Anote

Human Centered AI

The Problem

Enterprises have millions of rows of unstructured text data. Currently, they spend a lot of time and money manually going through these documents, by hand, to classify text, extract entities, and answer questions.

What's in the Market?

Manually Labeling

Label data in a spreadsheet, build internal tooling, or send to manual data annotators

Problem: Time Consuming, Iterative Relabeling

Labelbox · appen · amazon mechanical turk · scale

Gen AI APIs

New LLM technology enables enterprises to call General Purpose Models via API

Problem: Not Domain Specific, Not Accurate

ChatGPT · Claude · Perplexity · Bard

Document Processing

AI model handles specific document types (Ex: Receipts, invoices, K-1s, W-2, W-9)

Problem: Can't do new document types, Limited Functionality

VERYFI · UiPath · Abbey · impira

In-house Foundational Model

Enterprises want to train AI models on their own data for domain specific insights

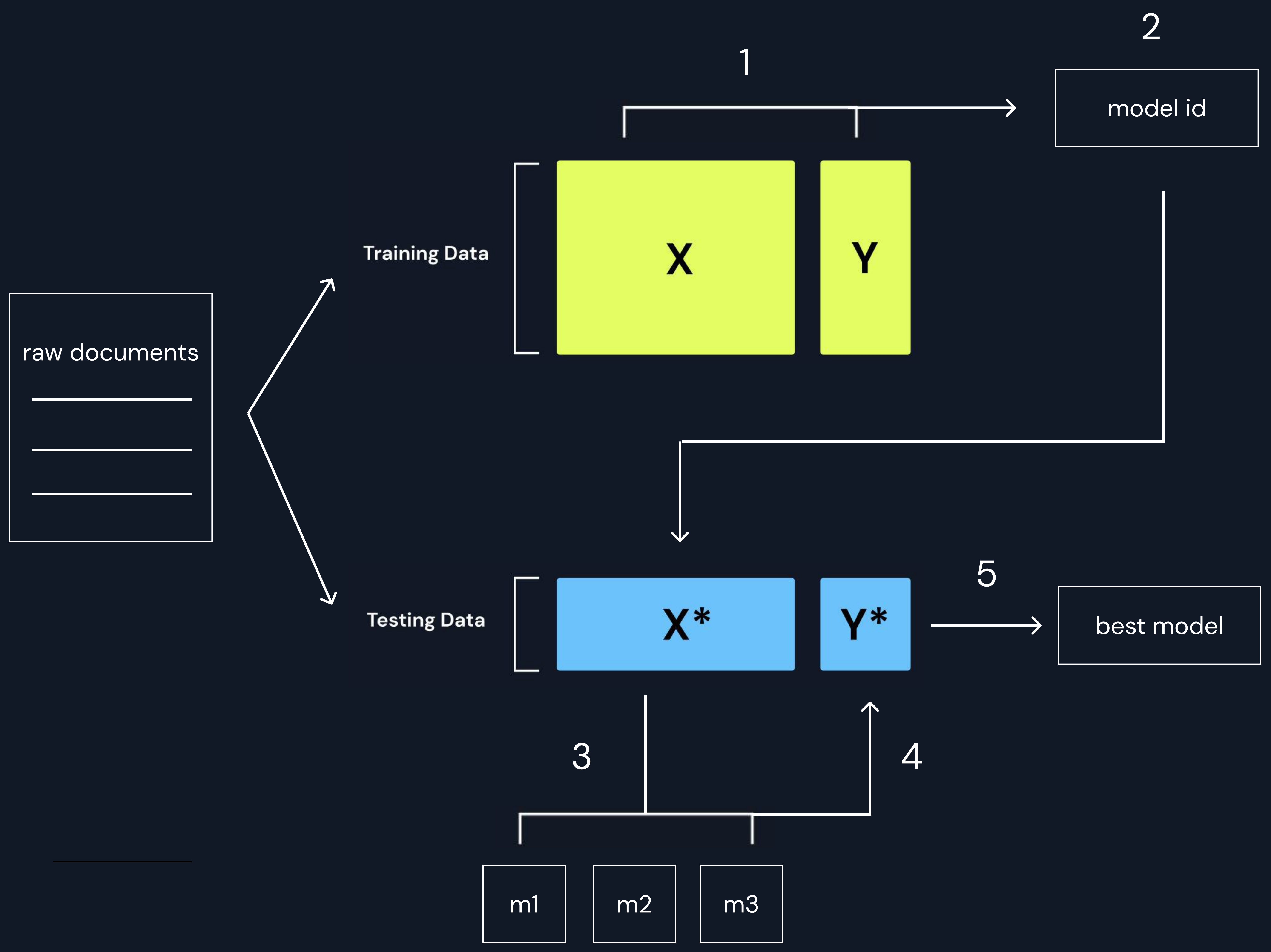
Problem: Requires expertise, Costly

LAMINI · Predibase · accenture · Deloitte.

As a result, many enterprises have been looking for solutions to fine tune or re-prompt large language models on their own data, for tailored, accurate and domain specific results.

Our Solution

1. Label Data
2. Train Model
3. Make Predictions
4. Evaluate Results
5. Choose Best Model



Data Labeler

State of the art few shot learning to make high quality predictions with a few labeled samples.

Supervised Fine-tuning

Fine-tune your model on your labeled data

Unsupervised Fine-tuning

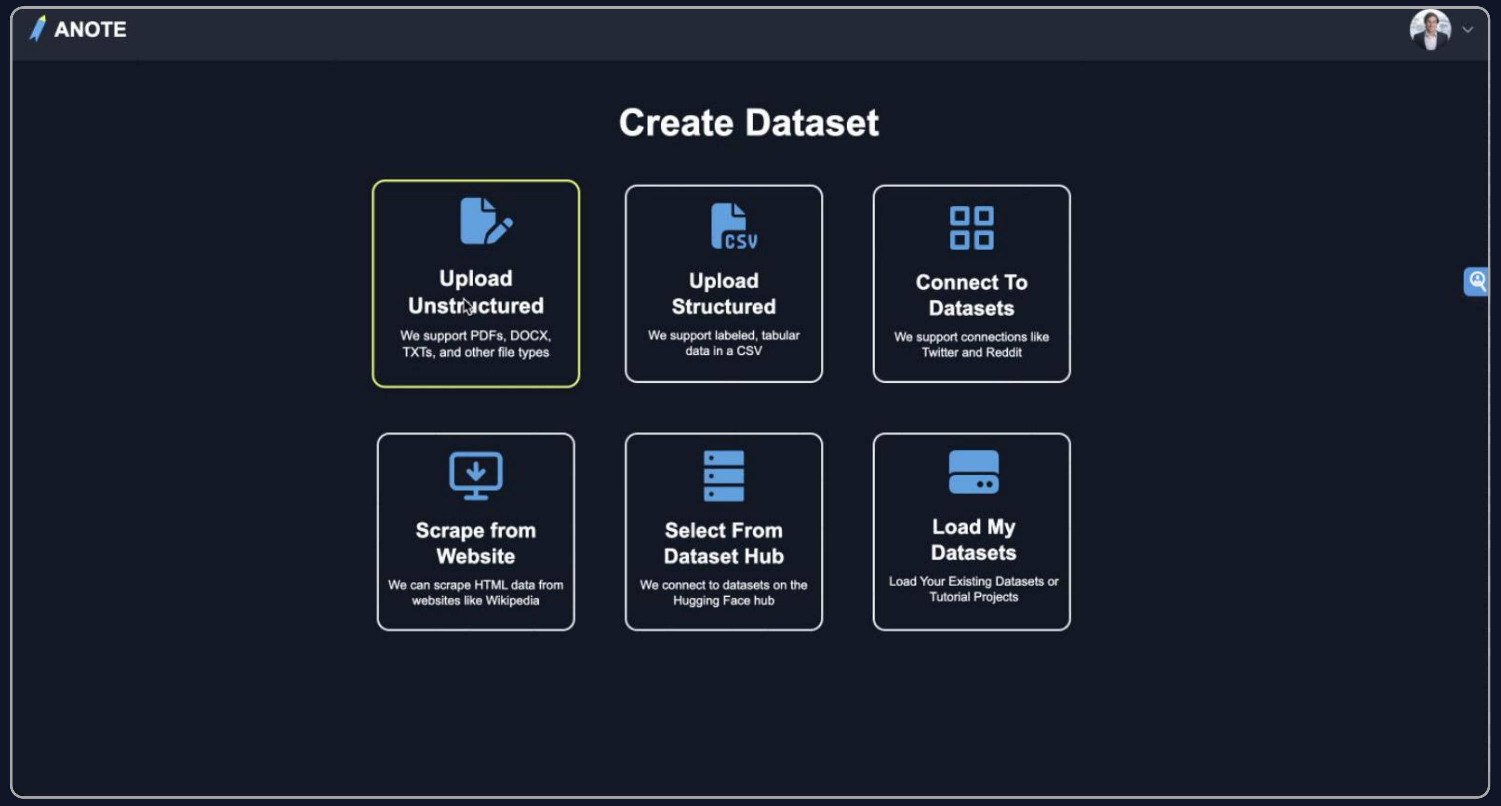
Fine-tune your model from your raw unstructured documents

