

QUICK START GUIDE



Reexpress one

Welcome! You've come to the right place for reliable data analysis!

Reexpress one is the most advanced on-device natural language data analysis platform ever created. It works quite differently than any other data analysis or AI platform. Fortunately, as a user, there are only a handful of new things to learn, and the result is an overall simplification of your data analysis pipeline with new capabilities not available anywhere else. This **Quick Start Guide** will get you up and going with a high-level overview. After reading through this guide, we recommend watching our **video tutorials** and experimenting with the downloadable **sample data**. The auxiliary document **Reference** provides a high-level conceptual overview for interpreting and analyzing the data partitions and their associated probabilities; provides the specification for the input data format; and explains how to enable reranking of semantic searches. These resources are available via our website <https://re.express/>. Additional assistance is provided within the application itself in popovers wherever you see the following:



Introspection.

Updatability.

Uncertainty.

Similarity  Distance  Magnitude 

Reexpress AI for your data

Overview

Reexpress one is an on-device natural language processing training, prediction, and data analysis system.

Reexpress one is centered around the task of **Document Classification**. At its most basic, the model takes as input a document, optionally with a prompt, and predicts one element of a pre-determined set of labels (e.g., determining whether the sentiment of a product review is positive or negative). *But the overall system is much more than just predicting a discrete output label and is very general in application.* By structuring the overall task around classification, we can produce well-formed, principled probability estimates. This then unlocks a host of additional capabilities that are not readily available with other AI systems or cloud APIs. Training the model for classification enables semi-supervised feature detection, dense matching capabilities, and semi-supervised semantic search and reranking that make **Reexpress one** more effective than using a generative model alone for many settings. In many cases, a separate generative model is no longer needed.

Reexpress one is great for analyzing customer feedback, product reviews, and enterprise data. Analyze and search your IP-rich machine learning training data. Boost your productivity by filtering and searching scientific abstracts from arXiv, conferences, and journals. Analyze and constrain the output from your generative model. **Reexpress one** can add uncertainty quantification and feature analysis to essentially any natural language task, so it's an indispensable component of your data analysis toolkit. We're excited to see the possibilities. And since all of the processing happens right on your Mac, you never have to second-guess whether you've lost to the cloud the data your customers entrust to you for safeguarding, and the hard-earned data that gives your company or organization a competitive edge.

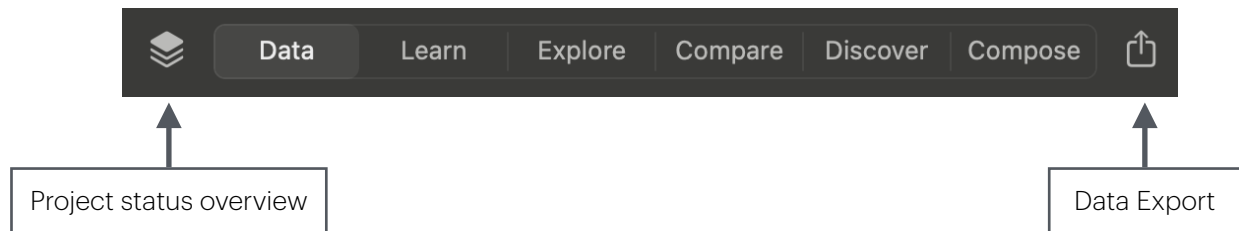
Underneath the hood is a set of large encoder-decoder language models, which are a unique fusion of a subset of weights originally derived from Flan-T5 (xl, large, and base) and mTO-base, combined with additional layers and parameters. The on-device models are very effective for typical natural language document classification tasks. And since **Reexpress one** produces reliable uncertainty estimates, you will know if the chosen on-device model is ever not sufficient for your task and an alternative is needed. You can always combine the on-device models with your existing enterprise language models or cloud services via adding **Reexpression Attributes** to a document: Real-valued features that can be, for example, the output logits from your existing model, or a categorical indicator derived from the output of a generative model.

Getting started.

The overall process starts by adding data via a simple JSON lines format. Next, you train and compress the model. **Reexpress one** can then make uncertainty-aware estimates of document-level labels. The prediction step involves running a dense (vector) search and partitioning the data in order to provide a reliable uncertainty estimate of the prediction. Unique to **Reexpress one**, these calibrated probabilities are accompanied by an estimate of the reliability of the calibration process itself. Once the predictions have been made, you can run feature analyses, perform document- and feature-level matching, run semantic searches, view interactive visualizations of your data, and more. The following sections provide an overview of the overall application and these capabilities.

Note: If you start the application and you are notified that the program has entered Experimental mode, you will only be able to use a subset of the available models. See the note at the end of this document for additional information.

