

Please find the details of the tasks below and create a Github repo per task then reply to this email with the links.

Work as hard as you can and deliver the results you got by the deadline.

Task 1 - Getting to Philosophy

Please write a Python script to check the "Getting to Philosophy" law.

https://en.wikipedia.org/wiki/Wikipedia:Getting_to_Philosophy

Clicking on the first link in the main body of a Wikipedia article and repeating the process for subsequent articles would usually lead to the article Philosophy.

The program should receive a Wikipedia link as an input, go to another normal link and repeat this process until either Philosophy page is reached, or we are in an article without any outgoing Wikilinks, or stuck in a loop.

A "normal link" is a link from the main page article, not in a box, is blue (red is for non-existing articles), not in parentheses, not italic and not a footnote. You don't have to check style tables or other fancy things, it is enough that the script works with the current Wikipedia style (for example you can use 'class' attribute in Wikipedia tags). For easy validation, please print all visited links to the standard output.

Use a 0.5 second timeout between queries to avoid heavy load on Wikipedia (sleep function from time module).

You can use <https://en.wikipedia.org/wiki/Special:Random> to check this hypothesis at home.

Task 2 - Binary Classification Problem

Given the training and validation datasets, <http://bit.ly/widebot-new-binc1f-data> , Create and train a machine learning model using the training set that performs well on the validation set. You should decide on the metrics of "performance" yourself, We will assess your decision.

It is up to you to use any of the following languages: [Python, Scala, Java, R]. We appreciate a small write up of the observations and your thoughts to follow your thought process.

- Bonus points if you can design and create a microservice that will serve predictions of that model.

We wish you enjoying the tasks.

Thanks for your time and consideration.