Assignment 1B

CAB420, Machine Learning, Semester 1, 2020

This document sets out the three (3) questions you are to complete for CAB420 Assignment 1B. The assignment is worth 20% of the overall subject grade. All questions are weighted equally. Students are two work either individually, or in groups of two. Students should submit their answers in a single document (either a PDF or word document), and upload this to TurnItIn.

Further Instructions:

1. Data required for this assessment is available on blackboard alongside this document in *CAB420_Assessment_1B_Data.zip*. Please refer to individual questions regarding which data to use for which question.

2. Answers should be submitted via the TurnItIn submission system, linked to on Blackboard. In the event that TurnItIn is down, or you are unable to submit via TurnItIn, please email your responses to `cab420query@qut.edu.au`.

3. For each question, a short written response (approximately 3-5 pages depending on the nature of the question, approach taken, and number of figures included) is expected. This response should explain and justify the approach taken to address the question (including, if relevant, why the approach was selected over other possible methods), and include results, relevant figures, and analysis.

4. MATLAB or Python code, including live scripts or notebooks (or equivalent materials for other languages) should be included as appendices. Figures and outputs/results that are critical to question answers should be included in the main question response, and not appear only in an appendix. Note hat MATLAB Live Scipts, Python Notebooks, or similar materials will not on their own constitute a valid submission and a written response per question is expected as noted above.

5. Students who require an extension should lodge their extension application with HiQ (see `http://external-apps.qut.edu.au/studentservices/concession/`). Please note that teaching staff (including the unit coordinator) cannot grant extensions.

Problem 1. Person Re-Identification. Person re-identification is the task of matching a detected person to a gallery of previously seen people, and determining their identity. In formulation, the problem is very similar to a typical biometrics task, however large changes in subject pose and their position relative to the camera, lighting, and occlusions make this a challenging task.

Person re-identification is typically evaluated using Top-N accuracy and Cumulative Match Characteristic (CMC) curves. Top-N accuracy refers to the percentage of queries where the correct match is within the top N results. A CMC curve plots the top-N accuracy for all possible values of N (from 1 to the size of the dataset).

You have been provided with the VIPeR [1] dataset (see `Q1/VIPeR.tar.gz`, a widely used dataset for person re-identification. This data has been split into two segments:

- Train: consists of the first 482 identities. Each identity is visible in both cameras, and there is a single image from each camera for each subject.
- Test: consists of the remaining 150 identities. Each identity is visible in both cameras, and there is a single image from each camera for each subject.

Your Task: Using this data, you are to:

1. Develop and evaluate a **non-deep learning** method for person re-identification. The method should be evaluated on the test set by considering Top-1, Top-5 and Top-10 performance. A CMC (cumulative match characteristic) curve should also be provided.
2. Develop and evaluate a **deep learning based** method for person re-identification. The method should be evaluated on the test set by considering Top-1, Top-5 and Top-10 performance. A CMC (cumulative match characteristic) curve should also be provided.
3. Compare the performance of the two methods. Are there instances where the non-deep learning method works better? Comment on the respective strengths and weaknesses of the two approaches.

In completing your answer you may also wish to consider the following:

- You may wish to resize images to reduce computational burden. This is acceptable, but should be documented.
- You may wish to fine-tune a pre-trained network. This is acceptable, but should be documented.
- A high level of accuracy alone will not guarantee a high mark for the question. The approach you choose and the rationale for it, and the quality of your evaluation and discussion are far more important.

Problem 2. Clustering and Recommendations. Recommendation engines are typically built around clustering, i.e. finding a group of people similar to a person of interest and making recommendations for the target person based on the response of other subjects within the identified cluster.

You have been provided with a copy of the MovieLens small dataset[1], which contains movie review data for 600 subjects. The data is contained in the `Q2` directory within the data archive, and is split over several files as follows:

- `ratings.csv`: Contains the movie ratings, and consists of a user ID, a movie ID, a rating (out of 5), and a timestamp.
- `movies.csv`: A list of all movie ID's, alongside the movie titles and a list of genres.
- `tags.csv`: A list of tags applied to movies by users. Each entry consits of a user ID, a movie ID, the text tag, and a timestamp.
- `links.csv`: Contains IDs to link the MovieLens dataset to IMDB and TMBD.

