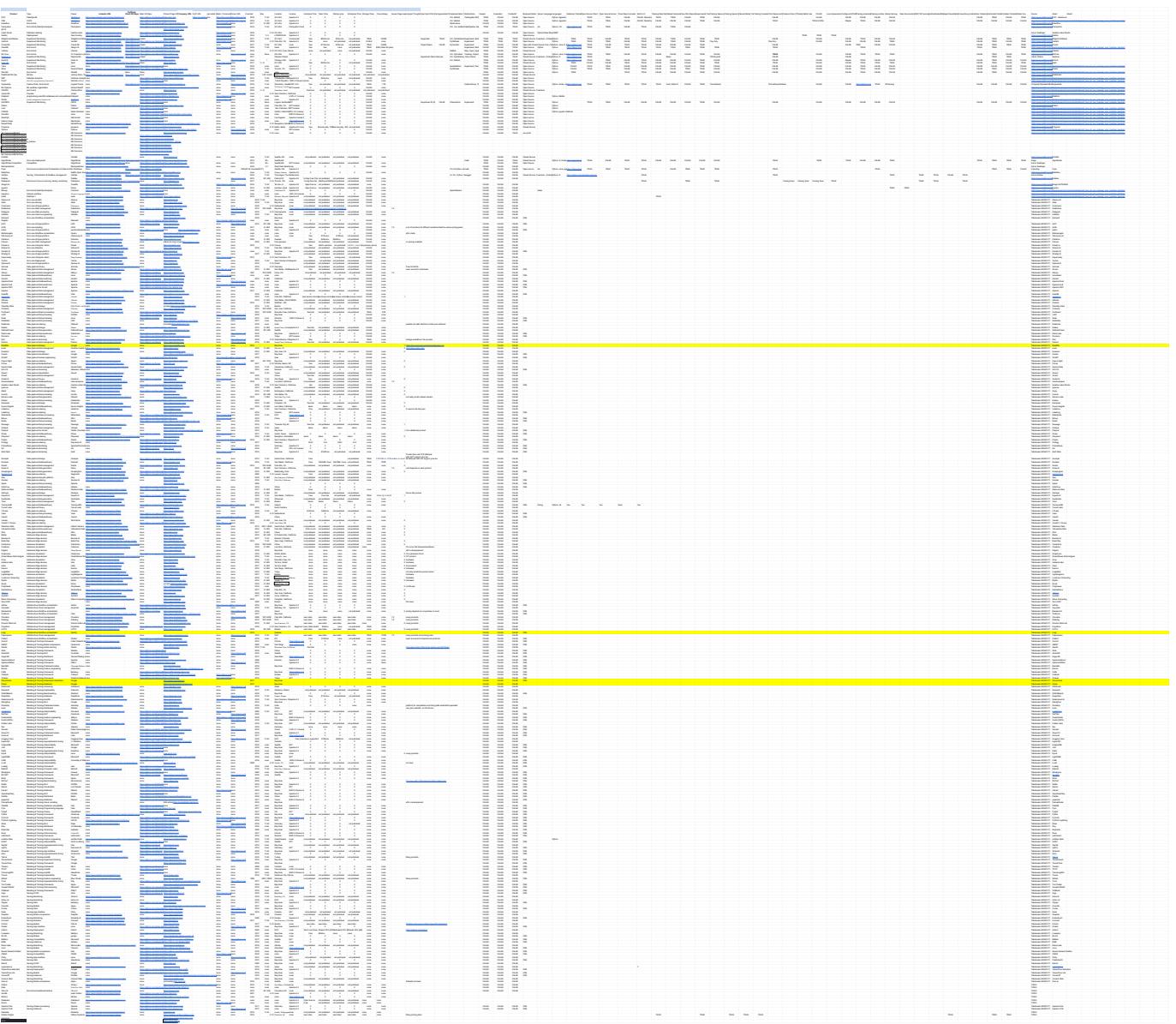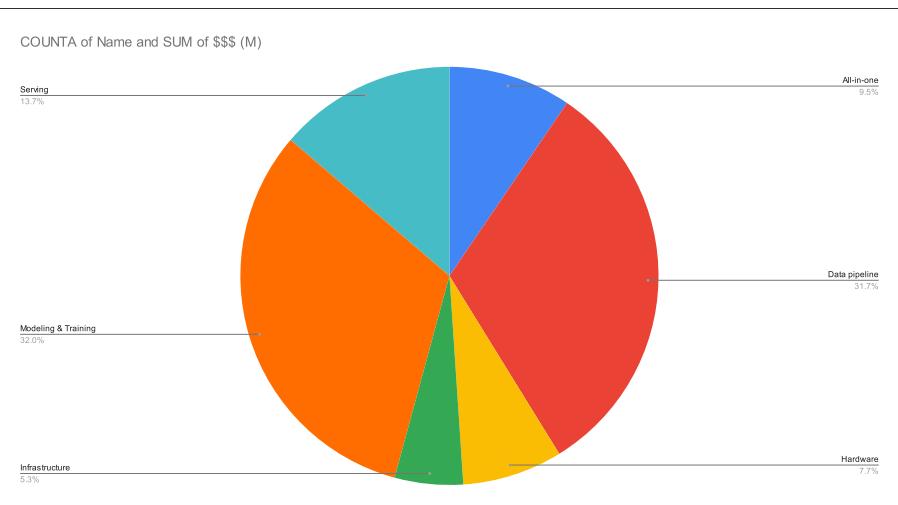| Name	Category	Type	Stage	Size	Year / Location	OSS	Raised	URL	Description
Cambricon	Hardware	Accelerator	B	200	2016 China			http://www.cambricon.com/	Cambricon Technologies builds core processor chips for intelligent cloud servers, intelligent terminals, and intelligent robots.
Cerebras	Hardware	Accelerator	C	112	2016 Bay Area			https://www.cerebras.net/	AI insights, faster Cerebras is a computer systems company dedicated to accelerating deep learning. The pioneering Wafer-Scale Engine (WSE) – the largest chip ever built – is at the heart of our deep learning system, the Cerebras CS-1.
Graphcore	Hardware	Accelerator	E	682	2016 UK		Raised 372M	https://www.graphcore.ai/	Graphcore has built a new type of processor for machine intelligence to accelerate machine learning and AI applications for a world of intelligent machines.
Groq	Hardware	Accelerator		62.3	2016 Bay Area		Raised unknown	https://groq.com/	The Next Generation of Computing is here.
Lightelligence	Hardware	Accelerator	A	36	2017 Boston		Raised 26M	https://www.lightelligence.ai/	Accelerate AI, Neuromorphic, AI Chip, Optical Computing, Lightmatter
Luminous Computing	Hardware	Accelerator	A	9	2018 Bay Area		Raised unknown	https://luminous.co/	Hardware is bottlenecked by data movement & compute. We use photonics to solve both
Nuvia	Hardware	Accelerator	B	293	2019 Bay Area		Raised 240M	https://nuviainc.com/	Silicon design reimagined for a compute-intensive world.
SambaNova	Hardware	Accelerator	C	465.3	2017 Bay Area		Raised 250M	https://sambanova.ai/	SambaNova Systems is a computing startup focused on building machine learning and big data analytics platforms.
Wave Computing	Hardware	Accelerator	E	203.3	2008 Bay Area			https://wavecomp.ai/	Wave Computing is revolutionizing AI and deep learning with its dataflow-based systems and embedded solutions.
Alectio	Modeling & Training	Active learning		0	2019 Bay Area			https://alectio.com/	Not all data is created equal You can build better models with less data. We can show you how.
CleverHans	Modeling & Training	Adversarial robustness	Google		2017 Bay Area	OSS		http://www.cleverhans.io/	An adversarial example library for constructing attacks, building defenses, and benchmarking both
Cubonacci	All-in-one	AI Apps platform		0	2018 Netherlands			https://www.cubonacci.com/	Machine learning lifecycle management Cubonacci enables organizations to focus on developing custom machine learning models without having to worry about peripheral matters. The Cubonacci platform manages deployment, versioning, infrastructure, monitoring and lineage for
Dataiku	All-in-one	AI Apps platform	D	246.8	2013 NYC		Raised 100M	https://www.dataiku.com/	Dataiku's single, collaborative platform powers both self-service analytics and the operationalization of machine learning models in production.
DataRobot	All-in-one	AI Apps platform	F	750.6	2012 Boston		Raised 320M	https://www.datarobot.com/	DataRobot combines a trusted enterprise AI platform and a trusted AI-native strategic partnership for global enterprises that want to harness the power of AI and their existing teams to succeed in today's Intelligence Revolution.
H2O	All-in-one	AI Apps platform	D	151.1	2012 Bay Area	OSS		https://www.h2o.ai/	H2O.ai is the creator of H2O the leading open source machine learning and artificial intelligence platform trusted by data scientists across 14K enterprises
Iguazio	All-in-one	AI Apps platform	C	72	2014 Israel			https://www.iguazio.com/	The Iguazio Data Science Platform automates your machine learning pipeline, transforming AI projects into real-world business outcomes.