RLHF / RLAIF

Actively improve your models from human / AI feedback

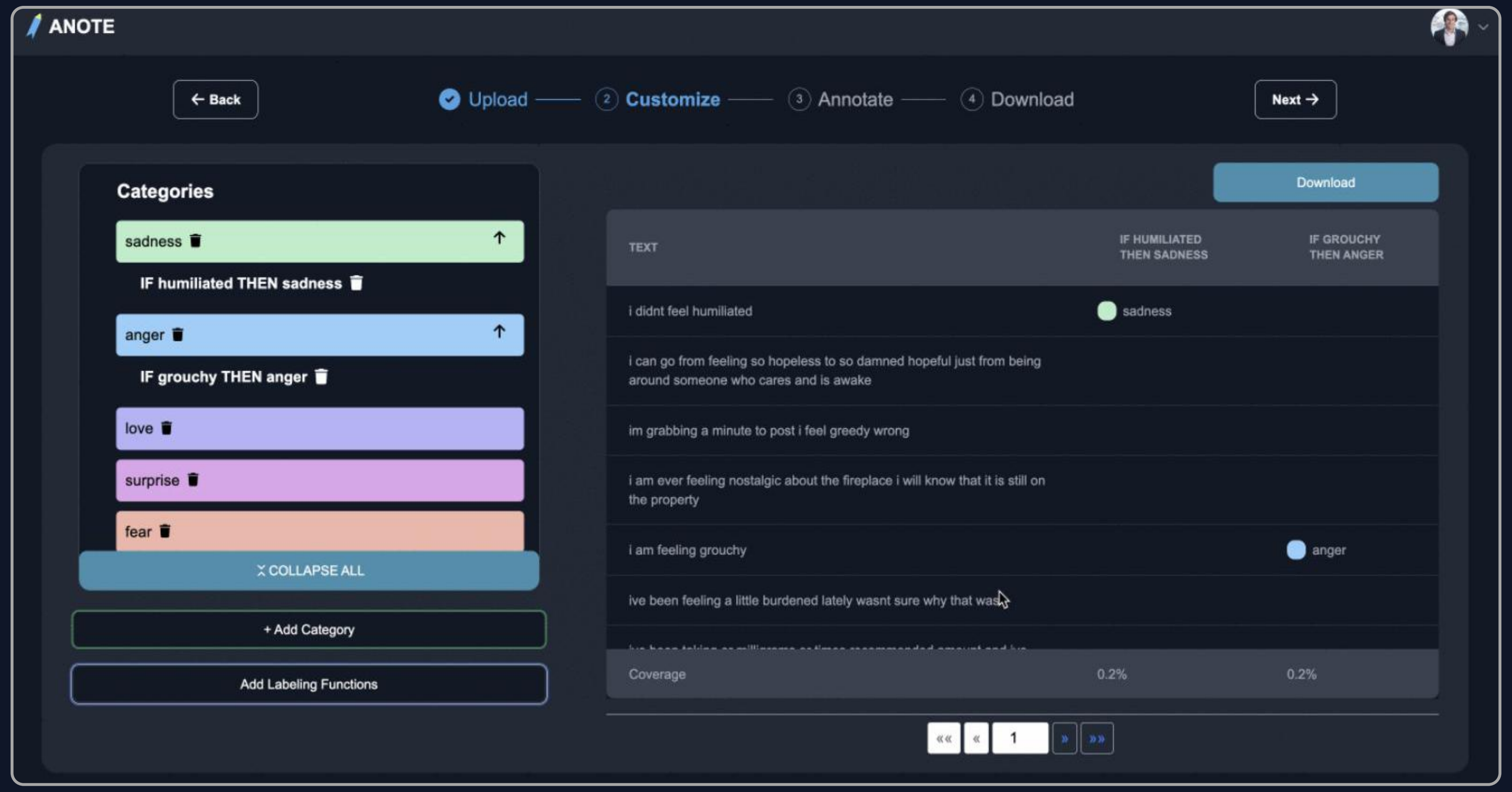
Upload

Create a new text based dataset.



