# **Document Classifier**

# Problem Owner Details

|  |  |  |
| --- | --- | --- |
| Role | Expectations (do not edit) | Mail IDs |
| Problem Owner | Who has identified and elaborated the problem | Joe George |
| Account, if any | If it is relevant to a specific account, please mention | NA |
| Mentors | Technical team who will be able to provide clarifications on the problem, define the score and guide the hackathon participants on the solution approach | <TBD> |

# Problem Description

# Business Context (Describe the problem and its impact)

# A Compliance Auditor needs to scan tons of physical documents into digital documents every day – ranging from purchase orders, invoices, goods receipt/delivered memos, invoice payment cheques. These documents need to be associated with each other before they are audited. For example – Once purchase order is generated, many goods receipt/delivered memos will be generated with each associated delivery of the ordered item on the purchase order. Once item is received per the Purchase order a consolidated invoice for the said amount is generated, payment checks are generated against the purchase order

# Solution Features (Must Have, Good to Have, Optional)

The challenge that needs to be resolved is : How to associate digital documents via common denominator – in this the Purchase order number.How to validate the PO values with Invoice Value and Pay check values, Similarly, PO item quantity with Invoice Qty, Goods receipt Qty / Delivered memo Qty.

# Software Required, if any (Mention suggested technologies to be used)

Potential solutions – OCR, Computer Vision, ML

# Hardware Required, if any (will need to be funded by the account)

NA

# How to test (How can the solution be tested and what will be expected in the demo)

Demo

# Input datasets (Problem Owner needs to share the datasets relevant to the problem)



# Evaluation Criteria (Add other solution specific parameters that will be measured to judge the solution)

|  |  |  |  |
| --- | --- | --- | --- |
| Creativity | How creative or innovative is the idea within the given challenge? Has this been done before, or is this something completely new and original idea (architecture, design and implementation)? |  |  |
| Innovative Technology & Quality | How you have utilized the existing technologies to their solution and quality of code? |  |  |
| User experience and functionality | Is the overall user experience intuitive? Does the flow make sense? |  |  |
| Simplicity | How elegantly does it solve the problem (smart vs harder)? |  |  |
| Feasibility/Impact/Reuse/Extensibility/ Modularity | Does your solution work? Can it be implemented at scale? Can it be extended to other similar problem statements? |  |  |
| Progress and Execution | How much accomplish in a day? |  |  |
| Demo & Presentation | How well did the team present? |  |  |