Computational Content Classification of Traditional Chinese poems

Ally Lo, Grace Tong, Steven Luo, Olga Chan

Artificial Intelligence Project Group
St. Paul's Co-Educational College
sp20166401, sp20197521, sp20156401, sp20165171@spcc.edu.hk

Abstract

Analysis of Chinese poetry according to content is a challenging task in the field of NLP. On the other hand, it is a very difficult and time-consuming task for poems to be classified manually and objectively. While for other languages, like English, automatic poetry classification has been investigated using several NLP techniques, it is not that common for Chinese poetry. The goal of this paper is to develop an ideal model for Chinese poem genre classification. To achieve our objectives, we have prepared a corpus consisting of a variety of Chinese poems representing a variety of linguistic features. This project utilizes 2 computational methods namely LSTM and Transformer models in Deep Learning, to classify poems based on the top seven typical themes that inheres in Chinese poems. The results of these models are compared, so as to find the best model for accurate classification of poems. Evaluation on the effectiveness of these different systems is done through comparing the accuracies and loss, distinguishing the impacts brought by the differences between the models and then further optimize the model based on the information, finally establishing an architecture with the most satisfactory results with a training accuracy of 97% and a validation accuracy of 82%.

List of Abbreviations

NLP	Natural Language Processing
SVM	Support Vector Machines
KNN	K-Nearest Neighbor
RFC	Random Forest Classifier
DTC	Decision Tree Classifier
RNN	Recurrent Neural Network
LSTM	Long Term Short Memory

1 Introduction

Poetry is an aesthetic and rhythmic form of literature that expresses emotion with words. The word "poet" originated from "poiein", which means "to make" in Greek. In particular, traditional Chinese poetry emphasizes on beauty and musicality. Ambiguous words and sentences may bear multiple meanings and have deeper contextual implications instead of their literal meaning. Therefore, it is complicated for computation to classify the poems into different genres accurately and efficiently. This makes investigating the interpretation of traditional Chinese poetry by computational methods intriguing.

There are a number of works related to the analyzation of poems using computational linguistics. In 2014, the SPARSAR system [1] is developed to analyze poetic features and literature techniques in poems such as poetic structure and rhythm with syntax, semantics and phonology and rhetoric devices by using NLP tools like tokenizers, sentence splitters, NER (Name Entity Recognition) tools, and taggers. It includes the production of expressive reading by a TTS module through sentimental analysis on a line by line basis. The results were encouraging through manual evaluation on 5 parameters with a total percentage error of around 5% only. In another study in 2018[2], automatic poetry analysis, focusing on rhyme, diction, and metaphor, is investigated. The data is processed using methods such as syntactic parsing, Entities / n-grams / word-based features and word embeddings before it can be utilized. A number of Machine Learning methods, like Logistic Regression, DTC, SVM and Deep Learning with CNN, word embeddings, Bag of Words, etc. are used. The Deep Learning networks produced encouraging results, which proved that computational language analysis is possible.

This project explores and experiments on a less discussed topic in NLP — the use of Deep Learning in classifying Chinese poems based on content. Generation of Chinese poems is also a widely investigated topic in NLP. Microsoft's XiaoIce[3] is an AI chatbot which has the function of generating Chinese poems based on an image or keyword. On May 15, 2018, XiaoIce published the first AI-created Chinese poem album in history. Each identified object generates a sentence, forming a poem using a hierarchical RNN which models the structure among the words and sentences. Tsinghua University has also developed Jiuge[4], a system that generates traditional Chinese poems based on images or keywords. Jiuge allows users to revise the unsatisfied parts of a generated poem draft repeatedly, achieving collaborative poem creation with the user. The working model comprises an encoder, a decoder and memory components. Despite the in-depth exploration of Chinese poetry generation, traditional Chinese poetry analysis is less touched upon and it is possible a more difficult task than poetry generation as the emotions and themes of the poems, which may be expressed in various ways, have to be identified.

This project brings about 3 major contributions to the field of poem analysis in NLP: 1) formulation of own dataset of categorized traditional Chinese poems available for other related studies and investigations; 2) evaluation on the effectiveness of LSTM networks and Transformer models on poem analysis; 3) production of a relatively ideal model capable of classifying traditional Chinese poems in terms of the 7 genres.

