國 立 成 功 大 學

資 訊 工 程 研 究 所

碩 士 論 文

使用關係與情境特徵進行社群文章電影預告推薦

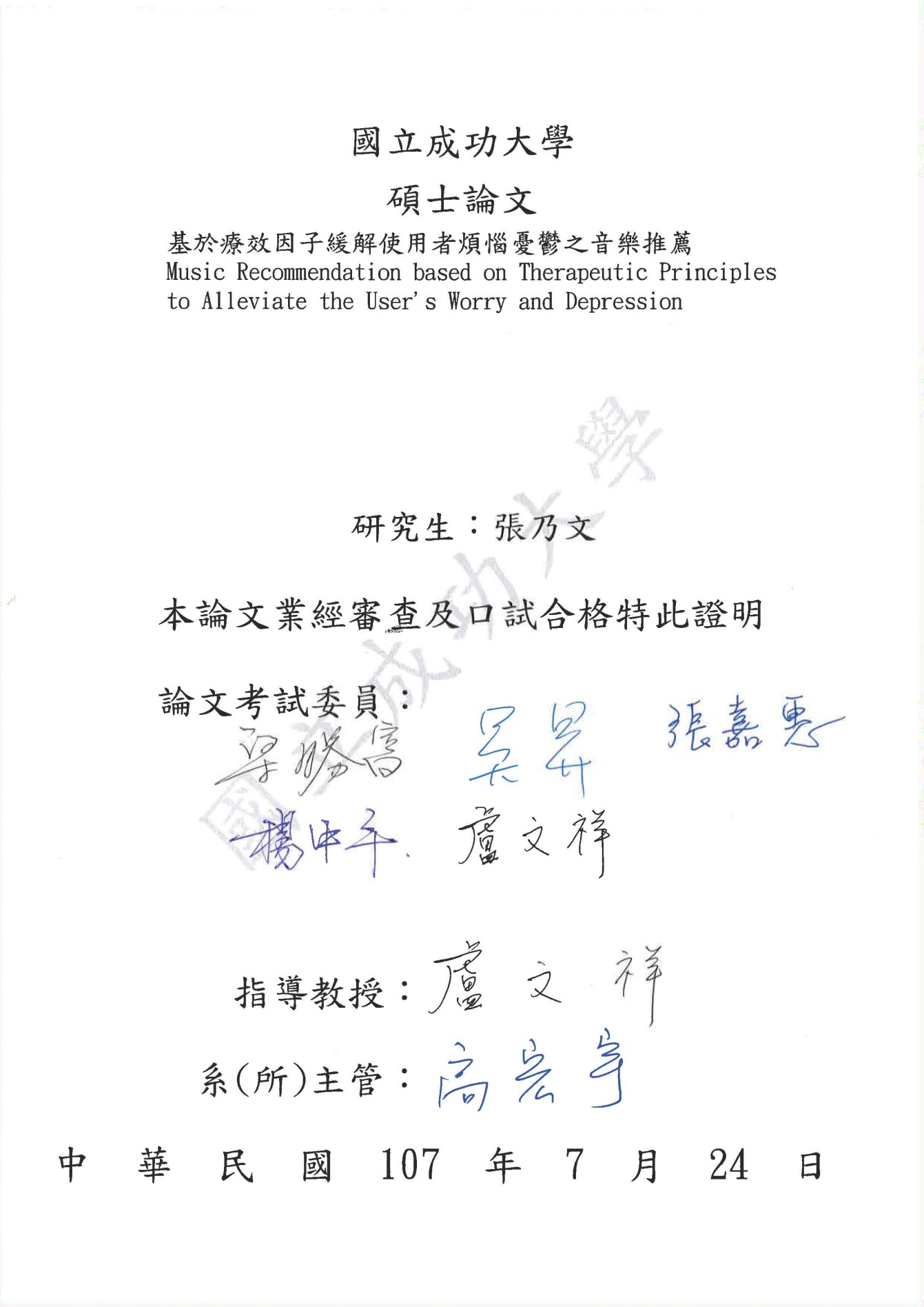
Using Relationship and Scenario Features of

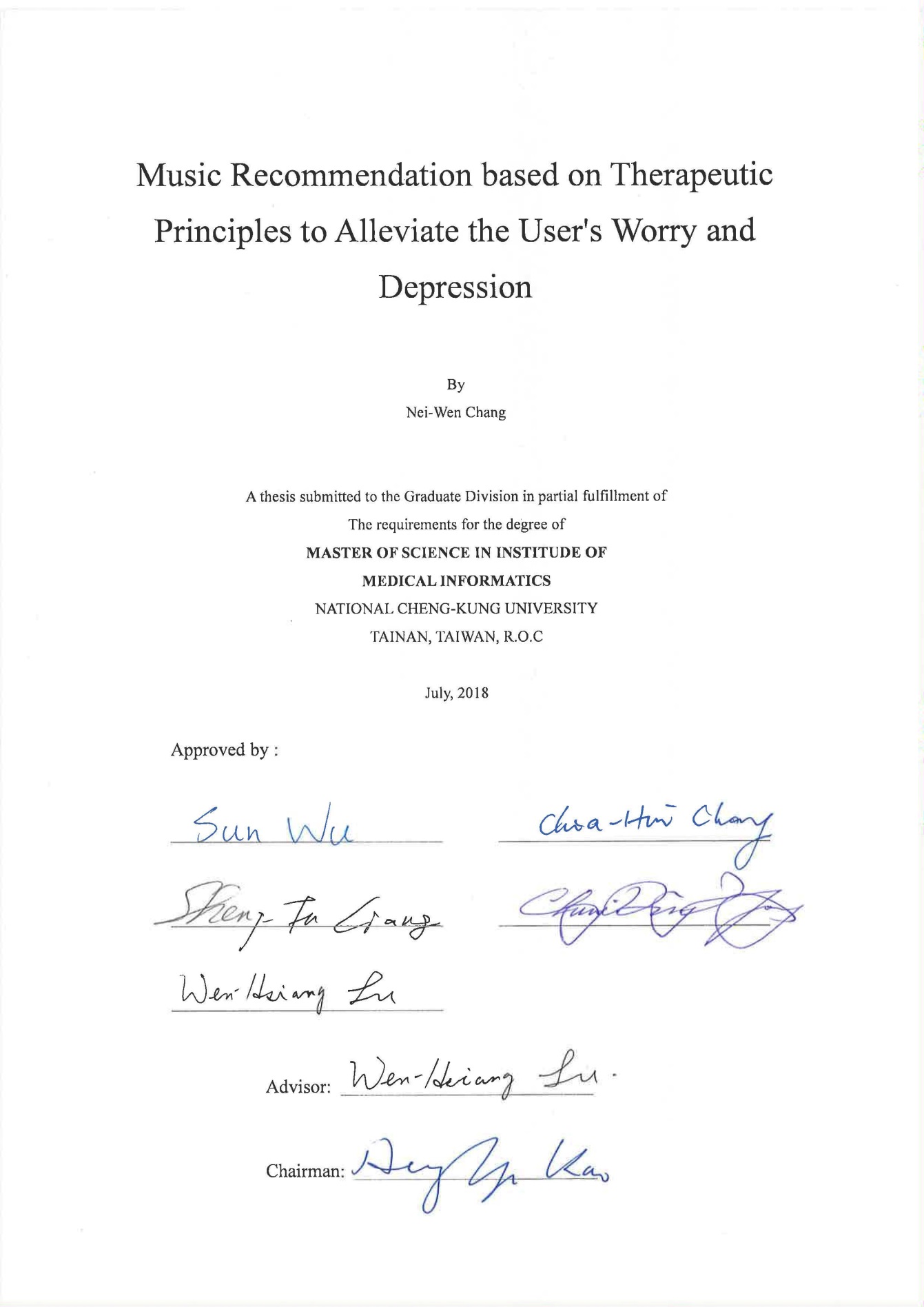
Plot Summaries for Social Article Trailer Recommendation

研究生：簡君聿

指導教授：盧文祥 博士

中華民國 一０八 年 七 月





摘要

推薦使用關係與情境特徵進行社群文章電影預告推薦

簡君聿\* 盧文祥\*\*

中華民國國立成功大學資訊工程研究所

社交平台上發表文章是年輕人最喜歡的活動。隨著電影產業的潛力，開發自動電影推薦引擎成為一個熱門話題。在社交媒體上，在共享相關預告片與關於日常生活在線社交平台的用戶生成文章的場景中，用戶傾向於選擇考慮其抒情主題的預告片。為了解決上述問題，我們提出了一種基於關係 - 場景的預告片推薦系統，該系統可以通過分析抒情主題來推薦預告片列表到輸入文章。我們認為抒情主題是關係和情景的結合，是情節總結的主觀和客觀視角。通過利用關係情景數據庫（Extend-HowNet作為知識庫），我們提取情節摘要和文章的關係和情景特徵。關係特徵表示為人物，情感，事件，地點和時間實體的實現。場景特徵表示為情感和事件實體的實現。

因此，我們表明，使用關係和場景特徵提供更好的推薦結果，而不僅僅考慮其中一個特徵，最後我們的推薦系統在用戶偏好和系統性能的兩個實驗中都優於新的W2V基線。我們還考慮用戶對系統關於不同關係類的偏好。

關鍵字：預告片推薦、情節摘要分析、文章分析、Word2Vec、變壓器雙向編碼器表示、卷積神經網絡、支持向量機、隨機森林分類器

\*作者 \*\*指導教授

Abstract

Using Relationship and Scenario Features of

Plot Summaries for Social Article Trailer Recommendation

Chun-Yu Chien\* Wen-Hsiang Lu\*\*

Institute of Computer Science and Information Engineering

National Cheng Kung University, Tainan, Taiwan, R.O.C

The post articles on the social platform is the favorite activity of young people. With the potential of digital movie industry, developing automatic movie recommendation engines becomes a popular issue. On social media, in the scenario of sharing related trailers with user-generated articles about daily life on line social platforms, users tend to choose trailers considering their lyrical theme.

To solve the above problem, we present a Relationship-Scenario-based Trailer Recommendation System which can recommend list of trailers to an input article by analyzing lyrical theme. We consider lyrical theme as a combination of Relationship and Scenario, the subjective and objective perspective of plot summaries. By utilizing relationship-scenario Database (Extended-HowNet as Knowledge base), we extract relationship and scenario features of plot summaries and articles. Relationship feature is represented as character, emotion, event, location and time entity relation. And scenario feature is represented as emotion and event entity relation.

Consequently, we show that using both relationship and scenario features provide better recommendation results than merely consider one of the features, In the end our recommender system outperforms a novel W2V baseline in both experiments of user preference and system performance. Also we consider user preference on our system about different relationship class.

Keywords: Trailer Recommendation, Plot summaries Analysis, Article Analysis, Word2Vec, Bidirectional Encoder Representations from Transformers, Convolutional Neural Network, Support Vector Machine, Random Forest Classifier

\*The Author \*\*The Advisor

致謝

首先真的很謝謝我的指導教授盧老師，老師以他心中懷有的崇高教育理念作為他的領導風格，帶領實驗室研究具有創意且與其他教授不同且前衛的觀點，故在我心目中我非常的尊敬他也非常感謝他，也希望接下來的學弟妹能夠繼續秉持堅持著老師帶領我們的精神繼續把自己的研究做到最好最亮。

在學習過程中與論文撰寫部分，非常謝謝在我身邊幫忙的人，包括老師的指導與國豪學長的幫忙跟觀念釐清，還有謝謝朋友佳純與敬濠的協助，過程中除了興奮以外還有感動，當然也謝謝在我的論文實驗中那些幫我標註的朋友們，未來也希望大家都可以順順利利朝著自己的夢想與目標前進，重點是要快樂的做自己。

最後我要謝謝我的父母親及其他愛我的家人，尤其是我的母親，讓我沒有經濟上的壓力可以完成碩班的學業，在我最後的求學生涯上當我最強大的後盾，最後，要感謝的太多了，那就謝謝老天吧！

君聿 八月于台南

Table of Contents

[摘要 IV](#_Toc13345198)

[Abstract V](#_Toc13345199)

[致謝 VII](#_Toc13345200)

[Chapter 1 Introduction 1](#_Toc13345201)

[1.1 Background 1](#_Toc13345202)

[1.2 Motivation 3](#_Toc13345203)

[1.3 Method 4](#_Toc13345204)

[1.4 Contribution 6](#_Toc13345205)

[1.5 Organization of this Dissertation 6](#_Toc13345206)

[Chapter 2 Related Work 7](#_Toc13345207)

[2.1 Studies on Sentiment analysis 7](#_Toc13345208)

[2.2 Studies on Film Trailer Topic Detection based on Plot Summaries 7](#_Toc13345209)

[2.3 Studies on Film Trailer Recommendation 8](#_Toc13345210)

[Chapter 3 Method 10](#_Toc13345211)

[3.1 System Framework 10](#_Toc13345212)

[3.2 Preliminaries 12](#_Toc13345213)

[3.2.1 CKIP Parser 12](#_Toc13345214)

[3.2.2 Extended-HowNet 13](#_Toc13345215)

[3.2.3 Data Sets and Preprocessing Steps 14](#_Toc13345216)

[3.3 Feature Generation 16](#_Toc13345217)

[3.3.1 Named Entity Recognition 16](#_Toc13345218)

[3.3.2 Training 20](#_Toc13345219)

[3.3.3 Relationship & Scenario Database 22](#_Toc13345220)

[3.4 Relationship Classification 25](#_Toc13345221)

[3.4.1 Relationship Term Extracting 26](#_Toc13345222)

[3.4.2 Relationship Model Training 27](#_Toc13345223)

[3.5 Scenario Classification 29](#_Toc13345224)

[3.5.1 Scenario Model Training 30](#_Toc13345225)

[3.6 Scenario Classification 32](#_Toc13345226)

[3.7 Relationship-Scenario based Recommendation 35](#_Toc13345227)

[Chapter 4 Experiments 36](#_Toc13345228)

[4.1 Dataset 36](#_Toc13345229)

[4.2 Experiment of Entity Embedding Quality 37](#_Toc13345230)

[4.2.1 Dataset for word Embedding 37](#_Toc13345231)

[4.2.2 Evaluation Metrics 37](#_Toc13345232)

[4.2.3 Evaluation Result 38](#_Toc13345233)

[4.3 Experiment of Relationship Classification 40](#_Toc13345234)

[4.3.1 Dataset for Relationship Classification 40](#_Toc13345235)

[4.3.2 Evaluation Metrics 42](#_Toc13345236)

[4.3.3 Experiment Result 43](#_Toc13345237)

[4.4 Experimental Data of Lyrics Pronoun Generation Model 44](#_Toc13345238)

[4.5 Analysis of Experimental Results of Lyrics Pronoun Generation Model 45](#_Toc13345239)

[4.5.1 All Three Annotators Labeled as Correct 48](#_Toc13345240)

[4.5.2 All Three Annotators Labeled as Incorrect 49](#_Toc13345241)

[Chapter 5 Conclusions 50](#_Toc13345242)

[5.1 Conclusion 50](#_Toc13345243)

[5.2 Future Work 50](#_Toc13345244)

[Reference 52](#_Toc13345245)

[Appendix A 54](#_Toc13345246)

[Appendix B 61](#_Toc13345247)

[Appendix C 63](#_Toc13345248)

# Introduction

## Background

A recommender system is a type of information filtering system which attempts to predict the preferences of user and make suggestion based on these preferences. There are vast variety of applications for recommender systems. In recent years, with the rapid development and astonished achievement on AI field. more and more researcher using the AI technique to develop more user-friendly and user-closed recommender system. Due to the advances in recommender system, users constantly expect good recommendation.