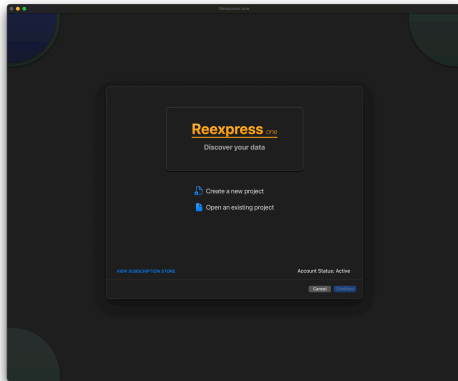
Navigation Overview



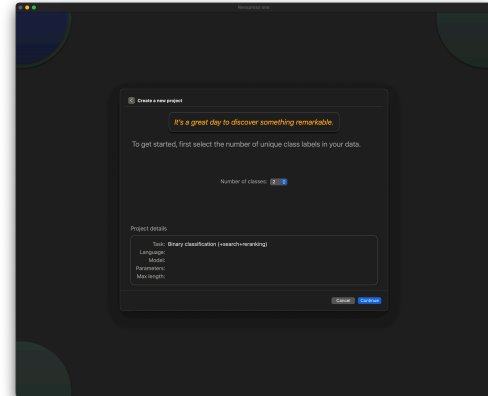
The Navigation Bar allows you to quickly switch between the core sections of the application. These sections are semantically structured and ordered to reflect the temporal progression of a typical data analysis pipeline, starting with adding **Data**, **Learning** model weights, **Exploring** the data and predictions (including running semantic searches), visually **Comparing** subsets of the data and predictions, **Discovering** key features and label errors, and finally, **Composing** new inputs. Unique to **Reexpress one**, since everything is on your device, with a continual learning state machine keeping track of changes to your model and data, you can quickly jump from section to section, interacting with your data and model in ways not previously possible. And amazingly, you don't need to write a single line of analysis code.

Let's begin!

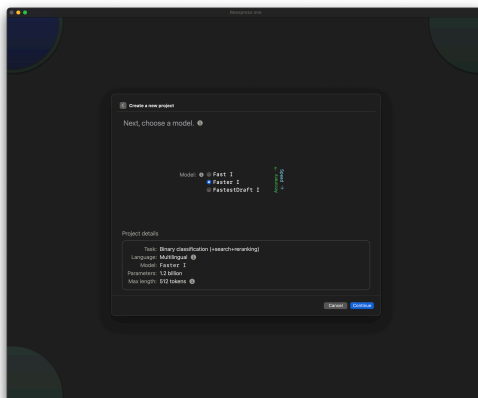
Sign up for a subscription using your Apple ID. Create a new project, choosing a model, the number of classes in your data, default prompt (if any), and a location to store the project file (extension .re1) on your local Mac hard drive.



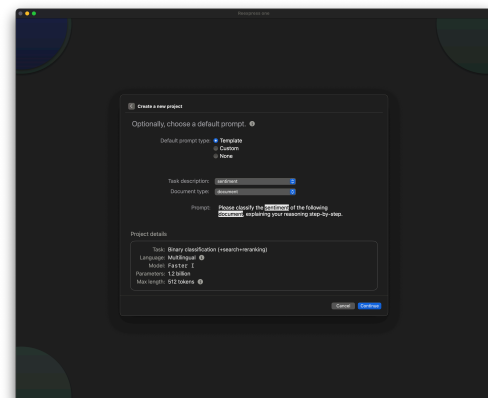
Once you've signed up for a subscription, you can create a new project.



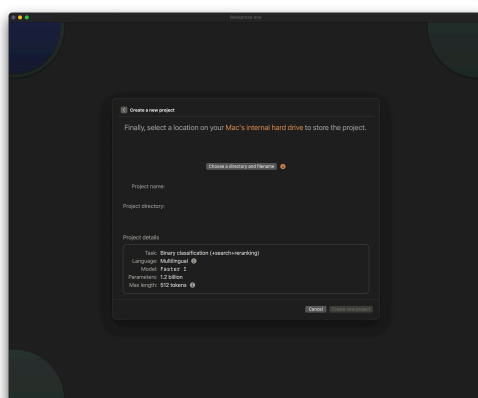
Select the number of classes in your data.