Kedro	All-in-one	AI Apps platform	McKinsey		2019 UK	OSS			Kedro is an open source development workflow tool that helps structure reproducible, scaleable, deployable, robust and versioned data pipelines.
Obliviously AI	All-in-one	AI Apps platform	Seed	0	2018 Bay Area		Raised unknown	https://www.obviously.ai/	The entire process of running Data Science - building Machine Learning algorithm, explaining results and predicting outcomes, packed in one single click.
Peltarion	All-in-one	AI Apps platform	A	36.8	2005 Sweden			https://peltarion.com/	A single AI platform, for real world deployments, without code. Fast & Efficient Production of AI Applications. Rich data capability. Develop AI Services fast. Usable & Affordable AI.
Snorkel AI	All-in-one	AI Apps platform	A	15.3	2019 Bay Area	OSS	Raised 15M	https://snorkel.ai	Programmatically Building and Managing Training Data
Stradigi AI	All-in-one	AI apps platform	A	40	2017 Canada		Raised 40M	https://www.stradigi.ai/	Stradigi AI's powerful AI business platform, Kepler, fuels tangible results for enterprises. No AI or machine learning experience required.
Xpanse AI	All-in-one	AI Apps platform			2015 Ireland			https://xpanse.ai/	The power of AI at the click of a button. Xpanse AI brings easy to use and lightning fast analytics to your business.
Sisu	Data pipeline	Analytics platform	B	66.7	2018 Bay Area			https://sisudata.com/	Sisu is the fastest, most comprehensive augmented analytics platform letting you You can't keep up with changing metrics using manual data exploration.
Streamlit	Modeling & Training	App interface	A	27	2018 Bay Area	OSS	Raised 21M	https://www.streamlit.io/	Streamlit is an open-source app framework for Machine Learning and Data Science teams. Create beautiful data apps in hours, not weeks. All in pure Python.
Dash	Serving	App interface	Plotly		2015 Canada			https://plotly.com/dash/	Dash Enterprise is the end-to-end development & deployment platform for low-code AI Dash applications.
Gradio	Serving	App interface		0	2018 Bay Area	OSS	1.0 released	https://gradio.app/	Gradio allows you to quickly create customizable UI components around your TensorFlow or PyTorch models, or even arbitrary Python functions. Mix and match
Plotly	Serving	App interface	B	19.8	2013 Canada		Raised 8.1M	https://plotly.com	Plotly is a data science and AI company that makes it easy to create and deploy interactive web apps in any programming language.
Abacus AI	All-in-one	AutoML	B	53.3	2019 Bay Area		Raised 48M	https://abacus.ai/	Abacus.AI makes it effortless to create large-scale customizable deep learning systems. Accurate predictions generated by our system can be easily and securely incorporated into all aspects of your customer experience and business processes
Determined AI	Modeling & Training	AutoML	A	13.6	2016 Bay Area			https://determined.ai/	Our AutoML platform streamlines your deep learning workflows, tracks your work, and manages your GPU clusters.
Tazi.ai	Modeling & Training	AutoML	Seed	1.22	2015 Turkey			https://www.tazi.ai/	TAZI's Automated Machine Learning is understandable continuous machine learning from data and humans, enables business domain experts to use machine learning to make predictions and take actions. It also helps data analysts and scientists for their daily model creation and c
TPOT	Modeling & Training	AutoML	UPenn		2016 Pennsylvania	OSS			A Python Automated Machine Learning tool that optimizes machine learning pipelines using genetic programming.
TransmogrifAI	Modeling & Training	AutoML	Salesforce		2017 Bay Area	OSS			an AutoML library for building modular, reusable, strongly typed machine learning workflows on Apache Spark with minimal hand-tuning
DAWNBench	Modeling & Training	Benchmarking	Stanford		2018 Bay Area	OSS			DAWNBench is a benchmark suite for end-to-end deep learning training and inference.
MLPerf	Modeling & Training	Benchmarking			2018 Bay Area	OSS			Fair and useful benchmarks for measuring training and inference performance of ML hardware, software, and services.
Enflame	Hardware	Chips		278.5	2018 China			https://www.enflame-tech.com/	Enflame's AI development focuses on deep learning approaches used in cloud data centers, and are aimed at cloud AI training applications.
Argo	Serving	CI/CD	Intuit		2018 Bay Area	OSS		https://argoproj.github.io/	Get stuff done with Kubernetes. Open source Kubernetes native workflows, events, CI and CD
RelicX	Serving	CI/CD		0	2020 Bay Area		New	https://relicx.ai	RelicX is a venture funded startup building an AI DevOps platform that brings CX intelligence into the CI/CD pipeline to ensure software release readiness based on real user behavior and customer experience.
Anyscale	Infrastructure	Cloud management	B	60.6	2019 Bay Area		Raised 40M	https://www.anyscale.com/	From the creators of Ray, a framework for building machine learning applications at any scale originating from the UC Berkeley RISELab.
Cloudera	Infrastructure	Cloud management	IPO	1000	2008 Bay Area			https://www.cloudera.com/	Cloudera delivers an Enterprise Data Cloud for any data, anywhere, from the Edge to AI.
Datadog	Infrastructure	Cloud management	IPO	147.9	2010 NYC			https://www.datadoghq.com/	See inside any stack, any app, at any scale, anywhere.
Domino Data Lab	Infrastructure	Cloud management	E	123.6	2013 Bay Area		Raised 43M	https://www.dominodatalab.com/	Deliver winning models. One place for your data science tools, apps, results, models, and knowledge
FloydHub	Infrastructure	Cloud management	Seed	4	2016 Bay Area			https://www.floydhub.com/	FloydHub is a zero setup Deep Learning platform for productive data science teams.
HYCU	Infrastructure	Cloud management		0	2009 Boston			https://www.hycu.com/	Keep hyper-converged infrastructure running with HYCU's powerful, simple backup & recovery and monitoring solutions. Deploy in seconds for superior results.
Paperspace	Infrastructure	Cloud management	A	23	2014 NYC			https://www.paperspace.com/	GPU cloud tools built for developers. Powering next-generation workflows and the future of intelligent applications.
MMdnn	Serving	Compatibility	Microsoft		2017 Seattle	OSS			MMdnn is a set of tools to help users inter-operate among different deep learning frameworks. E.g. model conversion and visualization. Convert models between Caffe, Keras, MXNet, Tensorflow, CNTK, PyTorch Onnx and CoreML.
ONNX	Serving	Compatibility			2018	OSS		https://onnx.ai/	ONNX is an open format to help represent machine learning models. ONNX defines a common set of operators - the building blocks of machine learning and deep learning models - and a common file format to enable AI developers to use models with a variety of frameworks, tools,
Picsell.ia	Modeling & Training	Computer Vision		0	2020 France		Raised 18.6M, new	https://picsellia.com/	Picsell.ia is a development platform dedicated to Computer Vision. From open-source to business, you can create and review datasets, track your experiments and follow your project in a Lean AI mode.
Supervisely	All-in-one	Computer vision		0	2017 Bay Area			https://supervise.ly/	First available ecosystem to cover all aspects of training data development. Manage, annotate, validate and experiment with your data without coding.
Matroid	Modeling & Training	Computer vision	B	33.5	2016 Bay Area		Raised 20M	https://www.matroid.com/	Computer vision made simple. Deploy computer vision solutions in minutes, not months.
DefinedCrowd	Data pipeline	Data generation	B	63.6	2015 Seattle		Raised 50.5M	https://www.definedcrowd.com/	Leverage machine learning technology and human intelligence to source, structure, and enrich high quality training data in speech, NLP, and computer vision.
Kimono Labs	Data pipeline	Data generation	Palantir	5	2014 Bay Area			http://www.kimonolabs.com/	Kimono Labs is [ACQ]
Scale AI	Data pipeline	Data generation	D	277.6	2016 Bay Area		Raised 155M	https://scale.com	Trusted by world class companies, Scale delivers high quality training data for AI applications such as self-driving cars, mapping, AR/VR, robotics, and more.
Scrapinghub	Data pipeline	Data generation		0	2010 Ireland	OSS		https://scrapinghub.com/	Turn websites into ELI5
Synthetaic	Data pipeline	Data generation		4.5	2019 Wisconsin		Raised 4.5M, new	https://www.synthetaic.com/	We grow high-quality data that unlocks impossible AI. What if edge cases no longer existed? What if training data was no longer a constraint?
Databricks	All-in-one	Data management	F	897	2013 Bay Area			https://databricks.com/	Unified Data Analytics Platform - One cloud platform for massive scale data engineering and collaborative data science.
Petuum	All-in-one	Data management	B	108	2016 Pennsylvania			https://petuum.com/	Petuum accelerates and simplifies AI solutions so your enterprise can deploy it easily and maintain it effortlessly.
Alluxio	Data pipeline	Data management	B	23	2015 Bay Area	OSS	Raised 7M	https://www.alluxio.io/	an open source data orchestration layer that brings data close to compute for big data and AI/ML workloads in the cloud.