Trailer has become a phenomenal trend in movie industry over the last decade, according to the report from '中國產業信息網' survey, the china cinema revenue has grown from 1 billion CNY dollars in 2002 to 45 billion CNY dollars in 2017(Fig 1.1.1).On the other hand, With the increased popularity of film, people gradually change their way of watching movie. and watching movies online is now very popular, according to 'ARO / MMX' survey, watching movies field online up to 45.2% in 2017(Fig 1.1.2)

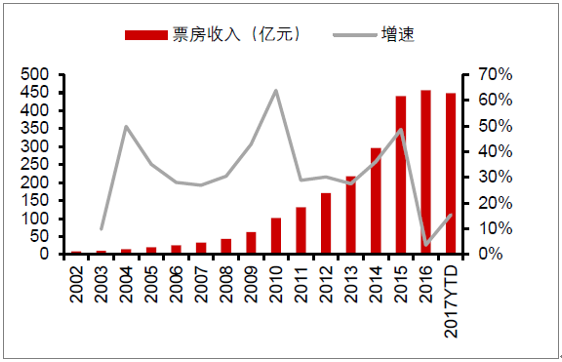


Figure 1.1.1 China Cinema Revenue from 2002 To 2017

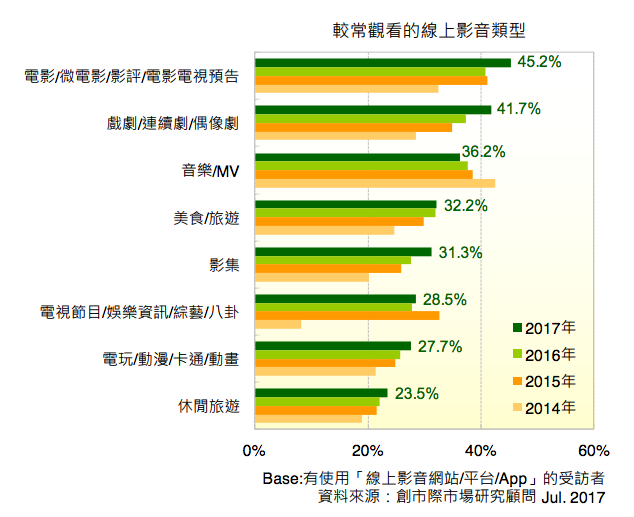


Figure 1.1.2 User Need from 2014 To 2017

## Motivation

With the growth of social network platforms, users can easily share personal stories online. Mood article, often it is the relationship and scenario between the life of a PO writer, whether it is a PO writer or a reader who likes this article, they want to be involved in the article. The storyline or movie trailer, often the relationship and scenario between the protagonist in the play, it also the trigger for the audience to want to go to the theater or watch movies online. Sometimes, in order to help the readers to better understand these stories, if also share a trailer, maybe will impact users watching the film and then increase the cinema revenue.

For example shown in Fig 1.2, a user posted an article and recommender the trailer "生生". The subject of this article matched the main idea of trailer which process of growth.

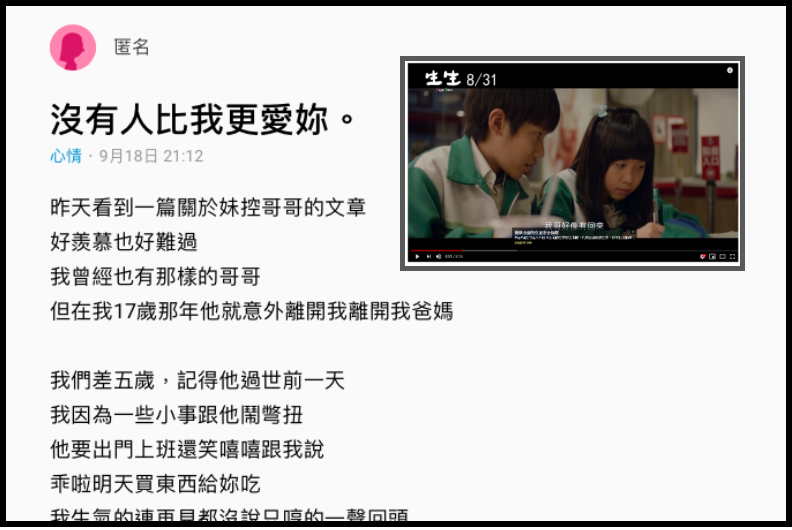


Figure 1.2 Example of User Post Articles with A Related Trailer

Above example, we observe that when users try to convey their stories, if we can recommend the trailer that are relationship and scenario similar, maybe it is a good idea. However, it is not an easy task for users to find the most thematically related trailer of their stories. Actually, this is state of the art idea, it doesn't currently exist. Therefore, in this paper, we introduce a recommendation system to help overcome the difficulty of finding relationship and scenario similar trailer of an article.

## Method

A recent research focus on film review [1, 2], it's direct sentiment analysis or opinion mining, but in this paper, we explore implicit user intention and we utilize relation between articles and film storyline. Two keys of “Storyline Themes” are pointed out in the article, "character relationship express" and “event and emotion description”. Storytelling is said to “be a common device among film”. It can either be in a narrative style of story or be related to a central theme of a person’s life. On the other hand, a film can fit into different moods that help the trailer “keeping the theme intact”.

Every user has different preferences and likes. In addition, even the taste of a single user can vary depending on a large number of factors, such as mood, season, or type of activity the user is doing. according to this problem, if you like this article, also you will like this trailer based on plot. In the domain of mining movie textual contents, recommender system has analysis film review and user preference [3, 4, 5]. However, to the best of our knowledge, the potential of combining both Relationship and Scenario elements for storyline analysis have not been explored yet.

Consequently, we extend the above ideas and further define the “Storyline Themes” of a trailer is composed of two elements: “Relationship” and “Scenario”. “Relationship” is the way in which two or more people or things are connected, or the state of being connected. In this paper, we view 5 features (character-object, event, time, location, emotion) of each film storyline and article as a probability distribution over several relationship by applying pre-trained model BERT [24] and W2V-SG[6] and then we use CNN Model Architecture [7] that implement classification. “Scenario” is a written outline of a film, novel, or stage work giving details of the plot and individual scenes which is the objective observation of human affairs which include the people and the things that are involved. We represent scenario as a feature vector that composed of several emotion and event concepts to capture the word level meaning of context. we take Extended-HowNet [8, 9, 10, 11, 12] as our knowledge base in our work. Last, we suppose social articles share the same feature space with film storyline and experiment the usability of our analysis result by developing a Relationship-Scenario-based Trailer Recommendation system for social articles.

## Contribution

The main contributions of our work are listed below:

* We use trailer storyline(plot summaries) and article as data source and analyze content in relationship and scenario perspectives.
* We generate relationship features utilizing Extended-HowNet, a Chinese knowledge base.
* We classify trailer storyline(plot summaries) into genres based on a variety of relationship-scenario feature extracted from the storyline(plot summaries).
* We proposed an application of Relationship-Scenario-based Trailer Recommendation system for recommending trailers to social articles using storyline(plot summaries) features.

## Organization of this Dissertation

The rest of the paper is organized as follows. In Chapter 2, we introduce several related works that are prior researches on analyzing movies or film recommender system. In Chapter 3, we give the details of our observation and method. In Chapter 4, we evaluate our system and show analysis results. Finally, In Chapter 5, we draw the conclusion and give some insights for future work.

# Related Work

In the chapter 2, we briefly review a number of researches that are relevant to our topic, including studies on sentiment analysis, studies on film trailer topic detection based on plot summaries and studies on film trailer recommendation.

## Studies on Sentiment analysis

Martineau and Finin (2009) [18] weighted bag-of-words in employing a delta TF-IDF function for training SVMs to classify the reviews. Maas et al. (2011) [19] introduced a model to catch sentiment information and word meanings. Dai Quoc Nguyen (2004) [20] present a new feature type named rating-based feature and evaluate the contribution of this feature to the task of document-level sentiment analysis, however, they only analyze polarity or shallow sentiment, it's very rough. our paper consider reader intention on social article, and then infer to similar movie plots. Learn more about readers' hidden intentions, and further analyze emotion deeply.

## Studies on Film Trailer Topic Detection based on Plot Summaries

Some previous studies showed their interested on analysis plot summaries topics. Hoang, Q et al. (2018) [16] presented a Predicting Movie Genres Based on Plot Summaries, this project explores several Machine Learning methods to predict movie genres based on plot summaries which like utilized Naive Bayes, Word2Vec+XGBoost, Recurrent Neural Networks.

Ali Mert Ertugrul et al. (2018) [17] expected to reflect the genre of movies since many spectators read the plot summaries before deciding to watch a movie and the project perform movie genre classification from plot summaries of movies using bidirectional LSTM (Bi-LSTM). However, considering human understanding on viewing radar chart, they only used normal topics which were too clear for people to tell what storylines were really about. In our work, we do relationship classification and specific scenario classification and further achieve real life expression.

## Studies on Film Trailer Recommendation

Some studies by using movie reviews for recommendation and showing how are they different from our work. XS Vu et al. (2017) [4] presented Mining User/Movie Preferred Features Based on Reviews for Video Recommendation System, the statistic from 70% customers consult reviews or ratings before purchasing. Along with this, viewers also check movie reviews before making decision to buy movie tickets. Apply LDA for finding hidden aspects for addressing user preference aspects and movie feature aspects. After user preferences based on aspects and movie aspects are addressed, KL divergence is used for measuring similarity between movie and user. Top K movies that close to user preferences are recommended to user. it's direct sentiment analysis or opinion mining, it's a polarity research but in this paper, we further explore implicit user intention and we utilize relation between articles and film storyline. Two keys of “Storyline Themes” are pointed out in the article, "character relationship express" and “event and emotion description”. Storytelling is said to “be a common device among film”. It can either be in a narrative style of story or be related to a central theme of a person’s life. On the other hand, a film can fit into different moods that help the trailer “keeping the theme intact”.

# Method

## System Framework

In this paper, we purpose a Relationship-Scenario-based Trailer Recommendation System using relationship and scenario features of plot summaries. The proposed system framework is shown in Fig 3.1.

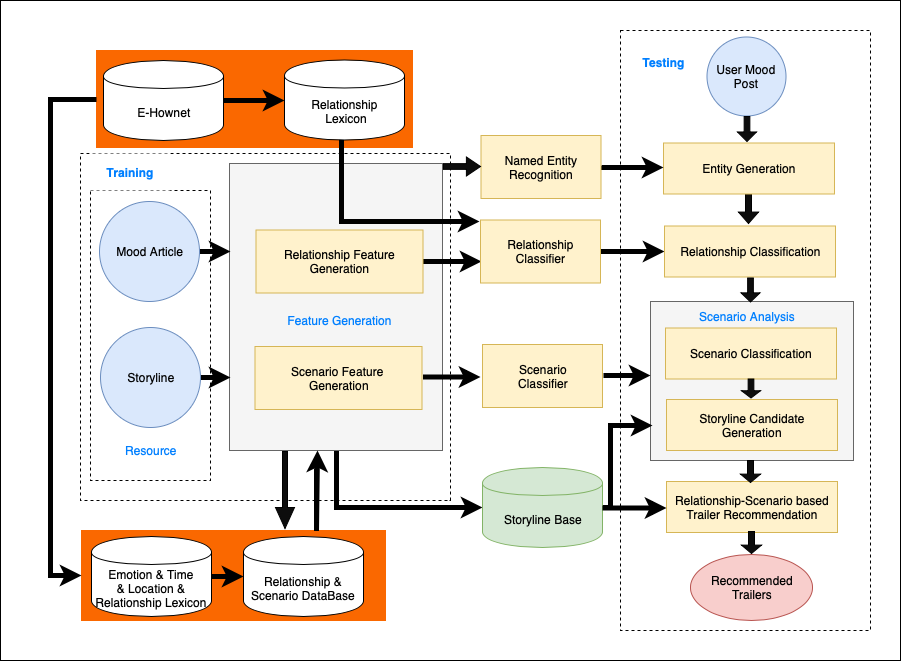
****

Figure 3.1 System Framework of Relationship-Scenario-based Trailer Recommendation

We defined a set of documents D = {P, S}, where P is a set of social mood articles and S is storyline(plot summaries) dataset. When a user submits a social article p ∈ P, we want to return a recommended list of trailers where each storyline s ∈ S. To complete this task, we want to represent each document d ∈ D as a composition of two features 𝑓 = {𝑓r, 𝑓𝑠}. Feature 𝑓r and 𝑓s denotes the relationship and scenario concepts of the given document. The entire system contains four major processing steps:

(1) Features Generation

Feature Generation is mainly divided into two parts, the first is the relationship feature generation, and the second is the scenario feature generation, Given a corpus c ∈ C, we will extract the character-object, time, location, events and emotions related entity using named entity recognition application and CKIP parser [13]. With the name entity recognition, we will then generate the Entity2Vec Model by applying bert algorithm or skip-gram algorithm.

(2) Relationship Classification

In relationship classification, we will extract the relationship related terms using a predefined relationship lexicon. The relationship lexicon is built with the help of a Chinese knowledge base Extended-HowNet [8]. we will define 7 common classes of article relationships. With the relationship feature 𝑓r, we will build a multi-class relationship classifier. Applying the classifier, we can classify each document into single class of relationship and thus narrow down our film trailer candidates for recommendation output.

(3) Scenario Classification

In scenario classification, according to different relationship we will define several common classes of film scenario. With the scenario feature 𝑓𝑠, we will build a multi-class scenario classifier. Applying the classifier, we can classify each document into single class of scenario and thus narrow down our film trailer candidates for recommendation output.

(4) Relationship Classification

Recommendation step formulates the relevance between a given social article p and each storyline(plot summaries) s ∈ S using both relationship feature 𝑓r and scenario feature 𝑓𝑠. Then our system will return a recommended list of trailers according to the ranking of relevance score.

## Preliminaries

In this section, first of all we are going to introduce CKIP Parser and we give a brief introduction of Extended-HowNet which we use as our knowledge base. We also describe how we prepare our datasets and our preprocessing steps.

### CKIP Parser

Before analyzing the patterns of entity, we must obtain the word segmentation results and semantic features of the articles. We use CKIP Chinese Parser to help us obtain the segmentation results of the articles. CKIP Chinese Parser [13] is a tool developed by the CKIP(Chinese Knowledge and Information Processing Group). It can automatically analyze articles and obtain a lot of useful information, such as word segmentation, POS tag, syntax tree and semantic role. The relevant standards and rules for segmentation, syntax tree and semantic role are documented in the CKIP technical report [13][14][15]. In Fig 3.2.1 we show the result parsed by CKIP Chinese Parser. It can be seen that the sentences are parsed into a tree structure according to the grammar. The words that have the same parent form a phrase. Each word and phrase has a POS tag and semantic role. In following chapter, we will take advantage of POS tag and semantic role to build our pattern of entity. The detailed usages are described in Section 3.2.1.

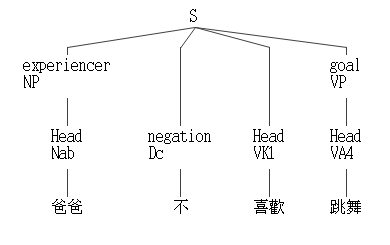


Figure 3.2.1 The example of result parsed by CKIP Chinese Parser

### Extended-HowNet

Extended-HowNet (E-HowNet) is a frame-based entity relation model [8, 9, 10, 11, 12] which expended from HowNet. Each word sense can be decomposed into its simplest concepts and defined by E-HowNet. The taxonomies of concepts are organized to hierarchical structure. A part of ontology map of E-HowNet taxonomy structure is shown in Fig 3.2.2

In this paper, we take the advantage of E-HowNet's hierarchical representation and utilize it in two ways. First, we build lexicons by extracting all terms defined by some concepts. Second, by mapping a term to its corresponding E-HowNet hypernym, we can replace the word level representation of a document to the concept level representation. The detailed usages are described in Section 3.4.1.

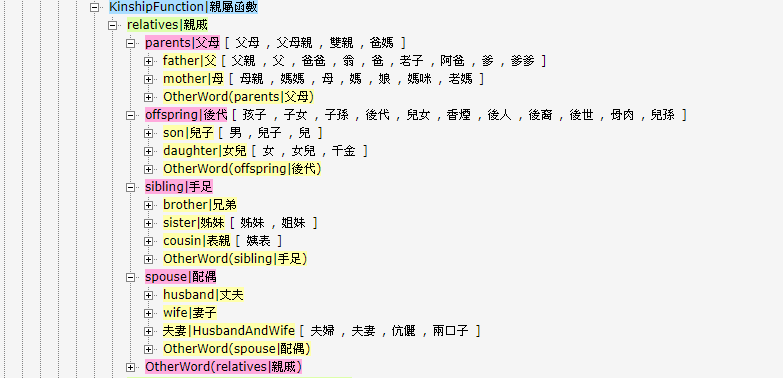
****

Figure 3.2.2 A chip of E-HowNet taxonomy structure

### Data Sets and Preprocessing Steps

We use two types of datasets which are social mood articles and film storyline(plot summaries). In regard that we want to recommend trailers utilizing storyline features, trailer storyline and social articles are used as our training data. On the other hand, social mood articles are used as our testing data. We describe the characteristics of our datasets as follows.

