

# Scheduler Documentation

## Primary modules

- A) Mentor Scheduling
- B) Tableau dashboard
- C) Calendar notifications

## Important data files

### 1. Faculty-Scientist Scheduling Matrix.xlsx

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

### 2. Program Schedule.xlsx

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

### 3. Topics-Proficiency-Modified-3.csv

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

#### 4. Ideal Topic Flow (PGP\_Topic\_Flow.csv)

Day	Week	Module	Module Text	Topics
Day2	Week1	CSE 7315c	Foundations of Probability and Statistics for Data Science	Foundations of Probability and Statistics for Data Science
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

#### 5. Contact List

Name	emailid	Location	Skills/Responsibilities	Contact#	Association
Manas Mohan Mishra	manasmohan.mishra@insofe.edu.in	Hyderabad	PGP - ML & BigData	8501943811	INSOFE
Pavan Srungaram	pavan.srungaram@insofe.edu.in	Hyderabad	PGP - ML	9493107946	INSOFE
Ramesh Melapu	ramesh.melapu@insofe.edu.in	Hyderabad	PGP - ML & BigData	9000169678	INSOFE
Surya Kompalli	surya.kompalli@insofe.edu.in	Hyderabad	PGP, Research	9989780170	INSOFE
Dakshinamurthy Kolluru	murthy.kolluru@insofe.edu.in	Hyderabad	PGP, Consulting	8008457777	INSOFE
Dr. Ananda Jayaraman	anand.jayaraman@insofe.edu.in	Hyderabad	PGP, Consulting	9959466177	External
Dr. Sreerama Murthy	smurthy@quadratyx.com	Hyderabad	ML & BigData	9866661197	External

### Important Scripts

#### 1. fetch\_schedule.py

**Purpose** – Melt faculty scheduling matrix for easy querring and aggregation of data.

**Approach** – Melt data

**Inputs:**

- Faculty Scheduling Matrix (*Faculty-Scientist Scheduling Matrix.xlsx*)

**Output:**

- Melted Mentor Schedule (*melted\_mentor\_schedule.csv*)

	Mentor 1	Mentor 2	Mentor 3
Date 1			
Date 2	Allocation		
Date 3		Allocation	Allocation



Date	Mentor	Allocation	Type	Hours	4W

*Faculty-Scientist Scheduling Matrix.xlsx*

*melted\_mentor\_schedule.csv*

#### Variable description:-

Type – Nature of allocation (Academic/Corporate)

Hours – Session length

4W – Past 4 weeks aggregate

## 2. ds\_utilization.py

**Purpose** – Summarize Data Scientist’s work efforts in term of total work hours.

**Approach** – 8 hours if allocation is a full-day-lab/examination/viva  
4 hours if just the afterternoon session.

**Inputs:**

- Faculty Scheduling Matrix (*Faculty-Scientist Scheduling Matrix.xlsx*)
- Program Schedule (*Program Schedule.xlsx*)

**Outputs:**

- Melted DS Schedule (*melted\_ds\_schedule.csv*)

## 3. mentor\_competency.py

**Purpose** – Convert categorical mentor profeciencies into a score that facilitates ranking.

**Parameters to be configured -**

- a) *near\_factor* – The factor by which the score of a local mentor is scaled
- b) *internal* – The factor by which the score of an internal mentor is scaled

**Approach** –  $\text{final\_score} = \text{base\_score} * \text{near\_factor} * \text{internal}$

**Note** – Since this scoring is location dependent, two seperate matrices are calculated based on the location.

**Inputs:**

- Faculty-Topics-Proficiency-Matrix (*Topics-Proficiency-Modified-3.csv*)

**Output:**

- Faculty-Topics-Score-Matrix (*HYD Mentor Topic.csv*)

*Topics-Proficiency-Modified-3.csv*

	Mentor 1	Mentor 2	Mentor 3
Topic 1	Proficient	Expert	Not-Comfortable
Topic 2	Comfortable	Expert	Not-Comfortable
Topic 3	Comfortable	Comfortable	Expert



*HYD Mentor Topic.csv*  
*BLR Mentor Topic.csv*

	Mentor 1	Mentor 2	Mentor 3
Topic 1	7	9	0
Topic 2	4	9	0
Topic 3	4	4	9

#### 4. schedule.py

**Purpose** – Generate candidate mentor allocations for a new batch

**Approach** –

1. Construct a mentor availability matrix by filtering the available mentors for each date-topic pair and rank them according to the proficiency score.
2. Iterate through each row (date) of the topic flow and allocate the top mentor for each date-topic pair.

**Inputs:**

- Mentor Schedule (melted mentor schedule.csv)
- Mentor-Topic proficiency matrix  
(HYD Mentor Topic.csv or BLR Mentor Topic.csv)
- Course start date
- List of Holidays
- Topic flow (Note – separate for PGP and Rennes)

**Output:**

- Batch schedule

**Sample output:**

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
2019-06-16	SVM and Summary on Architecting ML Solutions	Dr. Surya Kompalli	13.44	Dr. Anand Jayaraman	13.44	Dr. Manish Gupta	12.6

#### 5. block\_mentors.py

**Purpose** – Lock the best mentor and block their schedule in melted mentor schedule.csv

**Approach** – Consider the mentor with the maximum score

**Inputs:**

- Mentor Schedule (melted mentor schedule.csv)
- Schedule generated by schedule.py

**Outputs:**

- Melted Mentor Schedule (melted mentor schedule.csv)

**Note** – This has to be repeated for each newly generated schedule

Date	Mentor	Allocation	Type	Hrs	4W
2018-09-07	Dr. Sridhar Pappu			0	46
2018-09-08	Dr. Sridhar Pappu			0	46
2018-09-09	Dr. Sridhar Pappu			0	46
2018-09-10	Dr. Sridhar Pappu			0	46



Date	Mentor	Allocation	Type	Hours	4W
2018-09-07	Dr. Sridhar Pappu			0	46
2018-09-08	Dr. Sridhar Pappu	PGP49-HYD	Academic	4	46
2018-09-09	Dr. Sridhar Pappu	PGP49-HYD	Academic	4	46
2018-09-10	Dr. Sridhar Pappu			0	46

## 6. aux\_function.py

**Purpose** – Supplementary functions to read excel sheets, generate dates, flatten lists, etc

## 7. send\_weekend\_notifications.py

**Purpose** – Send calendar notifications to Mentors and DS for weekend class schedule along with other relevant information like topics to be taught and ROTe details.

**Approach** – Cross-reference the Program-Schedule and Faculty-Scheduling-Matrix to figure out details for each class like Mentor, DS, Topics, ROTe and maintain a dataframe for each weekend. Now using Google Calendar API generate event for each class and store eventId for future references.

### Inputs:

- Contact list
- Faculty Scheduling Matrix (*Faculty-Scientist Scheduling Matrix.xlsx*)
- Program Schedule (*Program Schedule.xlsx*)

### Outputs:

- Event DataFrame (*stores eventId and corresponding info for each class*)