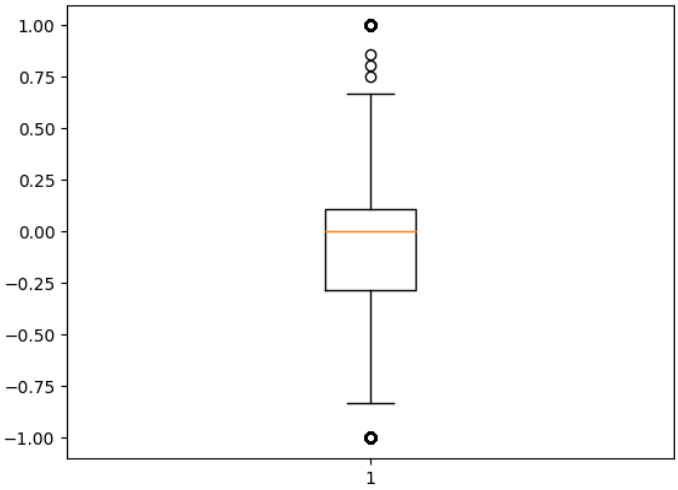
class DataLoader

- For the language generation model, I used bert tokenizer, and set max_length as seq_length + 1, because we use a bigram model. The input is from the first element in the sequence until the second to last, and the target is from the second element in the sequence until the last one.
- For the regression model, I first averaged msEngagement, is_like and is_dislike over content_id. I created a column `interaction` by (`is_like` - `is_dislike`)*1.5 + `msEngagement` (weigh likes a little bit more than engagement), and then use MinMaxScaler to scale the values to be from -1 to 1. The distribution of interaction is as shown below



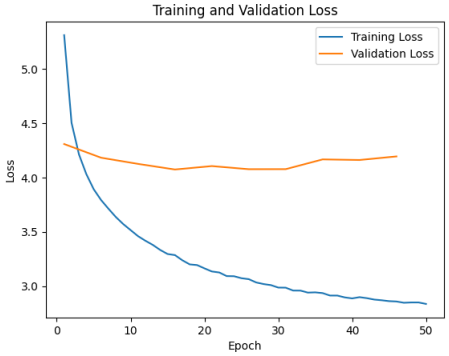
- The default train/validation/test split is 0.64/0.16/0.2.
- generate_next_token() first encodes the input, and generates logits, and finally decodes to output tokens with specified length.

class RegressionModel

- The regression model takes parameters from the pretrained language generation model, and adds a regression head nn.Sequential(nn.Linear(n_embd, 1), nn.Tanh()) to output a single regression score.

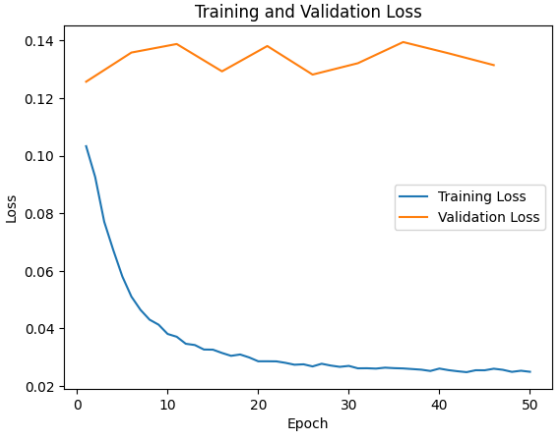
train_generation_model() uses a masked_cross_entropy() loss function to exclude padding tokens to prevent overfitting. train_regression_model() uses MSE as loss function because it is more suitable for regression tasks.

This is the performance of pretrained language generation model, and it's stored here:

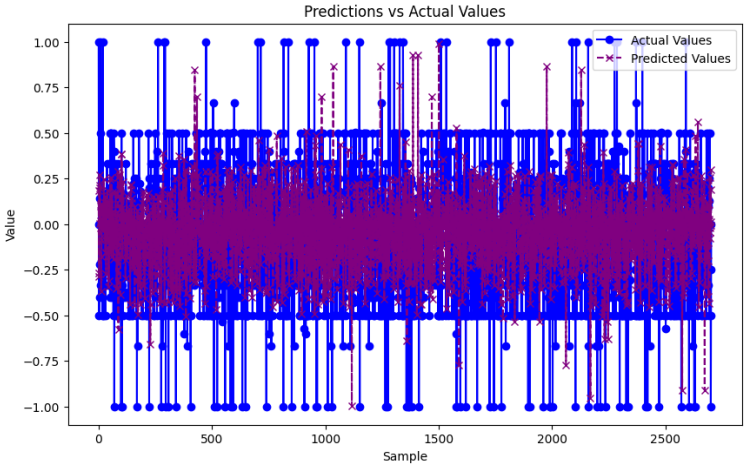


When inputting “8k”, the model generates “8koverthestreet,aphoto##real##istic,highlydetailed,digitalpainting,arts##tation,conceptart,smooth,sharpfocus,illustration,”

This is the performance of finetuned regression model, and it’s stored here:



When testing on test data, MSE is 0.1249.



For example, when inputting the prompt “Peter steps into a dingy kitchen. A dirty coffee cup and a plate sit on the counter near the sink. At the other end of the kitchen, past the refrigerator, is a door leading outside. Peter approaches the door, glances through the window. He then turns and walks through a doorway.\\\\n. In the style of the movie Lost-Souls. dim volumetric cinematic lighting, 8k”, the predicted interaction score is 0.8273, which means the user will likely like it.