**Social Mood Articles Data**

Social articles are user-generated essays that are published on social websites. We choose social articles from Dcard, a popular online social platform for teenager in Taiwan, as our social article source for the reason of its simplicity in article categorization. Since we aim at recommend trailers that are related to user posts about daily life, we main collect posts from “Mood (心情版) ” boards. We don’t use “Movies (電影版)” board because most of the articles on Movie board are about sharing movie review, not sharing one’s daily life. Each social article contains post title and user generated text.

**Plot Summaries Data**

For film storyline data, we extracted from the most popular Mandarin website yahoo.com(奇摩電影) and pixnet.com(痞客邦電影).  Film Storyline are usually composed of short pieces of paragraphs. Some paragraphs may attract movie lovers to watch movies. Table 3.1 shows and example of a film plot summaries. We record film title and plot summaries and link for each trailer.

|  |  |
| --- | --- |
| Trailer Title | 高年級姐妹會 Book Club |
| Trailer Link | https://www.youtube.com/embed/l2Uo0jv77Ug |
| Plot Summaries | 黛安在結婚40年後成為了寡婦；薇薇安周旋在和男人之間不需承諾的感情關係中；雪倫在離婚幾十年後成了工作狂；凱倫則是在維持了35年平淡無奇的婚姻後生活開始出現危機…這四位「高年級」的好姐妹們組成了一個讀書會，平時固定交流各自讀過的好書，直到有天薇薇安帶來了風靡全球的情慾小說《格雷的五十道陰影》，推薦給其他三位姐妹。在一口氣讀完《格雷的五十道陰影》的刺激下，她們發現自己過去已經錯過許多人生的精采之處，為了有所改變，這四位熟齡姐妹們決定，要從今天開始踏上既瘋狂又感性的全新人生旅程... |

Table 3.1 Example of Film Storyline

**Data Preprocessing**

A chinese parser, including word segmentation/POS tagging/parsing/role assignment, has been completed. For parser, both storyline dataset and social mood article set are parsed using CKIP Parser system [13]. Besides, a selected list of stopwords is removed from both mood article and film storyline in dataset. Finally, through the observation, a dictionary was also established for filtering.

## Feature Generation

The first part of our system is Feature Generation. In this part, we focus on finding a representation that can showcase the relationship state and scenario state of a given document. The model generating step is shown in Fig 3.3.

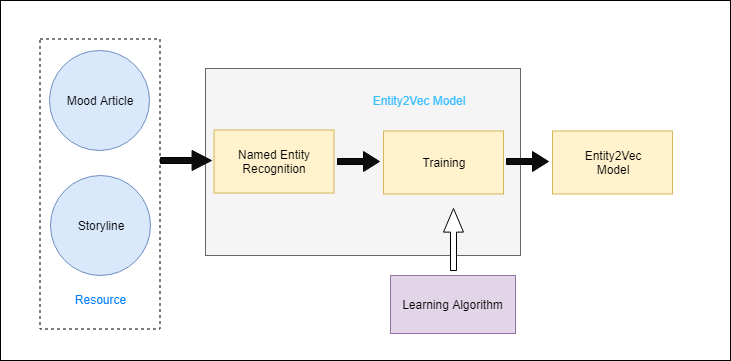


Figure 3.3 Structure of Feature Generation

### Named Entity Recognition

We use articles and storyline for observation and then define POS tagging and semantic role dictionary. The rule based, we can extract the character-object, time, location, events and emotions related entity using named entity recognition application and CKIP parser [13]. For example shown in Fig 3.3.1.

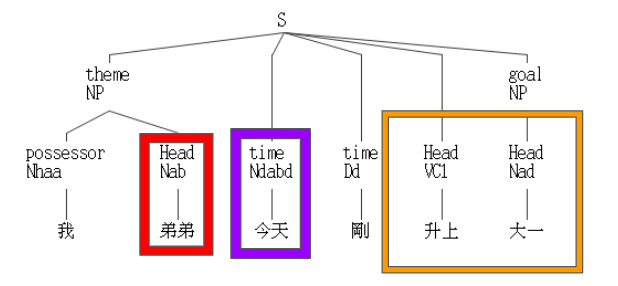


Figure 3.3.1 Name Entity Recognition Using CKIP Parser

How we extract character-object candidates, location candidates, time candidates, emotion candidates and events candidates? we use CKIP parser and exploit specific POS tagging and semantic role. it achieve feature extraction application. We randomly sample 50 articles and 50 storyline for observation and coverage calculation for all data are Shown in Table 3.2, Table 3.3, Table 3.4, Table 3.5 and Table 3.6.

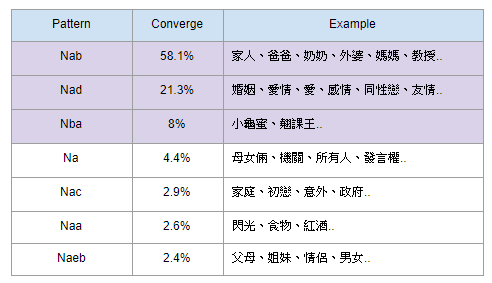


Table 3.2 The distribution and example of character-object feature

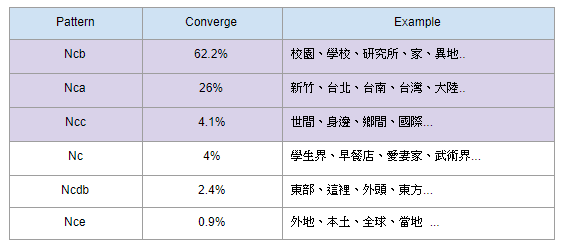


Table 3.3 The distribution and example of location feature

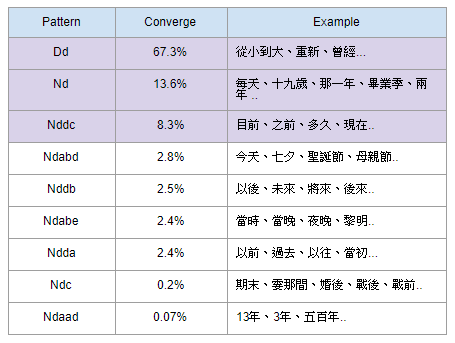


Table 3.4 The distribution and example of time feature

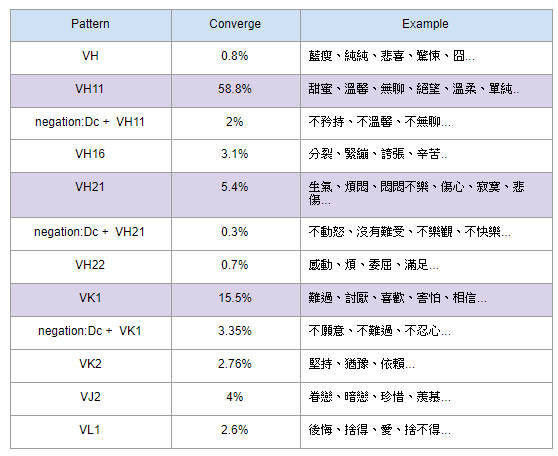


Table 3.5 The distribution and example of emotion feature

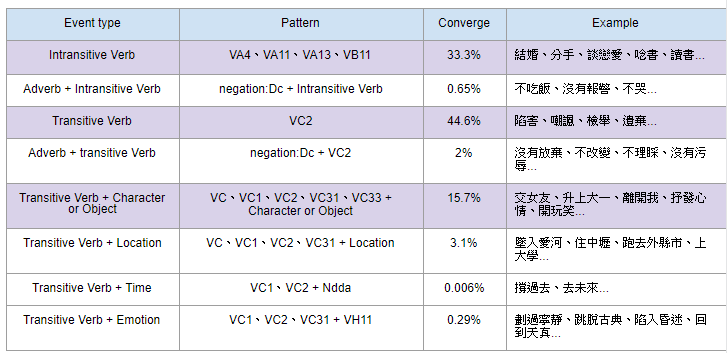


Table 3.6 The distribution and example of event feature

### Training

Now having character-object candidates, emotion candidates, time candidates, location candidates, and events candidates of articles and storyline, we want to identify the relationship state and scenario state of documents. We consider the relationship state and scenario state of a document to be represented as a probability distribution over a number of relationship topic related and scenario topic related. And we use BERT and Word2Vec embedding method.

**BERT**

A new language representation model called BERT[24], which stands for Bidirectional Encoder Representations from Transformers. Unlike recent language representation models, BERT is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task-specific architecture modifications and bert architecture shown in Fig 3.3.2.1.

BERT is conceptually simple and empirically powerful. It obtains new state-of-the-art results on eleven natural language processing tasks, including pushing the GLUE score to 80.5% (7.7% point absolute improvement), MultiNLI accuracy to 86.7% (4.6% absolute improvement), SQuAD v1.1 question answering Test F1 to 93.2 (1.5 point absolute improvement) and SQuAD v2.0 Test F1 to 83.1 (5.1 point absolute improvement).



Figure 3.3.2.1 Model for BERT

**Word2Vec**

Word embeddings were revolutionized by Mikolov et al. [6] who proposed the CBOW and skip-gram models. CBOW computes the conditional probability of a target word given the context words surrounding it across a window of size k. On the other hand, the skip-gram model does the exact opposite of the CBOW model, by predicting the surrounding context words given the central target word. The context words are assumed to be located symmetrically to the target words within a distance equal to the window size in both directions. In unsupervised settings, the word embedding dimension is determined by the accuracy of prediction. As the embedding dimension increases, the accuracy of prediction also increases until it converges at some point, which is considered the optimal embedding dimension as it is the shortest without compromising accuracy. Let us consider a simplified version of the skip-gram model where only one word is considered in the context. This essentially replicates a bigram language model.

As shown in Fig 3.3.2.2, the skip-gram model is a simple fully connected neural network with one hidden layer. The input layer, which takes the one-hot vector of target word has V neurons while the hidden layer has N neurons. The output layer is softmax of all words in the vocabulary.

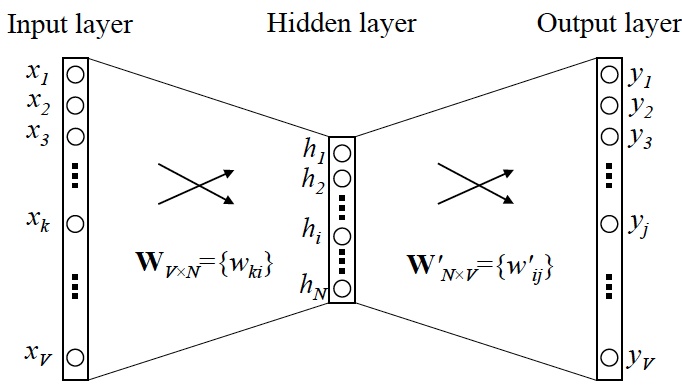


Figure 3.3.2.2 Model for skip-gram

### Relationship & Scenario Database

**Data Base**

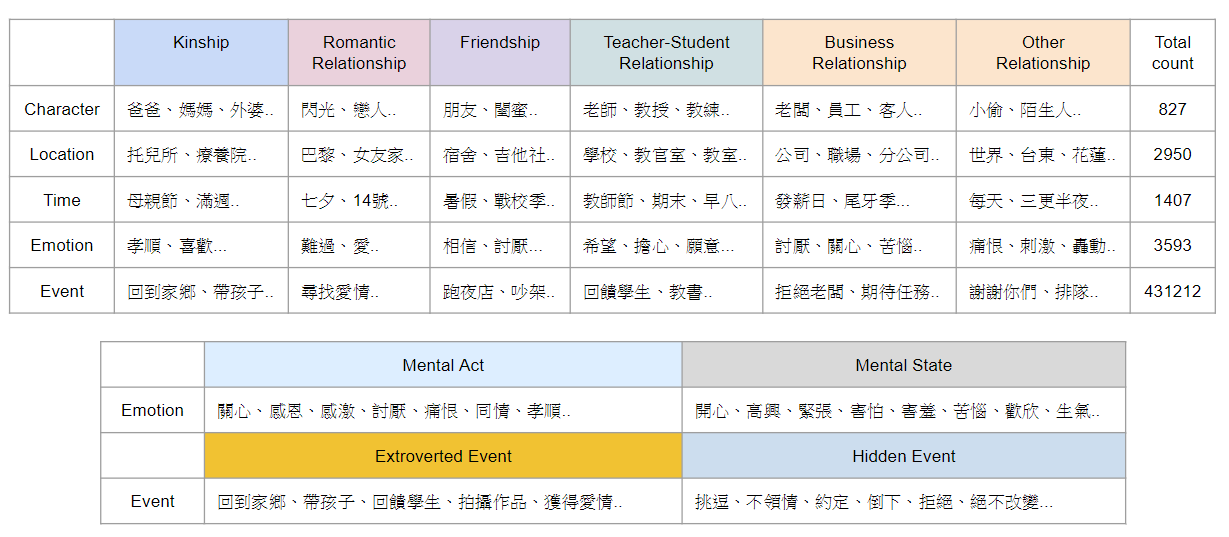


Table 3.7 Relationship & Scenario Database

* Character

E-HowNet Relationship Lexicon And One Annotator

* Location

E-HowNet  Place Lexicon And One Annotator

* Time

E-HowNet Time Lexicon  And One Annotator

* Emotion

E-HowNet Emotion Lexicon

* Event

Article Event

**Relationship Feature**

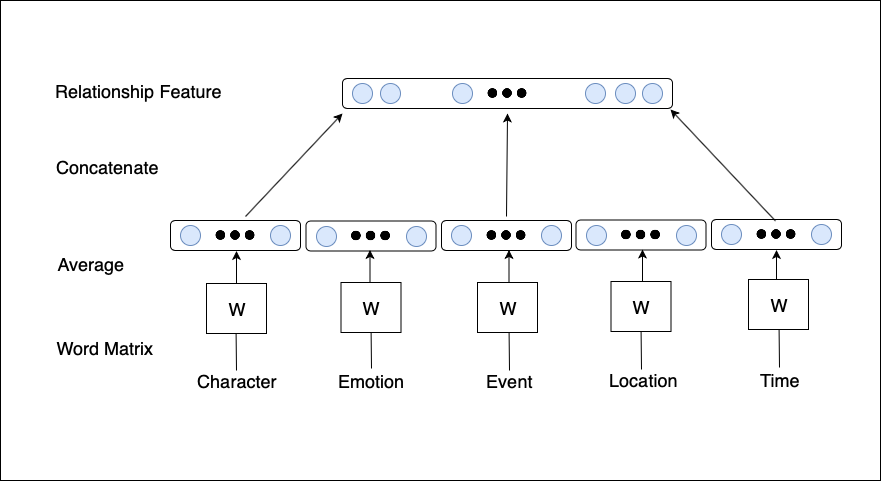


Figure 3.3.3.1 Relationship Feature

We consider character, emotion, event, location and time entities about relationship feature. In this framework, every entity is mapped to a unique vector, represented by a column in a matrix W. The column is indexed by position of the word in the vocabulary. The average of the vectors is then used as entity feature and then concatenate every entity feature will become relationship feature. As shown in Fig 3.3.3.1.

**Scenario Feature**

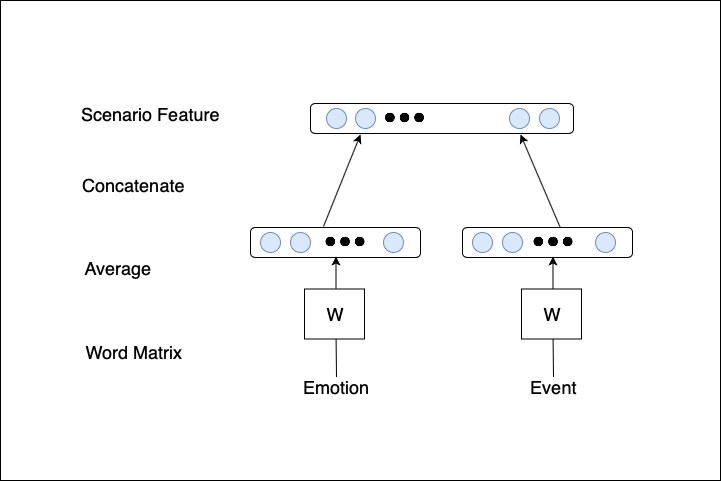
****

Figure 3.3.3.2 Scenario Feature

We consider emotion and event entities about scenario feature. In this framework, every entity is mapped to a unique vector, represented by a column in a matrix W. The column is indexed by position of the word in the vocabulary. The average of the vectors is then used as entity feature and then concatenate every entity feature will become scenario feature. As shown in Fig 3.3.3.2.

