Technical Assignment (Interview Round 2)

The following assignments are designed to evaluate real-world problem-solving capabilities, architectural thinking, and ability to deliver production-grade **GenAI**, **Computer Vision**, and **Machine Learning** solutions. Candidates are expected to submit a working prototype, along with a short document explaining the approach taken, challenges faced, and improvements planned if given more time

Please find the files related to the assignment in this drive location:
Senior AI engineer Assignment

Assignment 1: Resume Standardization – GenAI Use Case

Problem Statement

Build a **Generative AI-based application** that takes resumes in **any format (PDF or DOCX)** and converts them into the given **standardized resume format** provided by us.

The application should:

- Accept resumes in .pdf /.docx formats.
- Extract all relevant information such as name, experience, education, skills, etc.
- Generate a new resume document strictly following the **given standard template**.
- Output the final resume in **editable .docx format.**

Key Evaluation Criteria

- ✅ Accuracy and completeness of extracted resume information.
- ✅ How well the **final resume format** matches the **provided standard format**.
- ✅ Efficiency and optimization of **LLM API calls** (e.g., batching, context reuse, cost control).
- ✅ Scalability and modularity of the overall architecture.

Bonus Points

- 🌟 Ability to **inject external instructions or tone guidelines** (e.g., more formal vs. casual phrasing) into the output resume.
- 🌟 Good working UI

Assignment 2: Ice Coverage Checking – Computer Vision Use Case

Problem Statement

Analyze a **5-minute surveillance video** capturing factory workers loading crates of shrimps into tall utility carts for shipping. The goal is to develop a model that detects and flags crates with insufficient ice on the top layer, If the surface is not covered with ICE properly in long cold storage, the quality gets degraded resulting into export rejection.

Your task is to:

- Make a model to detect each **crate** as it's being loaded.
- **Check if the top layer** of the shrimp is **sufficiently covered in ice**.
- **Set a threshold** for ice coverage to decide if a crate is ready for proper storage.
- If ICE is not covering the top layer fully then, **flag the crate** / **trigger a signal/alert** for manual inspection.

Note: The scope of the analysis is restricted to the **top surface layer** only – not the inner layers of the crates.

Key Evaluation Criteria

- ✅ Accuracy of **crate count and ice detection on each crate**.
- ✅ Effectiveness of the **threshold-based classification** for flagging.
- ✅ Stability across different frames and lighting conditions.
- ✅ Technical depth of the **computer vision or deep learning model** used.
- ✅ Real-time or near-real-time applicability of the system.

Bonus Points

- 🌟 Implementation of an additional module to detect **worker safety compliance**, i.e., whether all workers are wearing:
    - Safety coat
    - Gloves
    - Protective hat
- 🌟 Multi-object detection & visual annotation output for debugging.

Assignment 3: Sales Forecasting – Traditional ML Use Case

Problem Statement

You are provided with **sample sales data** that captures **customer sales across different plants**. Your objective is to **predict both the amount ($) and quantity of sales** (LBS) for the **upcoming 12 months**, at both:

- **Customer level**, and
- **Plant level**

This task involves designing a robust forecasting pipeline using **classical or modern ML techniques**, grounded in **solid exploratory data analysis (EDA)** and **business logic**.

Note: Due to the nature of the business, you might face the fact that enough information is not available to train a perfect model. That is the challenge.

Key Expectations

- A thoughtful **EDA** that highlights seasonality, trends, anomalies, customer/plant behaviors, and other influential factors.
- Clear **feature engineering** and **group-based modeling** strategies (e.g., by plant, customer segment, or region).
- Selection of appropriate forecasting or regression models with **comparative evaluation**.
- Metrics such as **MAPE, RMSE, or custom business-defined accuracy metrics**.

- Forecasting outputs for each of the next 12 months, for each **(Plant, Customer)** pair.

Key Evaluation Criteria

- Depth and quality of the **EDA**.
- Accuracy of predictions at both plant level and individual customer level at both sales amount ($) and quantity (lbs.)
- Expecting to see at least 80% accuracy for the plant level predictions
- Soundness of the **modeling strategy** – grouping, handling seasonality, segmentation, etc.
- Justification of **model choices** and **validation strategy**.
- Quality of forecasts, with visualizations and tabular summaries.
- Clean, maintainable, and reproducible code.

Disqualifying Approaches

- Using **naïve time series models** without understanding the domain or validating assumptions.
- Skipping EDA and jumping straight to model training without strategy.

Bonus Points

- 🌟 Identification of new, churned, and irregular ordering customers based solely on order history.
- 🌟 architecture for predicting demand for these customer types using patterns derived from past ordering behavior.

Submission Guidelines

- Candidates are expected to **solve any one of the above assignments in full detail**.
- For the remaining two assignments, candidates can optionally submit a **report/document describing the architecture, approach, or ideas** they would consider solving them.
- The final submission can be shared either:
    - As a **ZIP file**, or
    - As a **GitHub repository** with a **neatly organized project folder structure**.

For the Solved Assignment:

- ✅ Must include a **detailed explanation** of the solution — this can be in the form of a **Word document, text file, or well-commented code**.
- ✅ If the code does not have inline comments, a **separate explanation document is mandatory**.
- ✅ The solution should be production-ready in structure, with modular code, clear naming, and separation of logic.

For Optional Architecture Reports:

- ➕ Provide a brief document outlining your **thinking, design decisions, and high-level steps** you would take to solve the remaining two assignments

ALL THE BEST!!