Alteryx	Data pipeline	Data management	IPO	163	2011 LA			https://www.alteryx.com/	We are a leader in the self-service data analytics movement with a platform that can discover, prep, and analyze all your data, then deploy and share analytics at scale for deeper insights faster than you ever thought possible.
Aparavi	Data pipeline	Data management		0	2016 LA			https://www.aparavi.com/	Aparavi's highly scalable data intelligence and automation solutions enable organizations to easily discover, classify, protect, and optimize their data.
Ascend.io	Data pipeline	Data management	A	19	2015 Bay Area			https://www.ascend.io/	Experience continuously optimized data pipelines with less code and fewer breakages. Enter the new era of data engineering with Ascend's autonomous dataflow service.
AtScale	Data pipeline	Data management	D	95	2013 Bay Area			https://www.atscale.com/	Freedom of choice for the enterprise. Break free the complexities and security risks associated with cloud migration and self-service analytics with Intelligent Data Virtualization—no matter where dat.
Cazena	Data pipeline	Data management		38	2014 Boston		Raised unknown	https://www.cazena.com/	First Data Lake with a SaaS Experience. Cazena empowers enterprises to collect, store and analyze any data in the cloud, without any DevOps resources or admin time. Cazena's Data Lake as a Service includes everything, and is delivered as secure SaaS, ready to load, store and
Cohesity	Data pipeline	Data management	E	410	2013 Bay Area		Raised 250M	https://www.cohesity.com/	Eliminate mass data fragmentation with Cohesity's modern approach to data management, beginning with backup. Gain instant recovery. Learn more today.
Dremio	Data pipeline	Data management	C	126.5	2015 Bay Area		Raised 81.5M	https://www.dremio.com/	Get more value f founders of the Apache Arrow and Apache Drill
erwin	Data pipeline	Data management	Parallax Capital Partners		2016 NYC			https://erwin.com/	Integrated enter [ACQ]
Gemini Data	Data pipeline	Data management		0	2015 Bay Area			https://www.geminidata.com/	Gemini Data provides Data Availability for AI/ML driven analysis and applications to enable unified enterprise knowledge and access.
Gravi9	Data pipeline	Data management	Seed	10	2019 China			https://www.gravi9.cn/	Unstructured data management experts, provide AI developers and development teams with data hosting, version management, online visualization, data collaboration and other services. You can use developer tools online to integrate and use data in the cloud
Igneous	Data pipeline	Data management	Rubrik	67.5	2013 Seattle		Acquired	https://www.igneous.io/	Igneous Unstructured Data Protection offers the scalability to handle hundreds of file systems, billions of files, and exabytes of enterprise data requiring backup
Imply	Data pipeline	Data management	B	45.3	2015 Bay Area			https://imply.io/	Imply delivers real-time analytics powered by Apache Druid. ... Stream or batch load data into Druid for high performance, ad-hoc analytic queries.
Octopai	Data pipeline	Data management	B	6.2	2015 Israel			https://www.octopai.com/	An automated, centralized, cross-platform metadata search engine that enables BI groups to quickly and precisely discover and govern shared metadata.
Rubrik	Data pipeline	Data management	E	553	2013 Bay Area		Acquired Igneous	https://www.rubrik.com/en	We provide a powerful, policy-driven platform to simplify recovery and unlock insights from data residing in the data center and cloud.
Superb AI	Data pipeline	Data management	Seed	2.3	2018 Bay Area			https://www.superb-ai.com/	Create, label and manage ML training data efficiently so you can build AI faster. Fully managed workforce. Powerful labeling tools. Training data quality control.
Tamr	Data pipeline	Data management		69.2	2012 Boston		Raised unknown	https://www.tamr.com/	Tamr's leading data management system and services work to create a data migration strategy that simplifies your data unification process. Talk with us today.
Waterline Data	Data pipeline	Data management	Hitachi Vantara	37.5	2013 Bay Area			https://www.waterlinedata.com/	Waterline's enter [ACQ]
Anodot	Data pipeline	Data monitoring	C	62.5	2014 Israel		Raised 35M	https://www.anodot.com/	We monitor your business. Anodot monitors all your data in real time for lightning fast detection of the incidents that impact your revenue
Datagrok	All-in-one	Data processing		0	2019 Pennsylvania			https://datagrok.ai/	Datagrok. Swiss Army Knife for Data. A platform for turning data into actionable insights
cuDF	Data pipeline	Data processing	NVIDIA		2018 Bay Area	OSS		https://rapids.ai/	Built based on the Apache Arrow columnar memory format, cuDF is a GPU DataFrame library for loading, joining, aggregating, filtering, and otherwise manipulating data.
Dask	Data pipeline	Data processing	Seed	5	2015 Remote	OSS	Became a startup	https://dask.org/	Dask natively scales Python. Dask provides advanced parallelism for analytics, enabling performance at scale for the tools you love
Datatable	Data pipeline	Data processing	H2O		2017 Bay Area	OSS		https://datatable.readthedocs.io/en/latest/	Python library for efficient multi-threaded data processing, with the support for out-of-memory datasets.
Incorta	Data pipeline	Data processing	C	72.6	2013 Bay Area			https://incorta.com/	Incorta aggregates large complex business data in real time, eliminating the need to reshape it. No Data Warehouse. No Transformations. Real-Time Insight.
Koalas	Data pipeline	Data processing	Databricks		2019 Bay Area	OSS		https://github.com/databricks/koalas	The Koalas project makes data scientists more productive when interacting with big data, by implementing the pandas DataFrame API on top of Apache Spark.
Modin	Data pipeline	Data processing		0	2018	OSS		https://github.com/modin-project/modin	Modin uses Ray to provide an effortless way to speed up your pandas notebooks, scripts, and libraries. Unlike other distributed DataFrame libraries, Modin provides seamless integration and compatibility with existing pandas code. Even using the DataFrame constructor is identical.
Naveego	Data pipeline	Data processing	Seed	0.5	2014 Michigan			https://www.naveego.com/	A leading provider of cloud-first, distributed data accuracy solutions for seamless, end-to-end data cleansing, Naveego enables organizations to proactively manage, detect and eliminate data accuracy issues across all enterprise data sources in real-time–regardless of structure or
Spark	Data pipeline	Data processing	Apache		2009	OSS		https://spark.apache.org/	Apache Spark is a unified analytics engine for big data processing, with built-in modules for streaming, SQL, machine learning and graph processing.
Vaex	Data pipeline	Data processing		0	2015 Netherlands	OSS		https://vaex.io/	Power up your bi Lore
Amazon Redshift	Data pipeline	Data warehouse	Amazon		2012 Seattle			https://aws.amazon.com/redshift/	Amazon Redshift is a fast, fully managed, and cost-effective data warehouse that gives you petabyte scale data warehousing and exabyte scale data lake analytics together in one service. Amazon Redshift is up to ten times faster than traditional on-premises data warehouses.
Apache Hudi	Data pipeline	Data warehouse	Uber		2016 Bay Area	OSS		https://hudi.apache.org/	Apache Hudi ingests & manages storage of large analytical datasets over DFS (hdfs or cloud stores)
Delta Lake	Data pipeline	Data warehouse	Databricks		2019 Bay Area	OSS		https://delta.io/	Delta Lake is an open-source storage layer that brings ACID transactions to Apache Spark™ and big data workloads.
Yellowbrick Data	Data pipeline	Data warehouse	C	173	2014 Bay Area			https://yellowbrick.com/	The ultimate solution for your data warehouse. Quick to deploy, easy to expand, and simple to manage. Yellowbrick Data can solve your data problems.
Amundsen	Data pipeline	Database/Query	Lyft		2019 Bay Area	OSS			Amundsen is a metadata driven application for improving the productivity of data analysts, data scientists and engineers when interacting with data.
Apache Druid	Data pipeline	Database/Query	Imply		2012 Bay Area	OSS		https://druid.apache.org/	Apache Druid is a high performance real-time analytics database
AresDB	Data pipeline	Database/Query	Uber		2019 Bay Area	OSS		https://github.com/uber/aresdb	A GPU-powered real-time analytics storage and query engine.
Fluree	Data pipeline	Database/Query	Seed	7.2	2017 North Carolina		Raised 2.5M	https://flur.ee/	Welcome to better data management. The Fluree platform organizes blockchain-secured data in a highly-scalable, highly-insightful graph database.
Hammerspace	Data pipeline	Database/Query		0	2015 Bay Area			https://hammerspace.com/	Hammerspace allows data to move freely, like the air you breathe, across clouds and services. Make data accessible exactly where you need it, when you need it – on demand.
Kyvos Insights	Data pipeline	Database/Query		0	2015 Bay Area			https://www.kyvosinsights.com/	Kyvos accelerates BI on trillions of rows of data on the cloud and on-premise platforms with a semantic layer powered by its next-generation OLAP technology.
Milvus	Data pipeline	Database/Query	Zilliz		2019 China	OSS		https://milvus.io/	Milvus is an open source similarity search engine for massive feature vectors. Designed with heterogeneous computing architecture for the best cost efficiency. Searches over billion-scale vectors take only milliseconds with minimum computing resources.
