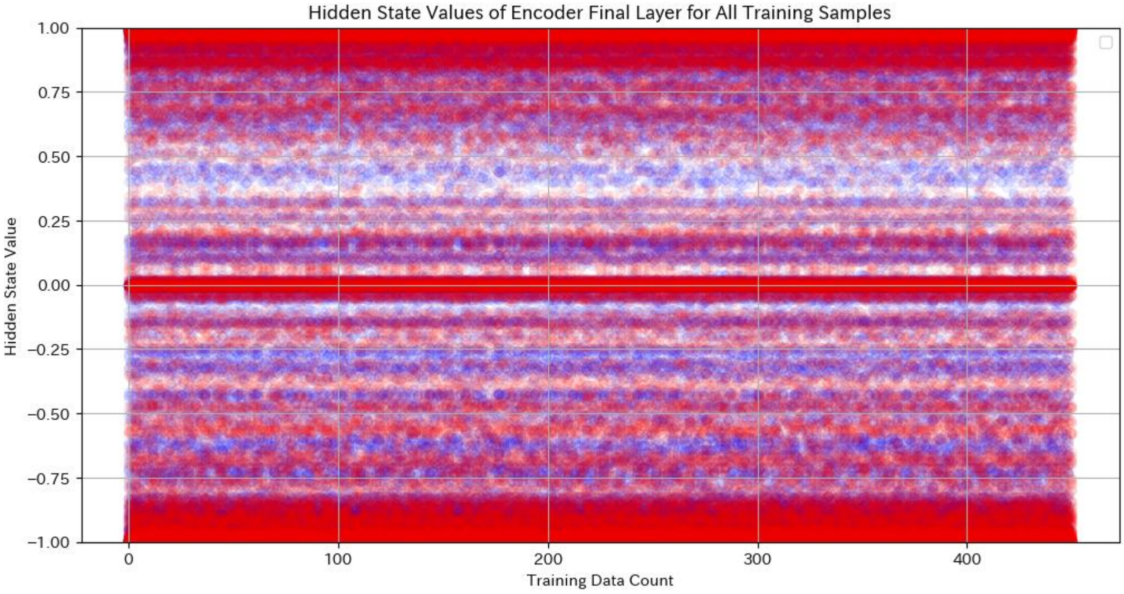
The texts with a frequency of 2 occurrences are "隠れ" (means "to hide"), "強く" (an expression meaning "strongly"), "散" (means "to scatter"), "自然" (means "nature"), "住民たち" ("residents" in plural form), "土" (means "soil"), and "大量" (means "a large amount"). In Figures 5.4(d), 5.4(e), and 5.4(f), among these, the percentages for all training data place "散," "自然," "住民たち," and "土" in the first half, while "隠れ," "強く," and "大量" are positioned in the latter half. None of these texts had percentages higher than 0% for tsunami training data.

Furthermore, in the landslide training data, the attention text "土" accounts for 100% of the percentage, and "住民たち" is the second highest, accounting for approximately 64%.

