



Inspire...Educate...Transform.

Improving organizational efficiencies

**Mentor Scheduling
Calendar Invites
Tableau Reports**

Sagar Patel
Data Scientist, INSOF



Mentor Scheduling – Problem Statement

Allocate mentors to classes

Date	PGP-47	PGP-48	PGP-49
27-Oct			
28-Oct			
3-Nov			
4-Nov			



Date	PGP-47	PGP-48	PGP-49
27-Oct	Sridhar	Kishore	Lab
28-Oct	Sridhar	Murthy	Lab
3-Nov	Sridhar	Anand	Kishore
4-Nov	Surya	Anand	Kishore

T&Cs ahead...



Conventional time-table scheduling problems

Day	Class 1	Class 2	Class 3
Mon			
Tue			
Wed			
Thu			
Fri			

INSOFE's mentor scheduling problem

Date	PGP-47	PGP-48	PGP-49
27-Oct			
28-Oct			
3-Nov			
4-Nov			

Weekly requirements are constant.
Eg. 3 stats class, 3 programming classes, etc

Requirements are dynamic and depend on the a lot of factors like **batch start date**, **topics dependency**, **mentor availability**, etc

Problem can be reduced to 5 days scheduling.
This generalized template is replicated throughout the duration of course.

Each week requirement is highly dynamic and hence one cannot reduce the problem into a smaller generalized form.

Usually the faculty is stationed at one campus and hardly is required to travel to another location.

Because of multiple satellite campus, the mentors are required to travel and technically serve multiple locations.

Most faculties are committed full-time and in terms of resources, have equal cost associated to them.

The mentors may have a full-time or part-time association. Moreover they are stationed at different locations. Hence each mentor-allocation has a variable cost associated with it.

Lets get an idea of how large the solution space is...

Date	PGP-47	PGP-48	PGP-49
27-Oct	Stats	Stats	ML



How many solutions are possible?

Dr. Sridhar Pappu
Dr. Anand Jayaraman
Dr. Manoj Duse
Dr. Anand Narasimhamurthy
Dr. Ventakesh Sunkad

Dr. Surya Kompalli
Dr. Kishore Konda
Dr. Sreerama Murthy
Dr. Rohit Lotlikar
Dr. Manish Gupta
Dr. Praphul Chandra
Dr. Manoj Duse

Lets get an idea of how large the solution space is...

Date	PGP-47	PGP-48	PGP-49
27-Oct	Stats	Stats	ML



How many solutions are possible?

140

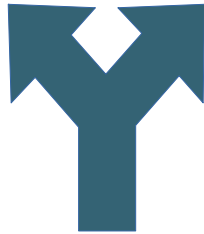
Date	PGP-47	PGP-48	PGP-49
27-Oct	5	4	7

Dr. Sridhar Pappu
Dr. Anand Jayaraman
Dr. Manoj Duse
Dr. Anand Narasimhamurthy
Dr. Ventakesh Sunkad

Dr. Surya Kompalli
Dr. Kishore Konda
Dr. Sreerama Murthy
Dr. Rohit Lotlikar
Dr. Manish Gupta
Dr. Praphul Chandra
Dr. Manoj Duse

How does it scale?

Date	PGP-47	PGP-48	PGP-49	PGP-50
27-Oct	Stats	Stats	ML	ML
28-Oct	Stats	Stats	ML	ML



4 Batches
2 Days

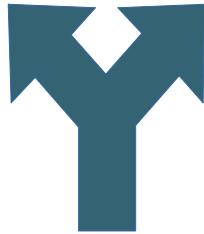
**How many solutions
are possible?**

Dr. Sridhar Pappu
Dr. Anand Jayaraman
Dr. Manoj Duse
Dr. Anand Narasimhamurthy
Dr. Ventakesh Sunkad

Dr. Surya Kompalli
Dr. Kishore Konda
Dr. Sreerama Murthy
Dr. Rohit Lotlikar
Dr. Manish Gupta
Dr. Praphul Chandra
Dr. Manoj Duse

How does it scale?

