

NLP coding round

The coding round is divided into 5 smaller tasks. The approaches taken in each of these tasks will be evaluated separately.

Task 1:

10 points

Use a learning-based approach for solving the task

Identify Question Type: Given a question, the aim is to identify the category it belongs to. The four categories to handle for this assignment are: Who, What, When, Affirmation(yes/no). Label any sentence that does not fall in any of the above four as "Unknown" type. You should come up with data structures to encapsulate these information as well as the code that populates the relevant data. You can use any machine learning technique and use the attached labelled data set for learning. The output should be driven keeping in mind the reply to the question.

Example:

1. What is your name? Type: What
2. When is the show happening? Type: When
3. Is there a cab available for airport? Type: Affirmation

There are ambiguous cases to handle as well like:

What time does the train leave (this looks like a what question but is actually a When type)

(Data Attached with the mail)

Expected output: When a question is given, the category (what, when, etc.) should be printed.

Task 2:

10 points

With respect to task 1, logically reason out why you have taken any particular approach, wherever you think necessary.

Example, explain why you have chosen a particular algorithm for vectorization, cleaning or classification.

Have comments in your code and explain the same.

Task 3:

10 points:

Document the code using tools like Doxygen. Please feel free to use any other tool you are already familiar with or are comfortable using.

Tools like Doxygen require you to adhere to a particular format. The documentation for the same could be found online.

Please note that this task is an overall documentation of the code as opposed to Task 2 that will test your approach to the problem.

Task 4: (BONUS ROUND)

15 points

This is an optional round. However, solving this challenge will give you an additional 15 points and an edge over others.

In this task you will be required to write a ROS wrapper around your code.

If you are not already familiar with ROS, you will have to go through ROS tutorial here:

<http://wiki.ros.org/ROS/Tutorials>

The task is as follows:

1. Create a node that accepts a sentence from the user (questions as in task 1) and continuously publishes the output, i.e the class (what, where, when etc.) the sentence is categorised into.

Expected outcome: The node should continuously publish the class the test question belongs to.

Task 5:

5 points

Push the code to your git repository. If you do not have a git account, we request you to make one and push your code into a repo.

Your final solution has to be sent to us in the form of the link to your git repo (sending the code separately through mail is not accepted).

In addition to just pushing your code, you will be required to have a readme.md file in the repo, with detailed instructions on libraries to be installed and steps to be taken for one to successfully run the code.