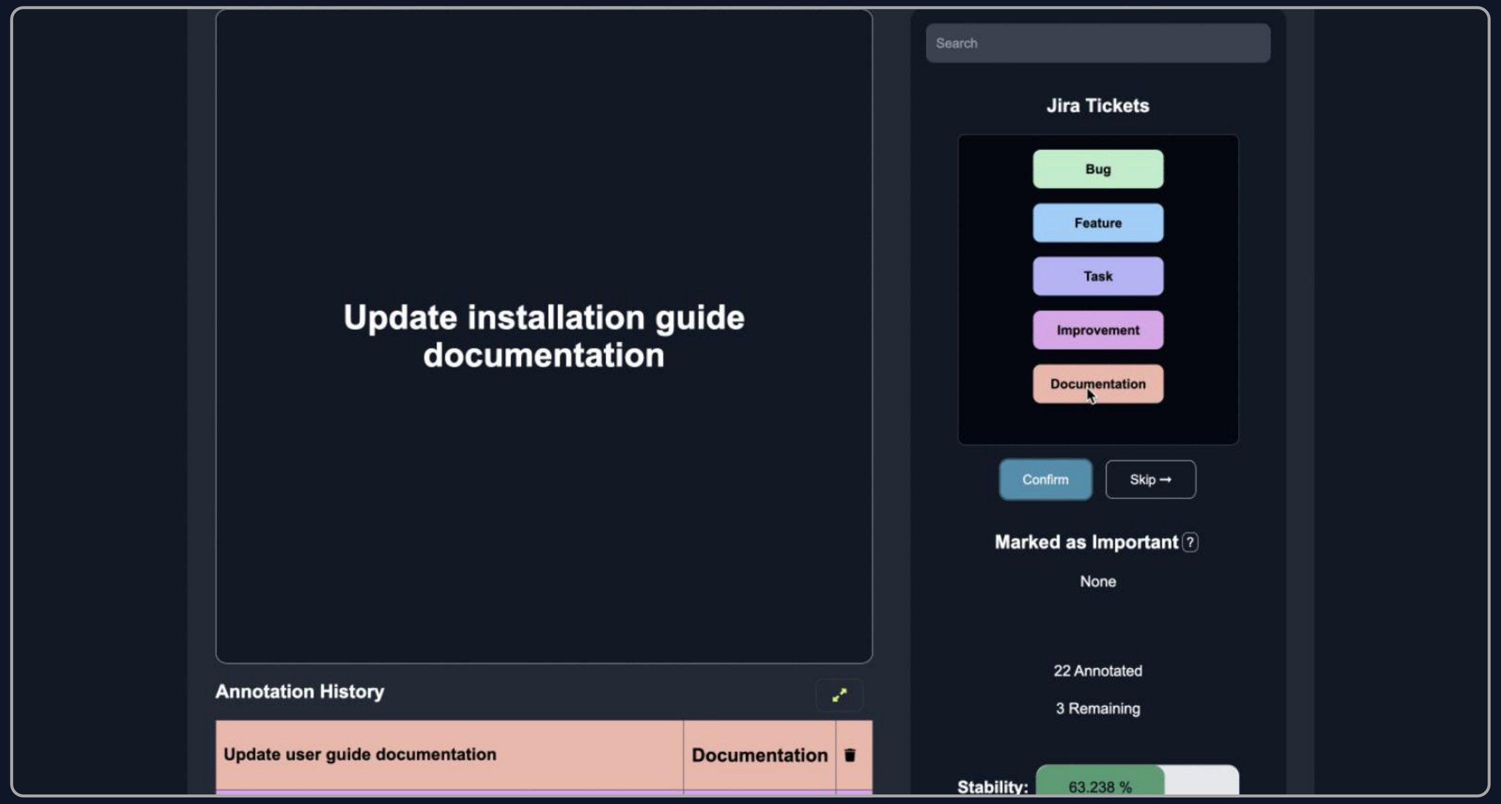
Customize

Add the categories, entities or questions you care about



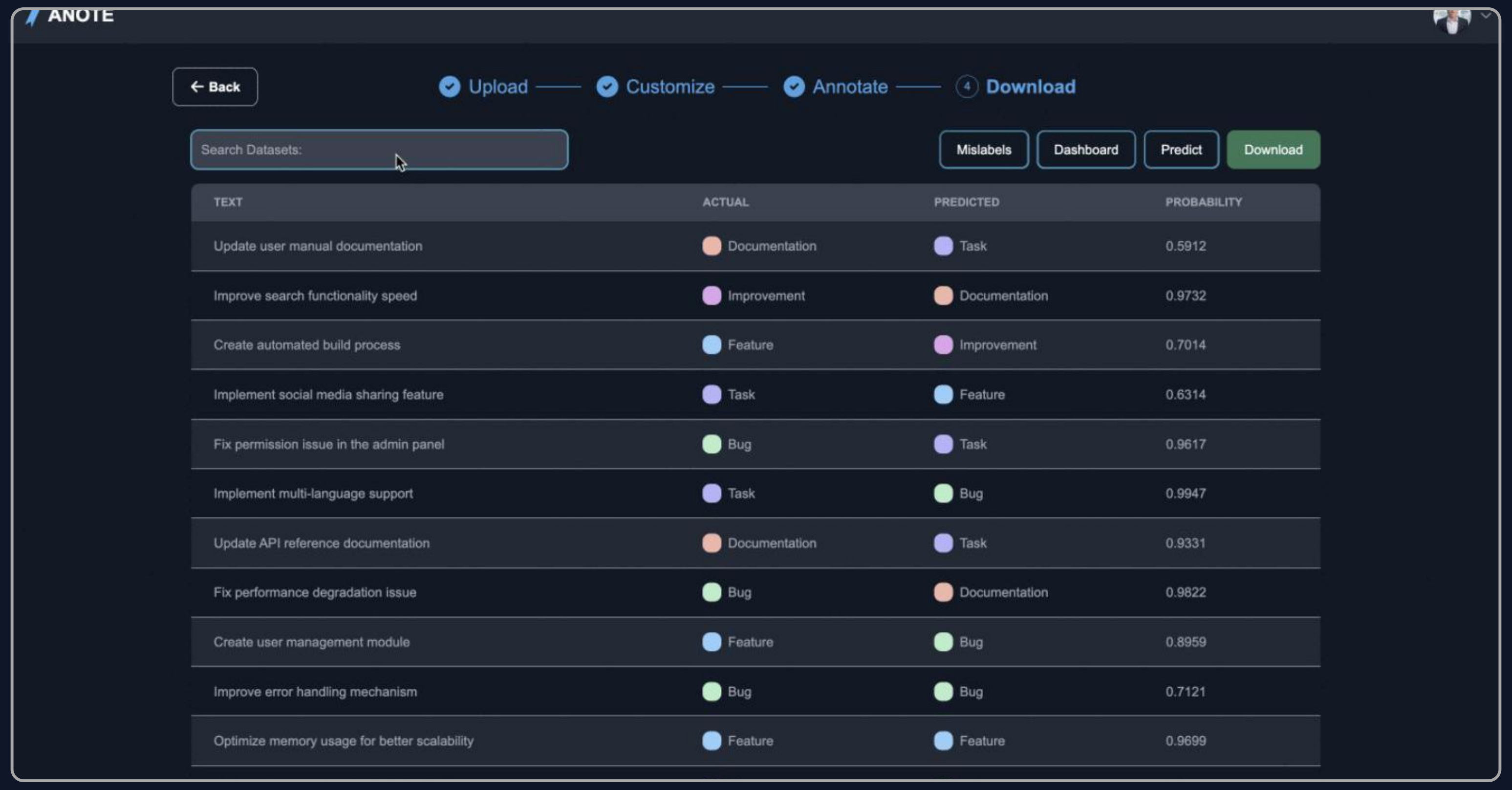