This project can provide insight to the capabilities of NLP in processing more abstract and aesthetic forms of Chinese language, giving progress to the study of understanding messages and thoughts hidden in complex imagery, history and philosophy in artificial intelligence. Also, seeing that some of the poems in those websites are not classified yet or do not have sufficient or accurate labels indicating their genres, we believe that our project can help to classify poems more effectively and accurately than the pre-existing websites and set a basis for the future development of other databases of poems.

2 Data Collection

Our dataset comprises over 2500 Chinese poetic text of varied lengths and of different dynasties and forms, including 唐詩, 元曲 and 宋詞, such that the Deep Learning models can have an elaborate understanding and analysis on Chinese poetry content. Between the 3 forms of poetry, 唐詩 has a significantly higher level of word bias, with stronger decisive words for most genres, compared to 元曲 and 宋詞.

The data needed in the models is collected by web-scraping some of the websites which already provided us with a large database of poems and its categories. Three pre-existing websites with large databases of poems are used for the data collection (see in reference list).

We have concluded 7 genres that classify the content of typical Chinese poems, namely farewell (送別), patriotism/appreciation of history (愛國/詠古), war (戰爭), homesickness/boudoir repinings (思鄉/閨怨), describing scenery/objects to express feelings (借景抒情/托物詠志), lament (慨嘆) and praise (讚頌). These 7 genres cover the main themes in traditional Chinese poems, and each poem can consist of more than one genre. The seven genres chosen for this classification project is based on some preexisting websites which indicates the different poems can be sorted based on their content.

All the poems are labelled according to the genres they belong to. For the output layer of data labelling, we tried using two different kinds of layers, namely softmax and sigmoid. A softmax function is more typically used for multi-classification in logistic regression models, which is what we are doing. For instance, softmax's effectiveness in multi-class classification is demonstrated in a study done by SMILE Lab[5], in which they optimized the softmax function into a Robust Softmax Regression. Using softmax, the probabilities of a grid adds up to one, which may provoke the competitiveness and comparison between the probabilities of the genres, causing a neglect or loss of information of a poem's content in the process. This is also proven in a study done by The Swiss AI Lab IDSIA[6] that tries to understand the local competitiveness in networks.

However, our results generated using softmax are inadequate, which is likely due to how our poems do not necessarily belong to one single class, unlike the Swiss study. Hence, we replaced the softmax function with a sigmoid function.

In sigmoid, each value is independently portrayed as the probabilities of a grid do not equal to 1. Instead, each value has an upper limit of 1 and a lower limit of 0, so that all genres with correlation and similarity to a poem can be presented, and the competitiveness between the genres can be eliminated. The results returned from the models with sigmoid layers are satisfactory. It can be concluded that sigmoid layers are more suitable for Chinese poetry genre classification tasks in relative to softmax layers.

3 Preliminary Investigation

Prior to this, we have made several attempts to create the best genre classification model for Chinese poetry. We used a simplified 5-genre dataset that consists of 唐詩 solely to make a genre classification model for it. The results are satisfactory, proving that genre classification of Chinese poetry based on content is possible before we move on to the more complex and profound genre classification task in this project — a 7 genre multi-form Chinese poetry classification.

To create a Chinese poetry genre classification model, we first made use of Machine Learning models. We made various attempts to optimize our models, using different data splitting methods, performing hyperparameters optimization and making use of different word embeddings and Machine Learning algorithms. Among all, word embeddings have the greatest impact on our models. There have been a few studies[7] [8] illustrating the benefits of word embeddings. Still, almost none has tested the effectiveness between different word embeddings methods on Chinese poetry classification which consist of a complicated use of archaic Chinese. Making use of various machine learning methods, including SVM, KNN, RFC and DTC, we compared various feature engineering approaches, proving that the effectiveness in improving genre classification models by using different feature engineering methods is observable but limited. We tested with two of the most popular word embedding methods, Word2Vec and Bag of Words. We concluded that word embedding methods can cut down the training time of models and solve underfitting and overfitting issues. However, accuracies are not significantly boosted by the optimization of feature engineering of textual data methods. After much effort in trying to optimize Machine Learning models without significant improvements, we decided that poetry classification is far too complex for simple Machine Learning algorithms and moved on to investigate Deep Learning models in the later experimentation part.

4 Experimentation

Neural Networks are computing systems that form probability weights between input and output after processing a set of examples, “learning” how to interpret the data to form the desired output. In this project, we have experimented on 2 types of Neural Networks, LSTM and Transformer, which perform well in a smaller range of data and more varied data respectively.