## Relationship Classification

The second part of our system is Relationship classification. In this part, we focus on finding a representation that can showcase the relationship class of a given document. The model generating step is shown in Fig 3.4.

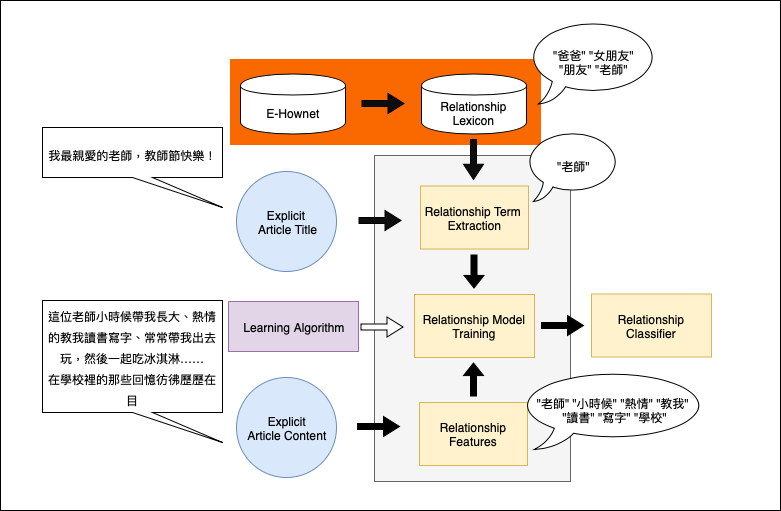


Figure 3.4 Structure of Relationship Classification

First of all, we define relationship lexicon built from E-HowNet. Using the lexicon, we can extract relationship terms from each explicit title article document for ground truth. Then we apply Convolutional Neural Network (CNN) [7] for relationship model training. Finally, based on the relationship classifier, the relationship class of each document can be represented as a probability distribution over the number of relationship topics, which we called relationship genre.

### Relationship Term Extracting

To extract relationship term, first we need to define which terms are related to relationship and further build relationship lexicon. We define seven types of relationships through articles and movie plots, which can be divided into family relationship, romantic relationship, friendship, teacher and student relationship, business relationship, others and no relationship etc.. For examples shown in Table 3.8, It is an example of a teacher-student relationship, we observe that no matter what kind of relationship will be affected by the character-object, emotion, time, location, and events.

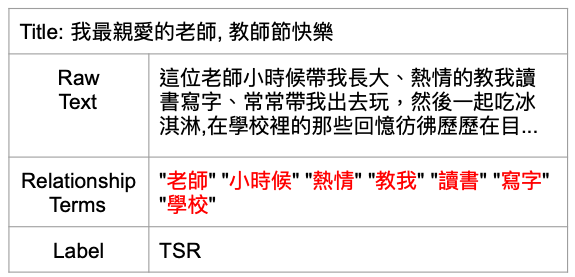


Table 3.8 Example of mood Article

On the other hand, the lexicon mentioned in the previous section is mainly used to use the title of the article, that is, to use lexicon for semi-automatic tagging is shown in Table 3.9. It can be seen from the table that there are mainly five types, which are just mentioned in the other Two types of artificial annotator.

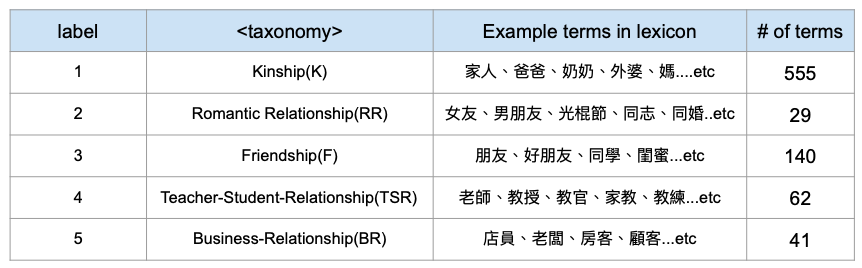


Table 3.9 E-HowNet Relationship Lexicon

### Relationship Model Training

Now having article genres and extracted relationship features (character-object, emotion, time, location, and events) of articles, we want to identify the relationship genre of documents. We consider the relationship state of a document to be represented as a relationship genre using supervised learning method.

**Convolutional Neural Network**

Convolutional neural networks (CNN) utilize layers with convolving filters that are applied to local features (LeCun et al., 1998). Originally invented for computer vision, CNN models have subsequently been shown to be effective for NLP and have achieved excellent results in semantic parsing (Yih et al., 2014), search query retrieval (Shen et al., 2014), sentence modeling (Kalch-brenner et al., 2014), sentence classification (Yoon Kim et al., 2014)[7], and other traditional NLP tasks (Collobert et al., 2011).

**Model Architecture**

Apply Convolutional Neural Network(CNN) (Kim, Yoon. 2014. Convolutional neural networks for sentence classification), based on this structure, we create fit our case model. As shown in Fig 3.4.2.

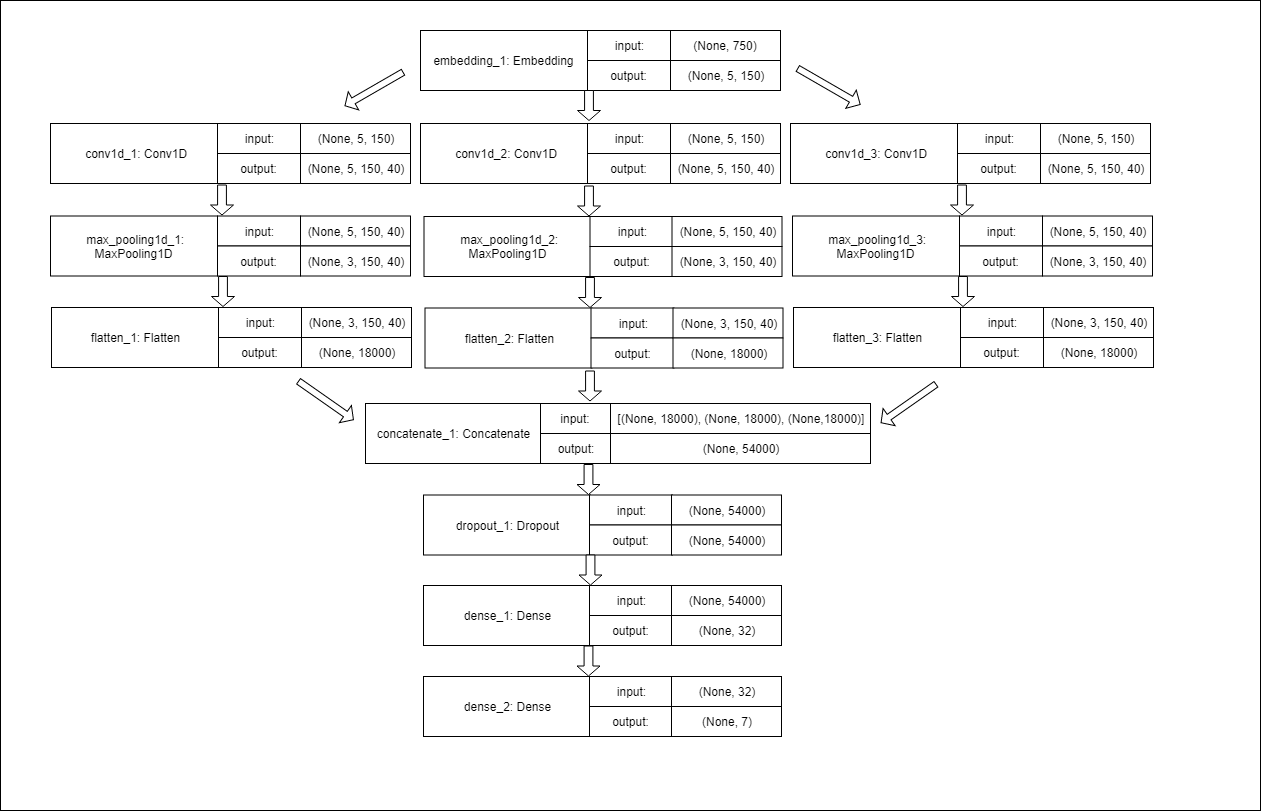


Figure 3.4.2 CNN Architecture

## Scenario Classification

The third part of our system is Scenario classification. In this part, we focus on finding a representation that can showcase the scenario class of a given document. The model generating step is shown in Fig 3.5.

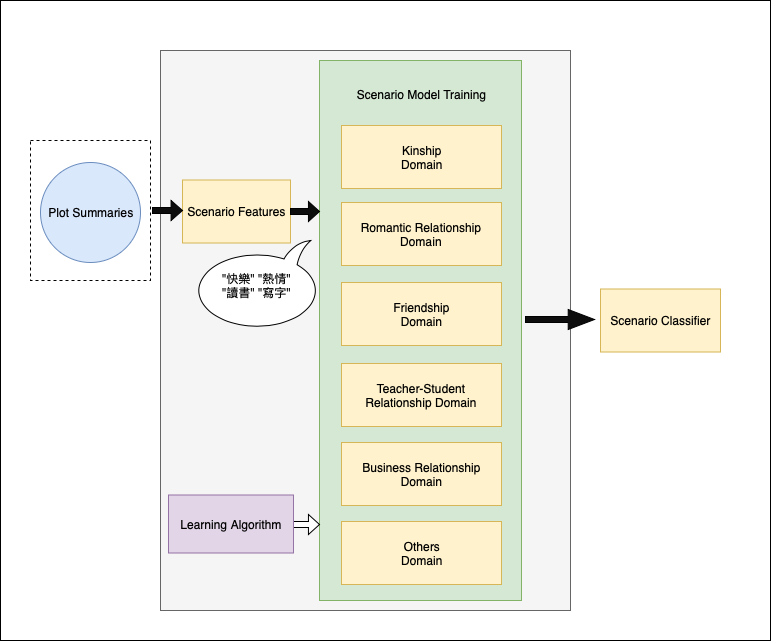
****

Figure 3.5 Structure of Scenario Classification

First we can obtain scenario features from each document. Then we apply Machine Learning (Random forest, SVM, Naive Bayes) for scenario classification. Finally, based on the scenario classifier in different relationship domain, the scenario class of each document can be represented as a probability distribution over the number of scenario topics, which we called scenario genre.

### Scenario Model Training

Now having trailer genres and extracted scenario features (emotion and events) of trailer, we want to identify the scenario genre of documents. We consider the scenario state of a document to be represented as a scenario genre using supervised learning method.

**Naïve Bayes(Baseline)**

In [machine learning](https://en.wikipedia.org/wiki/Machine_learning), naive Bayes classifiers are a family of simple "[probabilistic classifiers](https://en.wikipedia.org/wiki/Probabilistic_classifier)" based on applying [Bayes' theorem](https://en.wikipedia.org/wiki/Bayes%27_theorem) with strong (naive) [independence](https://en.wikipedia.org/wiki/Statistical_independence) assumptions between the features. Naive Bayes has been studied extensively since the 1960s. It was introduced (though not under that name) into the text retrieval community in the early 1960s,[21] and remains a popular (baseline) method for text categorization.



Figure 3.5.1 Naïve Bayes

**Support Vector Machine**

In [machine learning](https://en.wikipedia.org/wiki/Machine_learning), support-vector machines (SVMs, also support-vector networks[22]) are [supervised learning](https://en.wikipedia.org/wiki/Supervised_learning) models with associated learning [algorithms](https://en.wikipedia.org/wiki/Algorithm) that analyze data used for [classification](https://en.wikipedia.org/wiki/Statistical_classification) and [regression analysis](https://en.wikipedia.org/wiki/Regression_analysis). Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic [binary](https://en.wikipedia.org/wiki/Binary_classifier) [linear classifier](https://en.wikipedia.org/wiki/Linear_classifier) (although methods such as [Platt scaling](https://en.wikipedia.org/wiki/Platt_scaling) exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

**Random Forest Classifier**

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees habit of overfitting to their training set.

The first algorithm for random decision forests was created by Tin Kam Ho using the random subspace method [23] which, in Ho's formulation, is a way to implement the "stochastic discrimination" approach to classification proposed by Eugene Kleinberg.

## Scenario Classification

Build scenario classifier to generate storyline candidates for recommendation Q: What kinds of scenarios are most likely for user to want to post a trailer with?

* Separately observe 25 social articles and film storyline about daily life for kinship, romantic relationship, friendship, teacher-student relationship, business relationship and others respectfully.  Shown in Table 3.10, Table 3.11, Table 3.12, Table 3.13, Table 3.14 and Table 3.15.
* Movies by Genre: <https://www.imdb.com/feature/genre/>

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classes of Scenario | #articles | %(/100) | #movies | %(/100) | Definition |
| Filial(F) | 4 | 16% | 6 | 24% | relating to or due from a son or daughter |
| Love(L) | 19 | 76% | 15 | 60% | feel deep affection for (someone) |
| Betray(B) | 2 | 8% | 4 | 16% | to not be loyal to your family |

Table 3.10 Definition of scenario class for kinship and Percentage of labeled articles

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classes of Scenario | #articles | %(/100) | #movies | %(/100) | Definition |
| In Love(IL) | 10 | 40% | 14 | 56% | feel romantically attracted to them, and they are very important. |
| Lost Love(LL) | 14 | 56% | 7 | 28% | when you've lost someone that was your world. |
| Sex(S) | 1 | 4% | 4 | 16% | feel sexually attracted to them |

 Table 3.11  Definition of scenario class for romantic relationship and Percentage of labeled articles

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classes of Scenario | #articles | %(/100) | #movies | %(/100) | Definition |
| Support(S) | 22 | 88% | 21 | 84% | give assistance to, have memories with someone, or Face any difficulties together |
| Betray(B) | 3 | 12% | 4 | 16% | to not be loyal to your friends |

Table 3.12 Definition of scenario class for friendship and Percentage of labeled articles

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classes of Scenario | #articles | %(/100) | #movies | %(/100) | Definition |
| Learning(L) | 24 | 96% | 24 | 96% | the acquisition of knowledge or skills through study, experience, or being taught. |
| Others(O) | 1 | 4% | 1 | 4% | relationship other than learning for teacher-student |

Table 3.13 Definition of scenario class for teacher relationship and Percentage of labeled articles

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classes of Scenario | #articles | %(/100) | #movies | %(/100) | Definition |
| Interest(I) | 12 | 48% | 7 | 28% | get along with or compete with interests |
| Task(T) | 12 | 48% | 12 | 48% | a piece of work to be done or undertaken, or belief |
| Reveal(R) | 1 | 4% | 6 | 24% | make (previously unknown or secret information) known to others. |

Table 3.14 Definition of scenario class for business relationship and Percentage of labeled articles

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classes of Scenario | #articles | %(/100) | #movies | %(/100) | Definition |
| Pursuit of Self(PS) | 10 | 40% | 14 | 56% | pursuit of self-behavior or stick to your dream |
| Others(O) | 15 | 60% | 11 | 44% | not fit any of the scenario above, maybe it’s very complex scenario |

Table 3.15 Definition of scenario class for others relationship and Percentage of labeled articles

**Storyline Candidate Generation**

* Having the scenario classifier, the testing article will first be classified into specific class.
* Given the storyline base, with each storyline in has already classified into the storyline candidates for recommendation is

https://lh6.googleusercontent.com/1H7O-oqLxpyCmKkpL0RmFMIqGL1v1wI_IAkSytJp2IYWY7vLAq4JuzRAVEdFvjxDYX6GW42c_TczMSVvHZRKuNXC4XNCHytu3IAVA3vz48DTLirAp-MIxe3_eIUCMxsqMc89wnlD(1)

## Relationship-Scenario based Recommendation

Having the relationship features and scenario features for a social article p and a set of trailer storyline s, we now aim at finding trailer that are highly related to article p. In our work, we assume that a storyline s is suitable for article p if they are in similar relationship state and are stories about similar scenario.

