**PROPOSAL**

Glad to invite you to the interview session – Prototype Proposal.

In this cycle, you need to design a solution, prototype it and demonstrate it to your customer. During the demonstration, there will be two interviewers who will play the roles of the customer’s decision makers.

Please take some time to familiarize yourself with the detailed guidelines provided below.

**BRIEF**

Imagine that you meet with a company, we could call it ABC, is planning a product. You can choose any industry that you are particularly familiar with as this customer’s business.

You and your peers, as the Account Team, visited the customer’s office for the first time. During this visit, you learned the core knowledge of the product and identified the customer's needs.

The core feature and process of the product are as below:

* The core feature is “global urban air quality prediction”.
* For US cities, use 2 publicly available scientific datasets as source: 1) National Oceanic and Atmospheric Administration: Global Surface Summary of the Day, 2) OpenAQ. You can get detailed information from “Concepts and knowledge” section.
* Use multiple machine learning (ML) models to prediction. You can get detailed information from “Concepts and knowledge” section.
* The output of the model is "AQI Forecast for the next 1 days, 24-hour scale". For US Cities, follow U.S.AQI definition set by U.S. Environmental Protection Agency (Hereafter referred to as EPA). You can get detailed information from “Concepts and knowledge” section.
* There are two types of end users: enterprise and individual.
* Data scientists are responsible for the datasets and models.
* The R&D team is responsible for the front and back ends.
* The content for individual End-User is a picture that reflects the characteristics of the city, the theme of the day, AQI information and health information.

At present, the product is in its infancy and the team responsible for it is small. Team members want to focus on the business, not spend a lot of time and energy maintaining the infrastructure. Their next plan, related to information technology, are as below:

* Scaling to meet continuous increase of data.
* A web-based, integrated development environment (IDE) where you can perform all ML development steps.
* Use AutoML to evaluate and compare the models.
* Use GenAI to generate the picture for individual End-User.

Concepts and knowledge:

* National Oceanic and Atmospheric Administration (NOAA): Global Surface Summary of the Day

Global Surface Summary of the Day is derived from The Integrated Surface Hourly (ISH) dataset. The ISH dataset includes global data obtained from the USAF Climatology Center, located in the Federal Climate Complex with NCDC.

Dataset URL: https://registry.opendata.aws/noaa-gsod/

License: Open Data. There are no restrictions on the use of this data

Documentation: http://www.ncdc.noaa.gov/

* OpenAQ

Global, aggregated physical air quality data from public data sources provided by government, research-grade and other sources. These awesome groups do the hard work of measuring these data and publicly sharing them, and our community makes them more universally accessible to both humans and machines.

Dataset URL: https://registry.opendata.aws/openaq/

License: CC BY 4.0

Documentation: https://openaq.org

* The relationship between weather and air quality

The information comes directly from the United States Environmental Protection Agency (EPA):

*The weather can have a significant impact on air quality since different aspects of the weather affect the amounts of ozone and particulates that are present in a specific area. Sunshine, rain, higher temperatures, wind speed, air turbulence, and mixing depths all affect pollutant concentrations.*

1. *Sunshine can cause some pollutants to undergo chemical reactions, resulting in the development of smog*
2. *Higher air temperatures can speed up chemical reactions in the air*
3. *Rain typically results in less pollution since it washes away particulate matter and can also wash out pollutants that are dissolvable*
4. *Wind speed, air turbulence, and mixing depth all affect how pollutants disperse, or spread out from an area*

The weather dataset, Global Surface Summary of the Day from NOAA, contains a number of the above-mentioned attributes. Data scientists need to merge the weather data with air quality data from OpenAQ for a given location.

* U.S.AQI

Millions of people live in areas where air pollution can cause serious health problems. Local air quality can affect our daily lives. Like the weather, it can change from day to day. EPA developed the Air Quality Index, or AQI, to make information available about the health effects of the five most common air pollutants, and how to avoid those effects.

URL: https://www.airnow.gov/aqi/aqi-basics/

How is the AQI calculated?

URL: https://document.airnow.gov/technical-assistance-document-for-the-reporting-of-daily-air-quailty.pdf

* AutoGluon

A Python library for AutoML, you can use it to build machine learning (ML) models. Note: including this library, you can choose any method to build and evaluate your models.

URL: https://auto.gluon.ai/

During the visit, both you and the team leader agreed to prepare a prototype within a week, and you will demonstrate it to the data scientists and R&D team. The team leader emphasized that you would only have one opportunity, as they must decide the day after your demonstration.

**DELIVERABLES**

The prototype should include your code, your architecture/environment on AWS, and a design document about the code and architecture/environment. Once completed, please send the document to the NWCD Recruiter.

As a whole, recommend a usable, technologically advanced, competitive architecture that enables organic business growth. The architecture should specifically address the requirements and concerns outlined above.

A 1.5-hour demonstration is required to deliver your AWS-based solution to the team leaders. As the NWCD Rapid Prototyping Solutions Architect, you will have only one technical meeting with the customer’s team, as they must make a decision the day after your demonstration. Therefore, do not expect any follow-up calls or second meetings to clarify points you may have missed.

**CODE AND ENVIRONMENT Guidelines**

Provide your code via a public developer platform, such as Github.com.

Build up your environment on your laptop, run your code and show the result. Also, you can deploy your environment on AWS, but that incurs costs that NWCD can't reimburse. Other platforms are not accepted.

Use open-source library, commercial kits are not accepted.

**DOCUMENT Guidelines**

A WORD document (not PPT version) should be no longer than 6 or 12 pages.

Clearly and succinctly present an analysis of the requirements for why and how to meet the demand based on your understanding.

Provide a proposed architecture diagram using AWS product icons. Include a detailed description of your architecture and explain why you chose this solution.

Clearly state all assumptions and references made during the design, and explicitly identify the Amazon Web Services that were referenced.

**DEMONSTRATION Guidelines**

Imagine that you are already a member of the NWCD Solution Architect team.

First, summarize the customer’s status and analyze their needs.

Second, get your solution in place and familiarize yourself with every technical point in it, including code, environment and others that best suit your needs.

Third, deliver the demonstration at NWCD’s office to several NWCD employees, or conduct it via video. Scheduling will be coordinated with NWCD recruitment as needed.

The requirements for the demonstration are vague. NWCD wants to see how you would approach this customer with the goal of utilizing the AWS platform to address their business and technical needs (no template will be provided).

At a minimum, we expect your demonstration to include:

Architecture diagrams of the customer’s existing architecture and the proposed AWS architecture (please use AWS Product icons).

Clearly introduce the architecture, providing technical details to the best of your ability. Focus on how the solution can address their challenges.

Clarify all technology-related points. There will be no follow-up meetings or further discussions; this is it. Your goal is to ensure that customer executives are aligned and on the same page by the end of this meeting.