Pilosa	Data pipeline	Database/Query	Seed	3.7	2017 Texas			https://www.pilosa.com/	Pilosa is an open source, distributed bitmap index that dramatically accelerates continuous analysis across multiple queries.
Presto	Data pipeline	Database/Query	Linux Foundation		2012 Bay Area	OSS		https://prestodb.io/	Presto is an open source distributed SQL query engine for running interactive analytic queries against data sources of all sizes ranging from gigabytes to petabytes.
Rockset	Data pipeline	Database/Query	B	61.5	2016 Bay Area		Raised 40M	https://rockset.com/	Rockset: The Real-Time Indexing Database in the Cloud Rockset allows you to build data-driven applications on MongoDB, DynamoDB... AI. Test, validate and deploy models faster by analyzing live data in real-time.
SQLFlow	Data pipeline	Database/Query	Alibaba		2019 China	OSS		https://sqlflow.org	Extends SQL to support AI. Extract knowledge from Data. Currently support MySQL, Apache Hive, Alibaba MaxCompute, XGBoost and TensorFlow.
Starburst Data	Data pipeline	Database/Query	B	64	2017 Boston		Raised 42M	https://www.starburstdata.com/	Limitless Queries. Break boundaries and harness the power of the world's fastest SQL query engine.
TerminusDB	Data pipeline	Database/Query	Seed	1.2	2017 Ireland	OSS		https://terminusdb.com/	TerminusDB is an open source model and database for knowledge graph representation designed specifically for the web-age.
Vearch	Data pipeline	Database/Query		0	2019 China	OSS		https://vearch.github.io/	Vearch is the vector search infrastructure for deeping learning and AI applications.
Zilliz	Data pipeline	Database/Query	B	53	2017 China		Raised 43M	https://zilliz.com/	The company specializes in the development of open-source, AI-powered unstructured data analysis software, and is the initiator and primary contributor to the vector similarity search project Milvus.
Tecton	All-in-one	Deployment	B	60	2019 Bay Area		Raised 55M	https://tecton.ai/	The Data Platform for Machine Learning. Build a library of great features. Serve them in production. Do it at scale.
Inferrd	Serving	Deployment		0	2020 NYC		New	https://inferrd.com/	You build the model, we handle the deployment. Inferrd is the easiest, cheapest and the most performant hosting provider for ML models.
Kubeflow	Serving	Deployment	Google		2018 Bay Area	OSS		https://www.kubeflow.org/	The Kubeflow project is dedicated to making deployments of machine learning (ML) workflows on Kubernetes simple, portable and scalable. Our goal is not to recreate other services, but to provide a straightforward way to deploy best-of-breed open-source systems for ML to diverse
OctoML	Serving	Deployment	A	18.9	2019 Seattle		Raised 15M		Optimize machine learning and deep learning models for deployment. From the creators of Apache TVM, XGBoost and Apache MxNet, OctoML brings the cutting edge of AI, Systems, programming languages, compilers and architecture to make machine learning systems easier to
TensorFlow Extended	Serving	Deployment	Google		2019 Bay Area	OSS		https://www.tensorflow.org/tfx	TensorFlow Extended (TFX) is an end-to-end platform for deploying production ML pipelines
Angel ML	Modeling & Training	Distributed	Tencent/Peking University		2017 Shenzhen	OSS			A Flexible and Powerful Parameter Server for large-scale machine learning
euler	Modeling & Training	Distributed	Alibaba		2018 China	OSS			A distributed graph deep learning framework
Horovod	Modeling & Training	Distributed	Uber		2017 Bay Area	OSS			Horovod is a distributed deep learning training framework for TensorFlow, Keras, PyTorch, and Apache MXNet. The goal of Horovod is to make distributed deep learning fast and easy to use.
Paddle	Modeling & Training	Distributed	Baidu		2016 China	OSS			PArallel Distributed Deep LEarning: Machine Learning Framework from Industrial Practice
Ray	Modeling & Training	Distributed	UC Berkeley		2016 Bay Area	OSS			Ray is a fast and simple framework for building and running distributed applications.
Grid AI	Modeling & Training	Distributed training	A	18.6	2020 NYC		Raised 18.6M, new	https://www.grid.ai/	Seamlessly train hundreds of Machine Learning models on the cloud from your laptop. Focus on machine learning, not infrastructure.
Blaize	Hardware	Edge devices	C	65	2010 Sacramento			https://www.blaize.com/	Intelligence at the edge of everywhere. Blaize unleashes the potential of AI to drive leaps in the value that technology delivers to transform markets and improve the way we all work and live.
Boulder AI	Hardware	Edge devices		0	2017 Boulder			https://boulderai.com/	Human insight and decision making on a visual sensor.
BrainChip	Hardware	Edge devices	IPO	27.8	2006 LA			https://brainchipinc.com/	BrainChip brings artificial intelligence to the edge with a high-performance, small, ultra-low power solution that enables continuous learning and inference.
EdgeQ	Hardware	Edge devices	A	53	2018 Bay Area		Raised 51M, out	https://www.edgeq.io/	EdgeQ is an information technology company that specializes in the fields of 5G chip systems.
GreenWaves Technologies	Hardware	Edge devices	A	12.5	2014 France			https://greenwaves-technologies.com/	GreenWaves' GAP8 is the industry's first ultra-low-power processor enabling battery-operated AI in IoT applications.
Habana Labs	Hardware	Edge devices	Intel	75	2016 Israel			https://habana.ai/	Habana Labs was founded in 2016 to create world-class AI Processors, developed from the ground-up and optimized for training deep neural networks and for inference deployment in production environments.
Hailo	Hardware	Edge devices	B	87.9	2017 Israel		Raised 60M	https://hailo.ai/	The World's Top Performing AI Processor for Edge Devices Hailo offers a breakthrough microprocessor uniquely designed to accelerate embedded AI applications on edge devices. Breathe life into your edge AI product today with Hailo-8.
Kneron	Hardware	Edge devices	A	73	2015 San Diego		Raised 40M + unknown	https://www.kneron.com/	Kneron develops an application-specific integrated circuit and software that offers artificial intelligence-based tools.
LeapMind	Hardware	Edge devices	C	50	2012 Tokyo			https://leapmind.io/	Ultra-low power consumption AI inference accelerator IP specialized for inference arithmetic processing of CNN that operates as a circuit on FPGA device or ASIC device.
Mythic	Hardware	Edge devices	B	85.2	2012 Bay Area			https://www.mythic-ai.com/	An architecture built from the ground up for AI Mythic has developed a truly unique AI compute platform that enables smart camera systems, intelligent appliances, brilliant robotics, and more.
Prophesee	Hardware	Edge devices	C	65.3	2014 France			https://www.prophesee.ai/	With the world's most advanced Event-Based Vision systems, inspired by human vision and built on the foundation of neuromorphic engineering, PROPHESEE is the technology system that gives Metavision to machines, revealing what was previously invisible to them
SiMa.ai	Hardware	Edge devices	A	30	2018 Bay Area		Raised 30M	https://sima.ai/	Is your ML Green?TM We believe that the future of compute is high performance machine learning at the edge – and today, power is the limiter.
Syntiant	Hardware	Edge devices	C	65.1	2017 LA		Raised 30M	https://www.syntiant.com/	Always-On Voice powered by custom AI Silicon
Zero ASIC	Hardware	Edge devices		0	2020 Boston		Changed name for	https://www.zeroasic.com/	Removing the Barrier to Custom Silicon
MLFlow	All-in-one	Experiment tracking	Databricks		2018 Bay Area	OSS		https://mlflow.org/	An open source platform for the machine learning lifecycle
Allegro AI/TRAIN	Modeling & Training	Experiment tracking	A	11	2016 Israel	OSS		https://allegro.ai/	Deep learning platform tailored for computer vision. Allegro AI offers the first end-to-end machine learning product life-cycle management solution.
Comet	Modeling & Training	Experiment tracking	A	6.8	2017 NYC		Raised 4.5M	https://www.comet.ml/	Comet lets you track code, experiments, and results on ML projects. It's fast, simple, and free for open source projects.
Datmo	Modeling & Training	Experiment tracking	One Concern	1.2	2016 Bay Area		Acqui-Hired	https://www.datmo.com/	Be as effective as AI engineers at Google and Facebook. Workflow tools to help you experiment, deploy, and scale. By data scientists, for data scientists.
Neptune	Modeling & Training	Experiment tracking	A	4.7	2017 Poland		Raised 3M	https://neptune.ai/	All experiment-related objects relevant to your projects organized, ready to be analyzed, discovered and shared with your team.
Spell	Modeling & Training	Experiment tracking	A	15	2017 NYC			https://spell.run/	Spell is a powerful platform for building and managing machine learning projects. Spell takes care of infrastructure, making machine learning projects easier to start, faster to get results, more organized and safer than managing infrastructure on your own.
Tensorboard	Modeling & Training	Experiment tracking	Google		2015 Bay Area	OSS		https://www.tensorflow.org/	TensorBoard is a tool for providing the measurements and visualizations needed during the machine learning workflow. It enables tracking experiment metrics like loss and accuracy, visualizing the model graph, projecting embeddings to a lower dimensional space, and much more.