The trailer emotions and events is what we value most. Therefore, For each articles(p) state and storyline(s) state, we consider scenario state. To find the best trailers for p in storyline set s, for each storyline s ∈ S, we first calculate the relevance scenario score between p and s. The relevance scores are calculated as follows.

https://lh6.googleusercontent.com/XXhXPaOGWblsZJT0i94n47E2cAWvG4MRx76gb3VvdU-KNEB-BO1e8p6NM4K8a9u0BXNjsgXERBqPnTjAwGkTEbNWnT49_I2r8qOouJoJ5hLdI0SUSeY-xWjNxgb9fonJcxOnW-rd (2)

where https://lh4.googleusercontent.com/kkvqLOmzqA5WOJgpZVxLvu_VeKhvkVAkOb8NCKWWZZJ7dd_x0nbnpoUF4EzdrxSXhlHwzJHPYPSENfL3l-UAOWemJosAvOGUQcD4o0DT0D3vh0KOYE1XtChIenBhtpDB_mcwAHGG is scenario feature of article p, https://lh4.googleusercontent.com/MTo2pWwejAvUm5HW5ndZvE758pBJ89dg-3gUHnv4jLPSCdeE1NaV2EYcDOqnbQue-JxyDRDhE6FMNb-Qz6NFSKgVqNbeKcyEYMqBsPkxjalZZdElUJf6wsjWksLsp0MZ_i7xJCSf is scenario feature of storylines s. In the end, by ranking trailer storyline candidates with their recommendation scores, we can retrieve the recommended trailer list for an article.

# Experiments

## Dataset

(1) Social Articles

For social articles, We crawl social articles from “Mood (心情版)” boards of Dcard. The Social Articles - Dcard, a popular online social platform for young people in Taiwan. Therefore, we consider use it. Table 4.1 shows the number of articles we retrieved from board.

|  |  |
| --- | --- |
| Board Name | # Articles |
| Mood 心情版 | 215857 |

Table 4.1 Number of articles retrieved from Dcard board

(2) Storyline (Plot Summaries)

For trailer plot summaries data, we consider yahoo and pixnet movies, an online storyline database.  Table 4.2 shows the number of films we retrieved from movie web.

|  |  |
| --- | --- |
| Movie Web | # Film |
| Yahoo | 2926 |
| Pixnet | 796 |
| Total | 3722 |

Table 4.2 Number of films retrieved from movie web

## Experiment of Entity Embedding Quality

In entity embedding, We apply Mikolov et al. [6] who proposed the skip-gram models and Jacob Devlin et al. [24] who proposed BERT.

### Dataset for word Embedding

**Data Set**

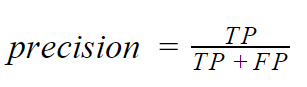
We use about one 3 million of sentences from our articles set and film plot summaries as training data.

### Evaluation Metrics

Two metrics are employed for Entity2Vec performance evaluation: precision (P@1) and normalized discounted cumulative gain (NDCG@k).

* **Precision at 1(P@1):**

The equation of precision is defined as follows:

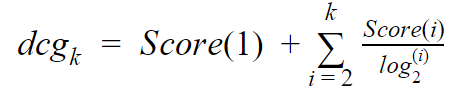


(3)

where the TP here is whether the rank order result retrieved by Enetiy2Vec model is the standard answer score.

* **Normalized Discounted Cumulative Gain at k (NDCG@k):**

With the relevance score, we also calculate NDCG to examine whether the terms of higher scores have the higher rank. We first calculate the DCG score as follows:



(4)

where 𝑘 is the number of top 𝑘 returned results, and https://lh4.googleusercontent.com/jnxGV2A7r0s7iyQXcbgTfmKFr4NMPvRRCLZe3m0e-dU_74Rr0zrViyuo6Oa_-nIuZXw5xfl0N8fzMnnXc6qv7aUdZVlSpN8UJ19RNeV6whCCM2G6Sefh4GuFQkx24jYmPLvRrfks is thehttps://lh3.googleusercontent.com/wkxHZMaHq3v2sKNBYprjF7oz3WP1xKJi_H1JojaHCBMf7Nvlfcujj65JtZC3Yr7e-rOebNe6VpehwUXQBdU75vylTQ-lFqeNGSP1bM3wrZEyZPfPUR94JP78kJXKYw5YTHLaIGfUscore of results. Then, we normalize DCG value with IDCG and forms the NDCG score as follows:

https://lh3.googleusercontent.com/J4LBGKiyLOMA0uA5AmI2F0YbVxZbVuJeDcXeRl3lAj1TFyTvn2cCoc0Yikos0Wq3rEGc4ByW87LIrbzic276HmnjN8eCY9net9aWJdGpRWL4DCVSmwMD0po_Cdvca1zBQO8DpMQr

(5)

where IDCG refers to the ideal DCG score by sorting the terms by their relative relevance score. On calculating NDCG, we also set k from 1 to 10.

### Evaluation Result

**Experiment Protocol**

* Entity2Vec
  + Skip-gram(W2V-SG-150) and Skip-gram(W2V-SG-300)
    - Training Set: Mood Articles and Storyline
  + Bidirectional Embedding Representations from Transformers(BERT-768)
    - Training Set: wiki data
* Testing Emotion and Event
  + Top 50 frequent unique emotion
  + Top 50 frequent unique event

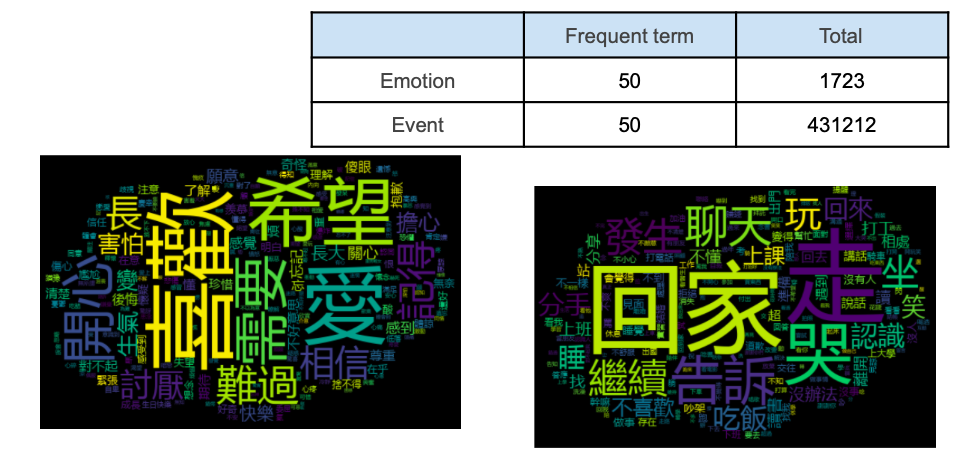


Figure 4.2.3.1 Frequent emotion and event count

* Three score criteria for entity embedding quality:
  + High relevance (score is 2): semantically correct or highly relevant
    - E.g. Event "聊天" > "抬槓", Emotion "記得" > "不會忘記"
  + Relevance (score is 1): semantically correct or relevant
    - E.g. Event "相處" > "不合拍", Emotion "開心" > "拿到禮物"
  + Not relevance (score is 0): semantically incorrect or less relevant
    - E.g. Event "不知道" > "沒這麼", Emotion "分手" > "復合"

We evaluate the performance of three different dimension:150(w2v-sg) and 300(w2v-sg) and 768 (bert). The result is shown in Fig 4.2.3.2.

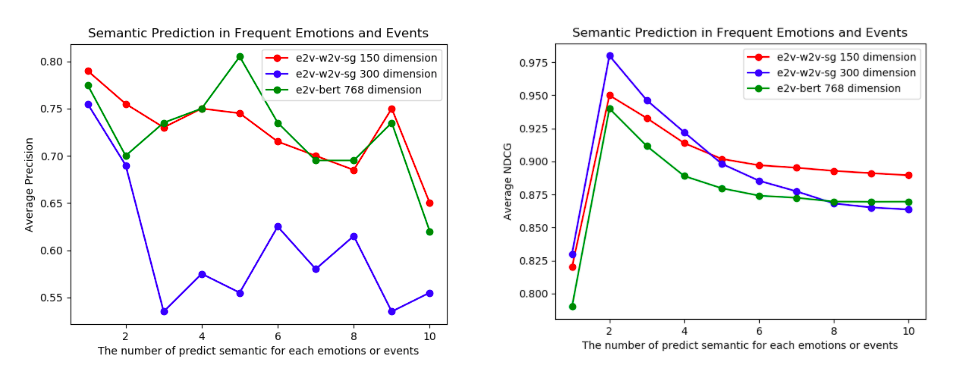


Figure 4.2.3.2 Precision and NDCG Result for w2v-sg-150 and w2v-sg-300 and bert-768

## Experiment of Relationship Classification

In relationship classification, we aim at finding a better embedding for multi-class classifier which can classify documents into one.

### Dataset for Relationship Classification

**Dataset**

We use about one fourth of articles from our articles set as training data. Percentage of labeled articles in each different annotator shown in Table 4.3 and Table 4.4.

|  |  |  |
| --- | --- | --- |
| Label Method | # Articles | Percentage(/100) |
| E-HowNet-Specific Relationship | 45678 | 96.30% |
| Annotator | 1756 | 3.70% |
| Total | 47434 | 100% |

Table 4.3 Percentage of labeled articles in each different annotator

|  |  |  |
| --- | --- | --- |
| E-HowNet | # Articles | Percentage(/100) |
| E-HowNet-Specific Relationship | 45678 | 97.91% |
| E-HowNet-Different Relationship | 976 | 2.09% |
| Total | 46654 | 100% |

Table 4.4 Percentage of labeled relationship articles in E-HowNet

**Label Strategy**

Beside semi-automatic tagging, we have artificial annotator. To label the training data, we employ two annotators who is interested in articles to annotate. Annotators label each article in training data to relationship labels by the following guideline shown in Table 4.5.

For each article, the relationship labels are then set to be the union set of labels labeled by the two annotators and semi-automatic tagging. Table 4.6 shows the percentage of articles in each class of relationship.

|  |
| --- |
| **Guideline for Labeling Relationship Classes** |
| Definition of Relationship are as follows:   * **Kinship(K):** Some typical characteristics of a family are support, mutual trust, regular interactions, shared beliefs and values, security, and a sense of community. * **Romantic Relationship(RR):** Romantic partnerships, including marriage, are close relationships formed between two people that are built upon affection, trust, intimacy, and romantic love. * **Friendship(F):** A friendship can be thought of as a close tie between two people that is often built upon mutual experiences, shared interests, proximity, and emotional bonding. * **Teacher Student Relationship(TSR):** Teachers are those that have the ability to maximise the learning potential of all students in their class. Developing positive relationships between a teacher and student is a fundamental aspect of quality teaching and student learning. * **Business Relationship(BR):** An association between individuals or companies entered into for commercial or responsibility purposes. Many senior corporate executives maintain a friendly business relationship with an extensive network of other executives. Or about task responsibility for employee. * **Others (O):** Articles that do not fit any of the relationship above (it mean other relationship) * **No Relationship(NR):** no fit any relationship |

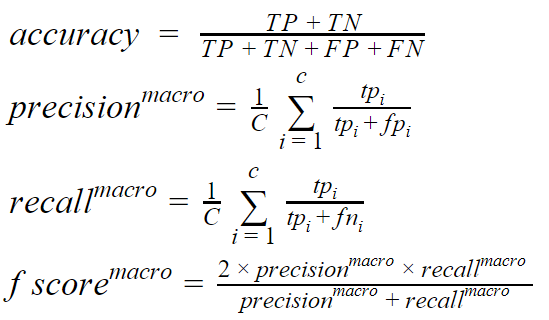
Table 4.5 Guideline for labeling Article relationship

|  |  |  |
| --- | --- | --- |
| Classes of Relationship | # Articles | Percentage(/100) |
| Kinship(K) | 20590 | 43.4% |
| Romantic Relationship(RR) | 11876 | 25.03% |
| Friendship(F) | 11668 | 24.59% |
| Teacher-Student Relationship(TSR) | 1359 | 2.86% |
| Business Relationship(BR) | 1542 | 3.25% |
| Others(O) | 268 | 0.56% |
| No Relationship(NR) | 131 | 0.27% |
| Total | 47434 | 100% |

Table 4.6 Percentage of labeled articles in each class of relationship

### Evaluation Metrics

To evaluate the performance of multi-class classifiers, We use 10-fold cross validation via our dataset and calculate the accuracy and hyperparameters are tuned by cross-validation on the task training data and evaluation test data. To further compare performance of different methods and we consider holdout data, I calculate the confusion matrix, accuracy, macro precision, macro recall, and F-score for each genre as well as for all of the test data (holdout data). These metrics are defined as:



(6)

(7)

(8)

(9)

where C is the number of target labels, TP, FP, TN and FN are true positive, false positive, true negative and false negative respectively. Note that, since we perform our experiments on a single dataset, micro precision, micro recall and micro f-score values are all equal and they represent the accuracy of the classifier. Accordingly, we only present the macro results.

### Experiment Result

We evaluate the performance of one multi-class classifiers: Convolutional Neural Network (CNN). We use 10-fold cross validation via our dataset and calculate the accuracy for different embedding method (e2v\_bert, e2v\_w2v\_sg and w2v\_w2v\_sg). The result is shown in Table 4.7.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 10-fold CV  Acc. | Accuracy | Macro  Pre. | Macro  Rec. | Macro  F1 |
| CNN-W2V-W2V-SG | 73.94% | 74.55% | 40.84% | 42.63% | 41.69% |
| CNN-E2V-W2V-SG | **84.12%** | **83.87%** | **66.75%** | **60.94%** | **62.00%** |
| CNN-E2V-BERT | 74.90% | 74.07% | 50.52% | 49.06% | 49.72% |

Table 4.7 Relationship Genre Classification Results (%) Measured by ACCURACY, PRECISION, RECALL and F-SCORE

## Experimental Data of Lyrics Pronoun Generation Model

In this study, the song list collection of Installation of Hope Song mainly refers to the user comments or articles of PTT Prozac Board and the top ten healing songs of Ministry of Health and Welfare, and David's most helpful decompression song in [[19]](#_Reference_45). As mentioned in 3.5.1, if the cosine similarity of lyrics and user’s article is lower than threshold, then the songs in the 10 decompression songs will randomly recommended. We set the threshold to 0.7, if the similarity after the comparison is <0.7, Decompression songs will recommended (decompression songs are almost all songs without lyrics, mainly based on song melody). Table 4.6 shows the sources and examples of songs. Please refer to Appendix B for the song list of all songs.

