

Rule-free Document Classification for Scanning Service Provider

THE CLIENT

Our renowned client has been offering business process outsourcing services to healthcare providers, the public sector and enterprise customers for over 50 years.

The client specializes in digitizing, indexing, and archiving documents and processes a volume of one billion pages annually.

DOCUMENT CLASSIFICATION KPIS

> 90 %

out of the box **automation rate**

> 80 %

cost savings by reducing manual correction efforts

THE CHALLENGE

Before implementing IDA, the client's daily processes involved scanning and reading incoming boxes of documents. Initially, a market-leading Intelligent Document Processing (IDP) product was used for **rule-based document classification**.

However, our client encountered significant difficulties in automating their document scanning and archiving processes. They struggled with an **automation rate of only 50% for document classification**. As a result, a substantial amount of manual work was required to correct and validate documents that fell below the necessary confidence thresholds.

Additionally, the client's existing rule-based document classification approach was resource-intensive and expensive to maintain, especially when it came to modifying existing document classes or creating new ones. The **introduction of regulatory changes** further emphasized the need for a more precise categorization system that their previous solution couldn't effectively deliver.

THE SOLUTION

During the evaluation phase, IDA, the new system being implemented, showcased a remarkable performance with an **automation rate of 90% for document classification out of the box**. This achievement can be attributed to IDA's state-of-the-art OCR engine providing outstanding accuracy and confidence scores.



Unlike the previous solution, IDA follows a **rule-free, few-shot learning approach**. This allowed for a swift initial setup in less than three days instead of the lengthy 240-day setup time with the previous solution. Regulatory changes necessitated a significant expansion in document classes, growing from approximately 150 to over 300. IDA's ability to quickly adapt to new document classes enabled the client to meet the requirements promptly.



The impact of implementing IDA was substantial, resulting in an **80% reduction in manual efforts** due to increased straight-through processing – even with the client's objective of minimizing false positive results to less than 1%.



Overall, the substantial improvement in automation empowers our client to significantly **scale their business**.

OUTLOOK

Scanning and processing large batches of documents can become incredibly challenging when dealing with PDFs that contain 100 or more consecutive forms. Automating the classification of these documents based on their categories seems nearly impossible. The process of separating the documents currently requires a significant amount of manual effort, which usually involves inserting separator sheets between each document.

However, in the future, the client plans to address this issue by leveraging **IDA's document splitting feature**. This will enable them to train a neural network to automatically detect the start and end of multi-page documents or split documents rule-based after a fixed number of pages.

By making use of IDA's **data extraction capabilities**, the client will have the opportunity to broaden their service portfolio and explore further expansion opportunities.