Choose a model

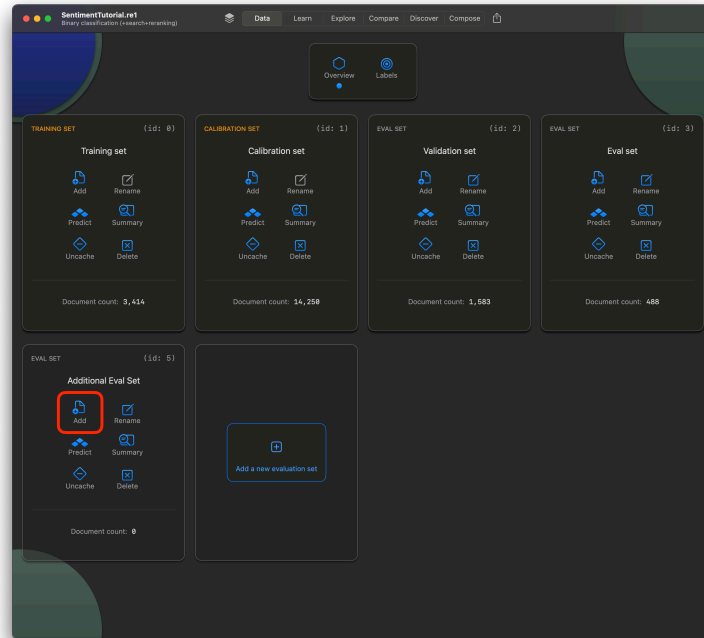


Optionally, select a default prompt

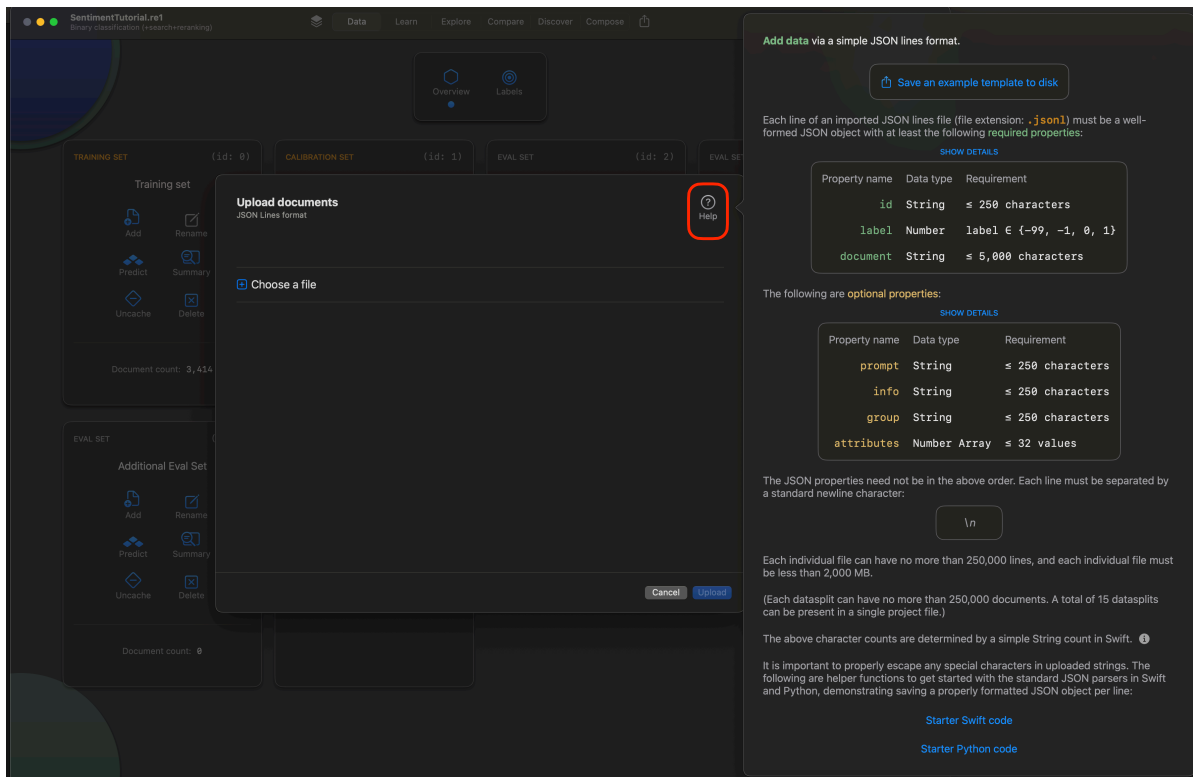


Choose a save location on your Mac's hard drive

Add data

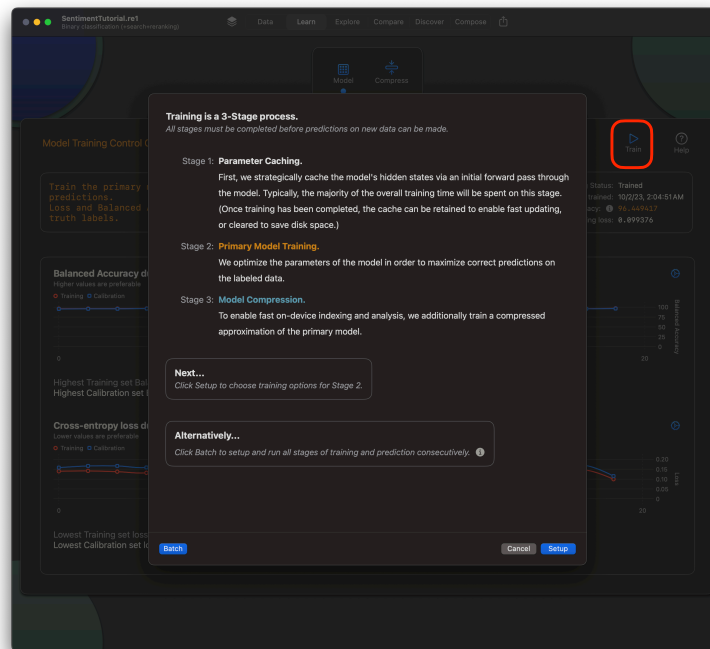


For any datasplit, click “Add” to upload additional documents.

