

TECHFEST 2022-23

API GENERATE-A-THON

PROBLEM STATEMENT

Mr. Jonty Rhodes is the COO of a multinational company. He attended an AI-focused seminar to evaluate opportunities of AI in his own business. When he looked at his global operations, he noted that every business unit and geography has its own process and data variation. He realised that this leads to various issues in bringing applications together to perform a unified designed function built around sharing data, executing pre-defined processes uniformly, and gaining economies of scale. Streamlining all the processes and data sharing is a mammoth task.

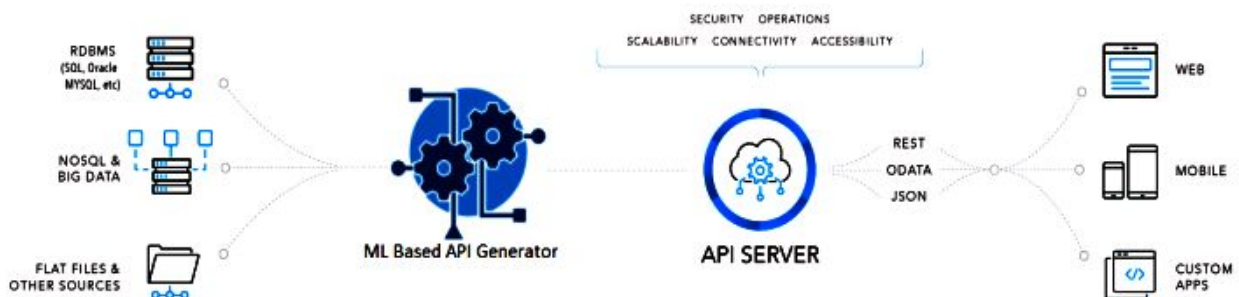
Mr. Rhodes wants to strategically leverage AI & cognitive technologies, which can work seamlessly across the range of databases, which might bring standardisation in data sharing, process and customer service.

As a first POC, he wants to start by analysing query logs on databases to extract useful analytics. Also, to expose these analytics via generic APIs (built around an ORM for persistence), enabling clients to access the metrics for interesting visualisation. When he discussed this idea with his CIO – he articulated the solution he wanted as a three-component build:

1. **Machine learning module** – It should train the model on query logs on databases and recommend useful API.
2. **API Generator Module** – It should be able to build an ORM and automatically expose a REST, GraphQL, or another kind of API for databases.
3. **Run-time component** – The solution should be a plug-and-play component for a platform to use as a library or an API.

Help Jonty in building a working solution.

SOLUTION DETAILS



STRUCTURE

There will be three stages to this competition – the solution design stage and the implementation stage -

Stage 1: Solution Design

1. Shortlist of ML algorithms to be tested with reasoning (participants can use Machine Learning, Deep Learning, NLP, NLU algorithms or combinations)
2. Front-end mock-ups – web interface to load training data, train the model, tune web interface to load training data, train the model, tune hyperparameters and perform quick testing
3. Solution flow/ documentation

Stage 2: Implementation (Teams shortlisted from stage 1)

Live demonstration of the working solution

Evaluation Criteria:

1. Implementing ML Algorithm – **20%**
2. Unified interface and API generation solution – **60%**
3. Number of databases and supported language for API generation– **20%**

Stage 3: Presentation and Working Demonstration

Working demonstration and presentation of the solution

Top 5 teams shortlisted from Stage 2 will present their solution at Techfest 2022-23.

TEAM SPECIFICATIONS AND ELIGIBILITY

- All students with a valid identity card from their respective educational institutions are eligible to participate in the competition
- One team can have a maximum of 4 members.
- Students from different institutes can be a part of the same team.
- Upto 3-year-old college pass-outs can participate in the contest.