Date	PGP-47	PGP-48	PGP-49	PGP-50
27-Oct	Stats	Stats	ML	ML
28-Oct	Stats	Stats	ML	ML



Dr. Sridhar Pappu
Dr. Anand Jayaraman
Dr. Manoj Duse
Dr. Anand Narasimhamurthy
Dr. Ventakesh Sunkad

Dr. Surya Kompalli
Dr. Kishore Konda
Dr. Sreerama Murthy
Dr. Rohit Lotlikar
Dr. Manish Gupta
Dr. Praphul Chandra
Dr. Manoj Duse

Date	PGP-47	PGP-48	PGP-49	PGP-50
27-Oct	5	4	7	6
28-Oct	5	4	7	6

4 Batches How many solutions
2 Days are possible?

$$= 5 * 4 * 7 * 6 * 5 * 4 * 7 * 6$$
$$= 705600$$

Note - Technically each day allocation is independent of the previous allocation.

Generalized form

Date	PGP-47	PGP-48	PGP-49	PGP-50
27-Oct	5	5	5	5
28-Oct	5	5	5	5

4 Batches

2 Days

5 Mentors*

***Approximating mentor**

availability for easier calculations

$(\text{Batches} * \text{Days}) ^ (\text{Available mentors})$

$= (4*2)^5$

$= 32,768$

Note – We are underestimating the available mentors



Generalized form

Date	PGP-47	PGP-48	PGP-49	PGP-50
27-Oct	5	5	5	5
28-Oct	5	5	5	5



4 Batches
2 Days
5 Mentors*

*Approximating mentor
availability for easier calculations
(Batches * Days) ^ (Available mentors)
 $= (4 * 2)^5$
 $= 32,768$

**Note – We are underestimating the available mentors.
So its a kind of lower bound.**

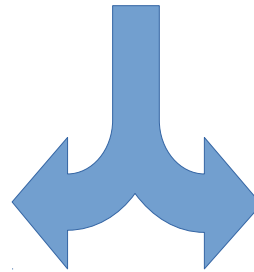
6 Batches
50 Weekend days*
5 Mentors

*Equivalent to 6 months of forward
planning.

$= (6 * 50) ^ 5$
 $= 2.43 \times 10^{12}$ (2.43 Trillion)
 $= 2430,000,000,000$

Assuming that you have the set of all possible solutions, how do you select the best solution??

Date	PGP-47	PGP-48	PGP-49
27-Oct	Stats	Stats	ML

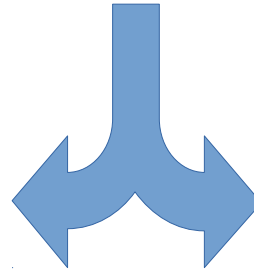


Date	PGP-47	PGP-48	PGP-49
27-Oct	Sridhar	Anand	Surya

Date	PGP-47	PGP-48	PGP-49
27-Oct	Kishore	Surya	Sridhar

Assuming that you have the set of all possible solutions, how do you select the best solution??

Date	PGP-47	PGP-48	PGP-49
27-Oct	Stats	Stats	ML



Date	PGP-47	PGP-48	PGP-49
27-Oct	Sridhar	Anand	Surya

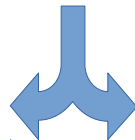
Date	PGP-47	PGP-48	PGP-49
27-Oct	Kishore	Surya	Sridhar

We need some kind of heuristic that gives us an idea of how good the solution is.

Mentor-Topic proficiency matrix

Module	Topics	Dr. Sridhar Pappu	Dr. Surya Kompalli	Dr. Kishore Konda
CSE 7302c	Statistics and Probability in Decision Modeling	Expert	Not Comfortable	Comfortable
CSE 7302c	Linear Regression	Expert	Proficient	Comfortable
CSE 7305c	Decision Trees	Comfortable	Expert	Expert
CSE 7305c	Association Rules	Comfortable	Proficient	Comfortable

Date	PGP-47	PGP-48	PGP-49
27-Oct	Stats	Stats	ML



Date	PGP-47	PGP-48	PGP-49
27-Oct	Sridhar (Expert)	Anand (Expert)	Surya (Expert)

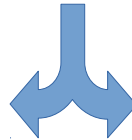
Date	PGP-47	PGP-48	PGP-49
27-Oct	Kishore (Comfortable)	Surya (Not Comfortable)	Sridhar (Comfortable)



However its better to have a number that quantifies the quality of the solution.