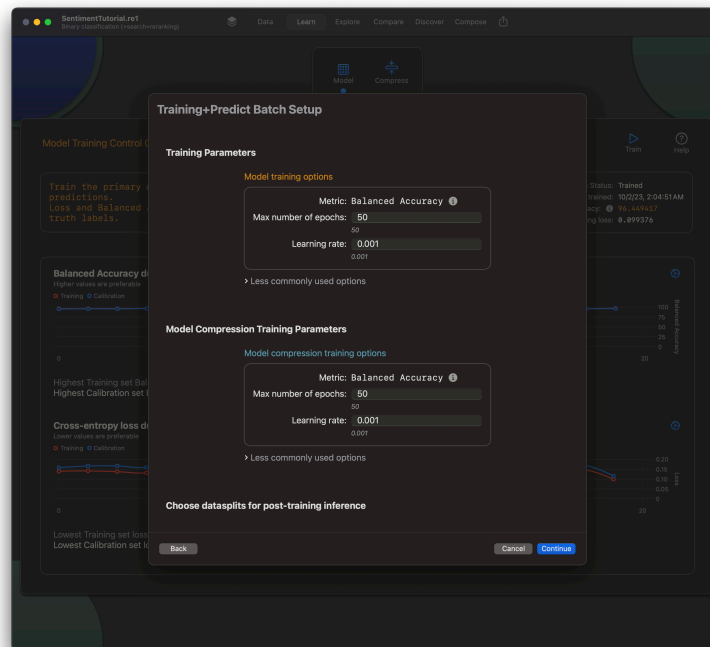


Click “Help” for information on the input data format. You can also download a JSON lines template for reference. The document *Reference* (available on our website) also provides a full description of the input data format.

Train and compress your model

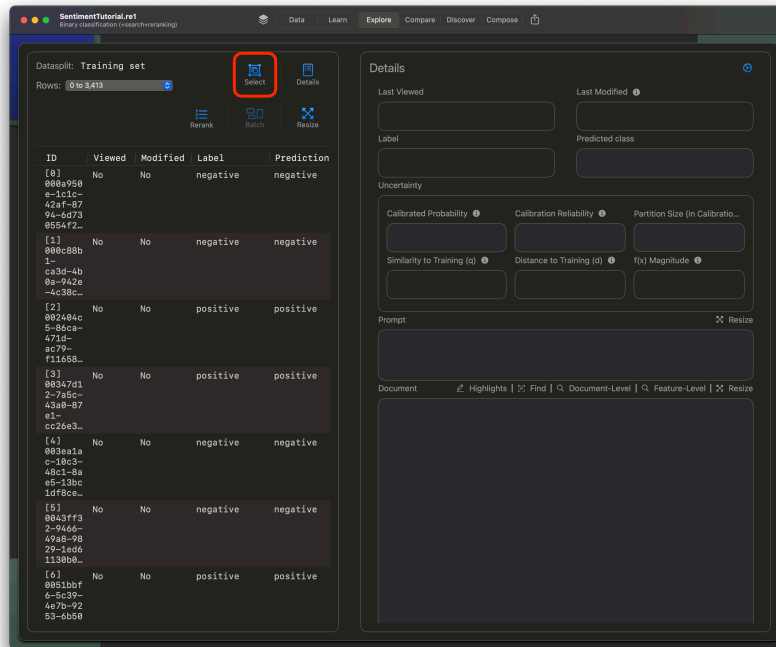


Go to Learn->Model and click “Train” to start model training. If you’re just starting out, we recommend choosing the Batch option and running all stages together. The Batch option also automatically runs the prediction step after training is complete.

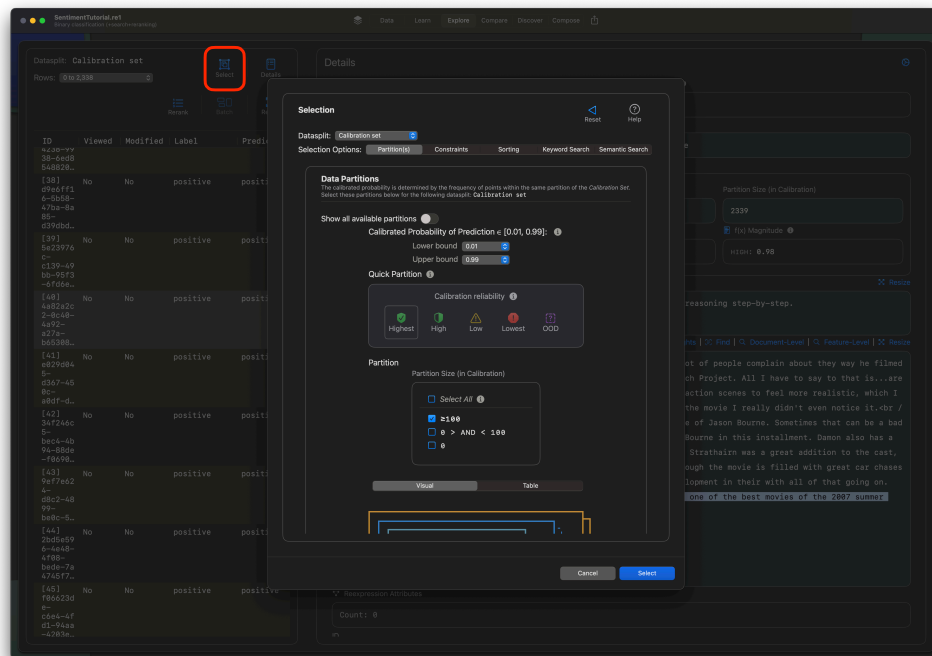


The default training settings are a good place to start. If you have a particularly small amount of data, consider reducing the Learning rate. You can always return to this section and retrain and/or continue training if your data changes.