RNN is a network that can “memorize” parts of the inputs and use them to make accurate predictions. The problem of vanishing-gradients often appears in standard RNNs. The network experiences difficulty in memorizing words far away in the sequence and makes predictions only based on the most recent ones. It is an obstacle prohibiting the models to achieve a higher accuracy. In many studies and researches, LSTM and Transformer models are seen as typical solutions to the problem.

(stage 1: LSTM)

To solve the problem of vanishing-gradient, we first made use of LSTM networks. LSTM networks can prevent gradients from vanishing with direct access to the forget gate's activations by an additive gradient structure that will update its memories after every learning process. This ability of LSTM models is proven by many studies, including a study done by The Technical University of Munich[9]. We experimented using solely 唐詩, Tang poems, with 5 or 7 words in each sentence, 4 or 8 sentences in each poem and of 5 major genres (farewell, patriotism/appreciation of history, war, homesickness/ boudoir repinings, describing

scenery/objects to express feelings) only in this stage.

Data processing

To reconstruct our dataset into computer readable data, our dataset is first tokenized by assigning random integers to each character. Then, we made use of two techniques to prevent a probable genre bias brought upon by our slightly imbalanced dataset. An effective solution for the imbalanced dataset is data augmentation. Two methods of data augmentation are used in this project: (1) replacing one random character in each poem with a blank and (2) switching the positions of 2 adjacent characters in each poem. Data augmentation balances the number of poems for each genre and increases the size of the dataset. This method not only lets the model consider poems of all genres to equal extents, but it also creates a more generic and quantitative dataset, allowing general features of different genres to be identified more easily and reducing overfitting.

Model Architecture

The LSTM model consist of multiple layers. The first layer is the embedding layer, where Chinese characters in our dataset are vectorized. Additional hidden layers are implemented to recombine the learned representation from prior layers and create new representations at high levels of abstraction, recognizing more complex patterns in the data than in single layer models.

In addition, Bidirectional LSTM is used as a replacement for basic LSTM. By this, additional context is added to the network and results are produced faster. The final layer of the models is a Dense layer with softmax activation. The softmax activation ensures that the sum of the components of the output vector is 1.

Experimental Results

The models with data augmentation perform best with a high resultant training accuracy and highest validation accuracy. The training and validation loss decreases continuously.

The classification of poems using the final LSTM model (data augmentation (1)) is most accurate and generally resembles the human interpretation of the poems' genres. The model is able to determine the extent of each genre in a poem objectively, in which usually the most significant genre has a probability of over 50%, the related genres, above 10%, and the unrelated ones, less than 3%.

However, the LSTM models struggle to distinguish between the genre "war" and "patriotism" precisely, likely due to the lack of distinction between "war" and "patriotism" genres. Many war poems contain ideas or messages about patriotism. Therefore, the model fails to identify features that are exclusive to war poems.

Another interesting result is the difference between the prediction of genres of separate sentences and of the whole poem. The model has been tested respectively on each sentence of a poem and the whole poem. The results came up vastly different. In a typical "describing scenery to express feelings" + "farewell" poem, 《夜雨寄北》(Appendix), the model predicted the first sentence to have 27.7% for the genre "farewell" and 65.7% for the genre "describing scenery to express feelings". For the second sentence, it predicted a majority of 81.8% for genre describing scenery to express feelings and only 6.45% for genre "farewell". However, the prediction for both sentences combined (the whole poem) results in 38.2% on genre "farewell", and 55.0% in genre "describing scenery to express feelings", which is different from the separate predictions of both sentences. This shows our model's ability to look at a poem as a whole when predicting genres, resembling human interpretation skills.

We have also experimented our models on poems of unseen Chinese poetic types (appendix 3) and varying lengths (not only with five/seven words per sentence and four/eight sentences).《虞美人・春花秋月何時了》(Appendix) is a Song poem that talks about the remembrance of the poet's fallen country and his desperation and sadness through describing how the scenery stays the same, but people have changed. The model predicted it to have a majority of genre "describing scenery to express feelings", followed by "homesickness" and "farewell". The model is able to predict the meanings of the poems quite accurately but not explicit enough, given that the text may contain themes other than our 5 genres.

(stage 2: Transformer)

To test the limits of Neural Networks on Chinese poem classification, other forms of traditional Chinese poetic literature (元曲, 宋詞 and 唐詩 with other structures and lengths) with an expanded set of genres with 7 typical genres have been added to the dataset. However, the LSTM model does not perform well when processing long sentences as it is impotent in capturing very long-term dependencies.