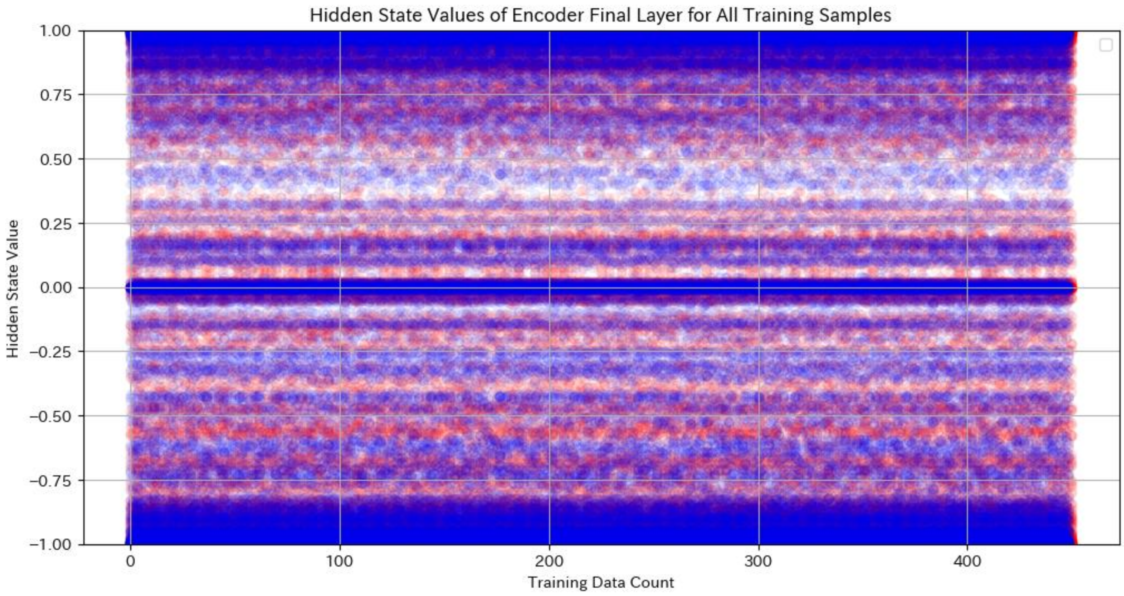
From the above, it was found that for landslide attention texts, texts with lower frequencies occupy a higher percentage of the overall training data than texts with higher frequencies. Additionally, attention texts with high percentages do not necessarily have relatively high frequencies.

5.6 Encoder's Output in Final Layer

In this study, we analyzed the encoder output from the final layer of the trained AttentionSeq2Seq model, which is a 256-dimensional hidden state vector. Figure 5.5 shows a plot of the vector values of the hidden state of the final layer when all training data, all training data for tsunami disaster site episodes, and all training data for landslide disaster site episodes are input to the trained AttentionSeq2Seq model for each number of training data.



(a) Plot of Encoder Final Layer Hidden State Vectors Using Tsunami Training Data with Overlaid Encoder Final Layer Hidden State Vectors Using Landslide Training Data



(b) Plot of Encoder Final Layer Hidden State Vectors Using Landslide Training Data with Overlaid Encoder Final Layer Hidden State Vectors Using Tsunami Training Data

Figure 5.5. Hidden State Vector of the Final Layer of the Encoder Using Training Data

The blue data points in Figure 5.5 represent the data points of the encoder's final layer hidden state vectors for the tsunami training data. The red data points represent the data points of the encoder's final layer hidden state vectors for the landslide training data. From Figures 5.5(a) and 5.5(b), it can be observed that both the blue and red data points are densely concentrated in the ranges of -1 to -0.9, around 0, and 0.9 to 1 for the values of the hidden state vectors. These regions represent the value ranges of features containing information from both the tsunami and landslide training data. Additionally, distinct horizontal band-like regions for the blue and red data points can be observed. It indicates that the encoder's final layer strongly retains the respective hidden state vectors for each training dataset. Furthermore, a white horizontal band-like region is visible, which indicates the absence of values for the encoder's final layer hidden state vectors in this region for the training data. The median of the hidden state vectors for each dimension of the encoder's final layer was extracted, and the difference between the median of the encoder's final layer for the tsunami disaster data and the landslide disaster data are shown in Figure 5.6.