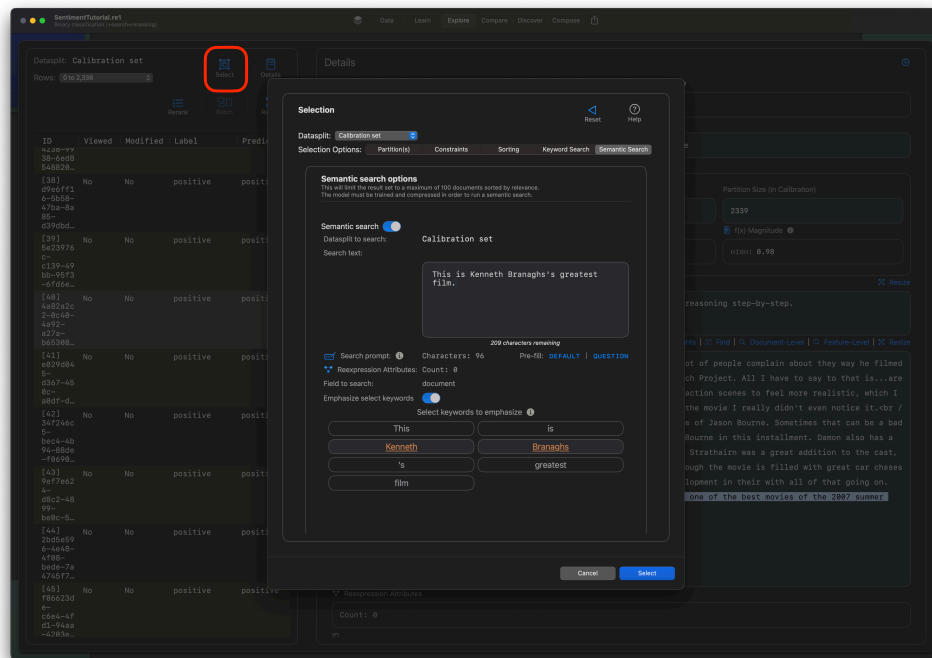
Explore



Explore provides a table-based view of your data, with a Document Details Navigator on the right. (By scrolling down, additional information is revealed.) Here and in Compare, you can filter and constrain the data by clicking “Select”.

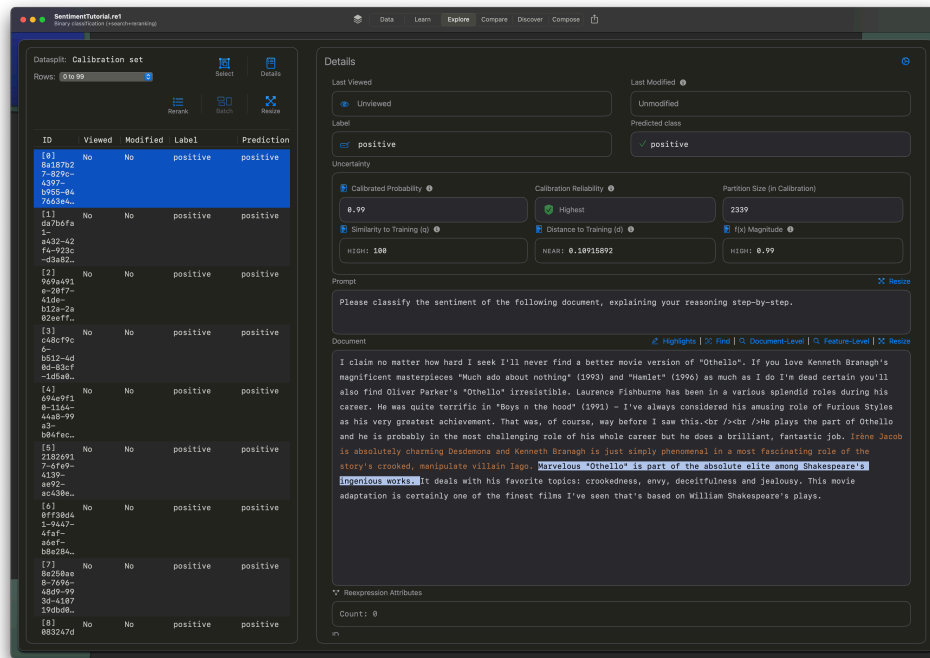


The Selection screen provides advanced data partitioning and constraints options.