Annotate

As you annotate a few edge cases, the model actively learns to predict the rest.



Download

Download the resulting labels as a CSV. Export fine tuned model as an API endpoint



Private Chatbot

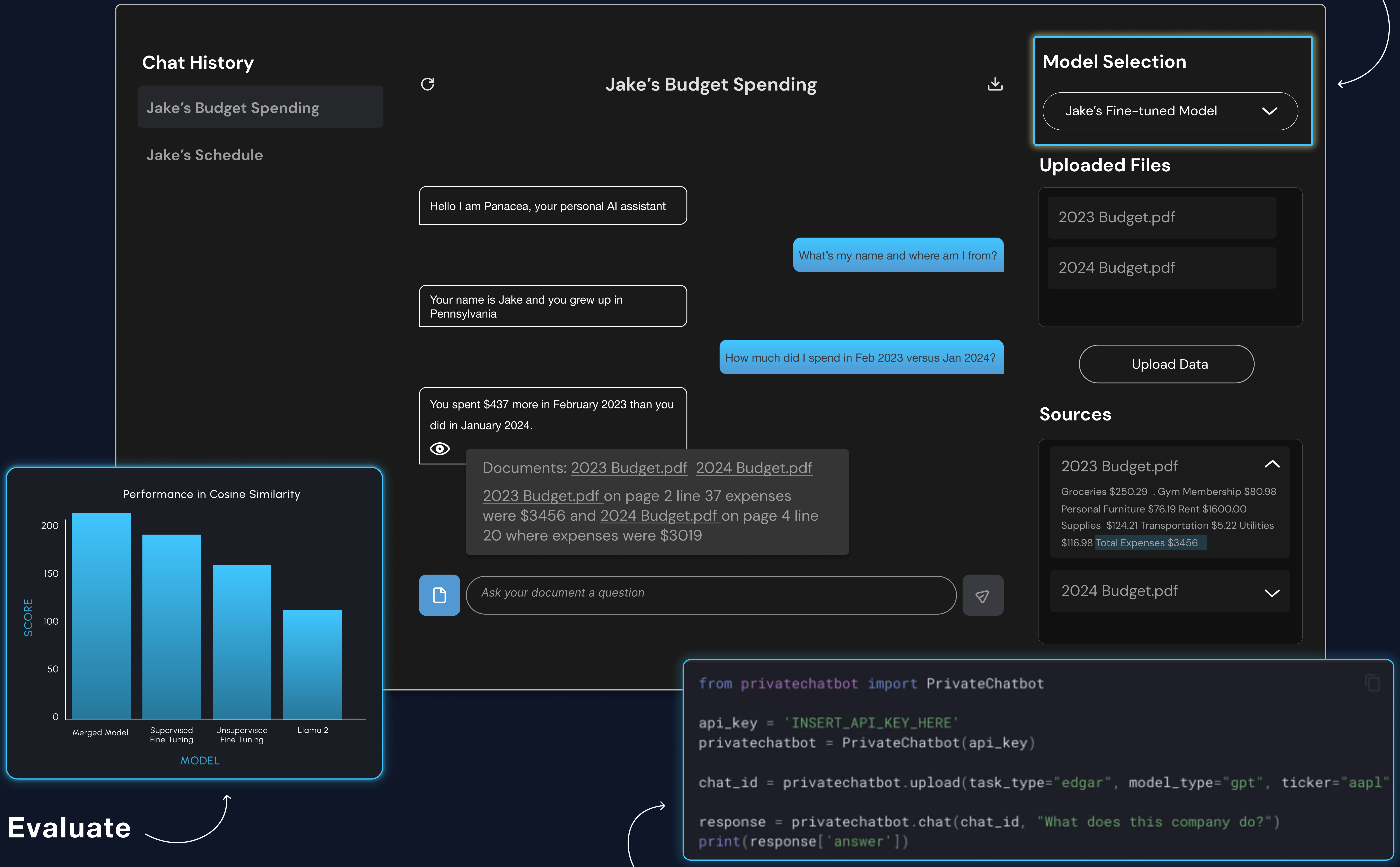
Your accurate private enterprise AI assistant

