The Final Project

Yanwei Fu

June 18, 2021

Abstract

(1) This is the final project of our course. The deadline is 5:00 pm, June 30, 2021. Please upload the report via elearning.

(2) The paper must be in NIPS format (downloadable from [1]). We do not need the double-blind review.

(3) The goal of your write-up is to document the experiments you've done and your main findings. So be sure to explain the results. Generate a single pdf file of your projects and turned in along with your code. Package your code and a copy of the write-up pdf document into a zip or tar.gz file and named as Final-Project-student-id1-student-id2.[zip|tar.gz]

(4) You are open to using anything to help you finish this task.

(5) About the deadline and penalty. In general, you should submit the paper according to the deadline of each mini-project. The late submission is also acceptable; however, you will be penalized 10% of scores for each week's delay. **The submission can only be delayed for up to three days.**

(6) For all the projects, we DO care about the performance on each dataset with the correct evaluation settings.

Note that:

(a) If the size of training instances is too large, you may want to apply some sampling techniques to extract a small portion of training data.

(b) If you think the dimension of features is too high, you may also want to use some techniques to do feature dimension reduction.

(c) The referring papers are listed as an introduction to the context of problems. It's not necessarily to exactly implement these papers, which is not an easy task.

1 Introduction

1.1 Collaboration Policy

You are allowed to work in a group with at most one collaborator. You will be graded on the creativity of your solutions, and the clarity with which you can explain them. If your solution does not live up to your expectations, then you should explain why and provide some ideas on how to improve it. You are free to use any third-party ideas or codes that you wish as long as it is publicly available. You must provide references to any work that is not your own in the write-up.

[1] `https://nips.cc/Conferences/2016/PaperInformation/StyleFiles`

1.2 Writing Policy

The final project (20%) is finished by one team. Each team should have up to 2 students. You will solve a real-world Big-Data problem. The final report should be written in English. The main components of the report will cover

1. Introduction to the background and potential applications (2%);
2. Review of the state-of-the-arts (3%);
3. Algorithms and critical codes in a nutshell (10%);
4. Experimental analysis and discussion of proposed methodology (5%).

Please refer to our latex example[2].

1.3 Submitting Policy

The paper must be in NIPS format. Package your code and a copy of the write-up pdf document into a zip or tar.gz file called Final-Project-student-id1-student-id2.[zip|tar.gz]. Include functions and scripts that you had used. Upload them to eLearning. In the submitted documents, you should include a readme.txt file to well explain the authors and co-workers of this project. The TAs can know the names of your works.

1.4 Evaluation of Final Projects

We will review your work on the following NIPS criteria:

Overview: You should briefly summarize the main content of this paper, as well as the Pros and Cons (advantages and disadvantages) in general. This part aims at showing that you had read and at least understand this paper.

Quality: Is the paper technically sound? Are claims well-supported by theoretical analysis or experimental results? Is this a complete piece of work, or merely a position paper? Are the authors careful (and honest) about evaluating both the strengths and weaknesses of the work?

Clarity: Is the paper clearly written? Is it well-organized? (If not, feel free to make suggestions to improve the manuscript.) Does it adequately inform the reader? (A superbly written paper provides enough information for the expert reader to reproduce its results.)

Originality: Are the problems or approaches new? Is this a novel combination of familiar techniques? Is it clear how this work differs from previous contributions? Is related work adequately referenced?

Significance: Are the results important? Are other people (practitioners or researchers) likely to use these ideas or build on them? Does the paper address a difficult problem in a better way than previous research? Does it advance the state of the art in a demonstrable way? Does it provide unique data, unique conclusions on existing data, or a unique theoretical or pragmatic approach?

[2] `http://yanweifu.github.io/courses/SLML/chap5/IEEE_TAC_2016.zip`

1.4.1 Minimum Requirements

For all the projects listed below, in general you should devise your own deep learning models which target each specific problem of each project. You should compare with the machine learning algorithms taught in this course/projects. Thus, the minimum requirements, as you can imagine, just apply and compare with these methods; and explain the advantages and disadvantages of using these methods for the project problem. Note that, your algorithms can be derived from one of these existing deep learning algorithms; and feel free to use any machine learning packages you like.

2 Potential Projects

2.1 Image Captioning

Given an image, the task of image captioning is trying to give a sentence to describe this image. In this task, we use the Flickr30K as the testbed for image captioning tasks.

You can find the Flickr30k dataset here[3]: there are two files and two folders in it. The folder "images" contains the original images of Flickr30k. There are four files in it. The compressed file "flickr30k_images.tar.gz" is the original images of Flickr30k. The file 'cap_flickr30k.json' contains the captions of images. Its order is corresponding to the key of 'images' of the file 'dic_flickr30k.json'. And the key of 'images' includes the training/testing split, id, and file name of each image. The original caption ground truth file is 'flickr30k_cleaned_class.json'.

