CS4242: PROGRAMMING LAB 3

Viral Item Prediction in Social Networks

Due date: 26 Mar 2018 (Mon, 1800 Hrs)
Student: Wang Jia (A0176605B / E0232209)

Executive Summary

A multi-modal integrated regressor is developed to predict the popularity of micro-videos based on image features, text features and social features. Four main progresses have been achieved from this project:

1. Popularity of the micro-video has been compared in terms of loops as well as integrated popularity index.
2. More social features and textual features are added into the program. Different combination among features have been demonstrated to verify the effectiveness of model integration and feature importance. Social features are observed to be vital in predicting viral micro-videos, with a nMSE of 0.7953.
3. The early fused modality has tried linear model, SVR and optimized with grid search. The best model gives nMSE of 0.9986, which is the same as baseline model.
4. The late fused model shows best performance with minimal nMSE in terms of predicted popularity index, comparing with early fusion model. nMSE of 0.8442 has been achieved from the multi-modal integrated program.

1. Introduction

With the revolution of social media posting and sharing, micro-videos recording people's daily life or special scenarios saw a pronounced increase in popularity. Vlog, as a shorten form of video blog or video log, often combines embedded video with supporting text, images and other metadata like user profile. The viral vlogs taking advantage of the short playing time and easily reposting, shed a light for new channels of advertising and propagating.
This project works on predicting the popularity of a micro-video based on multiple modalities including image, textual, social features.

2. Methodology

2.1 Dataset:

Dataset consists of 10000 micro-videos' information. It contains micro-videos' features and popularity index like loops and likes as ground truth.

2.2 Program Structure:

The baseline model took use of image feature and two attributes of social feature. The enhanced program further reduced the image features, added new features, and then adopted both early and late fusion methodology with optimized SVR model tuning by grid search. The program's output will only demonstrate the optimal results.

a. Image features
b. Text sentiment features
c. Social features

Each individual features' processing will be explained in the section below.

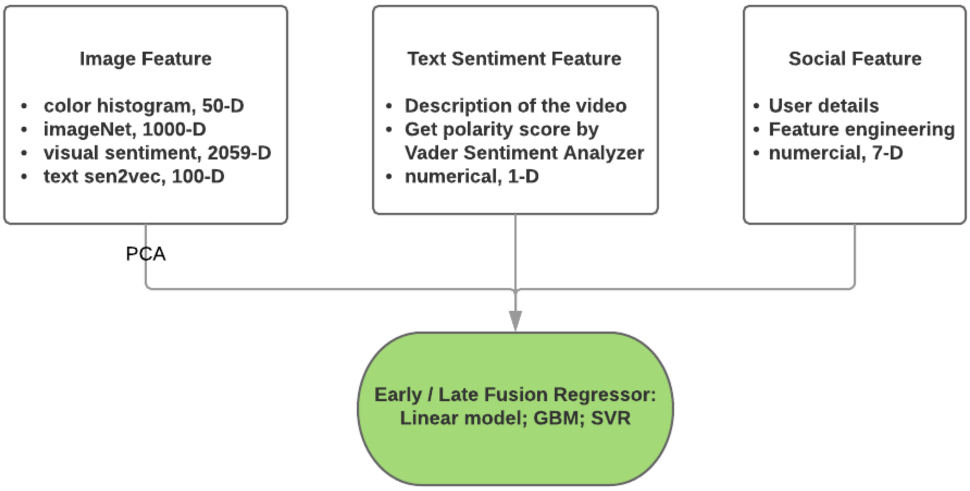

Figure 1 Program Structure

a. <u>Image Features</u>

Image features have been extracted and passed to the baseline model. Standing for the contents of the videos, they are supposed to be highly correlated to the popularity.

Table 1. Image features descriptions

Sub-feature	Description	Shape (no. samples, dimension)
Color histogram	Basic visual features i.e., intensity and mean of color channels	(10000, 50)
Object feature	Generated by DCNN to recognize object in the image	(10000, 1000)
Visual sentiment	Generated by DCNN to analyze sentiment of the image	(10000, 2059)
Text sentence2vec	Vector objects transferred from text sentences	(10000, 100)

All the 4 attributes are high-dimensional variables. The dimensions have been reduced with Principle Component Analysis (PCA) and then concatenated by column to consist a feature set.

b. Text sentiment features

Description of the videos consisted by words and emoji. A lexicon-based sentiment analyzer (from Lab2) is adopted to generate sentiment score for each description. Both English words and emojis are handled in this analyzer. The result is the polarity score ranged from -1 to 1 to denote negative to positive and stored in video_polarity.txt.

Table 2. Textual sentiment feature descriptions

Sub-feature	Description	Shape (no. samples, dimension)
Sentiment Score	Degree of positive/negative sentiment for each description. Ranged from –1 to 1, the larger the more positive it is.	(10000, 1)

c. Social features

User social feature is also an important aspect in predicting viral. Most viral is created by influencers on social media platforms. The character of an influencer could be higher loop count and like count, far more followers than following count. Two more sub-features are calculated for feature engineering.

