Block Diagram Short Description

Block A

User uploads various Reports related to Machinery and Maintainance. These reports are uploaded fortnightly. Reports might be in different format and sometimes OCR can also be required for scanning and extracting the text. There are around 5-6 types of reports and each one of them will have separate collection.

Block B

When new ships needs to be onboarded on aranti, we need few details of ships and we call it ship configurations. Various Ship configurations:

1. List of variables ship will be uploading
2. Limits(High and Low) of these variables
3. Static Data of Ship like (Length, Mfg Year etc) around 30

Also, historical data is uploaded for last ~2 years in the form on Excel/CSV files. This data will be extracted and stored in daily_data collection. To distinguish between historical data vs current daily data, flag `historical` will be set. Historical data can come into two files viz. Fuel and Engine data.

Block C

Users upload daily data (Engine and Fuel), which is parsed from source files such as xls/csv and stored in collection daily_data. Daily data can come up to 6 times per day. After daily data is received, subsequently block D is executed which processes all the data.

Block D

Block D is executed immediately after receiving daily data from block C. Block D encapsulates the entire ML pipeline, insights generator, data pre-processing, and predictions from stored models from Block E. This block also fetches API data from external services for weather and vessel position. Everything is stored in main_db.

Block E

Models are trained on scheduled basis like weekly/daily. After training, models are stored in the static storage like S3 in the pickle form which will fetched when needed from BLock D and G.

Block F

Will simply extract reports from the database.

Block G

On the basis of User Selection, on the fly prediction values will be calculated. Stored models will be used to predict

Block H

Extract Data from main collection and pass it to front end.