Genify soft dev | job application | take-home coding assignment

overview

In this assignment, you are asked to **build a REST API around a machine learning model that recommends banking products**. It is based on a solution to a Kaggle competition **[1]**: the model is already engineered for you (notebook **[3]**), you just need to run the training script to obtain the trained model, and then your work focuses on building from scratch an API that serves this model. Some bonus tasks to further test your API design and devops skills are also included.

goal

The goal of this assignment is for us to assess your API building skills, and the tasks in this assignment mirror the kind of work you would be doing at Genify as a Software Developer. Namely:

- inherit the work of a Data Scientist, with some but for sure incomplete explanations on the model built
- be tasked with building a system with little specifications, but instead some business needs for you to enquire/think about and translate into system requirements
- be given some messy references to earlier work
- have few opportunities to interact with the rest of the team given everyone is busy and the team is spread across time zones

context on the problem

The task this model solves is central to Genify's value proposition: helping banks recommend the right banking product (loan, deposit, credit card, etc.) to the right client at the right time. How these recommendations are delivered is equally important. Here, to do so, we opt for a REST API, and your task is to build it!

when to submit (a.k.a. deadline)

How many days from the receipt of this assignment should you send your submission? **If a deadline was not communicated to you**, this is left for you to estimate. Please communicate your timeline upfront to us (this is

your first task, in fact). The time you take to complete this assignment, and whether you meet the timeline communicated initially, will both be accounted for in your evaluation–so estimate the time wisely!

what to submit

- a link to a public repository on GitHub that contains all the code you wrote and the original source code from **[3]**, integrated together
- a video that demos the work you did, hosted on a service of your choice (OneDrive, Google Drive, YouTube, …)
- a 1-page document that shows 3 example query-response pairs
- an email to `alex@genfy.ai` or whoever else gave you this assignment, containing all references to the items you submit, with subject *Genify Soft. Dev. application | code submission | first name last name*

what tools to use

- Python 3 only
- Flask to build the API (exclusively)
  - if you do not know Flask yet, this will help us evaluate your ability to learn new tools on the job; if, please flag it in your submission
- whatever the source code **[3]** needs to run
- any other Python package of your choice

instructions

mandatory tasks

- read this assignement in full, estimate the time it will take you, and communicate this estimate to us
- download the source code **[3]** and data **[2]**
- install dependencies
- train the recommendation model and save the model
- **write a web app that serves the model via an API endpoint**
  - the API should be queried via a POST HTTP request (URL) and return a JSON file
  - to keep things reasonable, restrict the parameters sent to the API to the ones listed on Genify's online reco. sys. demo **[5]** (modulo what you deem sensible), and leave the other ones N/A or static / dummy when you input them into the model at inference time
  - you can get inspiration from the existing Genify API **[6]** which serves another type of model (categorizer)
  - when building this minimal API, you do NOT need to enable any form of user authentication / authorization
- run the web app (likely `python run.py`)

- **record your demo video**
  - walk through the high-level structure of your code
  - show a few example API queries and responses
  - 120 seconds max
  - show your screen only (not your face)
  - to grab it, you can use Windows 10's Game Bar or a similar app on other OSs
- record 3 example query-response pairs in a document
  - for the query, paste the URL
  - for the response, paste the JSON raw content

bonus task 1

optional

- dockerize your code and include relevant additional files in the repo
- let us know in your submission that you completed this, so we test this aspect too (we will build and run the Docker image)

bonus task 2

optional

- answer this question:
  how would you expand this API web app you built so it serves multiple models, e.g., one for each bank we sell this solution to?
- exclusively use a diagram to explain your system expansion

who to contact when in doubt

`alex@genify.ai` or whoever else gave you this assignment. Note, you won't receive help from us. Instead we might give you some clarifications, e.g., whether you are allowed to do this, or that.

references

- [1] overview of problem (Kaggle competition): https://www.kaggle.com/c/santander-product-recommendation/overview
- [2] data (Kaggle competition): https://www.kaggle.com/c/santander-product-recommendation/data
- [3] reference Kaggle notebook: https://www.kaggle.com/sudalairajkumar/when-less-is-more
- [4] details on ref. code: https://www.kaggle.com/c/santander-product-recommendation/discussion/25579
- [5] Genify recommender system demo: https://pfmv0.genify.ai/en/recosysdemo
- [6] Genify PFM API: https://docs.pfm.genify.ai/pfm-suite/v1/txn-data-api