

Snorkel | AI Made Practical

Snorkel AI

Founded by a team spun out of the Stanford AI Lab, Snorkel AI makes AI application development fast and practical by removing bottleneck of manually-labeled training data.

Snorkel AI has introduced Snorkel Flow, the first data-centric AI platform powered by programmatic labeling, which reduces the time, cost, and friction of labeling training data so data scientists can more easily build and scale AI models to deploy more meaningful applications while incorporating human judgment into the AI process through subject-matter experts leading to more ethical, responsible outcomes.

Snorkel AI's customers include two top US banks, global insurance, biotech, telecommunications providers, and several government agencies.

\$10M

Cost savings vs. manual labeling by F500 biotech

90%

Accuracy achieved in 24 hours by Top 3 US bank

4x

Faster model creation with SMEs by global insurer

Problem —

The Training Data Bottleneck

Today, enterprises are most often blocked from the successful deployment of AI not by lack of powerful models, algorithms, or infrastructure, but by the need for the massive manually-labeled training datasets that modern ML approaches require. These training datasets are costly and time-consuming to label, especially in industries where data is private, requires expertise to label, and changes frequently—blocking even the world's largest organizations from leveraging AI.

The Solution —

Snorkel Flow

Snorkel Flow is the first data-centric AI platform powered by the programmatic labeling pioneered by Snorkel AI.



01

Label & Build

Label and build training data programmatically in hours without months of hand-labeling.

02

Integrate & Manage

Automatically clean, integrate, and manage programmatic training data from all sources.

03

Train & Deploy

Train and deploy state-of-the-art machine learning models in-platform or via Python SDK.

04

Analyze & Monitor

Analyze and monitor model performance to rapidly identify and correct error modes in the data.

Benefits



Accelerated AI

Reduce development time by 10-100x with programmatic labeling.



Accurate AI

Increase predictive performance with massive training datasets.



Adaptable AI

Adapt to changing data or business goals without re-labeling from scratch and easily version and audit using code.



Private AI

Keep data in-house or label without humans viewing the majority of the data.



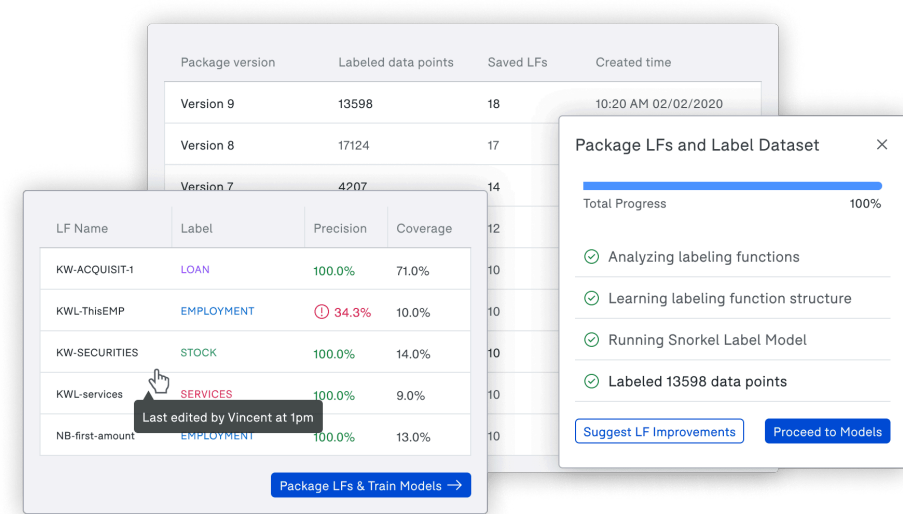
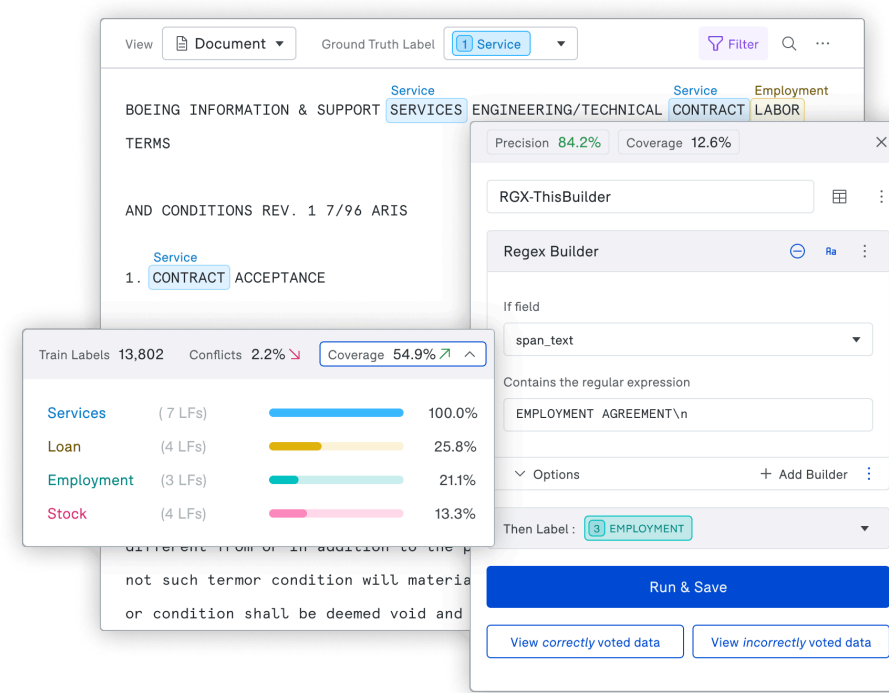
Collaborative AI