Weights & Biases	Modeling & Training	Experiment tracking	B	20	2017 Bay Area			https://www.wandb.com/	We're building developer tools for deep learning. Add a couple lines of code to your training script and we'll keep track of your hyperparameters, system metrics, and outputs so you can compare experiments,
DarwinAI	Modeling & Training	Explainability	Seed	3	2017 Canada		Raised unknown	https://www.darwinai.com/	DarwinAI's Generative Synthesis 'AI building AI' technology enables optimized and explainable deep learning.
Truera	Modeling & Training	Explainability	A	17.3	2019 Bay Area		Raised 17.3M	https://truera.com/	The Truera Model Intelligence Platform powered by Enterprise-Class AI Explainability eliminates the machine learning black box with Model Intelligence.
Truera	Modeling & Training	Explainability	A	17.3	2019 Bay Area		Raised 17.3M	https://truera.com/	The Truera Model Intelligence Platform powered by Enterprise-Class AI Explainability eliminates the machine learning black box with Model Intelligence.
dotData	All-in-one	Feature engineering		43	2018 Bay Area			https://dotdata.com/	When AutoML is enhanced with AI-powered feature engineering, the result is dotData. We focus on delivering data science automation for the enterprise. End-to-end data science automation platform accelerates, democratizes, and operationalizes the entire data science process.
FEAST	Data pipeline	Feature engineering		0	2019 Asia	OSS		https://feast.dev/	Feast (Feature Store) is a tool for managing and serving machine learning features. Feast is the bridge between models and data.
Boruta	Modeling & Training	Feature engineering	scikit-learn		2010	OSS		https://github.com/scikit-learn-contrib/boruta_py	Python implementations of the Boruta all-relevant feature selection method.
Featuretools	Modeling & Training	Feature engineering	Alteryx		2018 LA	OSS		https://www.featuretools.com/	An open source python library for automated feature engineering
scribble Data	Modeling & Training	Feature engineering		0	2016 India			https://www.scribbledata.io/	The feature store for your ML engineering needs
tsfresh	Modeling & Training	Feature engineering	Blue Yonder		2008 Germany	OSS		https://tsfresh.readthedocs.io/	Automatic extraction of relevant features from time series
Apache ORC	Data pipeline	File format	Apache		2013 Bay Area	OSS		https://orc.apache.org/	the smallest, fastest columnar storage for Hadoop workloads.
Parquet	Data pipeline	File format	Twitter, Cloudera		2013 Bay Area	OSS			Apache Parquet is a columnar storage format available to any project in the Hadoop ecosystem, regardless of the choice of data processing framework, data model or programming language.
Accord	Modeling & Training	Framework		0	2012 France	OSS	Archived	http://accord-framework.net/	Machine learning, computer vision, statistics and general scientific computing for .NET
Alink	Modeling & Training	Framework	Alibaba		2018 China	OSS			Alink is the Machine Learning algorithm platform based on Flink, developed by the PAI team of Alibaba computing platform.
Apache Mahout	Modeling & Training	Framework	Apache		2008 Remote	OSS		https://mahout.apache.org/	Apache Mahout(TM) is a distributed linear algebra framework and mathematically expressive Scala DSL designed to let mathematicians, statisticians, and data scientists quickly implement their own algorithms. Apache Spark is the recommended out-of-the-box distributed back-end
Apache MXNet	Modeling & Training	Framework	DMLC		2015	OSS			A flexible and efficient library for deep learning.
Caffe	Modeling & Training	Framework	Berkeley		2013 Bay Area	OSS			Caffe: a fast open framework for deep learning.
Catalyst	Modeling & Training	Framework		0	2018 Russia	OSS		https://catalyst-team.github.io/catalyst/	PyTorch framework for Deep Learning research and development. It focuses on reproducibility, rapid experimentation, and codebase reuse so you can create something new rather than write another regular train loop.
Chainer	Modeling & Training	Framework	Preferred Networks		2015 Tokyo	OSS		https://chainer.org/	A Powerful, Flexible, and Intuitive Framework for Neural Networks
FedAI (FATE)	Modeling & Training	Framework	Webank		2019 Shenzhen	OSS			FATE (Federated AI Technology Enabler) is an open-source project initiated by Webank's AI Department to provide a secure computing framework to support the federated AI ecosystem. It implements secure computation protocols based on homomorphic encryption and multi-party
Gensim	Modeling & Training	Framework	RaRe Technologies		2012 Czech	OSS		https://radimrehurek.com/gensim/	Topic Modelling for Humans. Gensim is a Python library for topic modelling, document indexing and similarity retrieval with large corpora. Target audience is the natural language processing (NLP) and information retrieval (IR) community.
JAX	Modeling & Training	Framework	Google		2018 Bay Area	OSS			Composable transformations of Python+NumPy programs: differentiate, vectorize, JIT to GPU/TPU, and more
LightGBM	Modeling & Training	Framework	Microsoft		2016 Seattle	OSS		https://github.com/microsoft/LightGBM	A fast, distributed, high performance gradient boosting (GBT, GBDT, GBRT, GBM or MART) framework based on decision tree algorithms, used for ranking, classification and many other machine learning tasks.
Ludwig	Modeling & Training	Framework	Uber		2019 Bay Area	OSS		http://ludwig.ai	Ludwig is a toolbox built on top of TensorFlow that allows to train and test deep learning models without the need to write code.
Mindspore	Modeling & Training	Framework	Huawei		2020 Shenzhen	OSS	New		MindSpore is a new open source deep learning training/inference framework that could be used for mobile, edge and cloud scenarios
ML.NET	Modeling & Training	Framework	Microsoft		2018 Seattle	OSS		https://dotnet.microsoft.com/	ML.NET is an open source and cross-platform machine learning framework for .NET
MLlib	Modeling & Training	Framework	Spark		2010	OSS		https://spark.apache.org/mllib/	MLlib is Apache Spark's scalable machine learning library.
Pythia	Modeling & Training	Framework	Facebook		2018 Bay Area	OSS			A modular framework for vision & language multimodal research from Facebook AI Research (FAIR).
PyTorch	Modeling & Training	Framework	Facebook		2016 Bay Area	OSS		https://pytorch.org/	Tools & Libraries. A rich ecosystem of tools and libraries extends PyTorch and supports development in computer vision, NLP and more
PyTorch Lightning	Modeling & Training	Framework	Grid AI		2019 NYC	OSS		https://www.pytorchlightning.ai/	The lightweight PyTorch wrapper for high-performance AI research. Scale your models, not the boilerplate.
scikit-learn	Modeling & Training	Framework		0	2010 Remote	OSS		https://scikit-learn.org/stable/	Machine Learning in Python
TensorFlow	Modeling & Training	Framework	Google		2015 Bay Area	OSS		https://www.tensorflow.org/	An end-to-end open source machine learning platform for everyone. Discover TensorFlow's flexible ecosystem of tools, libraries and community resources
Theano	Modeling & Training	Framework	MILA		2008 Canada	OSS			Theano is a Python library that allows you to define, optimize, and evaluate mathematical expressions involving multi-dimensional arrays efficiently
Turi Create	Modeling & Training	Framework	Apple		2018 Bay Area	OSS		https://apple.github.io/turicreate/docs/api/	Turi Create simplifies the development of custom machine learning models.
XGBoost	Modeling & Training	Framework	DMLC		2014	OSS			XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable. It implements machine learning algorithms under the Gradient Boosting framework.
PlaidML	Modeling & Training	Hardware compatibility	Intel		2018 Bay Area	OSS			PlaidML is a framework for making deep learning work everywhere.
HyperOpt	Modeling & Training	Hyperparameter tuning	U. Waterloo		2013 Canada	OSS		http://hyperopt.github.io/hyperopt/	Distributed Asynchronous Hyperparameter Optimization in Python - hyperopt/hyperopt
Katib	Modeling & Training	Hyperparameter tuning	Kubeflow		2018 Bay Area	OSS			Katib is a Kubernetes Native System for Hyperparameter Tuning and Neural Architecture Search. The system is inspired by Google vizier and supports multiple ML/DL frameworks (e.g. TensorFlow, MXNet, and PyTorch).
SigOpt	Modeling & Training	Hyperparameter tuning	Intel	8.7	2014 Bay Area		Acquired	https://sigopt.com/	SigOpt is a standardized, scalable, enterprise-grade optimization platform and API designed to unlock the potential of your modeling pipelines.
talos	Modeling & Training	Hyperparameter tuning	Autonomio		2018 Finland	OSS		https://autonom.io	Hyperparameter Optimization for TensorFlow, Keras and PyTorch
Tune	Modeling & Training	Hyperparameter tuning	Ray		2017 Bay Area	OSS		https://docs.ray.io/en/latest/tune.html	Tune is a Python library for hyperparameter tuning at any scale.
