**April 7th/2022**

* Finalize Progress Report 2
* Determine a time and date for our last meeting w/ CM [M/T 1-2pm]
* Update Gantt chart and determine what is missing and allocate tasks
* Determine leads and set clear goal for everyone on tasks

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**March 30th/2022**

* Profile report didn’t work (just try adding things to the algorithms and seeing the accuracy as we go)
* One person keeps adding the new variables to the algorithm while two others working on the feature
* See if it is possible to combine the UCIs?
* **Yutong** -> can do some feature engineering
* **Feyre** -> Do your EDA
* **Saul** -> Finished the EDA (includes the primary learning objectives)
  + Will upload the features, waiting on Yutong to send the feature correlations
* Break apart the Funnel specific
* **Everyone** -> Use SkLearn and try to get the highest prediction power
* Goal for tomorrow: Get some baseline for prediction, run it, and then focus on improving prediction

***Upcoming tasks to complete/start working on:***

* **Progress Report #2 [Create a template for after March. 17th]** - April 7th @8am
  + Everyone share your update
  + Template not needed
  + Primary Learning objective (Saul did it already)
* **Code** - April 12th @8am
  + 2 people start working on this
* **GUI** - April 12th @8am
  + 1 person start working on this (Muhammad &/ Yutong)
* **Report [create a template]** - April 12th @8am
  + 1 person create a template (Feyre)
* **(PPT)/Presentation** - April 14th @9:15am
  + 1 person create a template (Feyre)

**Next Meeting: Tomorrow, March 31st @8am SS347**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**March 17th/2022**

**Feyre:** Complete your EDA (Do your own analysis)

**Muhammad & Wes:** Work on the GUI (How will it be implemented/connected)

**Yutong**: Find “impact of number of creatives in market on performance” primary learning objective

**Saul**: Fix EDA and feature importance

***Have Weekly meeting (Saturday at 5pm)***

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**March 3rd/2022**

**Saul:** (Update on current status)

* We’ve encoded the data set using DVs
* Cleaned up the Dates w/ the KPI’s
  + Aligned the values by the weeks (so that they start at the same index)
* Next -> Start selecting features for modelling (recursive feature elimination, and user feature importance)

**Gunpreet:** Are you jumping into Modelling/are there an EDA (Exploratory Data Analysis - Analysing the data as received upfront is an important of the deliverable)

* **Wes:** Dr. Crost said just throw whatever you have on the ML Algorithm and figure out the importance

**Gunpreet:** Look at the data itself and decide what is happening, and determine if there are any outliers before jumping into Modelling (Go from future analysis to data analysis). Make sure Data Analysis is a key part of the project for you (do not ignore what the data looks like or means before and after you clean it.

**Gunpreet:** Are you thinking of KPI prediction as classification task/regression task?

* **Saul:** regression b/c we are trying to predict a numerical problem
* **Gunpreet:** [Pick both methods b/c there can be different performance, look at VIF (Variance Inflation(implaition??) features).
  + Compare what the overall gap is selecting one way/the other (they tend to typically be close to each other) you can also try to stack them. [There is no right/wrong method, just justify why you chose to do things the way you did]
* **Gunpreet:** Try regression method/other methods - Even though random forest would give you an answer, sometimes it can give you the wrong answer and will run longer…[Interpretability is a big component of selecting a model] Talk about what the gap is when selecting one model/another.
  + Google how to select the best models for key feature prediction…?

**Gunpreet**: How have you approached Feature Engineering:

* **Wes**: aggregated spending & impressions on the right hand side. Peaks of each one (Max values). Currently talking about grouping & making more interactive terms with the right hand side terms by utilising DVs. Growth of rate refers to change that is taking place on a weekly basis.

**Gunpreet:** Provide the explanatory (exploration part of the data) as part of the deliverable, and should be done either in-parallel/before the modelling (ties to our primary learning objective) [Up to you to decide whether you want to do it before/in-parallel to the modelling step]

* What is the data telling us at the publisher/other level
* Helps to form an informed KPI prediction

**Next steps:**

* **Feyre:** Send a bi-weekly meeting to gunpreet & Kimberly
* **Start working on Progress Report #1 due March 17th**
* **Show Gunpreet & Kim the cleaned data (screenshare)**
* **Get the updated data from Muhammad (do an exploration of the data separately then collaborate what we find)**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q for CM: Are all the advertisements made by one-client? - Watch the recorded meeting

* Send a progress report
* Meeting minute from last week

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Feb. 12th/2022**

**Team update:**

* Update Project Proposal
* Update Gantt chart
* Establish what to do for the next step
* Figure out days to meet over the reading break (Goal - have at least 1 model built)
* Create date variables (week 1, week 2, …)

**Data cleaning things to figure out?**

* What kind of model/models do we want to use?
  + None atm
* Encode variables or create DVs?
  + Done
* Launch KPI prior to launch (can we omit certain variables?)
  + Determine our predictor variable
    - All KPIs b/c we need funnel specifics
* Look into doing feature importance algorithm
  + Recursive feature elimination
* Funnel specific
* Do we want to view the project as a [regression/classification](https://www.javatpoint.com/regression-vs-classification-in-machine-learning) problem? (We can also break it down into multiple steps & apply a combination of both) ***\*Goal is to justify the thought process behind how we approached the solution\****
  + Not decided yet

**Questions for CM:**

* Further explain the following variables (Publisher, LOB, KPI audience)
* What is a reasonable maximum for CTR?

