CASE STUDY:
Automate Invoice Processing

NoviCap

Novicap Integrates Spot Intelligent's PDF Extraction Solution

> "Automatically extracting key data from invoices looked like an unsolvable problem until Spot Intelligence came along"
> — Nicolas, CTO Novicap

Executive Summary

Spot Intelligent enabled Novicap to automatically extract key information from invoices. This was previously being done manually.

Their own proprietary data was used to build a bespoke artificial intelligence system. An overall accuracy of 93.9% was achieved with 6.1% of the results being flagged for manual review and fed back into the self-learning system.

About Novicap

Novicap is a fin-tech company that allows businesses to receive early payments for their outstanding invoices. They have automated everything from risk to credit limits and providing financing.

The only exception in their process was reading the original invoices. Somebody still needed to extract this information.

Problem

Manually extracting information wasn't a problem when they were smaller but as they quickly grew, so did the problem.

How could they still provide one-click financing when information needs to be manually extracted from invoices?

"We reduced the workload from 4 hours per day to 5 minutes per day to process the flagged invoices.
-Nicolas, CTO Novicap

Alternative Solutions

1. Outsourcing: Provides an accurate solution but is rather slow. It doesn't provide a one-click solution.

2. Software Solutions: General extraction software has low accuracy or requires the user to highlight the relevant areas in the PDF.

Provided Solution

Spot Intelligent utilized Novicap's labelled dataset of 9 000 invoices to train a deep learning model to accurately extract company names, identification numbers, dates and invoice amounts.

PDF documents are complex to parse therefore state-of-the-art natural language processing techniques are combined with the location, size, and font of the text in order to produce a deep and human-like understanding of the document. The document is then analysed and the relevant fields are extracted.

Results

The self-learning algorithm has a high accuracy; over 93.9% of all invoices are being analyzed correctly. The remaining 6.1% are partially parsed or have a low confidence score and are therefore flagged for review. This greatly decreases the work required and allows them to provide their customers with a one-click financing solution.

The manually reviewed invoices are passed back to the system allowing the system to continuously learn from its past mistakes and adapt to new trends in the underlying data.

A sample of an invoice automatically parsed by Spot Intelligence.