REGISTRATION

The Participants have to register on the official Techfest Website and fill all the necessary details. www.techfest.org ->(Hover on) Competitions-> Competitions -> API Generate-a-thon -> Explore More -> Register -> Fill all your details -> Now you must create/Join a team

SUBMISSION

Participants are supposed to submit all the solution design components as a single PDF with the name: "TeamID_Solution Design.pdf" to apigenerator@techfest.org only

TIMELINE

Last date for registration	30th October, 2022
Last date for submission of Stage 1	30th October, 2022
Shortlisting of teams for Stage 2	7th November, 2022
Last date for submission of Stage 2	27th November, 2022
Shortlisting of teams for Stage 3	1st December, 2022
Final round - Stage 3 competition	16th-18th December, 2022

GENERAL RULES

1. Every team has to register online on the official Techfest website for the competition.
2. A Team ID will be allocated to the team on registration which shall be used for future references.
3. A team can register at any point of time before the deadline and submit final abstract
4. The decision of the organisers or judges shall be treated as final and binding on all.
5. No responsibility will be held by Techfest, IIT Bombay for any late, lost or misdirected entries.
6. The idea presented by the teams should be original (not protected by means of patent/copyright/technical publication by anyone).
7. Note that at any point of time the latest information will be that which is on the website. However, registered participants will be informed through mail about any changes.

INTELLECTUAL PROPERTY RIGHTS

The IP rights in the content(s) of the submitted entries and related prototypes shall be assigned to MMC without further consideration. MMC shall have the right to use it to whatever use it wishes.

Participants must send their signed Declaration Form to apigenerator@techfest.org to be considered eligible for participating in the MMC Innovate Competition. Otherwise, they will be disqualified.

The solution should not use any copyrighted component/ code.

GLOSSARY OF TERMS USED

1. **Hyper-parameter:** A value that is set before the ML Training process begins. These parameters are tunable and can directly affect how well the model trains. Some examples of hyperparameters in machine learning:
 - Learning Rate
 - Number of Epoch
 - Number of layers in a neural network
 - Regularisation constant
 - Number of branches in a decision tree
2. **Training Data:** The data set using which the algorithm is trained. In supervised learning problems, each training data point consists of an output variable and one or more input variables. Training data will be the query /activity log of the database. Participants can use various types of SQL DML and aggregate statements for generating query logs.
3. **Deep Learning:** Deep learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning.
Ref: <https://machinelearningmastery.com/what-is-deep-learning/>

FAQs

1. **What is the query log in the solution?**
A query log is a log of every SQL query received from a client in a database (SQL and NoSQL). It records every query the server receives, which can be used for analysis. If not already enabled in a database (MySQL, Postgres, SQL server, oracle, MongoDB etc.), this can be enabled and used for analysing the most frequently used SQL queries, categorised by user/client and in which all data tables are involved.
2. **What is the role of ML in the solution?**
The machine learning module analyses query/activity logs of different databases, presents data metrics for interesting visualisation and recommends generic rest API.
3. **What is the role of the web interface in the solution?**
A web interface is required to simplify the re-training and demonstration of the solution. It should allow the user to configure query log analysis while loading training data. It should have the functionality to call the API Generation module to generate and display the results.

4. What is the role of the API Generator Module in the solution?

This module will be responsible for generating, securing, and deploying a database-backed build of an ORM and automatically exposing a REST, GraphQL, or another kind of API based on ML module recommendation.

Below are some of the features you can implement in your API generation solution:

- Built-in Authentication and Authorization
- Automatic Swagger API specification generation.
- Row Filtering
- Column Filtering
- Pagination
- Sorting (ascending and descending)
- Access Control Lists or Role-Based Access Control (ACL/RBAC)

5. What is the Hyperparameter tuning?

Hyper-parameter tuning is required to fine-tune the training process to archive optimised results. The web portal should have a simple table displaying hyperparameters and allow users to change it for re-training.

Hyper-parameters are the only thing you can change while retraining with the query logs.

6. Front end Mock-ups

We expect a clear webpage outline, allowing users to train, test and demo the complete solution easily. Functionality is a critical evaluation criterion than look and feel.

NOTE

- Useful API can be crud/aggregate operation, but these API should be generated based on ML algorithm recommendation.
- Sample data with SQL queries will be provided to the final registrants.

CERTIFICATE POLICY

E-Certificates of participation will be given to the teams who will qualify for Stage 2. Qualifiers of Stage 3 will be provided with Certificate of Excellence.

PRIZE MONEY

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.

The Winners have to mail the following information (immediately after the announcement of the results) to **abhishek@techfest.org**

Subject: API Generate-a-thon, team id - your position (example API Generate-a-thon, AP12345678 – 1st position)

Body of mail-

1. Account Holder's Name
2. Account Number
3. Bank name and Branch name.
4. IFSC Code