

IDA Classification

Versatile feature for document categorization tasks

IDA Classification provides a **rule-free, few-shot learning feature** for document categorization tasks. Thanks to its **advanced machine learning capabilities**, the setup and maintenance of workflows are dramatically accelerated compared to rule-based or manual approaches. Even when minimal variation is apparent, IDA Classification offers exceptional accuracy by analyzing both textual and visual features. Combined with PLANET AI's **best-in-class OCR and ICR capability**, the need for manual classification corrections decreases significantly, resulting in enhanced straight-through processing.

KEY FEATURES

Few-shot learning capabilities

IDA Classification features advanced few-shot learning capabilities that consider both visual and textual features. This not only reduces the time to value but also minimizes maintenance efforts when document classes change or increase.

No-code training

IDA enables users without technical expertise to create, train, and customize classification models using a browser-based graphical interface.

Intelligent document splitting

Automating the separation of documents with 100 or more consecutive forms may seem impossible. However, with IDA Classification, it is possible to train a neural network to automatically detect document boundaries in large files, eliminating the need for manual separation during scanning.

PRODUCT CONFIGURATIONS

Input: JSON (proprietary "PAI File", e. g. from previous Recognition)

Output: PDF, PDF/A, JSON

SYSTEM REQUIREMENTS

For 64-bit systems

Linux: Ubuntu 18.04 - 25.10, Debian 11, 12; CentOS 8, Red Hat 8.x, 9; LEAP 15.x, SLES 15 SP 4-6

Windows: 10, 11

Windows Server: 2016, 2019, 2022

Docker

At least **12 GB hard disk storage** + varying sizes of self-trained models depending on settings (up to about 1 GB per full classification training)

At least **16 GB RAM**

Leveraging unmatched OCR quality

IDA Classification is based on **IDA Recognition**, an optical (OCR) and intelligent (ICR) character recognition engine that delivers outstanding results even when dealing with the most difficult scenarios. IDA Recognition captures machine-printed and handwritten text, checkboxes, tables, and historical scripts, even in poor-quality scans with rotated or skewed print.

When it comes to a downstream process like document classification, having high-quality input data is crucial because it directly affects the quality of the resulting output.

Easy deployment and integration

IDA is deployed either on-premises or in a (private) cloud as a Java application or containerization using Docker. The gRPC API (alternatively REST API) facilitates swift integration.

IDA CLASSIFICATION METHODS

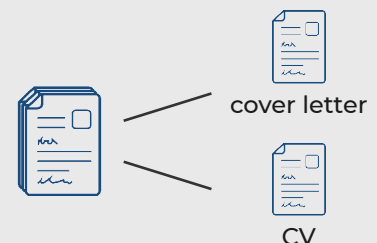
By feeding IDA a small set of tagged sample documents provided as directories, the system automatically recognizes significant characteristics and continues learning. IDA offers **two types of classification** applicable to both single and multi-page documents. Based on the resulting classes, documents can be directed to different downstream processes, such as a data extraction task.

Document Classification allows for the classification of entire documents by analyzing all the text within a given document. A possible use case includes identifying an e-mail attachment as an application.



application

Page Classification allows for the classification of individual pages within a document by analyzing the text of the current page as well as the previous and next pages, if available. One possible use case includes the analysis of all sections of an e-mail application, such as sorting the pages into cover letter and CV.



Both classification methods are based on either a trainable neural network or a dictionary (“bag of words”).

DOCUMENT SPLITTING

When scanning large batches of documents, it is common to encounter PDF files containing 100 or more consecutive forms. With IDA's document splitting feature, it is possible to **train a neural network** to automatically separate document batches containing multi-page documents.



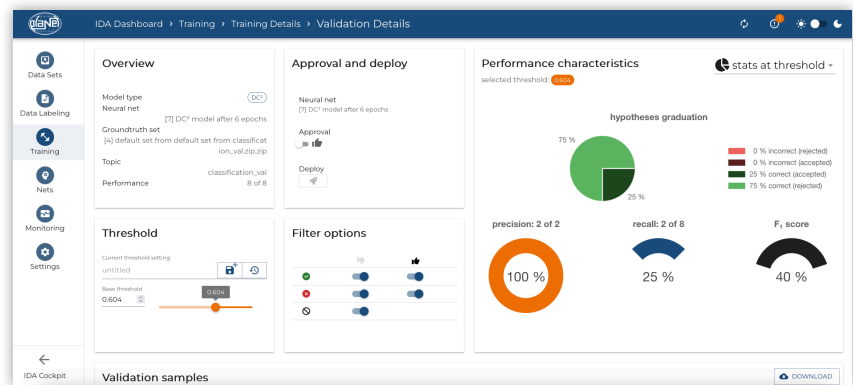
If the requirements align with a previously performed page classification, you can utilize the same model for document splitting.

Alternatively, documents can be split **rule-based** after a fixed number of pages instead of using machine learning. To split documents with layouts that do not frequently change, you only need to provide one blank document per class, e. g. a blank form. As of today, document splitting is limited to documents input in the correct order.

MODEL TRAINING

IDA provides a user-friendly graphical interface that allows for easy model training without requiring programming experience.

On the right: Training user interface



Neural Network

Training a neural network is the **recommended approach for most classification scenarios**. It considers both visual and textual features to create a model. During training, attention mechanisms learn to focus on the most relevant features.

For training purposes, a **minimum of 20 documents per class** is required, along with a minimum of ten documents for validation. However, to achieve higher training quality, it is recommended to provide at least 100 documents per class. When classifying structured documents with layouts that do not frequently change, providing just one blank document per class, such as a blank form, is sufficient.

Additionally, users have the option to select the **pre-trained, open-source LayoutLM** to enhance classification results for documents with similar layouts but different texts. LayoutLM is a large language model designed specifically for document layouts and context understanding. Note that utilizing this option requires GPU hardware.

Bag of Words

In contrast to the neural approach, IDA Classification offers a **keyword-spotting approach combined with the patented PerceptionMatrix capability**. The bag-of-words model primarily considers textual features, which makes it suitable for less complex classification use cases. IDA performs a **search within each document based on a customizable list of words** (incl. word groups, sentences). Additionally, it is possible to search within the PerceptionMatrix, which preserves all possible transcriptions of a given text without any loss of information.

For more information, please refer to the [software documentation](#).