Transformer[10] is another type of Neural Network that can build up on previous information throughout training, potentially solving the issue of vanishing-gradients[11]. It utilizes the Multi-Head Attention mechanism, which defines the weighted average, or amount of attention, given to each word when classifying, differentiating the genres more effectively than the LSTM model.

Model Architecture

The Transformer block consists of a self-attention layer and normalization layer, followed by a feed-forward layer. The self-attention mechanism assigns a value to each tokenized word of each poem, and changes as more examples are passed on. It recognizes the importance of each word on affecting the output effectively. Multi-head self-attention separates the embedding vector into multiple divisions for each attention head and generates keys, values and queries for each dimension, recognizing relationships between different combinations of words and the expected output.

The weights of each sequence is averaged in the global average pooling layer to give a value. Due to the permutational invariant nature of the network, position embeddings, which gives a value representing the position of each word in the sequence, are used in this task to give importance word order on the model's interpretation of the poem.

Unlike the common transformer using softmax to average the outputs in the final layer, sigmoid is used in the output layer as this task consists of 7 different categories and each poem can belong to more than one genre. Sigmoid allows each genre to achieve a high probability at the same time without affecting the percentages of other genres, giving more obvious probabilities closer to 100% for applicable genres.

Experimental Results

The results of the transformer model are relatively ideal with some limitations. Unlike the LSTM model, the imbalanced dataset does not affect the accuracy much as the model is more capable of recognizing and giving "attention" to relevant features in each sample. With a shorter training time than the LSTM model, the transformer model have reached a maximum training accuracy of 97% and validation accuracy of 82%.

In comparison with the LSTM model, the transformer model shows a better performance in giving high probabilities of above 80% to genres that are relevant to the poem. The irrelevance of other genres is also more distinct, generally giving probabilities of lower than 1% to obviously unrelated genres and values ranging from 10% to 50% for other related genres. For example, in Song poem 《浣溪沙・一曲新詞酒一杯》(Appendix), which illustrates the author's perception of time passing by through scenery such as the setting sun, withering flowers and swallows returning, the model predicted 94.6% for "lament" and 66.6% for "describing scenery to express feelings", with percentages for other genres lower than 5%. It seems to recognize the focus of the poem is on expressing feeling, bringing down the percentage of "describing scenery to express feeling". In 《聲聲慢・尋尋覓覓》(Appendix), a Song poem that describes the autumn scenery to express the sadness of losing the author's homeland after war, the model predicted 91.6% for "describing scenery to express feeling", 84.8% for "homesickness", 43.2% for "lament", and 30.9% and 14.3% for "patriotism" and "farewell" respectively, which is understandable given that the historic background and context of the poem is not provided. The model attempts to classify content with complex meanings in terms of the 7 genres, such as a majority of

"homesickness" and some extent of "farewell" and "patriotism" to understand the feeling of losing one's country in 《聲聲慢 • 尋尋覓覓》.

In the LSTM networks, the genres "war" and "patriotism" are often mixed up due to their similarities. However, Transformer models manage to distinguish between the two more accurately. For example, when predicting 《涼州詞》(Appendix), a representative Tang war poem, the LSTM model predicted a majority of 97.7% for "patriotism/appreciation of history", whereas the Transformer model predicts 72.2% for "war" accurately. Similarly, in another war poem《出塞（其一）》(Appendix), the LSTM model predicts it to have 68.5% for "patriotism/appreciation of history", and the Transformer predicted the highest percentage of 61.2% in "war". LSTM remembers dependencies from most recent sequences, while the Transformer considers each character equally and assigns weights to them, thus the Transformer can identify exclusive features of war poems more easily, differentiating them from genre "patriotism/appreciation of history" more effectively.

However, the Transformer model sometimes fails to recognize genre "homesickness/boudoir repinings", especially when interpreting Song and Yuan poems that convey varied ideas in more obscure ways. For instance, it only predicts 1.3% in genre "boudoir repinings" in 《春情》(Appendix), a Yuan poem that describes a girl's desperate feeling of missing her lover. Song and Yuan poems have more freedom and variety in structure and use of words, and usually express the main ideas indirectly as it emphasizes on the beauty and musicality of words, thus resulting in greater difficulty in identifying some genres.