It is recommended that you do not use the `tags.csv` and `links.csv` file, though they are contained here for completeness and you may choose to use them if you wish.

Your Task: Using this data, develop a method to cluster users based on their movie viewing preferences. Having developed this, provide recommendations for the users with the IDs 42, 314, and 444. Your answer should include:

1. A discussion of how you process and prepare the data, and what data you cluster.
2. A description of and justification for your clustering method. This should include why you select the clustering method you do, and why you select the parameters (i.e. number of clusters) that you do.
3. A brief discussion on the results of the clustering, including interpretation of the resultant clusters.
4. Recommendations for the three users with IDs: 42, 314 and 444; and a short discussion of these recommendations, including if the recommendations make sense.

[1] `https://grouplens.org/datasets/movielens/`

Problem 3. Semantic Person Search. Semantic person search is the task of matching a person to a semantic query. For example, given the query '1.8m tall man wearing jeans a red shirt', a semantic person search method should return images that feature people matching that description. As such, a semantic search process needs to consider multiple traits. You have been provided with a dataset (see `Q3/Q3.tar.gz`) that contains the following semantic annotations:

- Gender: -1 (unknown), 0 (male), 1 (female)
- Pose: -1 (unknown), 0 (front), 1 (back), 2 (45 degrees), 3 (90 degrees)
- Torso Clothing Type: -1 (unknown), 0 (long), 1 (short)
- Torso Clothing Colour: -1 (unknown), 0 (black), 1 (blue), 2 (brown), 3 (green), 4 (grey), 5 (orange), 6 (pink), 7 (purple), 8 (red), 9 (white), 10 (yellow)
- Torso Clothing Texture: -1 (unknown) , 0 (irregular), 1 (plaid), 2 (diagonal plaid), 3 (plain), 4 (spots), 5 (diagonal stripes), 6 (horizontal stripes), 7 (vertical stripes)
- Leg Clothing Type: -1 (unknown), 0 (long), 1 (short)
- Leg Clothing Colour: -1 (unknown), 0 (black), 1 (brown), 2 (blue), 3 (green), 4 (grey), 5 (orange), 6 (pink), 7 (purple), 8 (red), 9 (white), 10 (yellow)
- Leg Clothing Texture: -1 (unknown) , 0 (irregular), 1 (plaid), 2 (diagonal plaid), 3 (plain), 4 (spots), 5 (diagonal stripes), 6 (horizontal stripes), 7 (vertical stripes)
- Luggage: -1 (unknown), 0 (yes), 1 (no)

The *unknown* class can be considered either a class in it's own right (i.e. three classes of gender), or can be considered as missing data. Note that three colours are annotated for each of the torso and leg clothing colour, indicating the primary, secondary and tertiary colours. One or both of the secondary and tertiary colours may be set to unknown (-1) to indicate that there are only 1 or 2 colours in the garment.

In addition, the dataset contains semantic segmentation for each image in the training data, that breaks the image down into the following regions:

- Leg clothing
- Shoes
- Torso clothing
- Luggage
- Leg skin regions
- Torso/arm skin regions
- Facial skin regions

- Hair

Semantic segmentation information is supplied both as a single colour coded mask image, and as an individual mask for each component.

Your Task: Using this data you are to implement one or more classifiers that, given an input image, classify the traits:

- Gender
- Torso Clothing Type
- Primary Torso Clothing Colour
- Torso Clothing Texture
- Leg Clothing Type
- Primary Leg Clothing Colour
- Leg Clothing Texture, and
- Luggage.

Pose and the semantic segmentation data may optionally be used when developing your approach (though remember that semantic segmentation data is only available for the training set). Secondary and tertiary torso and leg colours should be ignored.

Your answer to this question should include:

- Any pre-processing that is performed on the data (cropping, resizing), or data augmentation that is used. Note that you may wish to crop and/or resize data to reduce the computational demands of your approach. This is completely acceptable, though the pre-processing should be explained.
- A description of your approach, including justification explaining why you selected this approach, and how the approach was trained.
- An evaluation of performance for each of the traits using the provided test set. The evaluation should also include an investigation of situations where the model performs poorly.

References

[1] D. Gray, S. Brennan, and H. Tao, "Evaluating Appearance Models for Recognition, Reacquisition, and Tracking," *Proc. IEEE International Workshop on Performance Evaluation for Tracking and Surveillance (PETS)*, 2007.