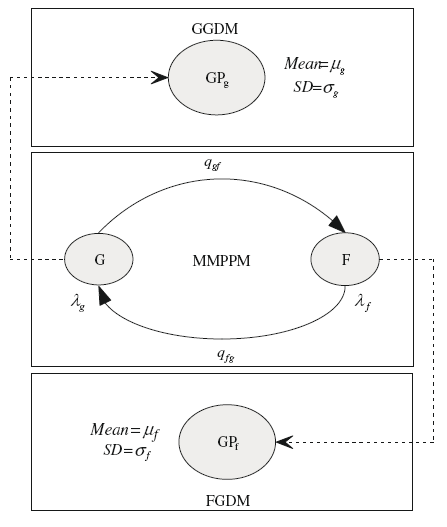
**MCP261 IE Lab I, Groups 1 and 2: April 4, 2019**

**Exercise 9: Analytics + Linear Programming**

1. (3 marks) For a credit-card fraud detection problem, genuine transactions and fraudulent arrive according to the Markov-modulated Poisson arrival process shown in the figure below (Fig. 2 from Panigrahi and colleagues (2009). Available at: <https://www.sciencedirect.com/science/article/pii/S1566253509000141>). Transaction amounts (genuine and fraudulent) are normally distributed.



Generate a dataset consisting with a total of 10,000 transactions with the following parameters of the MMPP simulator: arrival rate of genuine and fraudulent transactions: 10 per day and 2 per day, respectively; qgf = 0.2, qfg = 0.8; μg = 50, σg = 10; μf = 100, σg = 25. Transaction amounts cannot be negative. Use a seed of 1234 for your random number generator. Write the output to an Excel file (features as well as labels). Upload both the Matlab code for the synthetic data generator as well as the Excel file with the training and testing data set.

1. (4 marks) Once you generate the transaction dataset, use the Matlab classification learner to find the algorithm that works best (max. modified f1 score) in terms of predicting whether a transaction is genuine or fraudulent using transaction amount and time between transactions as features. Apply only the following algorithms: ensemble bagged trees, RUSboosted trees, quadratic discriminant, logistic regression, cubic SVM and linear SVM. Use the first 7000 transactions for training and 3000 for testing. The outputs of each algorithm must be in the same Excel file as for Problem 1, and in the same format (the output of each algorithm in a separate Worksheet) as prescribed for Exercise 8.
2. (3 marks) Consider the aggregate planning formulation including overtime you developed for the MCL361 Minor 2 exam. Assume that it is a 7-month problem with a single product, and the product is produced on 4 workstations. The following data is provided for the problem:

|  |  |  |  |
| --- | --- | --- | --- |
| lt |  |  |  |
| 10, 20, 5, 0, 5, 10, 15 |  |  |  |
| ut |  |  |  |
| 100, 150, 100, 200, 150, 160, 120 |  |  |  |
| c1t | c2t | c3t | c4t |
| 300, 250, 200, 400, 250, 180, 450 | 250, 250, 200, 400, 400, 300, 250 | 200, 200, 250, 200, 200, 150, 200 | 150, 200, 200, 250, 300, 250, 200 |
| a1 | a2 | a3 | a4 |
| 2 | 3 | 2 | 4 |
| r = 3 |  |  |  |
| b = 0.5 |  |  |  |
| h = 0.25 |  |  |  |
| O1 = 2.5 | O2 = 1.5 | O3 = 4.0 | O4 = 2.0 |

The symbols have their usual meaning, with the Oj referring to the unit cost of overtime on workstation j. Formulate and solve the above linear programming problem using Excel’s Solver add-in, and provide your solution in the same Excel file (separate worksheet named “P3”) as for Problems 1 and 2. The objective function, constraints and optimal solution must be clearly indicated in the Excel file. Validate the solution using Matlab’s **linprog** function.