Apache TVM	Serving	Inference		0	2017	OSS		https://tvm.apache.org/	Apache TVM (incubating) is a compiler stack for deep learning systems. It is designed to close the gap between the productivity-focused deep learning frameworks, and the performance- and efficiency-focused hardware backends. TVM works with deep learning frameworks to prov
MNN	Serving	Inference	Alibaba		2019 China	OSS			MNN is a lightweight deep neural network inference engine.
TensorRT	Serving	Inference	NVIDIA		2019 Bay Area	OSS			NVIDIA TensorRT™ is an SDK for high-performance deep learning inference. It includes a deep learning inference optimizer and runtime that delivers low latency and high-throughput for deep learning inference applications.
explainX.ai	Modeling & Training	Interpretability	Pre-seed	0	2020 NYC		New	https://www.explainx.ai/	ExplainX enables you to explain, pinpoint, and monitor how your AI models work. We make sure your models never fail in the real-world.
Fiddler Labs	Modeling & Training	Interpretability	A	13.2	2018 Bay Area		Raised unknown	https://www.fiddler.ai/	AI with trust, visibility, and insights built in. Fiddler is a breakthrough AI engine with explainability at its heart.
Fiddler Labs	Modeling & Training	Interpretability	A	13.2	2018 Bay Area		Raised unknown	https://www.fiddler.ai/	AI with trust, visibility, and insights built in. Fiddler is a breakthrough AI engine with explainability at its heart.

Company	Cat	SubCat	Series	$$$ (M)	Started	HQ	OSS	Change in 2020	Website	Description
InterpretML	Modeling & Training	Interpretability	Microsoft		2019	Seattle	OSS			Fit interpretable machine learning models. Explain blackbox machine learning
Kyndi	Modeling & Training	Interpretability	B	20	2019	Bay Area		Raised 20M	www.kyndi.com	
LIME	Modeling & Training	Interpretability	University of Washington		2016	Seattle	OSS			Lime: Explaining the predictions of any machine learning classifier
Lucid	Modeling & Training	Interpretability	Google		2018	Bay Area	OSS			A collection of infrastructure and tools for research in neural network interpretability
SHAP	Modeling & Training	Interpretability	University of Washington		2017	Seattle	OSS			A game theoretic approach to explain the output of any machine learning model.
HIVE	All-in-one	Labeling	B	20.2	2013	Bay Area			https://thehive.ai	Hive is a full-stack deep learning company focused on solving visual intelligence problems. Let us help you join the AI Revolution. End-To-End Solutions. Full-Stack Approach.
Dataturks	Data pipeline	Labeling	Walmart		2018	India			https://dataturks.com/	ML data annotation
Doccano	Data pipeline	Labeling		0	2018	Tokyo	OSS		https://doccano.herokuapp.com/	Text annotation for Human. Just create project, upload data and start annotation. You can build dataset in hours.
Figure Eight	Data pipeline	Labeling	Appen	300	2008	Bay Area			https://www.figure-eight.com/	Figure Eight com ACQ
Heartex Label St	Data pipeline	Labeling	Seed	0.15	2018	Bay Area	OSS		https://www.heartex.ai/	Label Studio is a multi-type data labeling and annotation tool with standardized output format
iMerit	Data pipeline	Labeling	B	23.5	2012	Bay Area		Raised 20M	https://imerit.net/	iMerit specializes in data labeling and annotation for purposes of training models for Machine Learning and Artificial Intelligence.
Labelbox	Data pipeline	Labeling	B	42.8	2018	Bay Area		Raised 25M	https://labelbox.com/	A complete solution for your training data problem with fast labeling tools, human workforce, data management, a powerful API and automation features.
Labelimg	Data pipeline	Labeling		0	2016	Canada	OSS		https://github.com/tzutalin/labelImg	Labelimg is a graphical image annotation tool and label object bounding boxes in images
Playment	Data pipeline	Labeling	Seed	2.5	2015	India			https://playment.io/	Build high-quality ground truth datasets with ML-assisted tools, sophisticated project management software, expert human workforce, and much more.
Prodigy	Data pipeline	Labeling	Explosion AI		2017	Germany			https://prodi.gy/	Prodigy is a scriptable annotation tool so efficient that data scientists can do the annotation themselves, enabling a new level of rapid iteration. ... With Prodigy you can take full advantage of modern machine learning by adopting a more agile approach to data collection.
Segments.ai	Data pipeline	Labeling		0	2020	Belgium		New	https://segments.ai/	Deep learning-fueled labeling technology with a focus on instance and semantic segmentation.
Snorkel	Data pipeline	Labeling	Snorkel AI		2016	Bay Area	OSS		https://snorkel.ai	Programmatically Building and Managing Training Data
V7Labs	Data pipeline	Labeling	Seed	3	2018	UK		Raised 3M	https://www.v7labs.com/	Create the Sense of Sight Label, train, and deploy artificial intelligence that effortlessly learns new objects from your data.
Voxel51 // Scoop	Data pipeline	Labeling	Seed	3.3	2018	Michigan			https://voxel51.com/scoop/	We build software that enables ML engineers to build better models, more quickly. Try FiftyOne, our powerful platform for dataset curation, analysis, and model
Core ML	Serving	Mobile	Apple		2017	Bay Area				Use Core ML to integrate machine learning models into your app. Core ML provides a unified representation for all models.
Fritz AI	Serving	Mobile	A	7	2017	Boston		Raised 5M	https://www.fritz.ai/	Fritz AI is the machine learning platform for iOS and Android developers. Teach your mobile apps to see, hear, sense, and think.
ML Kit	Serving	Mobile	Google		2018	Bay Area	OSS			ML Kit beta brings Google's machine learning expertise to mobile developers in a powerful and easy-to-use package.
ncnn	Serving	Mobile	Tencent		2017	Shenzhen	OSS			ncnn is a high-performance neural network inference framework optimized for the mobile platform
TensorFlow Lite	Serving	Mobile	Google		2019	Bay Area	OSS		https://www.tensorflow.org/lite	TensorFlow Lite is an open source deep learning framework for on-device inference.
AIMET	Modeling & Training	Model compression	Qualcomm		2020	San Diego	OSS	New	https://github.com/quic/aimet	AIMET is a library that provides advanced quantization and compression techniques for trained neural network models.
Deeplite	Serving	Model compression		0	2020	Canada		New	https://www.deeplite.ai/	Enabling faster, smaller and more energy-efficient DNNs to run on edge devices and in the cloud
Neural Network C	Serving	Model compression	Intel		2018	Bay Area	OSS			Distiller is an open-source Python package for neural network compression research. Network compression can reduce the memory footprint of a neural network, increase its inference speed and save energy. Distiller provides a PyTorch environment for prototyping and analyzing c
Xnor.ai	Serving	Model compression	Apple	14.6	2016	Seattle		Acquired 200M	http://xnor.ai/	Transform your ti ACQ
Dessa	All-in-one	Monitoring	Square	9	2016	Canada		Acquired	https://www.dessa.com/	Create more with ACQ
Prometheus	Data pipeline	Monitoring	Apache/PromLabs		2012	Germany	OSS		https://prometheus.io/	An open-source monitoring system with a dimensional data model, flexible query language, efficient time series database and modern alerting approach.
Arize AI	Serving	Monitoring	Seed	4	2019	Bay Area			https://arize.com/	Arize AI is the watcher, troubleshooter and the guardrail on deployed AI
Arthur AI	Serving	Monitoring	A	18.3	2018	NYC		Raised 15M	https://www.arthur.ai/	Always-on Explainability, Bias, and Performance Monitoring for AI, ML, and analytics. Get up and running in minutes and start sleeping better at night. Dedicated. Innovative.
Datatron	Serving	Monitoring	Seed	7.8	2016	Bay Area		Raised 1.4M	https://www.datatron.com/	Production AI Model Management at Scale. Automate the standardized deployment, monitoring, governance, and validation of all your models to be developed in any environment.
Evidently AI	Serving	Monitoring		0	2020	Russia	OSS	New	https://evidentlyai.com/	Open-source tools to analyze, monitor, and debug machine learning model in production
Losswise	Serving	Monitoring	Mathpix		2017	Bay Area			https://losswise.com/	Turn your GPUs into monitored build servers from a git push with Losswise. Interactive visualization, logs, smart notifications, and more. Start free today.
Mona Labs	Serving	Monitoring	Seed	3.9	2018	Atlanta		Raised 3.9M	https://www.monalabs.io/	PRODUCTION MONITORING.FOR.AI. With Mona, you gain complete transparency into how your data and models behave in the real world.
superwise.ai	Serving	Monitoring	Seed	4.5	2019	Israel		Raised 4.5M	https://superwise.ai	Monitor your AI from the moment it meets reality so you can finally trust every model
Unravel Data	Serving	Monitoring	C	57.2	2013	Bay Area			https://unraveldata.com/	Unravel provides AppDynamics, New Relic, Splunk, and DataDog
VertaAI	Serving	Monitoring	A	10	2019	Bay Area		Raised 10M	https://www.verta.ai/	Verta.AI is a Palo Alto-based startup building software infrastructure to help enterprise data science and machine learning (ML) teams rapidly develop and deploy ML models.