Bring together data scientists, developers, and domain experts to build solutions previously not possible.

Powered By Programmatic Labeling

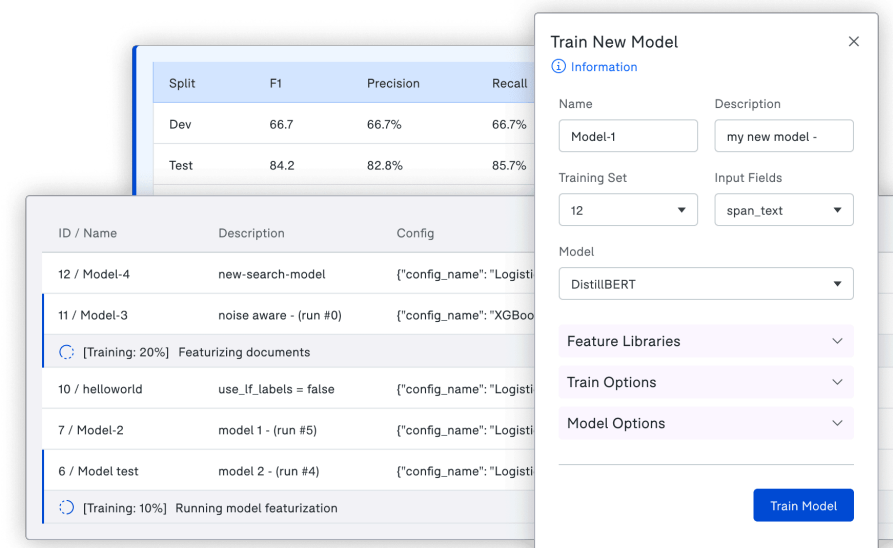
01 Label & Build

Instead of hand-labeling millions of data points by hand, automatically label vast amounts of training data using programmatic labeling functions—based on rules, heuristics, ontologies, legacy systems, and more—via a no-code UI or Python SDK.