Table 4.6 Sources and Examples of Songs.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Song of Installation of Hope | | | |
| Source of Songs | Health Promotion Administration | PTT Comments  or Articles | | Decompression Song from [19] |
| Number of Songs | 10 | 82 | | 10 |
| Similarity Threshold of  Recommended Song | 0.7 | | | 0.7 |
| Examples | 1. 讓我罩著你 2. 祝你幸福 3. 好的事情 4. 心內有數 5. 十年一刻 6. 傷心的人別聽慢歌 7. 欠一個勇敢 8. 讓我罩著你 9. 再出發 10. 那些你不敢解決的問題 | | 1. 為自己加油 2. 一起加油吧 3. 長途夜車 4. 如煙 5. 打開天空 6. 一定要相信自己 7. 繼續奔跑 8. 不要放棄 9. 一起出發 10. 讓全世界為我們加油   … | 1. Watermark 2. Mellomaniac 3. Electra 4. We Can Fly 5. Canzonetta Sull’aria 6. Someone Like You 7. Pure Shores 8. Strawberry Swing 9. Please Don’t Go 10. Weightless |

## Analysis of Experimental Results of Lyrics Pronoun Generation Model

In order to evaluate the Lyrics Pronoun Generation Model proposed in this study, except for Rule3, the accuracy of other rules is higher than 80%. The accuracy of Rule3 also higher than 50%. Table 4.7 shows all rules of each accuracy rate and examples:

Table 4.7 Examples and Accuracy Rate of Each Rules

|  |  |  |
| --- | --- | --- |
|  | Accuracy Rate | Example of Chat bot Response |
| Rule 1 | 75/69 (0.92) | 1. 我們感動十分，就有十分滿足 2. 我們就放個屁，大快人心，人生沒什麼，好來好去 3. 我們往自己的夢想前進，自己的感動只有自己能懂 4. 我們不要讓眼淚成為生活的客 5. 我們永遠驕傲和完美，彼此永遠不妥協 |
| Rule 2 | 60/51 (0.85) | 1. 嗯免驚，嗯免驚啦，我們是勇敢的小飛俠 2. 如果大海能夠帶走我們的哀愁就像帶走每條河流 3. 打開我們的天空，帶彼此看見世界有多麼遼闊 4. 一點點我們的微笑，已經讓彼此覺得溫暖 5. 也許我們該學習相信自己的方向感 |
| Rule 3 | 25/13 (0.52) | 1. 我願意提供快樂給你，交換我們隨意的笑容 2. 我默默的在那等著你，我們要好好的活著 3. 我們任風吹乾流過的淚和汗，總有一天你有屬於你的天 4. 星夜裡寂寞的人，看著同樣天空的，我才是你的依賴 5. 我們是人生的願望，你就去看看彼此眼中的光芒 |
| ~~Rule 4~~ | ~~1/1 (1)~~ | 1. ~~我要飛翔在你每個彩色的夢中陪著你~~ |
| Rule 5 | 41/34 (0.83) | 1. 為什麼我們總是那麼堅強，像黑暗中的光芒，不停的提醒著彼此不要放棄。 2. 我們難免曾經跌倒和等候，彼此要勇敢的抬起頭 3. 一點點我們的微笑，已經讓彼此覺得溫暖 4. 我們是人生的願望，你就去看看彼此眼中的光芒 5. 一點點我們的微笑，已經讓彼此覺得溫暖 |

But in fact, some of the songs in the positive lyrics will use multiple rules, such as the original lyrics: "你願意提供快樂給我，交換我隨意的笑容", where "你願意提供快樂給我" matches Rule3, and "交換我隨意的笑容" matches Rule2, so from the perspective of the Music Recommendation System, we must look at the context of positive lyric, so our following experiment is to recommend the whole song, and invited three annotators to mark the generated lyrics (Inspirational Quote) to label the correctness of the generated lyrics, and the majority of the correctness and rationality of the generated lyrics (Inspirational Quote). Table 4.8 below shows the accuracy rate of 91 songs in 102 songs (except for 10 decompression songs with almost no lyrics and 1 lyrics of Rule4) (labeled as 1 for correct, 0 for incorrect), all of correct and incorrect examples can refer to the appendix C.

Table 4.8 Examples of Accuracy Rate of 92 Songs

|  |  |  |
| --- | --- | --- |
|  | Accuracy Rate | Examples of each labeled |
| Labeled as 1 | 91/77 (0.847) | 1. 嗯免驚，嗯免驚，我們是勇敢的小飛俠 2. 如果大海能夠帶走我們的哀愁，就像帶走每條河流 3. 我們任風吹乾流過的淚和汗，總有一天你有屬於你的天 4. 我們還是期待，明日的新景色 5. 打開我們的天空，帶彼此看見世界多麼遼闊 |
| Labeled as 0 | 91/14 (0.153) | 1. 我們長出了新的模樣，就算跌跌撞撞，你要我從現在，開始懷抱希望 2. 我們擦乾眼淚不要哭了，我要像你一樣驕傲才對 3. 讓你陪著我，我們輕撫著脆弱的心 4. 我們有關人生的道理，現在你為我講解 5. 就算傷綁住我們的腳，你陪我，彼此往未來的街道奔跑 |

The following analysis is the result of the majority of the three annotators for the Inspirational Quote. The following section shows the label result of three annotators.

### All Three Annotators Labeled as Correct

In the example labeled as correct in Table 4.9, it can be observed from 1 and 2 that the chat bot response can achieve Emotional Contagion [[23]](#_Reference_54) more than the original lyrics, and in our Generation Rule, they all conform to Rule1, and the accuracy of Rule1 is the highest accuracy rate (0.92) among all Generation Rules. It can be observed from 3 that it conforms with Rule1 and Rule3, and the original lyrics "你願意提供快樂給我，交換我隨意的笑容", it is more like the chat bot is seeking feedback from the user, but our purpose is that chat bot encourages the user, so after replacing the pronoun to "我願意提供快樂給你，交換我們隨意的笑容", it has a positive effect.

Table 4.9 Correct Examples of Inspirational Quote

|  |
| --- |
| 1. 也許我們該學習相信自己的方向感 |
| 1. 我們如今不會明瞭一切，彼此也永遠不會放棄 |
| 1. 我願意提供快樂給你，交換我們隨意的笑容 |

### All Three Annotators Labeled as Incorrect

In the example labeled as correct in Table 4.10, we can observe that three examples all have errors in Rule3 in our Generation Rule; Rule3 is the lowest accuracy rate in our Generation Rule (0.52). In the original lyrics, it is already like the chat bot responding to the user (responding Inspirational Quote), but after being processed by our Lyrics Pronoun Generation Model, it becomes that the chat bot receives a response from the user, and therefore are labeled as incorrect.

Table 4.10 Incorrect Examples of Inspirational Quote

|  |
| --- |
| 1. 送我們一份愛的禮物，你祝我幸福 |
| 1. 其實你只願輕輕把我捧在手掌上，親吻我們的傷 |
| 1. 我們擦乾眼淚不要哭了，我要像你一樣驕傲才對 |

# Conclusions

## Conclusion

In the music recommendation, in addition to the universality and installation of hope in the Therapeutic Principles of Group Psychotherapy, the user can first feel that someone understands him, and installation of hope to the user can effectively alleviate the user's worries and depression. Using the concept of Natural Language Generation, first collect the song list about installation of hope, and then construct a Lyrics Pronoun Generation Model (LPGM), replace the unsuitable pronouns in the lyrics with suitable pronouns of generation, and then use the positive lyrics as Inspirational Quote. The positive lyrics in the recommended songs can achieve Emotional Contagion and are closer to the dialogue between people.

## Future Work

In the music recommendation, because the number of song list is not large, therefore the positive lyrics and negative lyrics in the songs of the installation of hope are select manually in current. Therefore, in the future, other methods can used in extracting the positive lyrics and negative lyrics from the songs, and then fill the song list of the songs of installation of hope, and the generation of the Inspirational Quote. At present, the rule only uses the part of speech or number of pronoun in the sentence to generate sentences. In the future, we can observe more sentences to find better rules.

# Reference

1. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V)*, Amer Psychiatric Pub, 2013.
2. WebMD, “Causes of Depression”, http://www.webmd.com/depression/guide/causes-depression.
3. World Health Organization, “The Internatinal Statistical Classfication of Diseases and Related Health Problem 11th Revision(ICD-11)”, http://www.who.int/classifications/icd/en/, 2018
4. Simandan, D., 2010. On how much one can take: relocating exploitation and exclusion within the broader framework of allostatic load theory. Health & Place, 16(6), pp. 1291–1293. <http://doi.org/10.1016/j.healthplace.2010.08.009>
5. Jhen-Jia Cai, Music and Emotional Reactions, Science Development, 2015

<https://ejournal.stpi.narl.org.tw/sd/download?source=10402/10402-03.pdf&vlId=A42C71F4-7A01-4056-9149-5BCA23CF3AA4&nd=0&ds=0>

1. Stress and Depression, Can stress cause depression?

<https://www.webmd.com/depression/features/stress-depression#1>

1. AMERICAN PSYCHOLOGICAL ASSOCIATION(2016), “Hard-hitting hormones: The stress-depression link”,

<http://www.apa.org/monitor/jan05/hormones.aspx>

1. Tiong-Thye Goh and Yen-Pei Huang, “Monitoring youth depression risk in Web 2.0”, The journal of imformation and knowledge management systems,2009
2. Ti-Ku Yao, “Design and Develop an Assistant System for Depression Quantization and Estimation”, Thesis of Institute of Computer Science and Information Engineering, Souther Taiwan University of Science and Technology, 2008
3. Hsieh-Cheng Chiang, “Automatic Capture and Detect Blog Message of Suicidal Ideation”, Thesis of Institute of Information Management, Yuan Ze University, 2011
4. Yalom, I. D. (1995). The theory and practice of group psychotherapy. Basic Books (AZ).
5. Sheng-Jia Hong, “Using Stressful Life Event-Driven Model to Predict Web User’s Depression Tendency”, Thesis of Institute of Medical Informatics, National Cheng Kung University, 2012
6. Shou-Nian Wang., “Predicting logger’s Tendency of Depression Using Multiple Passive Factors Analysis Model”, Thesis of Institute od Medical Informatics, National Cheng Kung University, 2011
7. W.H. Lu, R.S. Lin, Y.C. Chan, K.H. Chen, “Overcoming terminology barrier using Web resources for cross-language medical information retrieval.” in *AMIA Annu Symp Proc*. 2006, vol. 2006 ,pp. 519-523
8. W.H. Lu, R.S. Lin, Y.C. Chan, K.H. Chen, “Semi-automatic construction of the Chinese-English MeSH using Web-based term translation method.” in *AMIA Annu Symp Proc*. 2005, vol. 2005, pp. 475-479
9. Wikipedia(2018), “Power Law”, <https://en.wikipedia.org/wiki/Power_law>
10. 衛福部心理及口腔健康司「十大勵志療癒歌曲」,<https://anntw.com/articles/20150128-uidY>
11. Jaakko Erkkila, Individual music therapy for depression: randomised controlled trial, The British Journal of Psychiatry, 2011
12. David Lewis-Hodgson, Neuroscience Says Listening to This Song Reduces Anxiety by Up to 65 Percent(2016), MindLab
13. Wikipedia(2018), “Word2Vec”, <https://en.wikipedia.org/wiki/Word2vec>
14. Wikipedia(2018), “Cosine Similarity”, <https://en.wikipedia.org/wiki/Cosine_similarity>
15. Wang Xiao, “The Advance Research of Emotional Contagion Theory”, Business School, Nankai University, 2010
16. Wikipedia(2018), “Emotional Contagion”, <https://en.wikipedia.org/wiki/Emotional_contagion>
17. Wikipedia(2018), “Natural Language Generation”,

<https://en.wikipedia.org/wiki/Natural_language_generation>

Appendix A

|  |
| --- |
| **Lexicon of Negative Events** |
| 課業的問題、錢的負擔、小孩的問題、課業的事情、課業的壓力、異樣的眼光、現實的落差、高壓的狀態、房貸的狀況、論文的壓力、爸爸的狀況、社會的現實、憂鬱症、憂鬱症的原因、父母的壓力、傷口的膿、負面的事、醫生的話、差別的態度、藥的抉擇、課業的事情、異樣的眼光、錢的問題、藥的關係、前世的情人、最愛的人、命運的權利、負面的事、迴圈的生活、負面的想法、父母的壓力、高壓的狀態、機車的老闆娘、錢的問題、被誤會、被放鳥、被狂操、被死釘、搞我、傷害我、騙我、拋棄我、推開我、跟蹤我、嘲諷我、惹到我、揶揄我、圍剿我、壓迫我、凌虐我、數落我、壓抑自己、接受我、引發我、陪伴我、離開我、放棄我、折磨我、虐待自己、惹到我、騙我、吃定我、搞我、害你、傷害我、遇到妳、反諷自己、拋棄我、殺我、祝福你、不理我、住院、休學、分手、吵架、辭職、離婚、離家、熬夜、爭吵、作弊、絕交、考試、經濟壓力、工作噩夢、愛恨情仇、混蛋客人、論文壓力、酒鬼父親、財務問題、課業因素、功課壓力、腦殘老闆、情愛糾葛、課業問題、身材問題、社會現實、鳥事、感情犧牲品、煩惱論文、爸媽面子、無法承受、無法溝通、無法挽回、對我說謊、讓我哭、對我發脾氣、對我說教、被拋棄、被罵、被打、被欺負、被開除、被整死、被傷害、被念、被騙、被遺棄、被羞辱、被嘲弄、給趕走、被排斥、被排擠、被虐待、,被吵醒、受欺負、被當掉、被背叛、被丟下、被侷限、被折騰、受摧殘、被強暴、被傷到、被操、被控管、被甩、為了配合、被抓、被苛責、被淹沒、被強姦、受控制、被撞翻、被診斷出、被侵蝕、被凌虐、受傷害、被拋棄、被打、被排斥、被控管、被罵、被丟下、被綁死、被撞翻、被開除、被念、為了等、為了追、被背叛、被排擠、被騙、被嘲弄、被診斷出、被欺負、受欺負、被抓走、被偷走、被吵醒、被束縛、被殺、受傷害、被唸、被壓抑、被整死、被恥笑、被剝奪、受折磨、被壓抑、被束縛、被踐踏、被綁死、閱讀障礙、接近人群、做錯事情、丟工作、吊點滴、辦不了貸款、放棄論文、拿掉孩子、劈腿、發病、復發、離職、遲到、發酒瘋、酗酒、失戀、心碎、心寒、無家可歸、肄業、失業、失調、自作多情沒錢、沒收入、沒安全感、沒朋友、沒人緣、沒有社交、病情、惡夢、醫藥費、完美主義、重考、閒言閒語、家暴、第三者、重新去適應、惡意重傷我、自私朋友、不懂我、失去工作能力、配不上妳、融入新環境、對我大吼、婊我、嫁不出去、熬不下去、逃避、騷擾、開除、領養、當掉、客訴我、逼我、拒絕我、逼迫我、硬逼我、被貼上標籤、被同學欺負、不要收我、恐嚇電話、放我鴿子、掛我電話、惡意倒閉、不想上班、談論我、逼問我、不被了解、無法勝任、沒人要、大吵、當凱子、當傻瓜、不要我、霸凌我、克服不了、退學、扯後腿、討厭我、囂張的店家、嚴重的完美主義、針對我、被人討厭、被室友討厭、受夠了、被陰了、失去聯繫、搞砸一切、這樣對我、把我吵醒、趕出家門、回不了家、沒辦法支付房租、失去一切、呼我巴掌、沒心思讀書、離棄我、嘮叨個不停、罹癌症、急需錢、得罪、罹癌、工作壓力、人際方面、女朋友誤會、新舊傷口、性交易、社交障礙、廢物同事、論文問題、混帳同事、肉體痛楚、機車網路、工作時間、父母關係、垃圾心情、藥物負作用、垃圾天氣、鬼打牆、藥物問題、房東困擾、貸款壓力、家庭狀況、房貸問題、人際能力、工作時間、疤痕一輩子、經濟問題、混蛋客人、家庭害蟲、藍色情緒、情愛糾葛、社會現實、玻璃娃娃、老處女巫婆、強迫性記憶、家庭壓力、方面困擾、灰色心情、房租問題、社交壓力、感情困擾、課業壓力、工作惡夢、社交恐懼症、強盜小偷、人際關係、同事壓力、時間金錢、東西壓力、醫生談話、心理醫師、小孩眼淚、苦瓜臉、感情壓力、房東壓力、內心聲音、網路網誌、愛恨情仇、醫生壓力、肝指數、情人友人、情愛糾葛、老處女巫婆、垃圾同事、廢物老闆、渾身刀傷、藥物負作用、鬼打牆、藥物問題、方面困擾、藍色心情、父母壓力、課業問題、機車老闆、金錢問題、癌症問題、財務壓力、家暴問題、親情壓力、零零碎碎、心不在焉、離職、遲到、醒來、無家可歸、發酒瘋、心碎、酗酒、怎麼辦、不過如此、自作多情、復發、換醫生、出車禍、吃藥、治療癌症、掛掉閃光、避開傷口、面對學弟妹、閱讀障礙、做錯事情、出車禍、治療癌症、掛掉閃光 |