AllenNLP	Modeling & Training	NLP	Incubator		2016	Seattle	OSS		https://allennlp.org/	AllenNLP is an open-source NLP research library, built on PyTorch.
fastText	Modeling & Training	NLP	Facebook		2016	Bay Area	OSS			Library for fast text representation and classification.
flair	Modeling & Training	NLP	Zalando		2018	Germany	OSS			A very simple framework for state-of-the-art Natural Language Processing (NLP)
Hugging Face	Modeling & Training	NLP	A	20.2	2016	NYC			https://huggingface.co/	We're on a journey to solve and democratize artificial intelligence through natural language.
OpenSeq2Seq	Modeling & Training	NLP	NVIDIA		2017	Bay Area	OSS			Toolkit for efficient experimentation with Speech Recognition, Text2Speech and NLP
spaCy	Modeling & Training	NLP	Explosion AI	0	2014	Germany	OSS		https://spacy.io/	spaCy is a free open-source library for Natural Language Processing in Python. It features NER, POS tagging, dependency parsing, word vectors and more.
Dialogflow	Modeling & Training	NLU	Google		2014	Bay Area			https://dialogflow.com/	Dialogflow is a Google service that runs on Google Cloud Platform, letting you scale to hundreds of millions of users. Optimized for the Google Assistant.
NeMo	Modeling & Training	NLU	NVIDIA		2019	Bay Area	OSS			NeMo: a toolkit for conversational AI
Rasa	Modeling & Training	NLU	B	40.1	2016	Germany	OSS	Raised 26M	https://rasa.com/	Build contextual AI assistants and chatbots in text and voice with our open source machine learning framework. Scale it with our enterprise grade platform.
Colab	Modeling & Training	Notebook	Google		2017	Bay Area			https://colab.research.google.com/	Colab notebooks allow you to combine executable code and rich text in a single document, along with images, HTML, LaTeX and more.
DeepNote	Modeling & Training	Notebook	Seed	3.8	2019	Bay Area		Raised 3.8M	https://deepnote.com/	The notebook you'll love to use Deepnote is a new kind of data science notebook. Jupyter-compatible with real-time collaboration and easy deployment. Oh, and it's free.
nteract	Modeling & Training	Notebook	Nteract		2015	Texas	OSS		https://nteract.io/	nteract is an open-source organization committed to creating fantastic interactive computing experiences that allow people to collaborate with ease. We build SDKs, applications, and libraries that help you and your team make the most of interactive (particularly Jupyter) notebooks
papermill	Modeling & Training	Notebook	Nteract		2017	Texas	OSS		https://papermill.readthedocs.io/en/latest/	Papermill is a tool for parameterizing and executing Jupyter Notebooks.
River	Modeling & Training	Online learning		0	2017	France	OSS	Changed name f	https://riverml.xyz/	A Python package for online/streaming machine learning.
Vowpal Wabbit	Modeling & Training	Online learning	Microsoft		2010		OSS		https://vowpalwabbit.org/	Vowpal Wabbit provides a fast, flexible, online, and active learning solution that empowers you to solve complex interactive machine learning problems
BentoML	Modeling & Training	Pretrained models		0	2018	Bay Area	OSS		https://bentoml.com/	BentoML makes it easy to serve and deploy machine learning models in the cloud. It is an open source framework for building cloud-native model serving services. BentoML supports most popular ML training frameworks and deployment platforms, including major cloud providers a
Dockship	Modeling & Training	Pretrained models		0	2019	India	OSS		https://dockship.io/	Dockship.io is a marketplace for AI models and datasets. Publish your models on Dockship for people all over the world.
GluonCV	Modeling & Training	Pretrained models	Microsoft		2018	Seattle	OSS		https://gluon-cv.mxnet.io/	GluonCV provides implementations of state-of-the-art (SOTA) deep learning algorithms in computer vision. It aims to help engineers, researchers, and students quickly prototype products, validate new ideas and learn computer vision.
Aircloak	Data pipeline	Privacy		1.3	2012	Germany	OSS		https://aircloak.com/	Aircloak's unique approach ensures the existing primary database is not modified in any way. Aircloak handles all data types including unstructured text.
Gretel AI	Data pipeline	Privacy	A	15.5	2019	San Diego		Raised 15.5M	https://gretel.ai/	The first and only APIs to enable you to balance, anonymize, and share your data. With privacy guarantees.
Tumult Labs	Data pipeline	Privacy		0	2019	North Carolina			https://www.tmlt.io/	Unleashing the power of data with ironclad privacy protection
PySyft	Modeling & Training	Privacy	OpenMined		2017	UK	OSS			PySyft is a Python library for secure and private Deep Learning. PySyft decouples private data from model training, using Federated Learning, Differential Privacy, and Multi-Party Computation (MPC) within the main Deep Learning frameworks like PyTorch and TensorFlow.
Pyro	Modeling & Training	Programming language	Uber		2017	Bay Area	OSS			Pyro is a flexible, scalable deep probabilistic programming library built on PyTorch
Robust AI	All-in-one	Robotics	A	22.5	2019	Bay Area		Raised 15M	https://www.robust.ai/	Robust.AI: Creating a New Foundation for the Future of Robotics.
Formant	Serving	Robotics	A	6	2019	Bay Area			https://formant.io/	Deploy faster. Improve uptime. Achieve scale.
Able	All-in-one	Serving		26.4	2018	Bay Area		Raised 26.4M	https://www.aible.com/	Create AI that delivers impact, not accuracy, with cost-benefit tradeoffs & operational constraints, in a friendly, intuitive UI designed for real business.
Polyaxon	All-in-one	Serving		0	2016	Germany	OSS		https://polyaxon.com/	A platform for reproducing and managing the whole life cycle of machine learning and deep learning applications.
Algorithmia	Serving	Serving	B	37.9	2013	Seattle			https://algorithmia.com/	Algorithmia makes applications smarter, by building a community around algorithm development, where state of the art algorithms are always live and accessible to anyone
Seldon	Serving	Serving	A	13.7	2011	UK	OSS	Raised 9.4M	https://www.seldon.io/	Manage, serve and scale models built in any framework on Kubernetes. Take your ML projects from POC to production.
ClearSky Data	Data pipeline	Storage	B	59	2014	Boston			https://www.clearskydata.com/	ClearSky Data offers enterprise storage as a hybrid cloud service delivering on-demand primary storage, offsite backup, and DR as a single service.
Datera	Data pipeline	Storage	C	63.9	2013	Bay Area			https://datera.io/	Get sub-200μS latency & millions of IOPS with 100% software-defined data automation. Save up to 70% on data infrastructure total-cost-of-ownership.
Elastifile	Data pipeline	Storage	Google	74	2013	Bay Area			https://www.elastifile.com/	Elastifile's cloud- ACQ
Excelero	Data pipeline	Storage	B	35	2014	Bay Area			https://www.excelero.com/	Local NVMe performance at data center scale through true convergence. Software-defined block storage for Cloud and Enterprise applications at any scale.
Komprise	Data pipeline	Storage	C	50.7	2014	Bay Area			https://www.komprise.com/	In 15 minutes, our free data management software trial will show you how you can save 70% on data management costs, on-premises and in the cloud.
Quobyte	Data pipeline	Storage	B	0	2013	Germany			https://www.quobyte.com/	Quobyte is software defined storage that turns commodity servers into a reliable and highly automated data center file system.
Storbyte	Data pipeline	Storage		0	2014	DC			http://storbyte.com/	Storbyte designs and manufactures all-flash & hybrid flash enterprise storage arrays that offer performance, power management, availability, reliability, density, efficiency, flexibility, expandability, and affordability. Storbyte is providing innovative data storage solutions and has not lost
Vexata	Data pipeline	Storage	StorCentric	54	2014	Bay Area			https://www.vexata.com/	Vexata is an act ACQ
Confluent	Data pipeline	Stream processing	E	455.9	2014	Bay Area		Raised 250M, hir	https://www.confluent.io/	Confluent is a fully managed Kafka service and enterprise stream processing platform. Real-time data streaming for AWS, GCP, Azure or serverless. Try free!
Materialize	Data pipeline	Stream processing	B	40.5	2019	NYC		Raised 32M	https://materialize.com/	Materialize delivers SQL exploration for streaming events and real-time data. Incrementally updated materialized views - in ANSI Standard SQL and in real-time. Micro-batching.
Apache Flink	Serving	Stream processing	Apache		2011	Germany	OSS			Apache Flink is an open source stream processing framework with powerful stream- and batch-processing capabilities
Apache Kafka	Serving	Stream storage	Apache		2011	Bay Area	OSS			Apache Kafka is an open-source stream-processing software platform developed by LinkedIn and donated to the Apache Software Foundation, written in Scala and Java. The project aims to provide a unified, high-throughput, low-latency platform for handling real-time data feeds.
Company	Cat	SubCat	Series	$$$ (M)	Started	HQ	OSS	Change in 2020	Website	Description — IF ACQ
Dolt	Data pipeline	Versioning	Seed	2	2018	LA			https://www.liquidata.co/	Liquidata's mission is to make data move more efficiently. We built Dolt, an an open-source version-controlled SQL database with Git-like semantics.