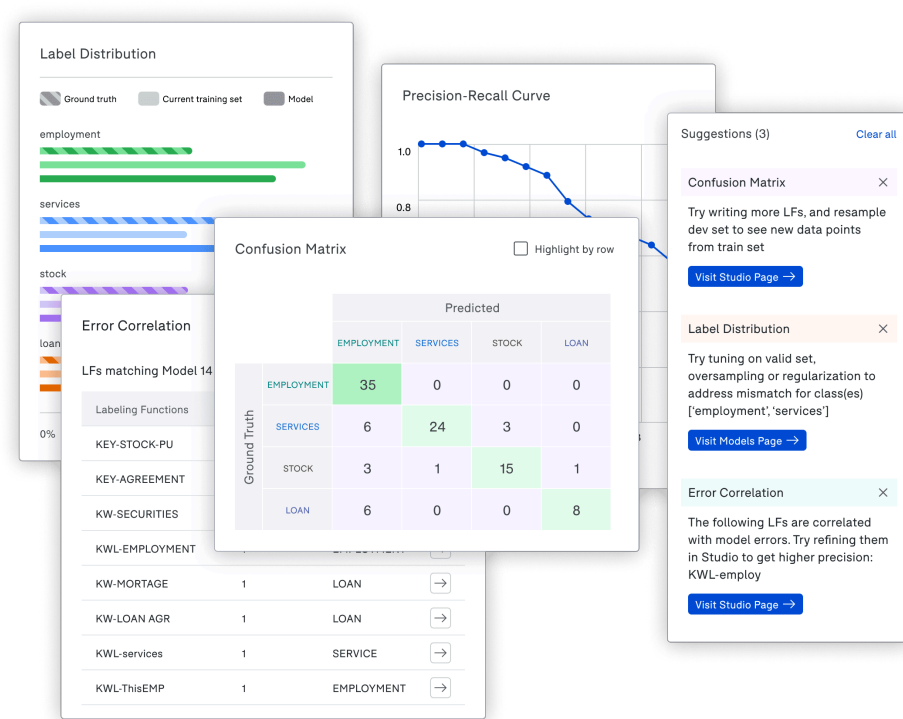
02 Integrate & Manage

Snorkel Flow automatically estimates the different labeling functions' accuracies, denoises and integrates them, and stores versioned training data. Unlike with hand-labeled data, you create training data in Snorkel Flow using code, so you can audit and modify it almost instantly.



03 Train & Deploy

Train state-of-the-art ML models with a button push or use the Python SDK to plug into your existing modeling pipelines. Deploy these high-accuracy models immediately as real-time or batch APIs or via SDK.



04 Analyze & Monitor

Identify errors to improve application quality with guided analysis. Monitor and rapidly adapt to changes without having to relabel from scratch. The result: AI development evolves from a one-shot exercise to a continuous improvement process.

Endless Use Cases

Classification

Text Classification

Classify documents and free text
Sentiment of Tweets, Contract Classification

Conversational Classification

Classify conversations and utterances
Intent Classification, Conversation Classification

Numeric-Only Classification

Classify numeric data
Classifying network traffic with metadata (e.g. packet loss)

Information Extraction

Numeric Information Extraction

Extract numeric values from text
Currency Extraction, Date Extraction

Structured Entity Extraction

Extract entities with a predefined structure
Address Extraction, Email Extraction

Proper Noun Extraction

Extract entities mapping to proper nouns
Competitor Extraction, Geopolitical Entities

Originated at the Stanford AI Lab

Snorkel's technology is based on novel research carried out the Stanford AI Lab. It has been co-developed with some of the world's leading organizations, represented in over 50 research papers, and taught in several machine learning courses at top academic institutions.

Technology developed and deployed with the world's leading organizations



AI Made Practical

Founded by a team spun out of the Stanford AI Lab, Snorkel AI makes AI application development fast and practical by unlocking the power of machine learning without the bottleneck of hand-labeled training data. Backed by Blackrock, Addition, Lightspeed Ventures, Greylock, GV and In-Q-Tel, the company is based in Palo Alto. For more information on Snorkel AI, please visit: snorkel.ai or follow [@SnorkelAI](https://twitter.com/SnorkelAI).