Expert	Proficient	Comfortable	Not-Comfortable
3	2	1	0

Date	PGP-47	PGP-48	PGP-49
27-Oct	Stats	Stats	ML



Date	PGP-47	PGP-48	PGP-49
27-Oct	Sridhar (Expert)	Anand (Expert)	Surya (Expert)
Avg=3	3	3	3

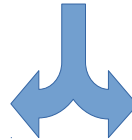
Date	PGP-47	PGP-48	PGP-49
27-Oct	Kishore (Comfortable)	Surya (Not Comfortable)	Sridhar (Comfortable)
Avg=0.6	1	0	1



Increasing the complexity...

Adding location and faculty association into consideration

Date	PGP-47 (BLR)	PGP-48 (HYD)	PGP-49 (HYD)
27-Oct	Stats	Stats	ML



Date	PGP-47	PGP-48	PGP-49
27-Oct	Sridhar (Expert)	Anand (Expert)	Manish (Expert)
	Travel		
			External

Date	PGP-47	PGP-48	PGP-49
27-Oct	Rohit (Comfortable)	Anand (Expert)	Surya (Expert)

Problem

Competency	Travel	Internal
Expert	0	1
Expert	0	0
Expert	1	1
Expert	1	0
Proficient	0	1
Proficient	0	0
Proficient	1	1
Proficient	1	0
Comfortable	0	1
Comfortable	0	0
Comfortable	1	1
Comfortable	1	0
Not-Comfortable	0	1
Not-Comfortable	0	0
Not-Comfortable	1	1
Not-Comfortable	1	0

Competency	Travel	Internal
Proficient	0	1
Expert	1	1



Solution

Competency	Travel	Internal	Rank	Score
Expert	0	1	1	15
Proficient	0	1	2	14
Expert	0	0	3	13
Proficient	0	0	4	12
Expert	1	1	5	11
Proficient	1	1	6	10
Comfortable	0	1	7	9
Comfortable	0	0	8	8
Expert	1	0	9	7
Proficient	1	0	10	6
Comfortable	1	1	11	5
Comfortable	1	0	12	4
Not-Comfortable	0	1	13	2
Not-Comfortable	0	0	14	1
Not-Comfortable	1	1	15	1
Not-Comfortable	1	0	16	0



Using Linear Regression to understand the intuition

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.8750	0.7925	2.366	0.039560	*
Competency_Comfortable	5.5000	0.9152	6.010	0.000130	***
Competency_Expert	10.5000	0.9152	11.474	4.45e-07	***
Competency_Proficient	9.5000	0.9152	10.381	1.13e-06	***
Travel	-3.7500	0.6471	-5.795	0.000174	***
Internal	2.0000	0.6471	3.091	0.011431	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.294 on 10 degrees of freedom

Multiple R-squared: 0.9537, Adjusted R-squared: 0.9305

F-statistic: 41.19 on 5 and 10 DF, p-value: 2.354e-06

Notice the beauty of this technique. It can be used pretty much anywhere to decode the factors behind any decision making.

Cracking Warren Buffet's strategy...

Imagine you have financial statistics (EPS, P/E Ratio, Dividend, etc) of a bunch of companies.

Now if you can get Warren Buffet to give his recommendation for each of them, after a sufficiently large number of samples you can come up with a model that replicates Warren Buffet's logical thinking.

