Infrrd

Infrrd Achieves Unbeatable Data Extraction Accuracy **For A Travel & Expense Management Company**

Use Case

"The simple task of identifying specific fields and extracting them manually was not only time consuming, but also required a lot of extra hands. Infrrd solved that issue for us with greater accuracy."
– Company's Chief Product Officer

- **Business**

  The company offers expense management system that automates the process of receipt to accounting.

- **Problems**

  • 3 fields from a receipt i.e. billing amount, billing date and currency needed to be extracted with near 100% accuracy

  • Manual entry of this data would not only consume a lot of time but also needed an excessive investment in workforce

- **Result**

  Infrrd's Machine Learning OCR was used to automatically extract the fields required by the client, with unbeatable accuracy; dramatically cutting down the errors, time and manual effort required to do the job.

Client Background

The client is a business travel and expense management company that aims to ease problematic and frustrating expense reporting by digitizing receipts and storing them on the cloud. The company has also extended its use with a new platform for booking travel and monitoring travel expenses that allow users to book all travel plans (such as airfare, car rentals, and hotel) from one space.

Challenges

• A general way of OCR measurement is in terms of Field- level accuracy, this means that if a receipt has 10 fields to be extracted and 10 such receipts need to be processed- it would amount to 100 fields to be extracted. And if 90 fields are extracted accurately, it is considered as a 90% accuracy. But the client proposed a completely different way of accuracy measurement. They had 4 important fields (date, currency, amount, merchant) in every receipt for which the accuracy was a must. Even if 1 of the 4 fields in a receipt was inaccurate, the accuracy for the entire receipt was considered 0. This made achieving high accuracy numbers, a challenge.

• Traditional OCRs can only extract words or characters that match the templates & a receipt may have more thanone matching field information. Infrrd's machine learning OCR had to be trained to understand the right field that needed to be extracted from a document. Example: It had to extract a string of characters most relevant to billing date out of the multiple date mentions in a receipt, along with a confidence score.

• The receipts were of varying formats, sizes, and quality.

• Performance of the OCR needed to be increasingly fine-tuned.

Infrrd

How Infrrd Helped?

• Infrrd's machine learning OCR extracted data from multiple receipts which were fed by the client. Since our solution isn't template-based it adapted to the data extracted earlier and learned from it with time, to provide high accuracies.

• With our solution, the company got the required information such as - date, time, currency time, bill amount, etc extracted directly to their system.

• A cross-referencing procedure was conducted post the extraction. Issues faced during this procedure were classified under different categories and changes were made in the algorithm to fix these outputs. This ensured that the future extracted results are more accurate and at the same time more automated than before.

Process of our platform

The client uploads documents which get analyzed for the list of desired entities.

Extraction of raw text is done using OCR.
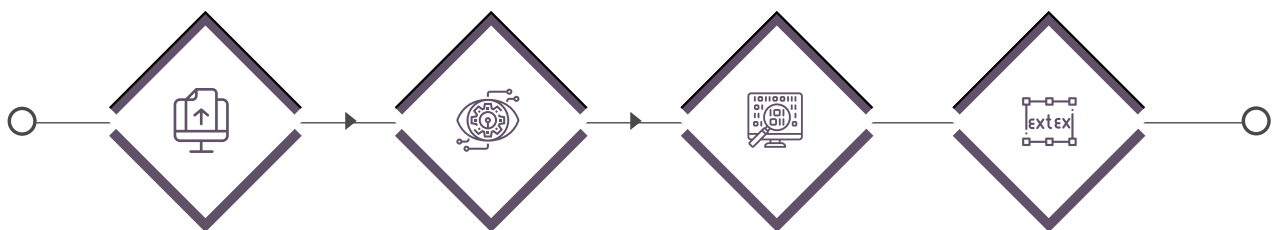
Visual entity detection uses deep learning model to analyze the list of desired entities. Here an algorithm takes a string of text as input and identifies the relevant nouns such as billing amount, date, currency etc., mentioned in that particular string.
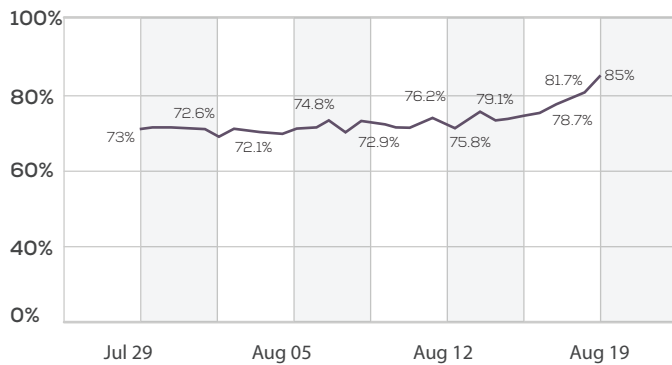
The entities and the raw text are matched to get the desired output for specific fields.
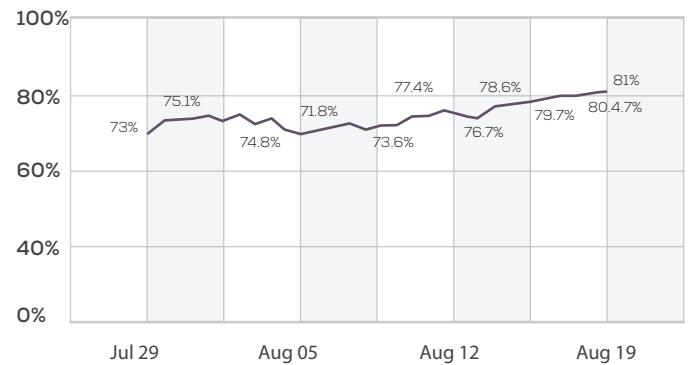
Extracted Fields

