

## TECHFEST 2021-22

### NO CODE ML CHALLENGE

#### TASK

Participating teams must develop ML models that can predict if a customer will discontinue the service or not depending on the data given, by using Knowledge Studio software which allows you to learn and develop ML models **without requiring a single line of code**.

#### INTRODUCTION:

The Telecommunication sector is made up of companies that make communication possible on a global scale, whether it is through the phone or the Internet, through airwaves or cables, through wires, or wirelessly. These companies created the infrastructure that allows data in words, voice, audio, or video to be sent anywhere in the world. In India, the Telecom sector is under consolidation. They are under pressure to maintain the bottom line and invest in the growth potential of technology disruption like 5G. Given the scenario, here is the problem statement through which an engineer can propose a value proposition to these companies based on AI/ML.

#### PROBLEM STATEMENT:

A telecom company wants to use its historical data to predict behavior to retain customers. You need to analyze the relevant customer data using Knowledge Studio. You have to focus on the People **Discontinued** = 'Yes' group (customers who discontinued using the service in the past months) and correctly predict them with reasonable accuracy and also correctly predict for the People **Discontinued** = 'No' group as well. In the given training data, a "1" in the Discontinued column implies the customer stopped using the services in the past month and vice versa.

Try different ML models that will help to identify the potential customer who has a higher probability to discontinue the service. It would help the company in strategizing customer retention.

You must use the **Knowledge Studio** software which is an easy-to-use machine learning and predictive analytics solution that rapidly visualizes data as it quickly generates explainable results - all without requiring a single line of code. This software is provided to all participants by the sponsors, ALTAIR. The steps to install and use this software are provided in the problem statement ([Click here](#)).

Extensive **tutorials** in the form of courses, on using the Knowledge Studio software and how Machine Learning works are available on the ALTAIR website. These will help you immensely if you have little or no prior knowledge of ML. Details on how to access and use them are provided [here](#).

## DATA PROVIDED:

Data Description:

7 Data set includes the information about:

- Customer who left within the last month – column is called **Discontinued**
- Services signed up by each customer
- Customer account information
- Demographic information about the customer.

The dataset can be accessed at: <https://bit.ly/NoCodeDataset>

You are advised to go through the Data Description document in the above folder to gain a better understanding of what each data label means.

## RULES:

- Every team has to **register online** on the Techfest website for the competition. A **Team ID** will be allocated to the team on registration which shall be used for future references.
- A team can register at any point of time before **9th January 2022** and submit the final abstract and video (as mentioned in the structure).
- The decision of the organizers and/or judges shall be treated as final and binding on all. Techfest has all the rights to verify the identity and accuracy of the details provided by the participants.
- Techfest, IIT Bombay will hold no responsibility for any late, lost, or misdirected entries.
- The idea presented by the teams should be original (not protected by patent/copyright/technical publication by anyone else).
- Note that at any point in time the latest information will be that which is on the website.
- All modes of official communication will be through the **Techfest e-mail**.

## JUDGING CRITERIA:

Try different ML models that will help to identify the potential customer who has a higher probability to discontinue the service. It would help the company to understand the pattern of customers and will increase the focus on strategizing customer retention.

1. **Import the data:** [5 Points]
  - a. Merge the data set and import
  - b. Explore the final shape and size of data.
2. **Data Cleanup:** [5 points]
  - a. Missing value treatment
  - b. Convert categorical columns to continuous columns
  - c. Drop attribute(s) if required.

3. **Data analysis and visualize:** [10 points]
  - a. Write the conclusion about the distribution of data and your understanding.
  - b. Determine which factors influence the churning of the data with evidence.
  
4. **Data pre-processing:** [5 points]
  - a. Segregate prediction and target attribute
  - b. Perform train test split
  - c. Check if the train and test data have similar statistical characteristics when compared with original data
  
5. **Model training, testing, and tuning:** [15 points]
  - a. Train and test all the supervised models present in Knowledge studio.
  - b. Display and compare all the models designed with their train and test accuracies
  - c. Select the final best-trained model along with detailed comments for selecting the model.
  
6. **Conclusion and improvisation:** [10 points]
  - a. Write a detailed conclusion on your results
  - b. Detailed suggestions or improvement on the quality, variety, etc. on the data points collected by the telecom operator to perform better data analysis in the future.

**Important Note:**

- Early Submission carries weightage in event of the same score.
- Submitting the Knowledge Studio Certification course certificate will carry weightage in the event of a similar score.
- Certification courses can be found after you [register](#).

**COMPETITION PLATFORM:**

The competition will be conducted on the **Knowledge Studio** software. After registering on the official Techfest website, hop onto the link given below and click on **Request Now**. After that you will be directed to the Sign-Up page where you have to fill in the details.

**Note:**

**Please use your university email address if you have one:** On doing so you will receive the email for verification instantaneously and following that you will receive the license, link for downloading software, and installation procedure.

**IN CASE you don't have a university email address:** You will have to provide a valid proof to prove you are a college student like college id, Aadhar card, Birth Certificate, etc. In such a case it can take upto ( 24 hours if you request on normal days and 48 hours if you request on weekends ) to receive the license via email, link for downloading software, and installation procedure. *Do check your spam folder for the license.*

**Software Request Link:**

<https://altairuniversity.com/free-altair-student-edition/>

**Resources and Certification Courses:**

Courses are available for understanding the software and knowledge about different ML models and other techniques.

The recommended course is provided [here](#) which you can access after you [register](#).

For certification of this course, you have to give a [quiz](#). You will be provided **3 attempts** for the quiz and you have to score more than **70%** for certification. So it is recommended to go through the course and then attempt the quiz.

**COMPETITION TIMELINE**

Last Date of Registration	9th January 2022
Submission Deadline	9th January 2022
Results Announcement	20th January 2022

**FINAL SUBMISSION:**

In the final submission, the participants have to share the **whole project folder** and not just the workflow model. This folder has to be submitted via [this google form](#). As a backup, you can also send it via mail to [nocode@techfest.org](mailto:nocode@techfest.org).

Link to the submission google form: <https://forms.gle/7etreTisBaPehuk59>

**TEAM SPECIFICATIONS:**

- A team may consist of a maximum of 4 members.
- Students from different educational institutes can form a team.

**Eligibility:** All students with a valid identity card from their respective educational institutes are eligible to participate.

**PRIZE**

The prize money will be awarded to the top 3 winners via **NEFT** and will be processed within 30 working days after receiving the prize money from sponsors. Winners have to mail their bank details (immediately after the announcement of results) to [kunal@techfest.org](mailto:kunal@techfest.org), in the format specified below.

**Format of Mail**

Subject: NO CODE ML , Team Id, Position (example - NO CODE ML , NC211003, 1st Position)



**Body of mail**

1. Account Holder's Name
2. Account Number
3. Bank name and Branch name
4. IFSC Code
5. Photograph of Bank Passbook as a proof