Apart from that, we have also tested out the models on modern free verse poems which are written in vernacular Chinese, which is very different from the archaic Chinese used in our dataset in terms of content and use of words. 《再別康橋》(Appendix) is a famous Chinese modern poem that expresses the poet's sorrow of leaving Cambridge and his love for the school. When an extract of it is passed on to the models, the LSTM model predicted it to contain the genre of "patriotism" significantly (94.2%), which is irrelevant, whereas the Transformer model predicted 96.7% for "describing scenery to express feeling", followed by 31.6% for "lament", which is more relevant than the LSTM model. Due to the completely different language patterns, rhythms, poetic structures, and the ways of expressing emotions of modern poems as compared to traditional poems, the LSTM model struggles to understand the poems. As the Transformer model is trained on Song and Yuan poems in addition to Tang poems with rigid structure, and can identify 7 genres in addition to the LSTM model's 5 genres, the results suggest that the Transformer model can understand larger variants of text with different wordings and structures as compared to the LSTM model, and it also has high potential to be trained on other types of text such as modern Chinese.

Future Works

To the best of our knowledge, this is the first work using artificial intelligence for the genre classification of Chinese poems. The experimental results of the project have created a satisfactory model for Chinese poem genre classification and provided a detailed analysis on the effectiveness of different models. Future work will focus on pushing the models' accuracies to perfection by creating a larger dataset. We will make use of pretrained word embeddings and Zero-shot learning to help our model figuring out new unseen genres of a poem and increase our model's proficiency in understanding a poem's content. We will also expand on the types of poems that the model can analyze by training it with modern poems, modern songs and Chinese poetry of even older times like 楚辭 and 詩經. We also anticipate investigating the genre classification of poems that are not written by an acknowledged poet, but by a learner of Chinese literature instead. Classification of poems that do not necessarily follow typical poetic rules, with less structure and expected literary behaviors, and are characterized by the zeitgeist of the present-day will also be explored in the future. We aim to overcome the challenges of creating an accurate poem genre classification model, without having fixed poetic structures and typical words and linguistic patterns that appear in a specific genre. In addition to that,

we will also attempt to create a Chinese poem generation model that can construct poems of manually selected genres, with the help of our genre classification model.

Conclusion

In this project, we have developed an LSTM model and Transformer model that could classify different types of poems into 5 and 7 main categories respectively efficiently by using an original database of classified poems. We have used multiple Machine Learning methods, and Deep Learning algorithms by restructuring different combinations of LSTM models and using Transformer models to experiment on the poem classification task. By using these NLP methods to classify poems, we can provide those who are studying Chinese poems with a more organized and systematic poem-genre database, so that they can know the genres of the different poems more easily and conveniently. This can also help the development of future related projects as we have provided a new, more systematic and objective approach towards the classification of poems as a basis for other related studies.

Reference list

[1] Delmonte and Prati, 2014: *SPARSAR: An Expressive Poetry Reader*

[2] Kesarwani, 2018: *AUTOMATIC POETRY CLASSIFICATION USING NATURAL LANGUAGE PROCESSING*

[3] Zhou, Gao, Li, Shun: *The Design and Implementation of XiaoIce, an Empathetic Social Chatbot*

[4] Guo, Yi, Sun, Li, Yang, Liang, Chen, Zhang, Li: *Jiuge: A Human-Machine Collaborative Chinese Classical Poetry Generation System*

[5] Yazhou Ren, Peng Zhao, Yongpan Sheng, Dezhong Yao, and Zenglin Xu: *Robust Softmax Regression for Multi-class Classification with Self-Paced Learning*

[6] Rupesh Kumar Srivastava, Jonathan Masci, Faustino Gomez, Jürgen Schmidhuber: *Understanding Locally Competitive Networks*

[7] Jinxing Yu, Xun Jian, Hao Xin, Yangqiu Song: *Joint Embeddings of Chinese Words, Characters, and Fine-grained Subcharacter Components*

[8] Mahnoosh Kholghi, Lance De Vine, Laurianne Sitbon, Guido Zuccon: *The Benefits of Word Embeddings Features for Active Learning in Clinical Information Extraction*

[9] Sepp Hochreiter: *THE VANISHING GRADIENT PROBLEM DURING LEARNING RECURRENT NEURAL NETS AND PROBLEM SOLUTIONS*

[10] Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser, Polosukhin, 6 Dec 2017: *Attention Is All You Need*

[11] Zhiwei Wang, Yao Ma, Zitao Liu, Jiliang Tang: *R-TRANSFORMER: RECURRENT NEURAL NETWORK ENHANCED TRANSFORMER*