• Billing amount, date, currency (on receipt)
• Locale (on receipt)
• Vat percentage and number (on receipt)
• Merchant ID (on receipt)
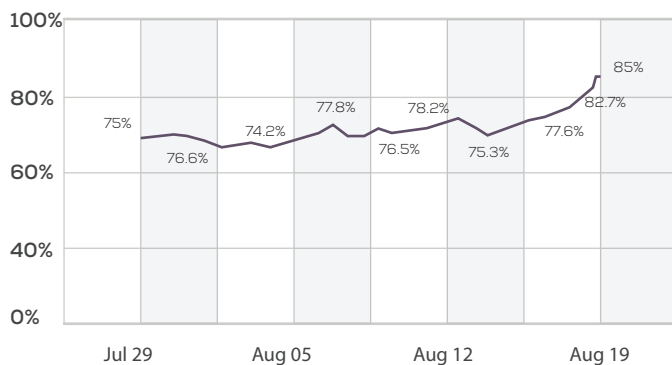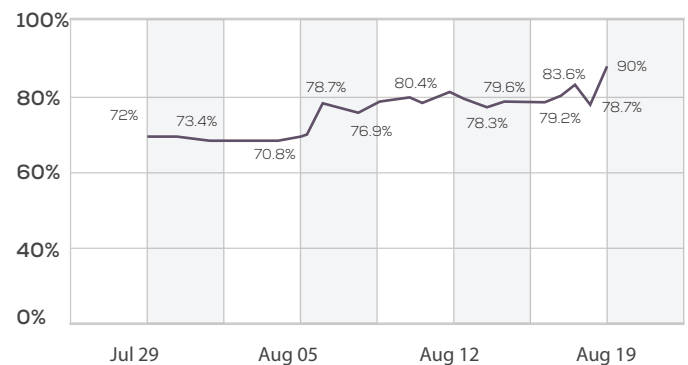• Category and category ID (on receipt)
• Extraction score

Infrrd

Learn more at www.infrrd.ai

Results

Currency Accuracy

73%	72.6%	74.8%	76.2% 79.1% 81.7% 85%
	72.1%	72.9%	75.8% 78.7%

Date Accuracy

75.1% 77.4% 78.6% 81%
73% 71.8% 79.7%
74.8% 73.6% 76.7% 80.4.7%

Amount Accuracy

75% 77.8% 78.2% 85%
74.2% 82.7%
76.6% 76.5% 75.3% 77.6%

Merchant Accuracy

72% 78.7% 80.4% 79.6% 83.6% 90%
73.4% 79.2% 78.7%
70.8% 76.9% 78.3%

In a matter of two months, we saw massive improvements in confidence scores.

• Currency Accuracy- improved from 73% to 85%

• Date Accuracy- improved from 73% to 81%

• Amount Accuracy- improved from 75% to 85%

• Merchant Accuracy- improved from 72% to 90%

>> Click here for Free Demo <<

About Infrrd

San Jose, CA, based Infrrd Inc is a leading Machine Intelligence partner to Banking, Financial services and Insurance industries across the globe. Infrrd's focus is on providing AI as a service and leveraging its homegrown machine learning platform to solve analytics and automation related problems for the customers. Infrrd's platforms and algorithms extract deep insights from big data based on artificial intelligence and deep learning and offer these insights to drive decisions & automate extraction for customers.

+1(844)-446-3773

San Francisco, Suite 360E, 2001 Gateway Place, San Jose, CA 95110