Note – T&Cs Apply. To be taken with a pinch of salt :p



Now that you have the formula to rank the mentors, all you need to do is get the topic-flow matrix and

- 1. Find the list of available mentors for each day**
- 2. Rank them according to their proficiency**

Day	Week	Module	Module Text	Topics
Day11	Week6	CSE 7302c	Statistics and Probability in Decision Modeling	Multiple Linear Regression and Feature Selection (StepAIC & VIF)
Day12	Week6	CSE 7302c	Statistics and Probability in Decision Modeling	Logistic Regression & Naive Bayes Classifier
Day13	Week7	CSE 7302c	Statistics and Probability in Decision Modeling	PCA & Regularization
Day14	Week7	CSE 7302c	Statistics and Probability in Decision Modeling	Time Series

Calculate the score for each mentor allocation and then rank them to get the final matrix...

Date	Topics	Best	Best Score	2nd Best	2nd Best Score	3rd Best	3rd Best Score
2019-03-23	Foundations of Probability and Statistics for Data Science	Dr. Sridhar Pappu	10.08	Dr. Anand Jayaraman	10.08	Dr. Surya Kompalli	3.36
2019-03-24	Introduction to Supervised Methods, Linear Regression	Dr. Anand Jayaraman	10.08	Dr. Surya Kompalli	6.72	Dr. Sreerama Murthy	4.2
2019-05-19	Logistic Regression & Naive Bayes Classifier	Dr. Anand Jayaraman	16.8	Dr. Surya Kompalli	13.44	Dr. Dakshinamurthy Kolluru	6.72
2019-05-25	Ensemble techniques: Bagging (RF) & Boosting (GBM)	Dr. Surya Kompalli	10.08	Dr. Anand Jayaraman	6.72	Dr. Manish Gupta	6.3



Disclaimer

- 1.The topics to be taught have dependency on each other (Statistics should be taught before ML). The current algorithm relies on a strict topic flow to handle this. However one can shuffle a couple of classes without hampering the dependencies. Such solutions aren't considered in this approach.
- 2.A good schedule is one which doesn't overburden a single resource. It is been observed that a carefully articulated Mentor-Topic proficiency matrix helps to converge at a reasonable solution. However it cannot be ruled out that this approach may fail in some rare scenarios leading to mentor over-utilization.
3. Given the dynamic nature of the ad-hoc requirements (corporate training, students requests, etc) one may have to manually improvise on the schedules generated.
- 4.For the above mentioned reasons, this system can be considered as an assistive system but not decisive system.




Is there anything better out there??

Genetic algorithms remain a popular approach in solving such type of problems.

Then why not?

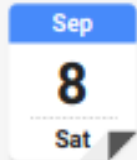
- Given the differences compared to traditional scheduling problems, it would require good amount to research and testing to come up with a complete solution.
- Development time for GA is very high and convergence time is always a challenge.
- GA has no transparency of logic and hence it becomes difficult to handle ad-hoc requests through manual intervention.
- In cases where rescheduling becomes inevitable, the global optimal solution becomes sub-optimal. The 5% improvement due to GA goes for a toss.



Invitation: Introduction to supervised methods, Linear regression @ P... @ Sat Sep 8, 2018 2pm - 6pm (IST)
(sagar.patel@insofe.edu.in)  Inbox x INSOFE x

INSOFE grader@insofe.edu.in [via](#) google.com

to me, Sajna, harshit.lamba ▾



Introduction to supervised methods, L...

[View on Google Calendar](#)

When Sat Sep 8, 2018 2pm - 6pm (IST)

Who Sajna Vilangapurath M, harshit.lamba@insofe.edu.in, grader@insofe.edu.in*

Agenda

Sat Sep 8, 2018

No earlier events

2pm Introduction to supervised i

No later events

Introduction to supervised methods, Linear regression @ PGP48-HYD

Batch: PGP48-HYD

Date: 08-Sep

ROTe: Day 7 Topics

ROTe Topics: Advanced commands in R, Begin with Data preprocessing

Module: CSE 7302c - Statistics and Probability in Decision Modeling

Topics to be taught: Introduction to supervised methods, Linear regression

Mentor: Dr. Anand Jayaraman

anand.jayaraman@insofe.edu.in - 9959466177

Data Scientists

Harshit - harshit.lamba@insofe.edu.in - 9158873720

SagarPatel - sagar.patel@insofe.edu.in - 9429830337

*This notification system is in beta phase. If you find anything suspicious or incorrect please write back at sagar.patel@insofe.edu.in

