



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



Gen AI APIs

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

Problem: Not Domain Specific, Not Accurate



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



In-house Foundational Model

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

Problem: Requires expertise, Costly



>>> 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

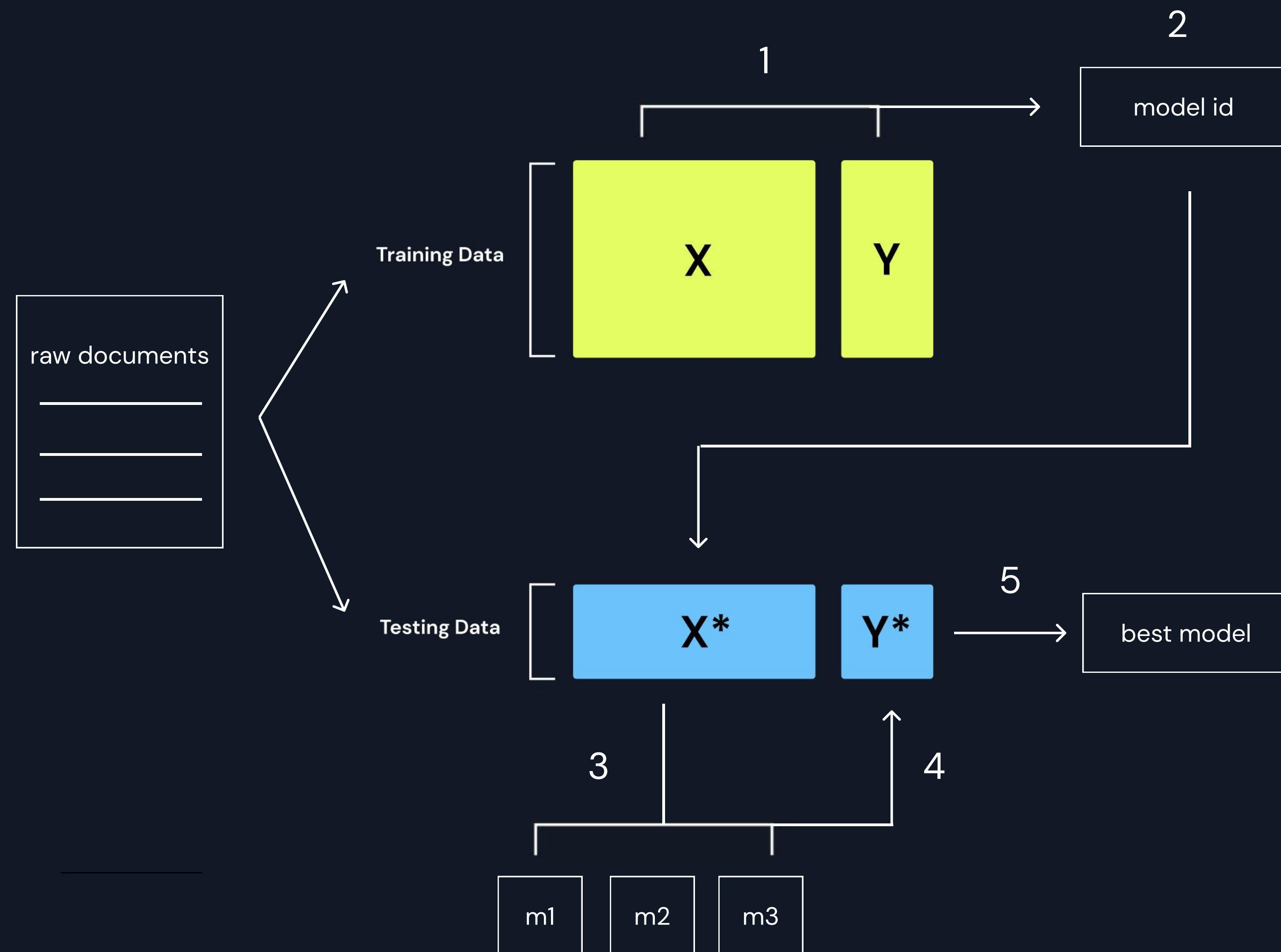
1. Label Data

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

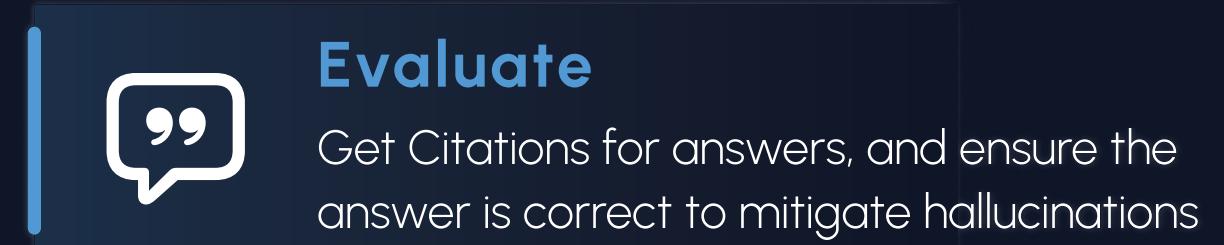
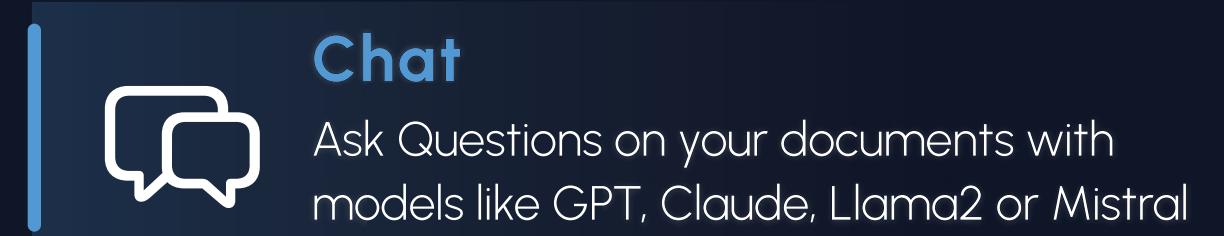
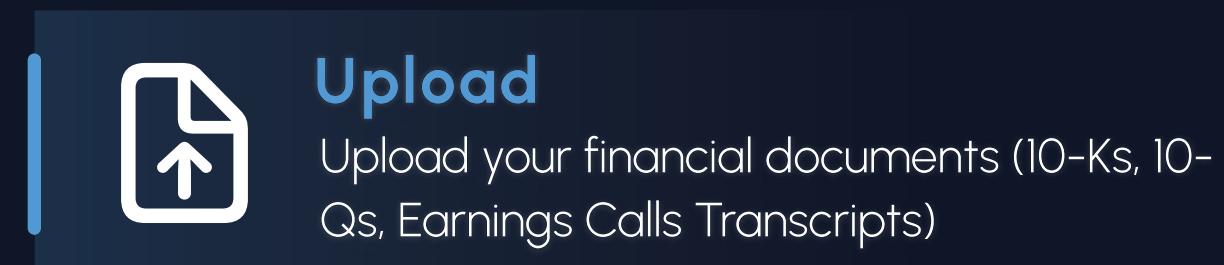
As you annotate a few edge cases, the model actively learns to predict the rest.

Download

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

Private Chatbot

Your accurate private enterprise AI assistant



The screenshot displays the Panacea AI application interface. At the top left is the "Chat History" section, which includes tabs for "Jake's Budget Spending" and "Jake's Schedule". The main area shows a conversation between the user and the AI. The user asks "Hello I am Panacea, your personal AI assistant" and the AI responds "Your name is Jake and you grew up in Pennsylvania". The user then asks "What's my name and where am I from?" and the AI replies "You spent \$437 more in February 2023 than you did in January 2024." Below this, a callout box provides document details: "Documents: [2023 Budget.pdf](#) [2024 Budget.pdf](#)". The user also asks "How much did I spend in Feb 2023 versus Jan 2024?" and the AI provides the comparison. To the right of the conversation is the "Model Selection" panel, which currently shows "Jake's Fine-tuned Model". Below it is the "Uploaded Files" section, listing "2023 Budget.pdf" and "2024 Budget.pdf", with a "Upload Data" button nearby. The "Sources" section shows the contents of