|  |
| --- |
| **Lexicon of Negative Emotions** |
| 不能自已、幹、不禁、不想、不快、不懂、不爽、不好意思、不平、不忿、不安、不解、不詳、不服氣、不順心、不開心、不高興、不耐煩、不快樂、不完整、不值得、不踏實、不滿、故意、幼稚、冷落、不是味兒、不是滋味、不寒而慄、不共戴天之仇、不共戴天、不可一世、不知所措、瞎了狗眼、不堪、不堪入目、不值一提、不屑、不屑一顧、不堪一擊、強顏歡笑、對不起、不起眼、不得已、寧願、自作自受、心浮氣躁、放蕩、雜亂、無知、無止盡、鬧脾氣、不要煩我、表裡不一、存什麼心、搞什麼鬼、蒙在鼓裡、死人骨頭、屁事、搞死、存心、沒心、該有多好、碎碎唸、碎碎念、捨不得、掙脫不了、掙脫、屁用、病態、很爛、廢話、雖小、三小、沙小、殺小、害的、任性、真煩、很煩、煩人、很賤、超賤、犯賤、馬的、媽的、靠北、哭霸、哭爸、哭腰、靠腰、哭么、靠么、傷人、想逃、發脾氣、飆髒話、裝模作樣、緊張兮兮、討厭鬼、怒罵、責備、打擊、罵、配不上、爛人、賤人、膚淺、爛透、指責、眼睜睜、一肚子火、王八蛋、死不完、你他媽、我他媽、他媽的、他馬的、操你媽、操他媽、去你媽、幹你娘、幹恁娘、幹拎娘、幹您娘、幹林娘、乾林良、承受不了、承受不住、人格分裂、變本加厲、狠下心、狠心、受不了、受不瞭、受不暸、受不鳥、喘不過氣、莫名、假面、搞砸、拜託、瘋了、受夠、哭訴、困難、麻煩、挖操、沈重、辛苦、有病、空虛、空洞、險惡、荒謬、煩死、遲鈍、錯誤、輸給、很糟、唉、挫折、煎熬、失控、做不到、熬不住、自己受、自己扛、蠢爆了、離我遠去、好傻、很傻、抽離、難纏、阻擾、默默、一無所有、無言以對、無言、走投無路、走頭無路、無路可走、冷嘲熱諷、心煩氣躁、自以為是、幸災樂禍、遍體麟傷、悲從中來、重蹈覆轍、假惺惺、自以為、諷刺、殘忍、拿走、折磨、拖累、憤慨、憤怒、惱火、氣不過、氣不忿、氣憤、悲哀、悲傷、沉痛、傷感、傷心、痛苦、痛心、心酸、膽怯、瘋狂、膽戰心驚、發怵、害怕、驚嚇、恐怖、恐懼、受驚、心有餘悸、無所謂、怨恨、仇恨、敵視、敵意、妒忌、嫉妒、反感、可恨、可惡、厭惡、憎恨、受傷、煩悶、難受、放空、討厭、窩囊、心煩、厭煩、冷血、憂愁、擔心、擔憂、發愁、犯愁、憂慮、壓抑、鬱悶、自卑、無能感、驕傲、高傲、狂妄、自大、委屈、抱屈、冤枉、著急、浮躁、急切、急躁、焦急、·、心急、心急如焚、心切、心慌、發慌、恐慌、心慌意亂、羞愧、慚愧、丟臉、丟人、害羞、可恥、虧心、愧疚、難堪、難看、怕羞、羞恥、羞辱、懊悔、悔悟、懺悔、後悔、抱歉、過意不去、寢食難安、自私、急促、感傷、內疚、吃驚、驚訝、震驚、警惕、疑惑、懷疑、可疑、困惑、迷茫、為難、狂亂、無所適從、可憐、假裝、困住、可惜、惋惜、心疼、思念、懷念、牽掛、想念、輕蔑、藐視、蔑視、輕視、失望、悲觀、沮喪、茫然、失落、失落感、不知不覺、奢望、無望、心寒、孤獨、孤單、孤立、寂寞、氣急敗壞、無奈、心虛、煩躁、苦悶、苦惱、納悶、羞怯、憂傷、淡漠、漠然、漠視、衝動、激動、緊張、低落、低沉、消沉、心灰意懶、心灰意冷、沉重、沉甸甸、無聊、沉思、解氣、惱羞成怒、不切實際、袖手旁觀、氣餒、消氣、喪氣、掃興、洩勁、厭倦、憤激、惱怒、激憤、氣惱、盛怒、悻悻、震怒、悲苦、悲酸、悲辛、哀傷、哀戚、哀痛、悲愴、慘苦、苦澀、淒慘、傷神、酸楚、痛心疾首、辛酸、誠惶誠恐、寒心、惶惑、驚駭、驚恐、懼怕、畏懼、畏怯、心驚膽戰、心驚肉跳、抱恨、可憎、痛恨、痛惡、嫌怨、嫌惡、嫌隙、嫌憎、憎惡、熬心、懊惱、憋悶、憋氣、煩擾、糟心、愁悶、窮愁、殷憂、沉鬱、陰鬱、自慚形穢、自餒、自滿、自恃、焦心、焦躁、焦灼、情急、心焦、煩亂、祈求、紛擾、如坐針氈、抱愧、害臊、愧恨、無地自容、羞人、羞澀、悔恨、失悔、痛悔、追悔、自怨自艾、負疚、歉疚、詫異、愕然、怪訝、駭怪、驚詫、驚異、迷惑、迷惘、彷徨、疑忌、哀憐、憐憫、憐惜、痛惜、掛念、牽腸掛肚、眷眷、眷戀、渴慕、貪戀、鄙視、鄙夷、侮蔑、失意、懊喪、抱憾、悵悵、惆悵、落魄、惘然、孤寂、落寞、落莫、消魂、銷魂、哀思、哀怨、悲憤、悲鬱、悵恨、悵惘、愁苦、仇怨、憤恨、憤懣、感憤、戒懼、驚疑、敬畏、愧痛、悶倦、惱恨、惱人、危懼、畏忌、銜恨、羞憤、疑懼、疑慮、憂煩、憂憤、憂懼、憂悶、怨憤、厭棄、索然無味、黯淡、暗淡、頹廢、頹靡、頹喪、頹唐、委靡、解恨、發飆、七竅生煙、上火、大發雷霆、心頭火起、火冒三丈、令人髮指、生氣、光火、腦羞成怒、含怒、狂怒、冒火、勃然大怒、怒不可遏、怒沖沖、怒髮衝冠、面有慍色、氣沖沖、氣呼呼、動火、動肝火、動怒、動氣、悻悻然、悻然、掛火、掛氣、發怒、發狠、感忿、惹氣、慍怒、義憤填膺、嗔怒、憤然、憤憤、暴跳如雷、激怒、觸怒、怏怏、怨聲載道、彆扭、遺憾、心如刀割、生不如死、如喪考妣、肝腸寸斷、兔死狐悲、哀哀、哀毀骨立、苦痛、絕不、淒切、淒涼、淒惻、淒愴、悲切、悲慼、悲痛、悲愁、悲慟、悲慟不已、破碎、椎心、椎心泣血、痛不欲生、痛切、痛定思痛、傷心欲絕、無助、傷悲、傷痛、腸斷、慘然、慘痛、酸辛、樂極生悲、斷腸、難過、亡魂喪膽、大驚失色、毛咕、毛骨悚然、失色、失魂落魄、生怕、生恐、屁滾尿流、忌憚、怯場、杯弓蛇影、畏縮、風聲鶴唳、望而生畏、喪膽、惶恐、惶惶、提心吊膽、猶有餘悸、發毛、虛驚、聞風喪膽、魂不附體、魂飛魄散、談虎色變、戰戰兢兢、擔驚受怕、懸心弔膽、懸心吊膽、懼色、驚愕、驚惶、驚惶失措、驚慌、驚慌失措、驚魂、驚魂不定、卑怯、怯生、怯生生、怯弱、怯懦、怕生、柔弱、畏首畏尾、畏葸、畏葸不前、苟且偷生、苟且偷安、脆弱、婆婆媽媽、貪生怕死、軟弱、愚懦、懦弱、縮手縮腳、縮頭縮腳、膽小、膽小如鼠、膽小怕事、膽寒、膽虛、薄弱、可怖、可怕、駭人聽聞、嚇人、嚇死、觸目驚心、入魔、沉迷、沉湎、沉溺、神魂顛倒、瘋魔、難分難解、難解難分、小心、介意、在心、在乎、在意、注意、留心、留神、留意、傾注、經心、經意、凝神、仇隙、切骨之仇、民怨、夙嫌、幽怨、怨毒、怨氣、冤仇、恩怨、宿怨、深仇大恨、惡意、惡感、悶氣、睚眥之怨、敵愾、積怨、切齒痛恨、含恨、抱恨終身、咬牙切齒、怨尤、怨艾、怨望、恨入骨髓、恨之入骨、記恨、恚恨、積掰、機掰、雞掰、深惡痛絕、飲恨、嫉恨、閉嘴、懷恨、臭嘴、吃醋、報應、打嘴砲、拍馬屁、妒火中燒、死八婆、忌克、忌妒、倒楣、忌刻、忌恨、眼紅、嫉賢妒能、討嫌、絮煩、該死、作嘔、嫌棄、膩味、膩煩、沉悶、悶悶不樂、悶悶的、悶悶、愁眉苦臉、煩惱、窩愁、鬱悒、鬱悒寡歡、鬱鬱寡歡、人心惶惶、六神無主、日坐愁城、多愁善感、作賊心虛、杞人憂天、芒刺在背、忐忑、哀愁、耿耿、做賊心虛、庸人自擾、掛心、掛慮、牽心、惴惴不安、惶惶不安、揪心、焦慮不安、想不開、愁腸百結、憂心、憂心如焚、憂心忡忡、憂悒、憂戚、憂憤成疾、憂鬱成疾、懸心、顧慮、鬱結、鬱積、束縛、牢籠、拘謹、拘束、枷鎖、桎梏、強壓、牽制、牽掣、壓制、檢束、關礙、自抑、自製、自持、克制、忍受、忍耐、忍氣吞聲、按捺、容忍、隱忍、妄自菲薄、自愧弗如、自慚、自輕自賤、自暴自棄、大模大樣、牛氣、妄自尊大、老氣橫秋、自我陶醉、自命不凡、自命清高、自高自大、夜郎自大、孤芳自賞、孤高、孤傲、恃才傲物、矜誇、旁若無人、唯我獨尊、盛氣凌人、趾高氣揚、虛驕、傲岸、傲慢、落落寡合、頤指氣使、驕矜、驕慢、自鳴得意、沾沾自喜、得意忘形、揚揚得意、搖頭晃腦、搖頭擺尾、顧盼自雄、叫屈、含屈而終、受屈、抱屈含冤、暗暗叫屈、窩氣、心焦如焚、火燒火燎、抓耳撓腮、性急、急赤白臉、起急、干急、乾瞪眼、發急、搓手頓腳、毛躁、坐立不安、急性子、狷急、粗暴、暴躁、褊急、操之過急、操切、手忙腳亂、手足無措、失措、周章、倉皇、倉皇失措、張皇、著慌、慌神、慌張、慌亂、羞赧、羞慚、赧然、赧顏、腆然、腆顏、愧汗、愧怍、愧恧、含羞、含羞答答、忸怩、怕醜、紅潮、紅臉、面有愧色、面有慚色、面紅耳赤、羞人答答、一事無成、羞答答、嬌羞、磨不開、臉皮薄、難為情、樣衰、醜陋、醜樣、仗勢欺人、吐剛茹柔、污辱、作踐、折辱、侮辱、玷污、玷辱、虐待、凌虐、凌辱、凌轢、辱沒、欺人太甚、欺生、欺侮、欺負、欺凌、欺壓、摧殘、輕侮、踐踏、蹂躪、糟蹋、褻瀆、自悔、悔不當初、悔之已晚、悔之無及、追悔莫及、嗟悔、愧悔、噬臍莫及、翻悔、悔改、悔罪、悔過、覺悟、大吃一驚、大驚小怪、呆若木雞、奇怪、奇異、納罕、瞠目咋舌、瞠目結舌、駭異、駭然、戒心、惕厲、警戒、鑑戒、半信半疑、生疑、多心、多疑、存疑、狐疑、起疑、將信將疑、猜疑、置疑、疑心、疑神疑鬼、疑問、質疑、闕疑、無疑、囫圇吞棗、沒譜兒、摸不著頭腦、漆黑一團、誤會、誤解、空落落、空廓、空蕩蕩、蒼茫、蒼莽、蒼蒼、依稀、恍惚、茫茫、渺茫、虛無縹緲、漫漶、影影綽綽、模糊、模糊不清、糊塗、縹緲、隱約、隱隱、朦朧、刁難、作對、作難、找岔子、找事、找茬、找碴、過不去、難為、同病相憐、其情可憫、哀矜、肉痛、歎惋、念念不忘、念舊、思鄉、思慕、思親、相思、記掛、惦念、惦記、牽念、眷念、朝思暮想、渴想、軫念、感念、感懷、懷古、懷想、懷舊、懷戀、顧念、小看、小視、小瞧、白眼、歧視、看不起、看輕、無足輕重、背叛、無視、菲薄、視如糞土、讓步、睇小、嗤之以鼻、微不足道、睥睨、鄙棄、鄙薄、賤視、瞧不起、差強人意、隨便、如饑似渴、大失所望、向隅、死心、事與願違、垂頭喪氣、絕望、萬念俱灰、廢然、心如死灰、意懶心灰、厭世、聽天由命、一籌莫展、力不從心、心有餘而力不足、心餘力絀、巧婦難為無米之炊、坐以待斃、束手待斃、束手無策、沒奈何、沒門兒、沒轍、奈何、孤掌難鳴、怎奈、迫不得已、望洋興歎、存亡、懶散、沒路用、無可奈何、無計可施、無能為力、無能、傷口、愛莫能助、萬不得已、萬般無奈、遠水救不了近火、獨木不成林、獨木難支、黔驢技窮、鞭長莫及、人情淡薄、世態炎涼、冰冷、冷冰冰、冷言冷語、冷若冰霜、冷淡、冷漠、冷漠無情、冷酷、冷酷無情、神情淡漠、淡薄、無動於衷、無情、漠不關心、漠然置之、沖沖、怒氣沖沖、灰心、利用、陷害、灰溜溜、沒精打采、洩氣、氣短、消極、得過且過、敗興、被動、無精打采、萎靡、傷氣、意志消沉、意興闌珊、槁木死灰、頹然、懨懨、黯然、黯然神傷、沉濁、深沉、悶沉沉、悶聲悶氣、濃濁、迷漫、漠漠、漫天、瀰漫、靉靆、入骨、沉沉、刻骨、徹骨、烏雲密佈、陰暗、陰冷、陰沉、陰霾、汍瀾、吞聲、呼天搶地、抽泣、抽咽、泣不成聲、泫然、哀泣、哀號、淚、流淚、哭天抹淚、哭泣、哭哭啼啼、哭鼻子、哭嚎、哽咽、唏噓、涕泣、涕泗滂沱、涕淚俱下、鬼哭狼嚎、乾號、啜泣、淚水、淚如雨下、淚如泉湧、淚汪汪、淚花、淚流滿面、淚珠、淚液、眼淚、掉眼淚、啼泣、啼哭、悲泣、悲咽、悲哽、揮淚、痛哭、痛哭流涕、飲泣、嗚咽、號哭、慟哭、漣洏、漣漣、撲簌、撕心裂肺、潸然、潸潸、熱淚盈眶、嚎啕、嚎啕大哭、簌簌、聲淚俱下、灑淚、心病、鄉愁、愁思、愁腸、愁緒、憂思、隱憂、離愁、打馬虎眼、自欺欺人、哄弄、哄騙、耍花腔、胡弄、欺哄、欺詐、欺蒙、欺瞞、隱瞞、欺騙、誆騙、爾虞我詐、蒙哄、矇混、蒙蔽、蒙騙、誘騙、糊弄、鬧玄虛、瞞天過海、瞞哄、妄求、苛求、強求、覬覦、悵然、惘然若失、惝怳、惝恍、愾然、風流、涼颼颼、清冷、清涼、陰涼、公憤、民憤、私憤、幽憤、義憤、鬱憤、陌生、乏味、平淡、平鋪直敘、味同嚼蠟、枯澀、枯燥、枯燥乏味、倒胃口、索然、乾巴巴、乾癟、無味、由不得、忍不住、忍無可忍、抒情、抒發、宣洩、洩恨、洩憤、發抒、發洩、禁不住、禁不起、縱情、難忍、難耐、灰暗、昏天黑地、昏沉、昏黑、昏暗、幽幽、幽暗、暗淡無光、慘淡 |