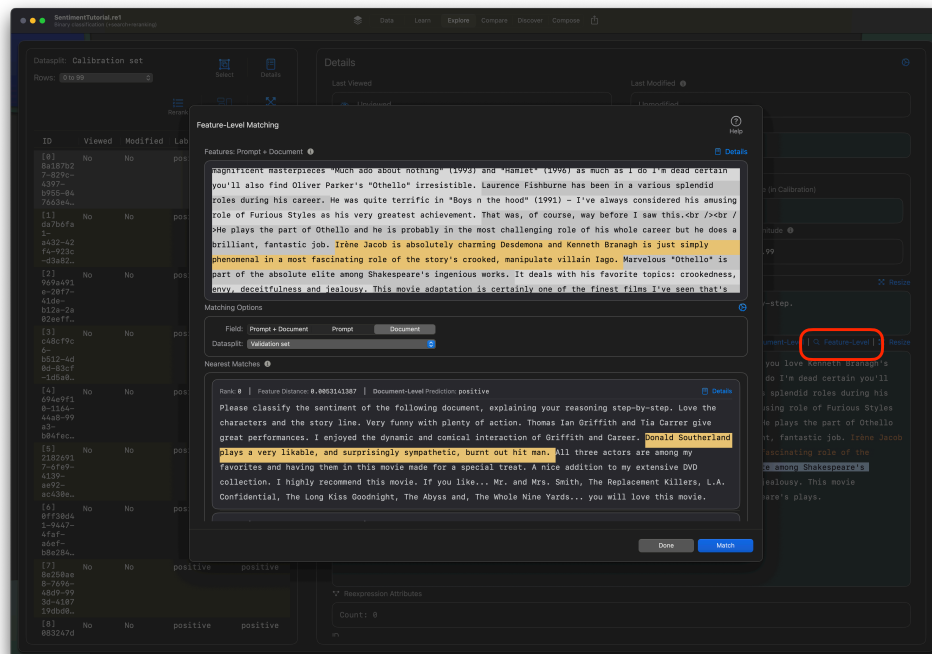


In Explore, you can run a semantic search via the Selection screen. Unique to **Reexpress one**, you can click on individual keywords to emphasize their contribution in the semantic search, providing a quick and easy way to modify your search intent. The semantic search is a novel combination of dense matching and information theoretic lexical weighting, and thus differs from the document-level and feature-level matching available via the Details view. Experiment with these various methods; each has a place in your data analysis toolkit.

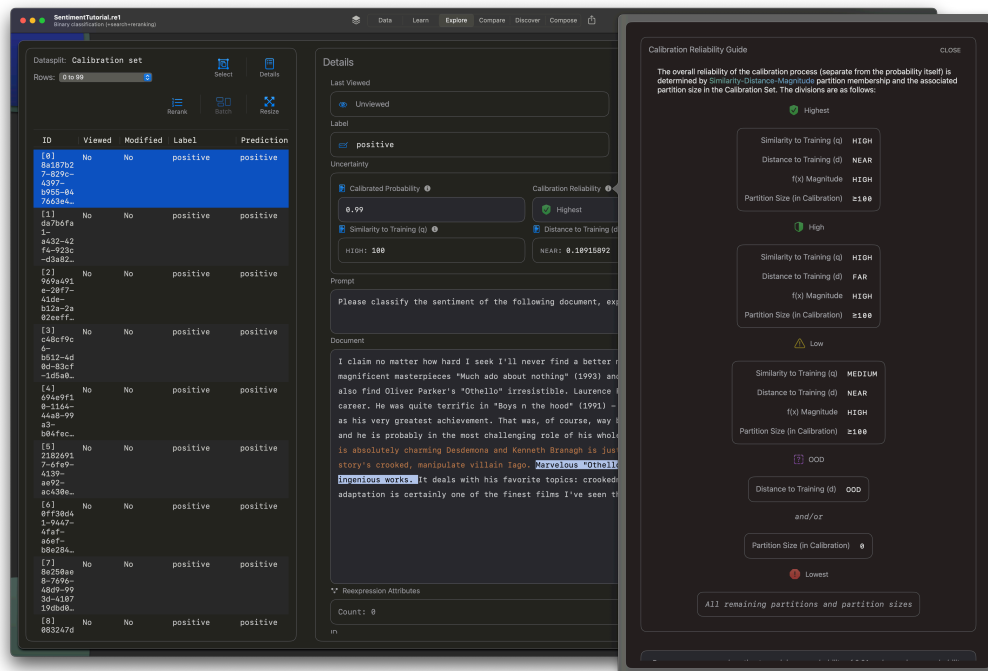
*(You can also rerank the search results with a cross-encoder if you train your model with multi-task labels and examples that resemble the structure returned by reranking. See Help assistance within the application and the document **Reference** for additional information.)*



A semantic search returns documents sorted by relevance. Additionally, the most relevant sentence *within a document* is highlighted in orange. The most prominent feature for the document's predicted class is also highlighted, here with a blue background (each class uses a different color). You can turn on/off such highlighting via the gear icon in the upper right.

