Hands on Exercise

Background

RightShip conducts vessel inspections for its customers (vessel owners) and comes up with a post inspection report detailing what deficiencies were found in the vessel. The report is shared with the vessel manager and the manager provides comments around each deficiency - what was the root cause of the deficiency, what preventive action was taken and what corrective action was taken.

RightShip SMEs manually reads through the inspection report and classifies Each deficiency as High, medium and low based on the risk severity. Risk severity denotes how severe a deficiency is and is an essential indicator towards determining the safety of a vessel.

Definition of each severity level has been given below.

High Risk Finding is a significant finding that poses a threat to personnel, the ship, or the environment and has the potential to cause economic and reputational harm.

Medium risk Finding are those that could be viewed as weaknesses in the organization's processes, procedures, or shipboard practices.

Low risk Finding: The findings are neither high nor medium risk; they are unlikely to result in an accident and, if they do, are likely to cause only minor damage.

Task

You are an AI engineer at RightShip.

You have been provided with a **Sample Inspection Report** which is in a PDF format. It contains 5 deficiencies. An SME at RightShip has given a risk severity label to each deficiency - High, Medium and Low provided in **Risk Severity** excel sheet.

Your task is to create a python script which ingests a **New Inspection Report**, extracts the deficiency details from it and labels each deficiency found in report as High, Medium and Low (which was earlier done manually by the SME).

Use the risk definitions and already labelled deficiencies to label new inspection report.

Feel free to use any LLM model of your choice.

Output

Share your code and any additional documents if needed.