2.1.1 Questions and Evaluation

As for this task, we ask several questions to inspire your work:

1. Can you exploit a model that can generate a sentence given an image? The generated sentence must be related to the given image.
2. Can you add Attention mechanism in your captioning model?

Minimum Requirements:

1. The captioning generation model is a requirement in Question-1.
2. The Attention mechanism is an extra point column.

Notes and Evaluation:

1. You can also extract any image features if you want.
2. You can evaluate the quality of generated sentenced by using the online script[4]. Note that they are the only objective metrics that we trust. We mainly test your sentence by the CIDEr metric.
3. We know there are lots of open-source codes and models for image captioning, and we know most of them. So we expect your model; and NO CHEATING here. If your model is modified from another model, please discuss the similarity and difference between your model and the referring model. The "*Originality*" is one evaluation metric.

[3] `http://www.sdspeople.fudan.edu.cn/fuyanwei/course/projects/final_project/Flicrk30k/`
[4] `https://github.com/tylin/coco-caption`

2.2 Learning to Score Figure Skating Sport Videos

We have one project of learning to predict the scores of figure skating sport videos. Please read our paper: Learning to Score Figure Skating Sport Videos. IEEE TCSVT 2020[5]. We have the Fis-V dataset which is download-able from here [6]. We upload the features of training and testing videos on the website. Please read our paper on how to evaluate the performance of the model.

For video-level dataset, you can download with this link [7] .

Note that, it would be VERY INTERESTING if you make significantly better results than those reported in our paper.

2.3 Identification and Synthesizing of Raphael's paintings from the forgeries

The following data is provided by Prof. Yang WANG from HKUST[8]. Since this link is in google drive, you can download the file from our course webpage[9]. The data contains 28 digital paintings of Raphael or forgeries. Note that there are both jpeg and tiff files, so be careful with the bit depth in digitization. The following file[10] contains the labels of such paintings.

2.3.1 Questions

1. Can you exploit the known Raphael v.s. Not Raphael data to predict the identity of those 6 disputed paintings (maybe Raphael)? The following papers[11] [12] might be some references for you.

2. We need to synthesize the painting of Raphael. That is, given one photo, we need to generate/synthesize a new photo that makes it look like Raphael. There are lots of recent works on this topic. You can try it by either (1) Generative Adversarial Networks [2], or Convolutional Neural Networks [1], or Attribute transfer [3]. The testing images are downloaded from [13]. You can use the images in [14] as well as other available images online to train your model. Your report should show the effects of synthesized testing images.

2.3.2 Minimum requirements

The minimum requirements include:

1. classification tasks in Question-1;

2. synthesizing tasks in Question-2.

[5] https://ieeexplore.ieee.org/document/8756030

[6] http://www.sdspeople.fudan.edu.cn/fuyanwei/course/projects/final_project/figure_skating or https://drive.google.com/file/d/1FQ0-H3gkdlcoNiCe8RtAoZ3n7H1psVCI/view

[7] https://pan.baidu.com/s/164eVQpg11KsBdUF5sR9czA(access code:nrd4)

[8] https://drive.google.com/folderview?id=0B-yDtwSjhaSCZ2FqN3AxQ3NJNTA&usp=sharing

[9] http://www.sdspeople.fudan.edu.cn/fuyanwei/course/projects/final_project/Raphael.zip

[10] https://docs.google.com/document/d/1tMaaSIrYwNFZZ2cEJdx1DfFscIfERd5Dp2U7K1ekjTI/edit

[11] http://www.sdspeople.fudan.edu.cn/fuyanwei/course/projects/final_project/artistic_poster.pdf

[12] http://dx.doi.org/10.1016/j.acha.2015.11.005

[13] http://www.sdspeople.fudan.edu.cn/fuyanwei/course/projects/final_project/test_images.zip

[14] http://www.sdspeople.fudan.edu.cn/fuyanwei/course/projects/final_project/Raphael.zip

2.4 Stock Prediction

In this task, you are provided 5-minute stock data from 2014 to 2020[15]. The overall goal is to predict the future stock market with a neural network. You should train your model based on the data from 2014 to 2019, and evaluate it in the year 2020. You can choose one of the following tasks or study any task you are interested in: (1) Predict the stock price on the eleventh day based on the data of the previous ten days; (2) Find strategies for maximizing rewards for an individual stock within 30 days. You should clearly define the task and show the importance of this task. The minimum requirements follow Sec. 1.4.1.

2.5 Other Projects

You can also try other projects. However, please let TA and me know first to get the approval. Note that if the project is too easy, it would affect your scores of final projects.

References

[1] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *CVPR*. 2017.

[2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial nets. In *CVPR*. 2017.

[3] Jing Liao, Yuan Yao, Lu Yuan, Gang Hua, and Sing Bing Kang. Visual attribute transfer through deep image analogy. In *ACM Siggraph*. 2017.

[15] `http://www.sdspeople.fudan.edu.cn/fuyanwei/course/projects/final_project/stock_data.zip`