# Assignment #3: COMP4434 Big Data Analytics

Due Date: **23:59pm, Thursday, 15 April 2021**

## Q1. Regularization

[15 points] We use polynomial regression for the prediction task of a dataset. The given dataset includes a train set (train.csv) and a test set (test.csv). To illustrate the effect of regularization, please first implement the following regression models using python language (third-party packages are allowed). Then, plot the data points of the train set and the regression lines of the trained models. Finally, compute the RMSE of the trained models using the test set and make a comparative discussion about underfitting and overfitting.

- Polynomial regression without regularization (polynomial to $5^{th}$ power)
- L1 Regularized polynomial regression: $\lambda = 1$ and $\lambda = 100$
- L2 Regularized polynomial regression: $\lambda = 1$ and $\lambda = 100$

The given datasets can be downloaded at:

https://drive.google.com/drive/folders/1LSZNIEWf6XKnQtRw8L01tS6yAB67Aad2?usp=sharing

## Q2. Recommender System

Build up a collaborative filtering-based recommender system to provide effective hotel recommendation. The training dataset as shown in the table below contains the ratings from 4 users to 3 hotels. The ratings range from 1 point to 5 points.

| | Hotel 1 | Hotel 2 | Hotel 3 |
|---|---|---|---|
| User 1 | 5 | 1 | ? |
| User 2 | 4 | ? | 3 |
| User 3 | ? | 4 | 5 |
| User 4 | 3 | 3 | 4 |

We use the gradient descent algorithm to solve cost minimization in the collaborative filtering model. Some settings are as follows.

- The constant learning rate $\alpha = 0.0002$
- The regularization parameter $\lambda = 0.02$
- The dimension for user/item feature vectors $K = 2$
- The initial values for parameters $x = \begin{bmatrix} 0.77 & 0.43 & 0.31 \\ 0.48 & 0.44 & 0.51 \end{bmatrix}$ and $\theta^T = \begin{bmatrix} 0.19 & 0.62 \\ 0.68 & 0.78 \\ 0.18 & 0.08 \\ 0.36 & 0.92 \end{bmatrix}$

a) [5 points] If we finally obtain $x^{(1)} = [1.268\ 0.994]^T$ and $\theta^{(3)} = [0.271\ 0.694]^T$ after the training procedure, what is the rating of user 3 on hotel 1?

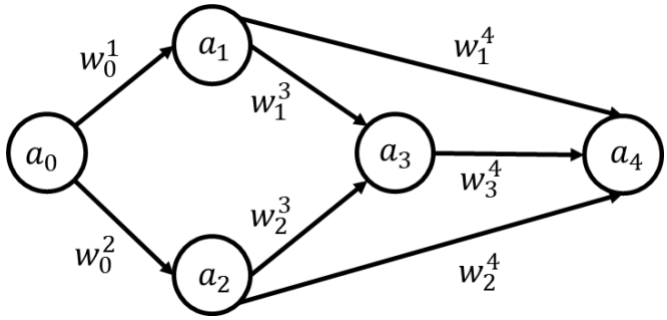
b) [10 points] Calculate the values of $x_1^{(1)}$ (i.e., the first element in the item feature vector of hotel 1) and $\theta_1^{(2)}$ (i.e., the first element in the user feature vector of user 2) after the first iteration.

c) [5 point] Implement the gradient descent algorithm to update the parameters $x$ and $\theta$ using python language. Please calculate the ratings of user 2 on hotel 2 after 50 rounds and upload the source code file.

ps. For a) and b), the detailed calculation process is required and the intermediate and final results should be rounded to 3 decimal places.

**Q3. Neural Network**

[10 points] Consider the following neural network:




Where $a_i = \sum_j w_j^i z_j$, $z_i = f_i(a_i)$ for $i = 1,2,3,4$, $z_0 = a_0$ (an input neuron), $f_3(x) = \text{relu}(x)$, and $f_1(x) = f_2(x) = f_4(x) = \text{sigmoid}(x)$. $\text{relu}(x)$ corresponds to a rectifier linear unit transfer function defined as: $\text{relu}(x) = \max\{0, x\}$. The cost function is defined as $J(w) = \frac{1}{2}(z_4 - y)^2$.

**(a)** Write a function $F$ to simulate the neural network.

**(b)** Assume that we are given a training data $x = 1.0, y = 0.1$, what is the value of $\frac{\partial J}{\partial w_3^4}$?