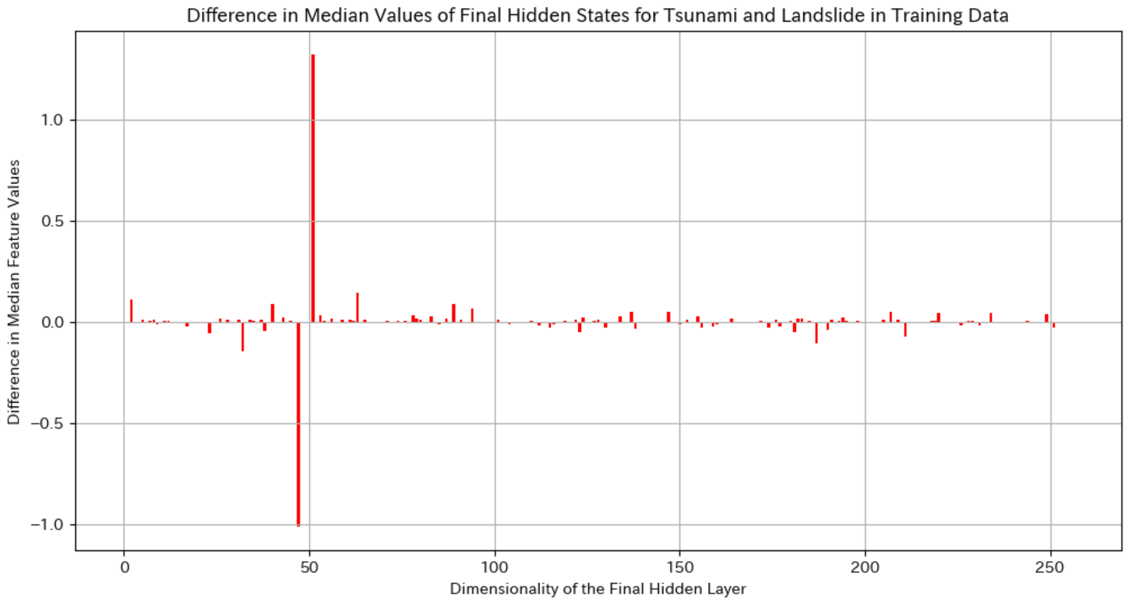


Figure 5.6. Difference in the Encoder Final Layer Median for Each Disaster Type

From Figure 5.6, the dimension with the largest positive difference in the encoder final layer median for each disaster type is the 51st dimension. The dimension with the largest negative difference is the 47th dimension. Additionally, several positive and negative bar graphs can be observed for dimensions other than these two.

6. Discussion

6.1 Generation Accuracy of the Pre-trained Attention Seq2Seq Model

Using the test data of the tsunami disaster site episodes and the landslide disaster site episodes, new episodes were generated by the trained model. The beginning of Guess Data generated as each new episode was relatively similar in content to Column Data. When training the AttentionSeq2Seq, the input text to the encoder is reversed. Furthermore, due to the structure of AttentionSeq2Seq, the final layer of the encoder retains information for all input strings. For these reasons, it is considered that the contents of Guess Data and Column Data are relatively similar in the first part. In addition, it is conceivable that the beginning part of the training data and the test data are similar.

Additionally, regarding the beginning and subsequent parts of the newly generated episodes, the tsunami disaster site episodes differ as they progress into the later parts. The landslide disaster site episodes exhibit relatively more differences around the middle, but the latter parts are relatively similar. Based on this result, it is considered that the tsunami training data has a higher degree of expressive freedom compared to the landslide training data. Furthermore, since this study uses a smaller AttentionSeq2Seq model, the higher expressive freedom of the tsunami training data's string representations likely results in suboptimal learning in tasks aiming to generate content similar to Column Data. However, when checking the Guess Data, it is considered that the model successfully outputs the characteristics of tsunami and landslide as found in their respective disaster site training data. This suggests that the AttentionSeq2Seq model has learned to distinguish and generate tsunami-like and landslide-like characteristics from the entire training dataset.

6.2 Regarding the Graph of Attention Text Proportions

By examining the attention texts for each disaster site, it is found that the characteristics of the sentence content differ. In the tsunami disaster site episodes, attention texts such as “押し寄せ,” “瓦礫,” “波に,” and “広がって”, which mean “surge,” “rubble,” “to wave,” and “expanding,” respectively, are generally more frequent. These words are used to express the destruction caused by the tsunami. In the landslide disaster site episodes, in addition to attention texts like “山” and “土” (mean “mountain” and “solid”, respectively) that describe the landslide, there is a relatively higher frequency of attention texts expressing the fear and emotions of residents affected by the landslide, such as “逃げ,” “助けを求め,” “には恐れ,” “隠れ,” and “心配そう” (mean “escape,” “call for help,” “fear,” “hide,” and “warried,” respectively).

These tendencies suggest that the AttentionSeq2Seq model learns to reflect the differences in the fear humans feel towards each disaster. From Figure 5.4 and Table 5.5, it was observed that even when attention texts have a high percentage in the training data, this does not necessarily result in a higher frequency of attention texts on the attention map. Moreover, even when the frequency of text appearances on the attention map is high, it does not always correspond to a higher percentage in the training data.