Table 3. Social features descriptions

Sub-feature	Description	Shape (no. samples, dimension)
Total loop count	Social feature from the user profile page of the social media app	(10000,7)
Follower count		
Following count		
Like count		
Post count		
Total loop/Follower count	Estimate the influence of the user. A viral micro-video may be reposted many times and the loops count	
Follower/Following	Estimate the attractiveness of the user and the quality of the users' posts. Influencers have far more followers than following counts.	

d. Popularity index

Besides number of loops, average value of the 4 attributes in the ground truth file is created as another evaluation for popularity.

Table 4. Popularity Index

Popularity index	Formula
Loops	n_{loops}
Avg.	$\frac{n_{loops} + n_{likes} + n_{post} + n_{comments}}{4}$

e. Early fusion

The above features are simply concatenated to a comprehensive feature set with a shape of (10000, 88). Several machine learning regression models including Ridge, Lasso, Gradient Boosting Machine, Support Vector Regressor are explored to get lower prediction error. The model is optimized via 10-fold cross validation and reported with normalized MSE (nMSE).

f. Late fusion

A regression model was developed to each feature set and returned the predictions of popularity index. Then the three predictions are concatenated to develop ensemble models as a late fusion scheme.
The model is optimized via 10-fold cross validation and reported with normalized MSE (nMSE). Predicted value is returned for late fusion.

3. Results and Discussion

3.1 Results of individual models:

a. Image Regressor

Dimension of four sub-feature are reduced with PCA:

Table 5. PCA parameter

	Color_hist	imgNet	vSenti	sen2vec
PCA(n_components)	30	20	20	10

Feature set shape: (10000, 80).

Table 6. Results of Image regressors

	nMSE /popularity = loops
Linear Model (Ridge)	1.1422
randomForest Regressor	5.0368
Gradient Boosting Regressor	55.6458
Untuned SVR	0.9985
Tuning SVR	**0.9892**

	nMSE /popularity = loops	nMSE /popularity = avg
Tuning SVR	**0.9892**	0.9894

Findings:

- Penalty term (C in parameter) is large, denotes the highly non-linear relationship of the image features. Tuning SVR with kernel RBF doesn't have significant improvement. Perhaps more flexible functions can be explored to create a better model for image feature sets.
- Performance of tree-based regressor like RF and GBM is random and is averaged to be bad.

b. Textual Sentiment Regressor:

Table 7. Results of Textual sentiment regressors

	nMSE /popularity = loops	nMSE /popularity = avg
Tuning SVR	0.9984	0.9983

Textual sentiment regressor returned the similar nMSE as image regressor. It is possible that sentiment information from the video description is already carried from the image sentiment feature and text sen2vec feature.

c. Social Regressor:

Table 8. Results of Social regressors

	nMSE /popularity = loops	nMSE /popularity = avg
Tuning SVR	0.7953	0.7953

Findings:

- Social regressor performs best in predicting popularity. It makes sense that influencers on social media platform play an important role in propagating. Intuitively, they have more followers, which enable their posts to be viewed and reposted by more people.
- Different popularity index has almost no effect towards the results. Loops and average of loops, likes, comments, reposts both can stand for popularity index. Meanwhile, from the distribution of the 4 attributes, it can be found that more than 80% videos have 0-4 comments or reposts, roughly 50% videos have 0-4 likes, the popularity index may be dominant by loops. Hence, it can be inferred that single video's loop count is related to the total loop count of the user.

3.2 Fused Model

Table 9. Comparison of fused model

Fused Feature	nMSE /popularity = loops
Baseline	0.9986
Early Fusion	0.9986
Late Fusion (LM)	0.8574
Late Fusion (SVR)	**0.8442**

The late fusion model of SVR gives the smallest nMSE. The early fusion model with additional social and textual sentiment features has no improvement on the baseline model. After PCA, image features set still has 80 dimensions while additional features only have 8 dimensions. The effect of new features is eliminated in the early fusion.

Note: In the program, the result demonstrate linear regression results instead of SVR for faster computing as the difference between these two is rather close.

4. Conclusion

- **The power of KOL (Key Opinion Leader)**

As social regressor performs best among all the models. The business value of big influencers can be verified. To create a successful business campaign via micro-video on social media channel, the company is supposed to cooperate with KOLs on social media channels.

- **Loop count as a key factor estimating viral for micro-video**

Given the 4 attributes of 'loops, likes, reposts, comments' in the ground truth file, count of loops is outstanding to represent popularity index. The results of predicting aggregated value (i.e., average) don't differ much from predicting count of loops.

- **Complex function for image feature processing can be explored for improvement**

Late fusion is a better way to combine multi-viewed modalities of micro-videos information. The key of challenge is to explore a good method to reduce and utilize the high dimensional image feature set. More complex defined functions and greater computation power to tune the parameter will improve the performance further.