# TECHNICAL CHALLENGE SOLUTION

Submitted by:

Aysegul Yalcinkaya-100841003

Basanta Dulal – 100851823

Hari Priya Ramasamy – 100901219

Sayeed Ahmed - 100853349

## WEB INTERFACE

Used Technologies:

* Python 3.8
* Flask
* MySQL

Demo site: <http://aysegulyalcinkaya.pythonanywhere.com/>

GitHub Repository: <https://github.com/aysegulYalcinkaya/MessageBoard>

Database Design

A picture containing text, font, line, screenshot

Description automatically generated

tb\_user: Table for registered users. Auto-generated id, email and encrypted password is saved in tb\_user table.

tb\_mail: Table for text messages. Auto-generated id, user id and message content is saved in tb\_mail table.

trigger\_table: This table is used for new messages. When a new message is inserted into tb\_mail table mysql trigger function writes the mail id into this table. Our forecasting model created in the second part of the project checks trigger\_table. If any record is found, this message is classified as spam or not. And the new\_mail id is deleted from the trigger table.

## ML MODEL

Data processing

* Import the required packages
* Loading the Dataset
* Remove the unwanted data columns (id, email)
* Preprocessing and Exploring the Dataset
* Build word cloud to see which message is spam and which is not.
* Remove the stop words and punctuations
* Convert the text data into vectors

Building spam classification model

* Split the data into train and test sets
* Use Sklearn built-in classifiers to build the models
* Train the data on the model
* Compare accuracy of models
* Save model and vectorizer to use for predictions

## REAL-TIME CLASSIFICATION

* Import the required packages
* Load model and vectorizer created in the previous step
* In infinite loop
  + Check trigger\_table if new mail is arrived
  + Predict if mail content is spam or not
  + Delete the record in trigger table
  + Display the result