Websites to collect data from:
http://www.gushice.com/list_1_0_0_0_0.html
https://www.gushici.com/p_p_1
http://www.haoshiwen.org/type.php?x=1

Appendix

LSTM model predictions

Poem	**送別**	**愛國/詠古**	**戰爭**	**思鄉/閨怨**	**借景抒情/托物言志**
《夜雨寄北》唐・李商隱	38.2%	<5%	<5%	<5%	55.0%
君問歸期未有期，巴山夜雨漲秋池。	27.7%	<5%	<5%	<5%	65.7%
何當共剪西窗燭，卻話巴山夜雨時。	6.45%	<5%	<5%	<5%	81.8%
《虞美人・春花秋月何時了》 五代・李煜 春花秋月何時了？往事知多少。小樓昨夜又東風，故國不堪回首月明中。雕欄玉砌應猶在，只是朱顏改。問君能有幾多愁？恰似一江春水向東流。	10.4%	<5%	<5%	13.9%	73.2%
《涼州詞》唐・王翰 葡萄美酒夜光杯，欲飲琵琶馬上催。 醉臥沙場君莫笑，古來征戰幾人回。	<5%	97.9%	<5%	0.974%	<5%
《出塞（其二）》唐・王昌齡 秦時明月漢時關，萬里長征人未還。 但使龍城飛將在，不教胡馬度陰山。	<5%	68.6%	18.5%	<5%	11.4%
《再別康橋》徐志摩 [i]	<5%	94.2%	<5%	<5%	<5%

Transformer model predictions

Poem	**送別**	**愛國/詠古**	**戰爭**	**思鄉/閨怨**	**借景抒情/托物言志**	**感嘆**	**讚頌**
《涼州詞》唐・王翰 葡萄美酒夜光杯，欲飲琵琶馬上催。 醉臥沙場君莫笑，古來征戰幾人回。	29.3%	<5%	72.2%	20.8%	<5%	<5%	<5%
《出塞（其一）》唐・王昌齡 秦時明月漢時關，萬里長征人未還。 但使龍城飛將在，不教胡馬度陰山。	22.5%	<5%	61.2%	6.5%	<5%	<5%	<5%
《浣溪沙・一曲新詞酒一杯》宋・晏殊 一曲新詞酒一杯，去年天氣舊亭臺。夕陽西下幾時回？ 無可奈何花落去，似曾相	<5%	<5%	<5%	<5%	66.6%	94.6%	<5%

識燕歸來。小園香徑獨徘徊。							
《聲聲慢・尋尋覓覓》 宋・李清照 尋尋覓覓，冷冷清清，淒悽慘慘戚戚。乍暖還寒時候，最難將息。三杯兩盞淡酒，怎敵他、晚來風急？雁過也，正傷心，卻是舊時相識。 滿地黃花堆積。憔悴損，如今有誰堪摘？守著窗兒，獨自怎生得黑？梧桐更兼細雨，到黃昏、點點滴滴。這次第，怎一個愁字了得！	19.3%	30.9%	5.4%	84.8%	91.6%	43.2%	<5%
《春情》元・徐再思 平生不會相思，才會相思，便害相思。身似浮雲，心如飛絮，氣若游絲。 空一縷餘香在此，盼千金遊子何之？證候來時，正是何時？ 燈半昏時，月半明時。	<5%	<5%	10.5%	<5%	97.7%	<5%	<5%
《再別康橋》徐志摩 [i]	<5%	19.7%	<5%	10.6%	96.7%	31.6%	<5%

[i]

《再別康橋》徐志摩

輕輕的我走了，
正如我輕輕的來；
我輕輕的招手，
作別西天的雲彩。
　　那河畔的金柳，
是夕陽中的新娘；
波光裡的艷影，
在我的心頭蕩漾。
　　軟泥上的青荇，
油油的在水底招搖；
在康河的柔波里，
我甘心做一條水草！

　　那榆蔭下的一潭，
不是清泉，是天上虹；
揉碎在浮藻間，
沉澱著彩虹似的夢。
　　尋夢？撐一支長篙，
向青草更青處漫溯；
滿載一船星輝，
在星輝斑斕裡放歌。
　　但我不能放歌，
悄悄是別離的笙簫；
夏蟲也為我沉默，
沉默是今晚的康橋！
　　悄悄的我走了，
正如我悄悄的來；
我揮一揮衣袖，
不帶走一片雲彩。