CSE 515 Multimedia and Web Databases

Phase #2

(Due October 15 2023, midnight)

Description: In this project, you will experiment with

- image features,
- vector models,
- dimensionality curse,
- graph analysis.

NOTES:

- You can use existing libraries for LDA decomposition.
- You can use existing libraries CP decomposition.
- You can use existing libraries for eigenvector and eigenvalue extraction.
- The tasks in this phase involve the feature models and similarity/distance functions developed in the previous phase.

PROJECT TASKS:

- **Task 0:**
  - **Task 0a:** Using pre-trained RESNET50 neural network model, map even numbered (labeled) images in the Caltec101 data set into 5 different feature spaces and store the resulting data vectors:
    - Color moments, CM10x10
    - Histograms of oriented gradients, HOG
    - ResNet-AvgPool-1024
    - ResNet-Layer3-1024
    - ResNet-FC-1000

    In the database, store not only the imageIDs and feature vectors, but also the original image labels.
  - **Task 0b:** Implement a program which, given (a) an (even or odd numbered) imageID or an image file, (b) a user selected feature space, and (c) positive integer k, identifies and visualizes the most similar k images, along with their scores, under the selected feature space.
- **Task 1:** Implement a program which, given (a) a query label, l (b) a user selected feature space, and (c) positive integer k, identifies and visualizes the most relevant k images for the given label l, along with their scores, under the selected feature space.
- **Task 2:**

  - **Task 2a:** Implement a program which, given (a) a query imageID or image file, (b) a user selected feature space, and (c) positive integer k, identifies and lists k most likely matching labels, along with their scores, under the selected feature space.
  - **Task 2b:** Implement a program which, given (a) a query imageID or image file and (b) positive integer k, identifies and lists k most likely matching labels, along with their scores, under the RESNET50 neural network model.

- **Task 3 (LS1):** Implement a program which (a) given one of the feature models, (b) a user specified value of k, (c) one of the four dimensionality reduction techniques (SVD, NNMF, LDA, k-means) chosen by the user, reports the top-k latent semantics extracted under the selected feature space.
  - Store the latent semantics in a properly named output file
  - List imageID-weight pairs, ordered in decreasing order of weights

- **Task 4 (LS2):** Implement a program which (a) given one of the feature models, (b) a user specified value of k, reports the top-k latent semantics extracted using CP-decomposition of a three modal (image-feature-label) tensor under the selected features space. Each latent semantic should be presented in the form of a list of label-weight pairs, ordered in decreasing order of weights.
  - Store the latent semantics in a properly named output file
  - List label-weight pairs, ordered in decreasing order of weights

- **Task 5 (LS3):** Implement a program which, (a) given one of the feature models and (b) a value k,
  - creates (and saves) a label-label similarity matrix,
  - performs a user selected dimensionality reduction technique (SVD, NNMF, LDA, k-means) on this label-label similarity matrix,
  - stores the latent semantics in a properly named output file
  - lists label-weight pairs, ordered in decreasing order of weights

- **Task 6 (LS4):** Implement a program which, (a) given one of the feature models and (b) a value k,
  - creates (and saves) an image-image similarity matrix,
  - performs a user selected dimensionality reduction technique (SVD, NNMF, LDA, k-means) on this image-image similarity matrix
  - stores the latent semantics in a properly named output file
  - lists image-weight pairs, ordered in decreasing order of weights

- **Task 7:** Implement a program which, given (a) an (even or odd numbered) imageID or an image file name, (b) a user selected latent semantics, and (c) positive integer k, identifies and visualizes the most similar k images, along with their scores, under the selected latent space.

- **Task 8:** Implement a program which, given (a) an (even or odd numbered) imageID or an image file name, (b) a user selected latent semantics, and (c) positive integer k, identifies and lists k most likely matching labels, along with their scores, under the selected latent space.

- **Task 9:** Implement a program which, given (a) a label l, (b) a user selected latent semantics, and (c) positive integer k, identifies and lists k most likely matching labels, along with their scores, under the selected latent space.

- **Task 10:** Implement a program which, given (a) a label l, (b) a user selected latent semantics, and (c) positive integer k, identifies and lists k most relevant images, along with their scores, under the selected latent space.

- **Task 11** Implement a program which (a) given a feature model or latent space, (b) a value n, (c) a value m, and (d) a label l
  - creates a similarity graph, $G(V, E)$, where V corresponds to the images in the database and E contains node pairs v_i, v_j such that, for each subject v_i, v_j is one of the n most similar images in the database in the given space
  - identifies the most significant m images (relative to the given label l) using personalized PageRank measure. See

    Huang, S., Li, X., Candan, K. S., Sapino, M. L. (2016). Reducing seed noise in personalized PageRank. Social Network Analysis and Mining, 6(1), 1-25.

Deliverables:

- Your code (properly commented) and a README file.
- Your outputs for the provided sample inputs.
- A short report describing your work and the results.

Please place your code in a directory titled "Code", the outputs to a directory called "Outputs", and your report in a directory called "Report"; zip or tar all off them together and submit it through the digital dropbox.