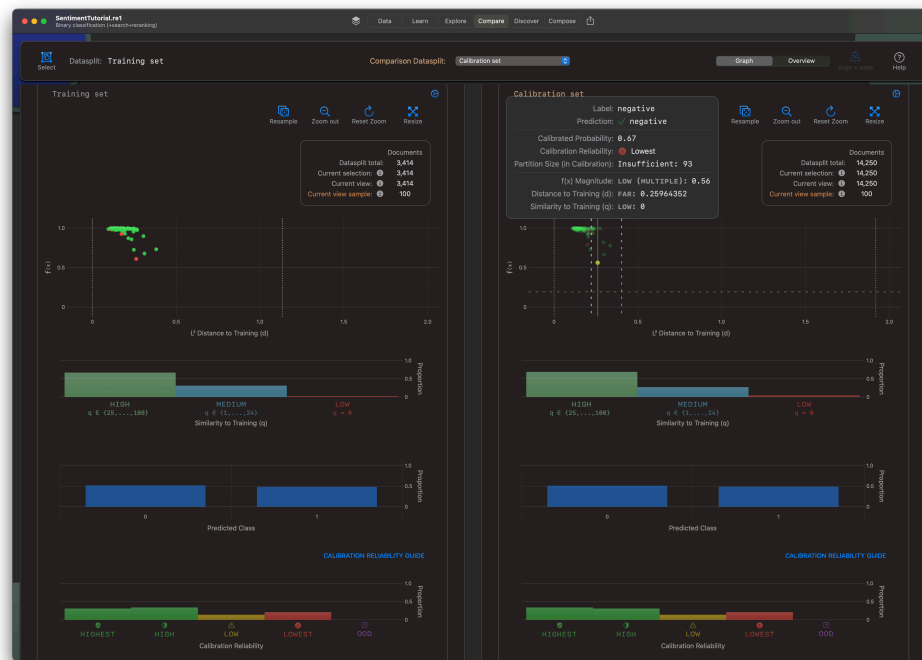


Dense matching is available to further analyze your data. Select a document and then select "Feature-Level" to run dense-vector matching against additional documents at the feature level. An analogous option is available for the document level.

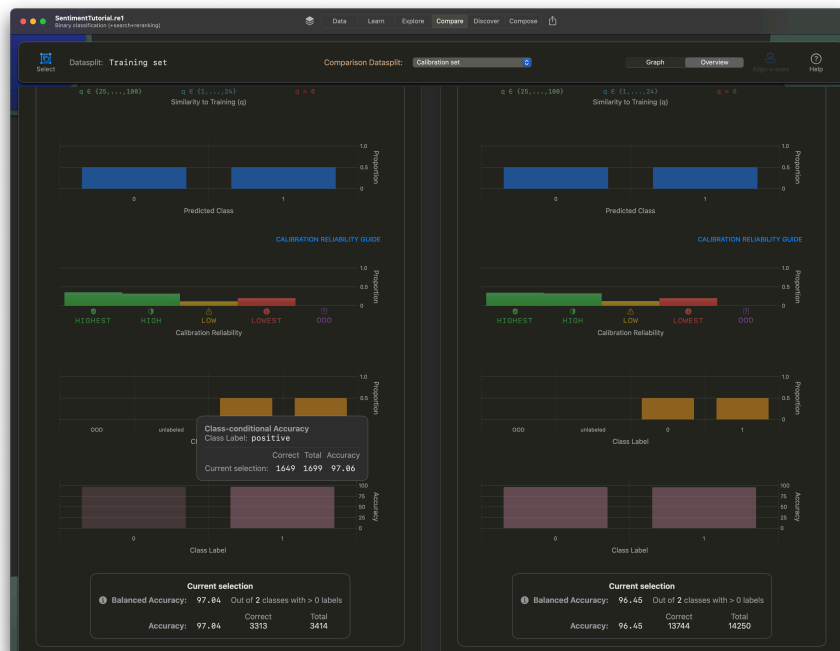


Reexpress one has a very advanced, complex, and robust uncertainty quantification system. **Reexpress one** handles the heavy lifting, and for users, the output is distilled into a small number of key indicators that can be quickly checked and compared, and easily interpreted. This includes a second-order estimate of how reliable the calibration process itself was in estimating the Calibrated Probability. The document **Reference** has additional details.

Compare

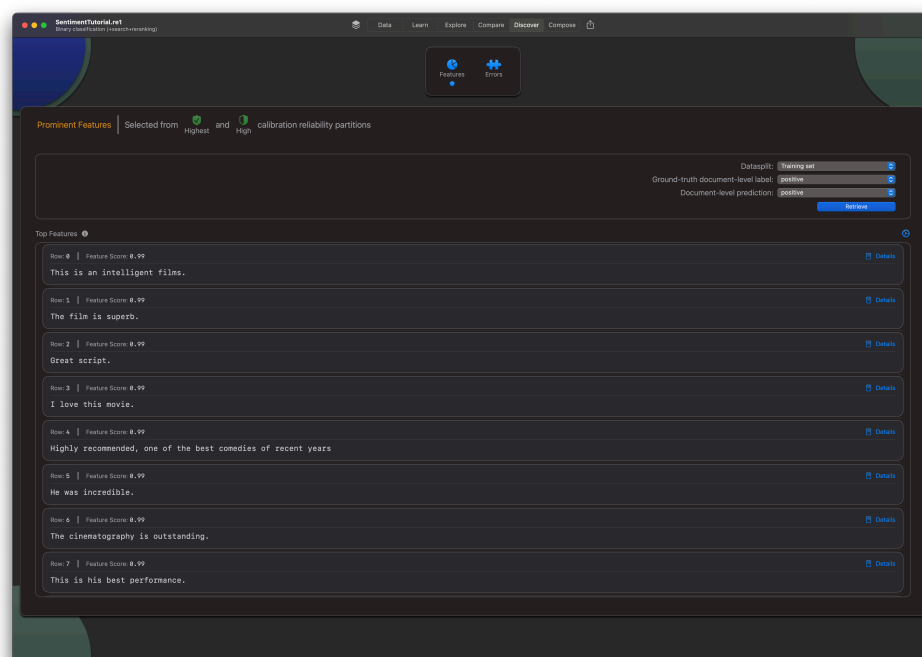


Compare provides a visualization of data selections. The scatterplot is interactive. Click on a point to display Details about a document. Click and drag to zoom. If you have two graphs side-by-side, you can click “Align x-axes” to redraw the graphs with a shared domain.