From these results, it can be inferred that when the AttentionSeq2Seq model generates new episodes, it does not simply rely on the quantity of words present in the training data. Instead, it is likely to generate episodes that reflect the characteristics of the fear of each disaster that humans feel as their own, based on the input data it receives.

6.3 Differences in Attention by Disaster Type

From Table 5.4, when generating new disaster site episodes with the trained AttentionSeq2Seq model using the test data, episodes containing the respective characteristics were generated. From Figure 5.5, the graph of the encoder's final layer values when the trained AttentionSeq2Seq model was input with the training data shows differences in the regions for the tsunami training data and the landslide training data. From Figure 5.6, the difference in the medians of the hidden state vectors at each dimension of the final layer of the encoder shows relatively large differences in the values of the hidden state vectors at specific dimensions. The resulting attention texts show that the stylistic features differ between the tsunami disaster site and the landslide disaster site. These results suggest that the hidden state vectors of the final layer of the encoder separate between the tsunami and landslide site training data and store the sentence features of each disaster site's training data. For the tsunami disaster site training data, it is thought that features emphasizing expressions of the tsunami destroying the town were learned, while for landslides, it is considered that features emphasizing expressions related to the fear and reactions of the residents after the landslide occurred were learned. It is assumed that these emphasized expressions were separated and stored in the final layer of the encoder. The trained AttentionSeq2Seq model generates new episodes through the encoder for new input data. By passing through the encoder final layer that separates the features of each disaster, it is considered to be possible to generate episodes that contain the characteristics of each disaster.

6.4 Episodes That Make One Feel as If It Is Their Own

In this study, a small language model, AttentionSeq2Seq, is used to generate episodes about each disaster site that viewers feel as their own, using relatively small amounts of training data. From the previous results, the elements that make viewers feel as though the episode is their own differ for each disaster site. In the tsunami disaster site episode, the emphasis is on words that highlight the destruction of the town by the tsunami, which is an element that makes viewers feel as their own. In the landslide disaster site episode, the emphasis is on words that highlight the fear and emotional expressions of the residents in response to the landslide, which is the element that makes viewers feel as their own. By training the model on data that contains these elements in abundance, it is considered that episodes in which viewers feel as their own can be generated for each disaster. Furthermore, by training with appropriate training data, it becomes possible to conduct learning that makes viewers feel as though it is their own even with a small amount of data.

7. Conclusion

An effective method to minimize the damage caused by disasters is evacuation training. However, it has been found that the participation rate in evacuation drills is low, especially among young people. One of the reasons for this may be that they do not feel as though disasters occurring elsewhere are their own. Various studies have been conducted using AR and VR in evacuation drills, however, none have examined whether participants feel as though the situation is their own. In the preliminary experiment of this study, an experiment was conducted during the viewing of tsunami disaster site videos, where frightening external factors were introduced, and through EEG and EDA data, the researchers examined whether the participant felt as though the situation was her own in different states. The results showed that simply watching the videos did not make the participant feel as though the disaster was their own. According to the participant’s feedback, they found that videos with a storyline were more memorable. Based on these results, this study aims to generate episodes that the viewers feel as their own using a small language model.

Generated episodes, those that the subjects feel as their own through a survey were extracted and used for training an AttentionSeq2Seq model. During training, the input text was reversed, and the labels were set starting from the second character of the input. Using AttentionSeq2Seq, new episodes related to tsunami and landslide disaster sites were generated. The attention text was extracted from the attention map during the new episode generation, and an analysis was conducted to determine how much of the attention text was present in the training data. Additionally, the final layer of the encoder was analyzed.

Based on the analysis, we discussed which elements of each disaster site episode make viewers feel as if the disaster is their own. In the tsunami disaster site episode, expressions depicting the tsunami destroying the town were the elements that made viewers feel as if the disaster was their own. Additionally, in the landslide disaster site episode, expressions depicting the residents' fear of the landslide were the elements that made viewers feel as if the disaster was their own. For each disaster, including these elements in the training data in large amounts can make it possible to generate episodes that viewers feel as their own. Moreover, this approach allows for a reduction in the amount of training data required.

