CS325 Final Project Proposal

Mohamed Ali, Menna Khaliel, Joseph Tadrous

Nov 24, 2021

1 Introduction

The goal of our project is to classify anime genres based on the wording of the synopsis. If time permits, we can then expand our algorithm to measure the similarities between different synopsis to see which animes are closely related.

2 Datasets and Related Work

Anime Dataset with Reviews:
https://www.kaggle.com/marlesson/myanimelist-dataset-animes-profiles-reviews
We will refer papers [4], [2], [3], and [1] in our work.

3 Techniques and Model Architecture

We will experiment with multiple techniques and evaluate them to see which combination of algorithms works better. We will have 4 models in total. For all the models, we will create GloVe word embeddings to account for word similarity. The models we will use are as follows:

- For the base model we will use logistic regression with average GloVe vectors. This will serve as a baseline to compare all the other models with and will give us a general idea on how difficult the task is.
- We will switch the logistic regression approach for an LSTM model. We will use the unbalanced dataset, mentioned in 2 to train the model
- We will train another LSTM model but this time we will curate and adapt a balanced dataset out of the one mentioned in 2. We believe this might lead to more fair results and hopefully higher accuracy.
- For the last model we will experiment with bi-directional LSTMs (BiLSTM), a sequence processing model that consists of two LSTMs: one taking the input in a forward direction, and the other in a backwards direction. We believe this will significantly increase the accuracy of the

model by increasing the amount of input and improving the context available to the algorithm (as it will know not only the words that precede a specific word, but also these that follow it).

4 Evaluation, interpretation, and visualization methods for results

After splitting the data and training the model with supervised learning techniques, we can evaluate the model using k-fold cross-validation where we split the data into k random subsets several times, training k-1 of them and testing on the remaining one. We will also consider applying Leave One Out evaluation to ensure generalization of the model.

- Evaluating & Visualizing the Results
  - We will create different confusion matrices in order to have a better idea of which genres are more likely to be confused with one another.
  - Using these confusion matrices, we will calculate the following metrics for each genre:
    - Precision ($\frac{\text{true positive}}{\text{actual results}}$) where true positives is defined as the number of animes that we succeeded in predicting its genre (the label could be a list of genres, but we count a correct prediction as predicting one of its genres correctly)
    - Recall ($\frac{\text{true positive}}{\text{predicted results}}$)
    - Accuracy ($\frac{\text{true positive} + \text{true negative}}{\text{total}}$)
    - Error ($\frac{\text{false positive} + \text{false negative}}{\text{total}}$)
  - Then, we will use the aforementioned evaluation metrics to compare between the performance of our models.
  - In addition, we will tune the model architecture and parameters and plot the accuracy and loss graphs to see which parameters give us the best results.

5 Timeline

- By December 3, Base model and first LSTM model; establish code for evaluation measures and plots.
- By December 9, Introduction and related work sections; code for the third and fourth models.
- By December 17, write up final report, including description, results and evaluation of all four models.

References

[1] Ali Mert Ertugrul and Pinar Karagoz. "Movie genre classification from plot summaries using Bidirectional LSTM". In: *2018 IEEE 12th International Conference on Semantic Computing (ICSC)* (2018). DOI: 10.1109/icsc.2018.00043.

[2] Akshi Kumar, Arjun Rajpal, and Dushyant Rathore. "Genre classification using word embeddings and deep learning". In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (2018). DOI: 10.1109/icacci.2018.8554816.

[3] Jeffrey Pennington. *GloVe: Global Vectors for Word Representation*. Aug. 2014. URL: https://nlp.stanford.edu/projects/glove/.

[4] Arjun Raj Rajanna et al. "Deep Neural Networks: A case study for music genre classification". In: *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)* (2015). DOI: 10.1109/icmla.2015.160.