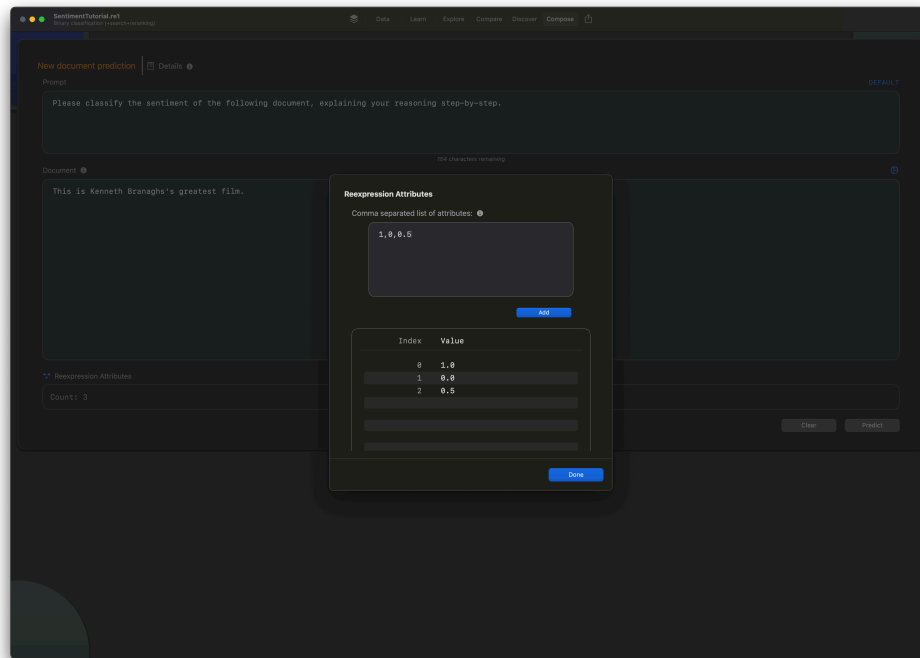
The “Graph” toggle option displays summary statistics for documents currently in view in the scatterplot, which may be a subset of documents in the current selection (e.g., as a result of zooming). Choose “Overview” to view summary statistics for all documents in the selection. If your data has labels, Balanced Accuracy and Accuracy are shown. Additional information is revealed when hovering your mouse pointer over a chart.

Discover



In Discover, you can quickly view prominent features for each class. Additionally, the “Errors” tab helps uncover possible ground-truth labeling errors.

Compose



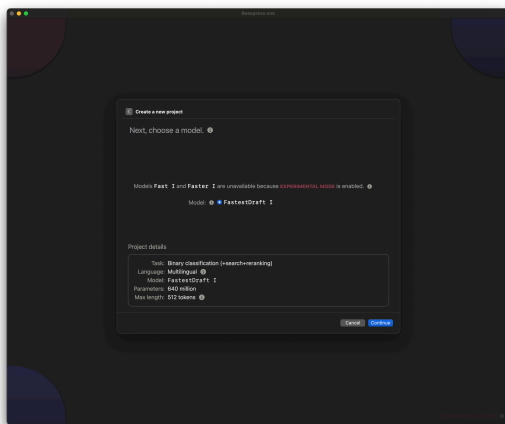
In Compose, you can make a prediction for a new document that you manually enter. As with a document added via the standard JSON lines format, you can add Reexpression Attributes, real-valued features that encode information from another model or other information you have about a document. Once the prediction step is complete, you can save the document.

(Reexpress one, version 1, and its underlying continual learning state machine, is optimized for batch prediction. If you plan to add multiple documents, it is generally recommended for efficiency reasons to add them via the JSON lines format in the Data section of the application and then run Data->Predict.)

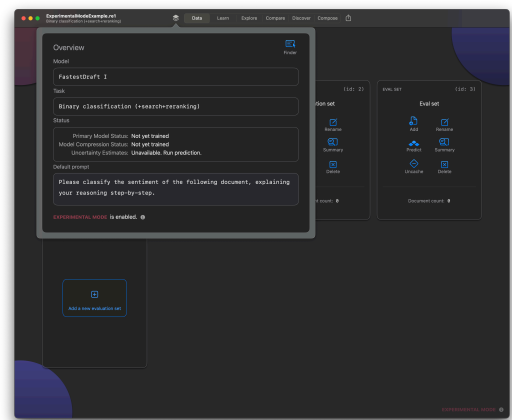
Experimental mode

If the program detects that your Mac falls below our baseline hardware system requirements (as noted on our website), it will automatically enter “Experimental mode”, which is Beta software. You will see a notice when you start the program. You will also know you are in Experimental mode by the background, which contains purple and red spheres when in Experimental mode instead of the standard green and blue. Additionally, the Status Overview will state that “Experimental mode is enabled”.

Our baseline hardware system requirements are relatively conservative. **Reexpress one** may well work on additional machines—particularly those near the boundary of our requirements, such as machines with less RAM—with the small(er) models. Experimental mode exists to test these additional settings. We would love to hear your feedback if you are interested in testing this Beta Software mode.



In Experimental mode, only a subset of the models are available for creating new projects and loading existing projects.



In addition to the notice on program start, you will know that Experimental mode is enabled if the background contains purple and red spheres and by the presence of the indicator in the lower right of the screen. Additionally, the Status Overview will indicate that it is enabled.

Discovery awaits.