Regarding future challenges, this study only conducted model training and generation for two types of disasters: tsunamis and landslides. it will be necessary to conduct experiments using the same approach as in this study for a wider range of disasters.

Acknowledgments

We greatly appreciate the valuable contributions of our community advisory committee members. We would also like to thank the every team member who took the time to participate in this study.

Authors' contributions

Mr. Shogo Matsumoto was responsible for study design, revisions, data collection, data analysis, manuscript drafting, and manuscript revision. Prof. Hiromitsu Shimakawa and Prof. Fumiko Harada contributed to the study design, revisions, and manuscript editing. Mr. Shogo Matsumoto, Prof. Hiromitsu Shimakawa, and Prof. Fumiko Harada engaged in discussions and jointly determined the direction of the study. The data analysis results obtained by Mr. Shogo Matsumoto were discussed by Prof. Hiromitsu Shimakawa, and Prof. Fumiko Harada. All authors read and approved the final manuscript.

Funding

Not applicable.

Competing interests

Not applicable.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

ANNnewsCH. (2020, January 17). 2011 nen 3 gatsu 11 nichi Higashi Nihon Daishinsai Miyako-shi o osotta "kuroi" kyodai tsunami【Mainichi Bōsai】Kono dōga ni wa tsunami eizō ga fukumareteimasu / Great East Japan Earthquake, Tsunami [Video]. YouTube. Retrieved January 27, 2025, from https://www.youtube.com/watch?v=4XvFFfgXwnw (in Japanese)

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods*, *190*(1), 80-91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877-1901.

Catal, C., Akbulut, A., Tunali, B., Ulug, E., & Ozturk, E. (2020). Evaluation of augmented reality technology for the design of an evacuation training game. *Virtual Reality*, *24*, 359-368. https://doi.org/10.1007/s10055-019-00410-z

Choi, J. S., Bang, J. W., Heo, H., & Park, K. R. (2015). Evaluation of fear using nonintrusive measurement of multimodal sensors. *Sensors*, *15*(7), 17507-17533. https://doi.org/10.3390/s150717507

Clayton, M. S., Yeung, N., & Cohen Kadosh, R. (2018). The many characters of visual alpha oscillations. *European Journal of Neuroscience*, *48*(7), 2498-2508. https://doi.org/10.1111/ejn.13747

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. *Handbook of psychophysiology*, *2*, 200-223. https://doi.org/10.1017/CBO9780511546396.007

Empatica Inc. (n.d.). *E4 wristband | Real-time physiological signals.* Empatica. Retrieved January 27, 2025, from https://www.empatica.com/research/e4/

Feng, Z., González, V. A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., ... Cabrera-Guerrero, G. (2020). An immersive virtual reality serious game to enhance earthquake behavioral responses and post-earthquake evacuation preparedness in buildings. *Advanced Engineering Informatics*, *45*, 101118. https://doi.org/10.1016/j.aei.2020.101118

FNN311. (2013, June 13). Kamaishi-shi yakusho fukin ni oshiyoseru tsunami [Shichōsha teikyō eizō] [Video]. YouTube. Retrieved January 27, 2025, from https://www.youtube.com/watch?v=aQj2zn5Axmk (in Japanese)

Gwynne, S., Amos, M., Kinateder, M., Bénichou, N., Boyce, K., van Der Wal, C. N., & Ronchi, E. (2020). The future of evacuation drills: Assessing and enhancing evacuee performance. *Safety science*, *129*, 104767. https://doi.org/10.1016/j.ssci.2020.104767

Hao, S., Lee, D. H., & Zhao, D. (2019). Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system. *Transportation Research Part C: Emerging Technologies*, *107*, 287-300. https://doi.org/10.1016/j.trc.2019.08.005

Herrmann, C. S., Strüber, D., Helfrich, R. F., & Engel, A. K. (2016). EEG oscillations: from correlation to causality. *International journal of psychophysiology*, *103*, 12-21.https://doi.org/10.1016/j.ijpsycho.2015.02.003

Hughes, J. R. (2008). Gamma, fast, and ultrafast waves of the brain: their relationships with epilepsy and behavior. *Epilepsy & Behavior*, *13*(1), 25-31. https://doi.org/10.1016/j.yebeh.2008.01.011