Upload
Upload your financial documents (10-Ks, 10-Qs, Earnings Calls Transcripts)

Chat
Ask Questions on your documents with models like GPT, Claude, Llama2 or Mistral

Evaluate
Get Citations for answers, and ensure the answer is correct to mitigate hallucinations

Input Fine-tune Model



Evaluate

Software Development Kit

Differentiators

ACCURATE PREDICTIONS

Fine tuning and enhanced RAG for more accurate and tailored predictions. Our AI models actively learns and rapidly improve from SMEs.

ACCURATE CITATIONS

Accurate sources (page number, chunk of text, important features) to explain the models predictions and mitigate hallucinations.

COMPREHENSIVE CAPABILITIES

Supervised, Unsupervised and RLHF / RLAIF fine tuning for classifying text, extracting entities, answering questions, and chatting with documents

EASY TO USE

Accessible UI similar to ChatGPT, and simple SDK for developers where you can input a fine tuned model for improved results.

PRIVATE VERSION

On premise enterprise-grade solution using Llama2 and GPT4All to leverage LLMs on your unstructured documents while keeping your data local and secure.

EVALUATION FRAMEWORK

Robust evaluation framework with metrics like Ragas Rouge-L, Cosine Similarity and Answer Relevance to show fine tuned model performance improvements

Why It Matters

	Before	After Anote
Data Labeling	Manually Labeling their data themselves in a spreadsheet	State of the art few shot learning to make high quality predictions with a few labeled samples.
	Tedious, Time Consuming, Costly	**Less time, less expensive, higher accuracy**
	Manual Iterative Relabeling	**Rapid flexibility for changing business requirements**
Document Processing	Given a raw unstructured documents, such as a 10-k or earnings call transcript, you can't get answers to the questions right if trying to extract info, where accuracy really matters.	After a few interventions, we go from 10 questions right, to 15 questions right, to 20 questions right, to enable insights that were otherwise impossible to obtain.
	Not accurate and largely manual extraction	**Higher accuracy for raw unstructured documents**
	Sub-optimal analytics for critical business decisions	**New insights that otherwise were not obtainable**

Thank You!

Anote is a startup in New York City, helping make artificial intelligence more accessible. We believe there is a massive gap between the tremendous power of AI models, and the everyday tasks that people care about.

Contact Info

https://anote.ai/

nvidra@anote.ai

https://www.linkedin.com/company/anote-ai/

https://www.youtube.com/@anote-ai

BENCHMARKING Q&A MODELS TALK

Feb 2024 | Katherine Jijo

BENCHMARKING TEXT CLASSIFICATION TALK

Jan 2024 | Katherine Jijo

FEW SHOT LEARNING TED TALK

April 2023 | Natan Vidra

IMPROVING RETRIEVAL FOR Q&A TALK

Feb 2024 | Spurthi Setty

FINE TUNING OF LLMS TALK

Jan 2024 | Spurthi Setty