|  |
| --- |
| **Lexicon of Symptoms** |
| 眼角酸痛、虛弱、胃口下降、慢性疲勞症候群、做噩夢、疼痛、吃藥、精神運動遲滯、體重增加、乾口症、妄想、厭食、病魔、幻聽、病痛、行動緩慢、耳鳴、食慾減少、消化不良、胃痛、生理時鐘節律障礙型失眠、胃口差、嗜睡、癱瘓、食慾下降、疲勞、體重下降、口乾、緊張型頭痛、體重減經、頸椎不適、打飽嗝、沈默少言、發炎、疲倦、口乾病、神經質、低潮期、眩暈、食慾變差、食慾不好、頸部痠痛、記憶混亂、情緒失調、食慾減低、食慾減退、頭痛、面色不華、錯覺、眼睛疲勞、疲憊、輕度憂鬱症、暴飲暴食、性情急躁、盜汗、多眠、口苦、舌紅、厭倦感、遲滯、健忘、頭昏、多夢、痠痛、酸痛、晚睡、胃口不好、睡眠障礙、頭好暈、胃口不振、腰酸痛、乏力、舌質淡、思慮過多、熱暈、胃口減退、記憶力消退、記憶力變差、口乾症、冷汗、失眠、口乾躁病、無法入睡、食慾不佳、眼睛紅、全身無力、全身乏力、勞疲、嘔吐、抽痛、失去活力、食慾減輕、拉肚子、睡得少、幻覺、癌細胞、精神委靡、中毒、口渴、睡眠大增、食慾大增、頭重感、褥瘡、憂鬱性情緒失調、憂鬱、精神疲倦、躁鬱、心神不寧、胸痛、胃酸過多、腰痠背痛、胃口減少、胃口減輕、性慾降低、疲累、呼吸不順暢、無力感、動作緩慢、欲嘔、心悸、頭暈、囊腫、緊張型頭痛、低落性情感疾患、噁心、頻尿、產後憂鬱症、身體酸痛、食慾增加、打嗝、便秘、呼吸不暢、早醒、抑鬱、精神疲乏、濕疹、妄想疾患、記憶力衰退、抑鬱症、腹瀉、胃口大增、記憶困難、神疲懶言、紅疹、行動遲緩、妄想性疾病、胸悶、體重減少、月經失調、胃口變差、活動量減燒、腹脹、腹瀉 |

|  |
| --- |
| **Lexicon of Negative Thoughts** |
| 燒炭、上吊、輕生、天堂、解脫、身亡、尋死、地獄、結束、奪走、傷害、想死、燒炭、跳海、殺人、離開、苦難、死胡同、放手、做傻事、死死、放棄、吃苦、尋短、吸毒、藥罐子、死亡、死刑、想死、絕處、自殘、上吊、多餘、殘殺、生死、棺材、脫離、往生、掛掉、死一死、遺物、消失、犧牲、死活、遺書、勒死、不想活、割腕、自殺、永別、活下來、噩夢、兇手、跳樓、累贅、往生、一了百了、放棄生命 |

Appendix B

|  |  |  |  |
| --- | --- | --- | --- |
| Singer | Song | Singer | Song |
| 鳳飛飛 | 祝你幸福 | F.I.R | 我要飛 |
| 黃士祐 | 心內有數 | 蕭煌奇 | 逆風飛翔 |
| 蘇打綠 | 十年一刻 | 蕭煌奇 | 天若光 |
| 五月天 | 傷心的人別聽慢歌 | 東方合唱團 | 奮起飛揚的心 |
| 棉花糖 | 欠一個勇敢 | 東方合唱團 | 太陽依然出現 |
| MP魔幻力量 | 讓我罩著你 | Amazarashi | 給你 |
| 任賢齊 | 再出發 | S.H.E | 你曾是少年 |
| 嚴爵 | 好的事情 | 五月天 | 第二人生 |
| 廖文強 | 那些你不敢解決的問題 | 亂彈阿翔 | 完美落地 |
| 張韶涵 | 淋雨一直走 | 莊鵑瑛 | 星之所向 |
| 張雨生 | 大海 | 吳青原 | 那就飛吧 |
| 動力火車 | 打開天空 | 林宥嘉 | 巨人的肩膀 |
| 蛋堡 | 過程 | 范瑋琪 | 在幸福的路上 |
| 盧廣仲 | 大人中 | 旺福 | 快樂的出航 |
| 張懸 | 無狀態 | 梁文音 | 哭過就好了 |
| 孫燕姿 | 相信 | 張惠妹 | 旅程 |
| 五月天 | 如煙 | JodeC | Moving |
| 五月天 | 王子麵 | 賴慈泓&TOZZ | 大城市小雞腿 |
| 輕鬆玩樂團 | 放一個屁 | 朴樹 | 平凡之路 |
| 布朗 Mr.Brown | 為自己加油 | 辛曉琪 | 我親愛的你 |
| 1976 | 方向感 | 陳奕迅 | 相信自己無限極 |
| 來吧！焙焙！ | 連繫和共鳴 | 楊培安 | 夢想從心開始 |
| 來吧！焙焙！ | 那些事情我都不在乎 | 動力火車 | 莫忘初衷 |
| 來吧！焙焙！ | 無所畏懼與寬容 | 汪峰 | 飛得更高 |
| 來吧！焙焙！ | 一起加油吧 | 周杰倫 | 夢想啟動 |
| Christina Aguilera | Fighter | 黃玠 | 下雨的晚上 |
| Christina Aguilera | Beautiful | 魏如萱 | Don’t cry Don’t cry |
| Christina Aguilera | The Voice Within | 郭頂 | 水星記 |
| Stars | Calendar Girl | 先知瑪莉 | 星夜裡的人 |
| 南方之星 | Tsunami | 伍佰 | 鋼鐵男子 |
| 孫燕姿 | 眼淚成詩 | Tizzy Bac | You'll see |
| 郁可唯 | 時間煮雨 | 黃玠 | 在一片黑暗之中 |
| 滅火器 | 長途夜車 | 舒米恩 | 旅途 |
| Mary See The Feature | Cheer (Winter) | 929 | 溫度 |
| 蕭煌奇 | 下個街角 | 李宗盛 | 希望 |
| 盧廣仲 | 一定要相信自己 | 陳昇 | 不再讓你孤單 |
| 莊鵑瑛 | 希望 | 許美靜 | 陽光總在風雨後 |
| 伍佰 | 活下去 | 盧廣仲 | 幾分之幾 |
| 張韶涵 | 隱形的翅膀 | 陳奕迅 | 讓我留在你身邊 |
| 張韶涵 | 有形的翅膀 | 黃玠瑋 | 面對明日的勇氣 |
| 舒米恩 | 不要放棄 | 鄭宜農 | 人生很難 |
| 棉花糖 | 100個太陽月亮 | Rue Du Soleil | We Can Fly |
| 棉花糖 | 向晚的迷途指南 | Mozart | Canzonetta Sull'aria |
| Tizzy Bac | 鞋貓夫人 | Adele | Someone Like You |
| 廖文強與壞神經樂團 | 一起出發 | All Saints | Pure Shores |
| 朱俐靜 | 光的定律 | Barcelona | Please Don't Go |
| 鄭智化 | 水手 | Coldplay | Strawberry Swing |
| 張心傑 | 繼續奔跑 | Enya | Watermark |
| Coldplay | Up & Up | DJ Shah | Mellomaniac |
| 周杰倫 | 蝸牛 | Airstream | Electra |
| 楊培安 | 讓全世界為我們加油 | Weightless | Weightless |

Appendix C

|  |  |  |  |
| --- | --- | --- | --- |
| **Correct Inspirational Quote** | | | |
| Song | Inspirational Quote | Song | Inspirational Quote |
| 十年一刻 | 可是換來成長，可是換來希望 | 蝸牛 | 我們任風吹乾流過的淚和汗，總有一天你有屬於你的天 |
| 傷心的人別聽慢歌 | 我們還是期待，明日的新景色 | 讓全世界為我們加油 | 讓全世界都為我們加油，彼此燃燒夢想絕不退縮 |
| 欠一個勇敢 | 雨下過後的晴天，光線做成了橋樑 | 我要飛 | 我們要用力飛，不管有多遠，彼此超越了極限，挑戰的冒險 |
| 再出發 | 嗯免驚，嗯免驚，我們是勇敢的小飛俠 | 逆風飛翔 | 總有人在我們身旁，為我們加油，彼此逆著風也要飛翔，很辛苦也要堅強 |
| 好的事情 | 我們感動十分，就有十分滿足 | 天若光 | 我們相信黑暗的背後總是有日頭 |
| 那些你不敢解決的問題 | 我們能逃避也該樂觀 | 奮起飛揚的心 | 把回憶拋向天際，我們勇敢地乘風而去 |
| 淋雨一直走 | 人都應該有夢，有夢就別怕痛 | 太陽依然出現 | 明天太陽依然出現在東方，心中的愛有個名字叫做希望 |
| 大海 | 如果大海能夠帶走我們的哀愁，就像帶走每條河流 | 給你 | 我們快點擦乾淚吧，對彼此的過去，一笑置之吧 |
| 打開天空 | 打開我們的天空，帶彼此看見世界多麼遼闊 | 你曾是少年 | 許多年前，我們有一雙清澈的雙眼，想看遍這世界，彼此去最遙遠的遠方 |
| 過程 | 我們忘不了的光，我們忘不了的安心，彼此那些都是過程 | 第二人生 | 我們期待一種永恆，如果命運註定彼此的誕生，如果世界末日始終沒有發生 |
| 無狀態 | 我們不要讓眼淚成為生活的客串 | 完美落地 | 我們會用盡所有力，我們奮力的躍起在天際，彼此迎著光明 |
| 相信 | 一點點我們的微笑，已經讓彼此覺得溫暖 | 那就飛吧 | 那就飛吧，飛吧，尋找我們的生活 |
| 如煙 | 我們永遠驕傲和完美，彼此永遠不妥協 | 巨人的肩膀 | 我們的方向不是幻想，那是回憶裡最感動的眺望 |
| 放一個屁 | 我們就放個屁，大快人心，人生沒什麼，好來好去 | 在幸福的路上 | 我們理直氣壯，是彼此支撐讓生命茁壯，我們快樂地走，有彼此的路就沒盡頭 |
| 為自己加油 | 我們往自己的夢前進，自己的感動只有自己能懂 | 快樂的出航 | 有好多地方在等著我們啊，等著彼此一起懶懶洋洋 |
| 方向感 | 也許我們該學習相信自己的方向感 | 哭過就好了 | 我們哭過就好了，傷都會好的 |
| 連繫和共鳴 | 你擁有我的心，隨時都能感受到我和你在一起 | 旅程 | 我們勇敢總是在左右，再困難的夢陪彼此一起做 |
| 那些事情我都不在乎 | 我們不說話也很舒服，和我在一起你很幸福 | Moving | 讓我們背對黑暗面向著陽光，丟棄那無助，彼此拋棄那緊張 |
| 無所畏懼與寬容 | 我願意提供快樂給你，交換我們隨意的笑容 | 大城市小雞腿 | 我們圓一個夢，我們選擇一個人的方向，彼此就得不怕寂寞 |
| 一起加油吧 | 仍然想要為了我們唱一首歌，彼此仍然想要在冬天裡擁抱 | 平凡之路 | 我們要走嗎，我們易碎的驕傲著，那也曾是彼此的模樣 |
| Fighter | 我們如今不會明瞭一切，彼此也永遠不會放棄 | 我親愛的你 | 時間是誠實的，回過頭才曉得，當我們真的懂了，感謝那些挫折 |
| Beautiful | 言語上的傷害，無法將我們打倒 | 相信自己無限極 | 我們要相信自己無限極，彼此相信做好自己，你相信你就是奇蹟 |
| The Voice Within | 很快的我們將看見更加燦爛的日子 | 夢想從心開始 | 讓心中沸騰的希望，帶著我們勇闖，所有堅固的城牆 |
| Calendar Girl | 我默默的在那等著你，我們要好好的活著 | 莫忘初衷 | 我們別忘了那一年，那一天出發時心中的夢 |
| Tsunami | 其實我們是個越挫越勇的人，彼此超越外表給人的印象 | 飛得更高 | 我們要飛得更高，彼此飛得更高，翅膀捲起風暴，心生呼嘯 |
| 眼淚成詩 | 誰把我們變美，我們的眼淚寫成了一首詩，無所謂 | 夢想啟動 | 我們微笑吧，就算不斷失敗，我們站起來再重來，彼此把脆弱推開 |
| 時間煮雨 | 我們手拉手也成舟，劃過悲傷河流 | 下雨的晚上 | 我們期待著明天的陽光，曬乾悲傷，彼此溫暖而奔放 |
| 長途夜車 | 我等你成功 | Don’t cry Don’t cry | 大雨裡的烏雲啊請帶我們離開，彼此就不用害怕靠不到岸的大海 |
| Cheer (Winter) | 想緊握溫暖的太陽，在我們的手中任憑發燙，也掩蓋不了彼此絢麗的光芒 | 水星記 | 當你還可以再跟我飛行，至少可以陪著我們 |
| 一定要相信自己 | 我們睜開了眼，明天會美麗 | 星夜裡的人 | 星夜裡寂寞的人，看著同樣天空的，我才是你的依賴 |
| 活下去 | 就算不可以，就算不願意，就算為了我們，也要活下去 | 鋼鐵男子 | 像個鋼鐵般的男子，我們像做堅強的山，彼此能抵擋風和雨 |
| 隱形的翅膀 | 我們知道，我們一直有雙隱形翅膀，帶我們飛，給彼此希望。 | You'll see | 一個人漸成熟，我們就會笑著淚流，我們總有些遺憾要學會放開，彼此活到這把年紀也該明白 |
| 不要放棄 | 我們遺憾可以釋懷，生命不會重來，彼此不要放棄自己 | 旅途 | 喜歡看著我們倚靠著窗，我們喜歡微笑帶來的力量，彼此喜歡一望無際的那海洋 |
| 100個太陽月亮 | 我們唱一首勇敢的歌吧，彼此做一場轟烈的夢吧 | 溫度 | 為什麼我們總是那麼堅強，像黑暗中的光芒，不停的提醒著彼此不要放棄 |
| 向晚的迷途指南 | 我們突然有天向晚，想起了夢想，才領悟了快樂是指南 | 希望 | 我們是人生的願望，你就去看看彼此眼中的光芒 |
| 光的定律 | 我們再微弱的力量，我們也能折射出光，彼此不怕阻擋 | 不再讓你孤單 | 路遙遠我們一起走，我要飛翔在你每個彩色的夢中陪著你 |
| 水手 | 風雨中，這點痛算什麼，我們擦乾眼淚，我們不要怕，至少彼此還有夢 | 陽光總在風雨後 | 我們難免曾經跌倒和等候，彼此要勇敢的抬頭 |
| 繼續奔跑 | 一個肩膀，讓我們把所有重量都輕輕放下 | 幾分之幾 | 那一天我走進了你的生命，謝謝我成為了你的幾分之幾，閉上眼睛也能看見彼此 |
| Up&Up | 我們現在就要振作起來，無論如何此刻就要打起精神 | 面對明日的勇氣 | 我給你勇氣，我給你信心，我給你光明 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Incorrect Inspirational Quote** | | | |
| Song | Inspirational Quote | Song | Inspirational Quote |
| 祝你幸福 | 送我們一份愛的禮物，你祝我幸福 | 有形的翅膀 | 我們隱形翅膀，帶著我們幻想，掠過那絕望，彼此找希望 |
| 心內有數 | 讓你繼續陪伴我，好不好 | 鞋貓夫人 | 我們擦乾眼淚不要哭了，我要像你一樣驕傲才對 |
| 讓我罩著你 | 你會為了我更努力，我們也要努力愛自己 | 一起出發 | 你都覺得我們很好，只是我們都不知道，我們要給人微笑 |
| 大人中 | 流眼淚的星星正在看著我們，我們說加油，讓你為我感到光榮 | 星之所向 | 讓你陪著我，我們輕撫著脆弱的心 |
| 王子麵 | 我們有關人生的道理，現在你為我講解 | 在一片黑暗之中 | 我們在一片黑暗之中，讓你牽著我慢慢走 |
| 下個街角 | 就算傷綁住我們的腳，你陪我，彼此往未來的街道奔跑 | 讓我留在你身邊 | 生命中所有的路口，我們絕不是盡頭，我們別怕，讓你留在我身邊，都陪彼此度過 |
| 希望 | 我們長出了新的模樣，就算跌跌撞撞，你要我從現在，開始懷抱希望 | 人生很難 | 其實你只願輕輕把我捧在手掌上，親吻我們的傷 |