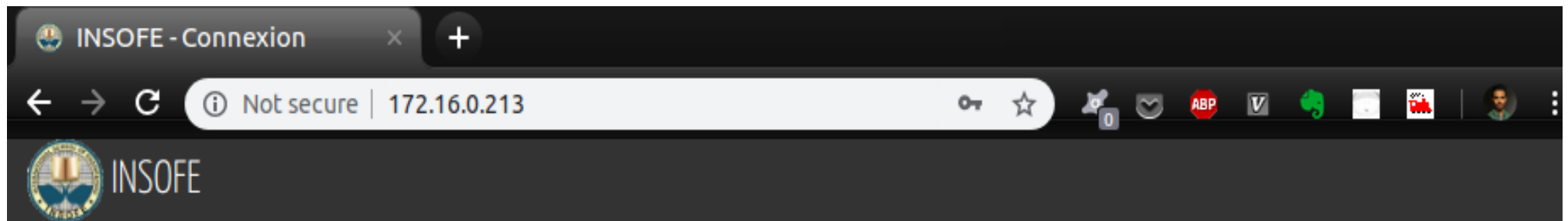


Calendar notifications

	DS 1	DS 2	DS 3	DS 4	DS 5
12-Sep					Corporate
13-Sep	Corporate				Corporate
14-Sep		Corporate			Corporate
15-Sep	PGP		PGP	PGP	
16-Sep	PGP		PGP	PGP	

Date	Module	Module Text	Mentor	ROTe	CUTe	Topics
2-Sep	CSE 7212c	Essential Engineering Skills in Big Data Analytics Using R and Python	Data Scientists			Orientation and Introduction to R
3-Sep	CSE 7315c	Foundations of Probability and Statistics for Data Science	Dr. Sridhar Pappu			Probability and Statistics-Day 1
9-Sep	CSE 7315c	Foundations of Probability and Statistics for Data Science	Dr. Sridhar Pappu			Probability and Statistics-Day 2
10-Sep	CSE 7305c	Methods and Algorithms in Machine Learning	Dr. Sreerama Murthy			KNN + Collaborative filtering
16-Sep	CSE 7305c	Methods and Algorithms in Machine Learning	Dr. Sreerama Murthy			Ensemble technique: RF, SVM





Hello! Welcome on the INSOFE platform.

INSOFE allows to securely grade code made by students.

Please log in to see the courses.

Email

sagar.patel@insofe.edu.in

Password

.....

[Register / Lost password?](#)

[Log in](#)



Solution	Before	Now
Grader	Exam submission done on piazza. DS had to download each student submission independently. Students engagement was less.	Submissions organized at one place. Single click submission download for whole class. Increased students engagement due to Kaggle like environment and leader-board scores.
Scheduler	Manual mentor scheduling. Tedious work of checking topic → available mentors → blocking the allocated mentor (repeated 50 times)	Sajna believes that her work is reduced by 50% (<i>I would dispute that :p</i>) 1 schedule generated in 1 min. All she has to do is verify the allocation and take care of ad-hoc requests.
Calendar Notification	Refer to EDMS to check schedule every week. Keep a check of changes in EDMS	You get a mail along with your allocation details. Event gets added to your calendar along with appropriate reminder alerts.
Tableau Dashboard	Sajna had to manually generate resource utilization reports. BD team would call up Sajna every-time to check mentor availability or feasible dates.	Automated reports generations. Tableau dashboards generated providing a easy UI to understand resource utilizations. BD team has a go-to tool and less dependence of Sajna or EDMS

**Remember its a crime to do redundant work.
Let the smartness prevail ...**

Thank You

