

Next-level accuracy: The impact of AI on OCR in auto loan underwriting



JESSICA GONZALEZ

InformedIQ

Optical Character Recognition or OCR is a technology enabling computers to “read” text off documents. It is usually priced per page and is often riddled with mistakes. The poor accuracy makes it of little practical value for organizations looking for a clear picture of the data in the documents. To get the true picture, you often need a team of people.

This team, also known as human in the loop (HITL) augments the accuracy and completeness of the result. And those humans receive a salary, so costs increase.

Lenders are notorious for being late with technology

Often the goal of technology is to free up staff, but with antiquated systems and backlog of IT requests, it can often seem like a daunting task to adopt new technology. However, the pandemic and years since has created an urgent push to look beyond simple contactless equipment to a comprehensive digital solution. This reduces reliance on paper-based systems and improves functionality. Furthermore, many of today’s automated digital systems are reinventing workflows and providing greater efficiency and accuracy for lenders and dealers.

Lenders today must also process loans quickly to allow the dealer to close the sale. This is important because the faster the transaction closes, the faster the lender can be paid. The right digital process not only accelerates adoption by digitizing cumbersome and expensive paper processes, but also adds a level of accuracy and security to a lender organization by reducing human error and centralizing and protecting customer data.

Lenders still struggle with OCR

OCR is a key piece of technology in any document automation system. However, OCR has a long way to go, especially in the auto industry where forms change constantly and new document types are added regularly. However, by combining state-of-the-art OCR technology, with applied intelligence “learned” from a purpose-built large language model (LLM) and subject matter experts, advanced AI-enabled solutions today can intelli-

gently “read” the document, turn it into actionable insights and perform complex calculations.

It is important to accurately parse the data. With the power of AI, you can get accuracy rates of 99% on known document types. For lesser-known document types, accuracy rises quickly over time. Until the “machine’s” accuracy is where you need, you control, based on your lending policies, which documents you feel comfortable accepting and others requiring manually review.

Proof of Income documents are so widely used that AI models have learned them well. They can extract data, classify the document, compare the data and determine if the document is based on a fraudulent template found on the dark web. In addition, they calculate income based on your credit policies.

AI helping with OCR speed and efficiency today

Quickly processing a consumer or auto loan is the ultimate goal of most lenders. The more documents that can be instantly verified, without human input, the sooner the lender can make their decision. Lenders can “automatically” pass a loan that historically would have needed a person to manage the entire process. That includes data entry, classification, verification, calculation, compliance, or a fraud alert. For the small number of documents that do require human review, the system shows exactly what pieces of information that need review. This includes things like missing a signature or required document.

Image quality is another complaint about OCR providers. If the image quality is bad the accuracy will be low. And lenders

don’t want to invest in new processes to manage this since more and more documents are becoming digitized. Offering better ways to instantly receive digitized documents improves the consumer and dealer experience and addresses image quality issues.

OCR also cannot determine loan defects. OCR capabilities with the addition of machine learning ensures contracts are complete and accurate. Solutions that collect clearer documents ensure the lender gets what they need without disrupting their current workflow and provides the best solution from one financial technology provider. More options lead to customized streamlined solutions with better business outcomes.

Transforming documents and data into actionable insights and decisions is no easy task, but today’s machine learning models are now trained on millions of documents to support the auto, consumer lending, and mortgage loan industries.

OCR systems are continuously being improved through advances in machine learning and artificial intelligence. Additionally, pre-processing steps, such as image enhancement and document cleaning, can be applied to improve OCR accuracy in auto lending and other applications.

Jessica Gonzalez is director of lending strategies for InformedIQ.com, an AI startup company serving the financial services industry with a sophisticated Software-as-a-Service (SaaS) platform that uses AI and machine learning models to classify, analyze, and extract data from documents used in consumer lending, mortgage, and bank account openings. For more information, visit www.informediq.com.