the uploaded budgets. At the bottom left is a "Performance in Cosine Similarity" chart comparing four models: "Model" (black), "Supervised Fine Tuning" (dark blue), "Unsupervised Fine Tuning" (medium blue), and "Llama 2" (light blue). The chart indicates that Supervised Fine Tuning has the highest similarity. On the far right, a code snippet in Python demonstrates how to use the PrivateChatbot API to upload a document and get a response.

Chat History

Jake's Budget Spending

Jake's Schedule

Hello I am Panacea, your personal AI assistant

Your name is Jake and you grew up in Pennsylvania

What's my name and where am I from?

You spent \$437 more in February 2023 than you did in January 2024.

Documents: [2023 Budget.pdf](#) [2024 Budget.pdf](#)

2023 Budget.pdf on page 2 line 37 expenses were \$3456 and 2024 Budget.pdf on page 4 line 20 where expenses were \$3019

Ask your document a question

from privatechatbot import PrivateChatbot

api_key = 'INSERT_API_KEY_HERE'

privatechatbot = PrivateChatbot(api_key)

chat_id = privatechatbot.upload(task_type="edgar", model_type="gpt", ticker="aapl")

response = privatechatbot.chat(chat_id, "What does this company do?")

print(response['answer'])

Model Selection

Jake's Fine-tuned Model

Uploaded Files

2023 Budget.pdf

2024 Budget.pdf

Upload Data

Sources

2023 Budget.pdf

Groceries \$250.29 . Gym Membership \$80.98
Personal Furniture \$76.19 Rent \$1600.00
Supplies \$124.21 Transportation \$5.22 Utilities
\$116.98 Total Expenses \$3456

2024 Budget.pdf

Performance in Cosine Similarity

Model	Supervised Fine Tuning	Unsupervised Fine Tuning	Llama 2
Model	Supervised Fine Tuning	Unsupervised Fine Tuning	Llama 2

MODEL

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

Data Labeling

Before

Manually Labeling their data themselves in a spreadsheet

After Anote

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