Technical Consultancy - Digitisation of Data - Machine Learning

CASE STUDY

IMRANDD
INTELLIGENCE WITH INTEGRITY

CLIENT	GLOBAL OPERATOR
SECTOR	UPSTREAM OIL & GAS, UKCS
ASSET	PLATFORM - TOPSIDES
DATE	2021-22
DURATION	2-3 MONTHS
OBJECTIVE	DIGITISATION & VERIFICATION OF LEGACY FILES & IMAGE CLASSIFICATION
SERVICES	TECHNICAL CONSULTANCY, DATA EXTRACTION

PROJECT OVERVIEW

Imrandd's expert team of consultants utilised EXTRACT and a newly-developed machine learning model to extract and classify legacy files from two North Sea assets.

THE CHALLENGE

Imrandd was approached by its client, a major operator in the North Sea, to support the implementation of their new IDMS (Integrity Data Management System) by digitising and classifying large quantities of inspection data. Before fully migrating over to the new IDMS, the operator wanted to ensure that all systems on the two assets in question were covered. To do this, they required a solution that would enable them to easily sift through the legacy information and cross-check against their records to confirm that they had captured all available data. The client had eight years of legacy data stored in various locations and in multiple formats. They had requested that this scope also included classifying isometric drawings within Excel files, from a mix of images including P&IDs, isometrics (some drawn within Excel, some drawn elsewhere and embedded) photographs, and detailed drawings. The operator was faced with a major investment in time to bring the records into a standardised and quality-checked condition.

THE SOLUTION – EXTRACT

At Imrandd, our approach begins and ends with data. EXTRACT is a tool developed in-house at Imrandd to digitise unstructured, decentralised silos of legacy inspection data. It was conceived to extract specific file types from datasets and group them to enable easier upload into IDMS or analytics systems like EXACT.

Using EXTRACT as a starting point, the Imrandd team investigated the requirements of the operator further and developed additional scripts that would expand the tool's capability to include the other file types that the client required. To overcome the variations in formatting and organisation within the Excel files, the new script included expanded keyword searching for a variety of columns and key-value pairs. The data captured from each file was evaluated to determine the completeness and correctness of the data which was then used to inform and streamline the QC process. The extracted data was standardised and normalised to the client's specification. Data from the client's equipment register and Maximo system was used to identify gaps within the inspection history, or within the client's systems.

A SMARTER CLASSIFICATION PROCESS

To tackle the challenge of the embedded isometrics, our data scientists used their expertise to take the solution further, developing a machine learning model to provide image classification for technical images. Beginning with an initial training set of 200 labelled drawings, the team trained a convolutional neural network to predict the drawing type of input image. Batches of 200 drawings were fed into the model, and after examining the results and reclassifying incorrect predictions, the model was retrained with the expanded labelled set. This approach allowed the team to complete the QC of the image classification while continuously improving the model's accuracy. Each iteration of the process reduced the amount of reclassifying required on new image sets and within days the model was accurate enough to significantly reduce the amount of manual effort required to classify the images.

BENEFITS

- Project completed ahead of time and under budget
- The data from the two assets was digitised, standardised and ready to input to their IDMS
- Analytics and future efficiencies – with the data extracted and digitised from the legacy files, the operator can explore further interpretation and analytics to support future decision-making
- A well-trained image classifier ready for use on future projects

CONCLUSION

Imrandd delivered the digitised, classified datasets in the format required for the IDMS, ahead of time and under budget. Using this data, the asset team was able to verify the completeness of the other data that was being fed into the IDMS and ensure that none of their systems were missed. The success of this extraction and classification scope has opened the potential to apply advanced analytics to the datasets, using a product like EXACT. This has the potential to unlock actionable insights, with the potential to significantly reduce OPEX and improve asset integrity management and plant reliability in the future.