InteraXon. (n.d.). Muse 2. Choose Muse. Retrieved January 28, 2025, from https://choosemuse.com/products/muse-2

Komatsu, Y., Kawai, R., & Sakaguchi, M. (2018). Enhancement of emotion using temperature stimulation inducing physiological response. *Robotics and Mechatronics Conference Proceedings 2018*, 1P1-M16. The Japan Society of Mechanical Engineers. (in Japanese). https://doi.org/10.1299/jsmermd.2018.1P1-M16

Li, G., Li, F., Ahmad, T., Liu, J., Li, T., Fang, X., & Wu, Y. (2022). Performance evaluation of sequence-to-sequence-Attention model for short-term multi-step ahead building energy predictions. *Energy, 259*, 124915. https://doi.org/10.1016/j.energy.2022.124915

Saia, R., Carta, S., Fenu, G., & Pompianu, L. (2022). Brain Waves and Evoked Potentials as Biometric User Identification Strategy: An Affordable Low-cost Approach. In *SECRYPT* (pp. 614-619). |https://doi.org/10.5220/0011297600003283

Schneider, T. R., Hipp, J. F., Domnick, C., Carl, C., Büchel, C., & Engel, A. K. (2018). Modulation of neuronal oscillatory activity in the beta-and gamma-band is associated with current individual anxiety levels. *NeuroImage*, *178*, 423-434. https://doi.org/10.1016/j.neuroimage.2018.05.059

Seguchi, M. (2022). Figuring Out Commercial Gimmicks Influencing Consumer Engagement from Psychological Change Points. *Advances in Social Sciences Research Journal*, *9*(2), 190-210.

https://doi.org/10.14738/assrj.92.11738

Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems, 27*.

Takahashi, H. (2020). *Psychological effects of thermal tactile stimulation on virtual pet therapy* (Doctoral dissertation, Chuo University). (in Japanese)

TBS NEWS DIG Powered by JNN. (2021, February 25). [3.11] Kyodaina tsunami ga bōchōtei o nori koeru Iwate - Kamaishi-shi [JNN ākaibu 311 ano hi no kiroku] [Video]. YouTube. Retrieved January 27, 2025, from https://www.youtube.com/watch?v=yPajUmZHpOE (in Japanese)

Uehara, D., Shimakawa, H., & Harada, F. (2024). Gimmick estimation in video advertisements by scene analysis with interpersonal theory. *International Journal of Internet Marketing and Advertising, 21*. https://doi.org/10.1504/IJIMA.2024.10064751

vientorio. (2011, July 22). Amazing power of Japanese Tsunami caught on video 19 HD [Video]. YouTube. Retrieved January 27, 2025, from https://www.youtube.com/watch?v=O259ajfJylM

WankoDon. (2011, April 11). Tsunami TSUNAMI - Earthquake (3/11) 20110411 [Video]. YouTube. Retrieved January 27, 2025, from https://www.youtube.com/watch?v=_m8jLnOYzfg (in Japanese)

Wēzā Nyūsu Sutadi. (2023, November 5). Higashi Nihon Daishinsai Tsunami ga sematte kuru kyōfu [Video]. YouTube. Retrieved January 27, 2025, from https://www.youtube.com/watch?v=eriaCmUbvqU (in Japanese)

Yomiuri Shinbun onrain dōga. (2011, March 14). Ōtsunami ni nomikomareru shunkan no Kesennuma shigai [Video]. YouTube. Retrieved January 27, 2025, from https://www.youtube.com/watch?v=Ann27T6JTek (in Japanese)

Zhang, S., Roller, S., Goyal, N., Artetxe, M., Chen, M., Chen, S., ... Zettlemoyer, L. (2022). Opt: Open pre-trained transformer language models. arXiv preprint arXiv:2205.01068. https://doi.org/10.48550/arXiv.2205.01068

Zhang, Y., Deng, X., & Zhang, Y. (2021). *An Attention-based Seq2Seq Model for Short Term Energy Consumption Prediction.* Energy. https://doi.org/10.46855/energy-proceedings-8604