***Meeting with CM on Monday 1-2pm MST***

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Feb. 7th/2022**

**Wes:** Speakers notes for ‘Partner Information’ and ‘Data’

**Saul**: Speaker notes for ‘Prototype Description’

**Feyre:** Finish the motivation/objective slide & submit by tomorrow 7am

**Meeting @5pm (same zoom link as usual) on Wednesday.**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Feb. 5th/2022**

**Team update:**

* Finished the Project Proposal
* Completing final additions to the PPT (make it pretty while we’re at it)
* Updated Gantt chart
* Merge the Project Proposal Plan with the Project Agreement

**HMWK:**

* **Feyre:** Send the project proposal & agreement to Gunpreet and Kim before Monday, Feb. 7th
* **Yutong:** Get started on Data Wrangling & add a sentence explanation of what pre-training is to the Project Proposal
* **Rest:** Continue to work on the PPT Presentation **(Add any additional information by Sunday afternoon**)
* **Wes**: Convert slides to powerpoint and give theme

**Task for Monday Feb. 7th**

* Finalize presentation plan & practice presentation
* **Yutong:** Continue working on the Data Wrangling
* **BY TUESDAY FEB. 8TH before 8am - submit everything**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Feb. 3rd/2022**

**Team update:**

* Started working on PPT presentation
* Project Proposal has been edited (minor comments (please look at them and add in other feedbacks)
* Gantt chart (Google sheet) completed, updated, and good to go
* **Overall an amazing and productive day :)**

**HMWK:**

* Sign and send back the project proposal agreement to [rosy Oh](mailto:ohrosy483@gmail.com)by Tonight
* **Feyre:** Send the project proposal & agreement to Gunpreet and Kim by tomorrow @6am MST
* **All: Continue to work on the presentation, should be concise, detailed, and ideally appealing**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Intellectual & Property right - Initiate it

Yutong’s questions:

Which ML model do we use? Regression like xgb， ANN？

How do we define the difference channel?

What is the feature screening method? PCA?

What does the output look like? Each Id correlate to a forecasred KPI?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Jan. 29th-Feb. 1st /2022**

**Team update -**

* **Saul:** Formatted project plan & wrote “how project creates value for the org”
* **Feyre:** wrote “how org defines success”
* **Yutong:** wrote “project objective”
* **Wes:** Finished the Gantt Chart

**HMWK FOR Monday Jan. 31st**

* **All:** Expand on your responsibilities to the team information table
* **Saul:** Write the “Prototype description”
* **Feyre:** Finish in details the “milestone & deliverables” for the duration of the semester
* **Yutong:** Write the “Prototype testing”
* **Wes:** Write the “Team information”
* **Muhammad:** Write the post-prototype extensions (**Deadline Feb. 1-3rd**)

**TASK ON MONDAY Jan. 31st**

* **All:** Edit everything on Monday
* Make a PPT presentation for Feb. 8th
* For Tuesday:
  + Start on data cleaning.
  + Update CM [Meeting on Thursday]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Jan. 24th**

**Feyre:** Create a recurring zoom invitation (bi-weekly Thursday 3-4pm EST until Feb. 28th)

**Muhammad & Wes:** Continue to work on the Gantt Chart

**Saul & Feyre:** Continue to work on the Project Plan

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Jan. 22nd**

**HMWK:**

* **Feyre:** Create slack and invite everyone
  + Draft email and send the confirmation to the group & confirm deadlines
* **Wes & Muhammad:** Customize the Gantt Chart
* **Saul:** Provide a template for project plan
* **Muhammad:** Make diagrams for 'use case'

**Answers to Gunpreet:**

* Primary contact - Feyre
* Check-ins (bi-weekly and once project starts move to weekly) the rest communicate via slack/email
  + Thursday 3-4pm MST
* Confirm deadlines

**Moving forward:**

* Move private channel to slack for group
* Data language -> primarily python
* Establish tasks:
  + **Project Plan (5-7pg single spaced) - Feb. 8th/2021**
* Progress Report 1 - **Due Thursday, March 10th/2022**
  + Test 1
  + Test 2
* Progress Report 2 - **Due Thursday, March 31st/2022**
  + Test 1
  + Test 2
* Final Report & Code - **Due Friday, April 29th/2022**
  + Test 1
  + Test 2
* Deliverables
* Post-Prototype Extensions
  + Incorporate GUI
* Establish roles:
* Establish project expectations:
  + Publishing rights (right to use codes later) -> establish prior to the project agreement

Thursday Jan. 20th/2022 -> Meet with Project Partners for a team-meeting

Saturday Jan. 22nd/2022 -> Began working on the Project Plan

Monday Jan. 24th/2022 -> Have Gantt Chart and tasks done

Tuesday Jan. 25th/2022 -> Establish and determine the roles

Thursday Jan. 27th/2022 -> Complete Project Plan up until Prototype Description

Saturday Jan. 29th/2022 -> Finish Plan for Prototype Testing

Monday Jan. 31st/2022 ->

Tuesday Feb. 1st/2022 ->

Thursday Feb. 3rd/2022 ->

Saturday Feb. 5th/2022 ->

Monday Feb. 7th/2022 ->

Tuesday Feb. 8th/2022 -> **Project Plan Deadline!!**