DVC - Iterative.ai	Data pipeline	Versioning		3.9	2017	Bay Area	OSS		https://dvc.org/	Open-source version control system for Data Science and Machine Learning projects. Git-like experience to organize your data, models, and experiments.
Git LFS	Data pipeline	Versioning	Atlassian, GitHub		2014	Remote	OSS		https://git-lfs.github.com/	Git Large File Storage (LFS) replaces large files such as audio samples, videos, datasets, and graphics with text pointers inside Git, while storing the file contents on a remote server like GitHub.com or GitHub Enterprise.
Pachyderm	Data pipeline	Versioning	B	28.1	2014	Bay Area	OSS	Raised 16M	https://www.pachyderm.com/	Data Lineage with End-to-End Pipelines on Kubernetes, engineered for the enterprise. And… It's open source!
Qri	Data pipeline	Versioning		0	2016	NYC	OSS		https://qri.io/	Bigger than a spreadsheet, smaller than a database, datasets are all around us. Use Qri to browse, download, create, fork, & publish datasets across a network of peers.
Quilt Data	Data pipeline	Versioning	Seed	4.2	2015	Bay Area	OSS		https://quiltdata.com/	Quilt is a versioned data portal for AWS
DAGsHub	Modeling & Training	Versioning	Seed	3	2019	Israel		Raised 3M	https://dagshub.com/	DAGsHub is a platform for data version control and collaboration for data scientists and machine learning engineers.
Replicate	Modeling & Training	Versioning		0	2020	Bay Area	OSS	New	https://replicate.ai	Version control for machine learning
PerceptiLabs	Modeling & Training	Visual modeling	Seed	2	2019	Bay Area			https://perceptilabs.readme.io/	PerceptiLabs takes the process of building and training a machine learning model to warp speed. We not only accelerate machine learning, we advance explainability in AI
Facets	Data pipeline	Visualization	Google		2017	Bay Area	OSS		https://github.com/PAIR-code/facets	Facets: An Open Source Visualization Tool for Machine Learning Training Data
Gluent	Data pipeline	Visualization	Seed	8.2	2014	Texas			https://gluent.com/	Data virtualization software eliminates data silos. Gluent's transparent data virtualization provides virtual access to all enterprise data, with zero code changes.
Netron	Modeling & Training	Visualization		0	2011	Seattle	OSS		https://netron.app/	Netron is a viewer for neural network, deep learning and machine learning models.
Clipper	Serving	Web	Berkeley		2017	Bay Area	OSS		http://clipper.ai	Clipper is a low-latency prediction serving system for machine learning. Clipper makes it simple to integrate machine learning into user-facing serving systems.
Cortex	Serving	Web		0	2019	Bay Area	OSS		https://www.cortex.dev/	Cortex is an open source platform for deploying machine learning models as production web services.
PredictionIO	Serving	Web	Salesforce		2013	Bay Area	OSS		https://predictionio.apache.org/	Apache PredictionIO is an open source machine learning framework for developers, data scientists, and end users. It supports event collection, deployment of algorithms, evaluation, querying predictive results via REST APIs. It is based on scalable open source services like Hadoop
ElementI	All-in-one	Workflow orchestration		1.8	2018	Bay Area	OSS	Raised 1.8M	https://www.elementl.com/	Building Dagster, the data orchestrator. Dagster is a data orchestrator for machine learning, analytics, and ETL.
Michelangelo	All-in-one	Workflow orchestration	Uber		2015	Bay Area				Michelangelo, Uber's machine learning (ML) platform, supports the training and serving of thousands of models in production across the company. Designed to cover the end-to-end ML workflow, the system currently supports classical machine learning, time series forecasting, and
Airflow	Infrastructure	Workflow orchestration	Airbnb		2015	Bay Area	OSS		https://airflow.apache.org/	Airflow is a platform created by community to programmatically author, schedule and monitor workflows.
Backend AI	Infrastructure	Workflow orchestration		0	2016	Seoul	OSS		https://www.backend.ai	Backend.AI: Minute-made GPU clustering solution for Machine Learning.
Cadence	Infrastructure	Workflow orchestration	Uber		2017	Bay Area	OSS			Cadence is a distributed, scalable, durable, and highly available orchestration engine to execute asynchronous long-running business logic in a scalable and resilient way.
Flyte	Infrastructure	Workflow orchestration	Lyft		2019	Bay Area	OSS	Prepared to beco	https://flyte.org/	Lyft's Cloud Native Machine Learning and Data Processing Platform, Now Open Sourced
Luigi	Infrastructure	Workflow orchestration	Spotify		2012	Sweden	OSS			Luigi is a Python module that helps you build complex pipelines of batch jobs. It handles dependency resolution, workflow management, visualization etc. It also comes with Hadoop support built in.
Metaflow	Infrastructure	Workflow orchestration	Netflix		2019	Bay Area	OSS		https://metaflow.org/	Metaflow makes metaflow.org
Prefect	Infrastructure	Workflow orchestration	A	14.1	2018	DC		Raised 11.6M	https://www.prefect.io/	The Global Leader in Dataflow Automation
Valohai	Infrastructure	Workflow orchestration	A	1.8	2016	Finland			https://valohai.com/	The MLOps platform for the whole team. Valohai takes you from POC to production while managing the whole model lifecycle.

Cat	COUNTA of Name	SUM of $$$ (M)
All-in-one	27	$2,553.80
Data pipeline	90	$3,805.55
Hardware	22	$2,752.70
Infrastructure	15	$1,375.00
Modeling & Training	91	$301.22
Serving	39	$223.60

COUNTA of Name and SUM of $$$ (M)

All-in-one
9.5%

Data pipeline
31.7%

Hardware
7.7%

Infrastructure
5.3%

Modeling & Training
32.0%

Serving
13.7%

Name	Cat	SubCat	Series	$$$ (M)	Started	HQ	OSS	Change in 2020	Website	Description	IF ACQ
Graphcore	Hardware	Accelerator	E	682	2016	UK		Raised 372M	https://www.graphcore.ai/	With the world's most advanced Event-Based Vision systems, inspired by human vision and built on the foundation of neuromorphic engineering, PROPHESEE is the revolutionary system that gives Metavision to machines, revealing what was previously invisible to them.	
SambaNova	Hardware	Accelerator	C	465.3	2017	Bay Area		Raised 250M	https://sambanova.ai/	AI insights, faster Cerebras is a computer systems company dedicated to accelerating deep learning. The pioneering Wafer-Scale Engine (WSE) – the largest chip ever built – is at the heart of our deep learning system, the Cerebras CS-1.	
Nuvia	Hardware	Accelerator	B	293	2019	Bay Area		Raised 240M	https://nuviainc.com/	SambaNova Systems is a computing startup focused on building machine learning and big data analytics platforms.	
Wave Computing	Hardware	Accelerator	E	203.3	2008	Bay Area			https://wavecomp.ai/	Human insight and decision making on a visual sensor.	
Cambricon	Hardware	Accelerator	B	200	2016	China			http://www.cambricon.com/	Always-On Voice powered by custom AI Silicon	
Cerebras	Hardware	Accelerator	C	112	2016	Bay Area			https://www.cerebras.net/	Intelligence at the edge of everywhere. Blaize unleashes the potential of AI to drive leaps in the value that technology delivers to transform markets and improve the way we all work and live.	
Hailo	Hardware	Edge devices	B	87.9	2017	Israel		Raised 60M	https://hailo.ai/	An architecture built from the ground up for AI Mythic has developed a truly unique AI compute platform that enables smart camera systems, intelligent appliances, brilliant robotics, and more.	
Mythic	Hardware	Edge devices	B	85.2	2012	Bay Area			https://www.mythic-ai.com/	BrainChip brings artificial intelligence to the edge with a high-performance, small, ultra-low power solution that enables continuous learning and inference.	
Habana Labs	Hardware	Edge devices	Intel	75	2016	Israel			https://habana.ai/	EdgeQ is an information technology company that specializes in the fields of 5G chip systems.	
Kneron	Hardware	Edge devices	A	73	2015	San Diego		Raised 40M + unknown	https://www.kneron.com/	Ultra-low power consumption AI inference accelerator IP specialized for inference arithmetic processing of CNN that operates as a circuit on FPGA device or ASIC device .	
Prophesee	Hardware	Edge devices	C	65.3	2014	France			https://www.prophesee.ai/	Accelerate AI, Neuromorphic, AI Chip, Optical Computing, Lightmatter	
Syntiant	Hardware	Edge devices	C	65.1	2017	LA		Raised 35M	https://www.syntiant.com/	Habana Labs was founded in 2016 to create world-class AI Processors, developed from the ground-up and optimized for training deep neural networks and for inference deployment in production environments.	
Blaize	Hardware	Edge devices	C	65	2010	Sacramento			https://www.blaize.com/	Hardware is bottlenecked by data movement & compute. We use photonics to solve both	
Groq	Hardware	Accelerator		62.3	2016	Bay Area		Raised unknown	https://groq.com/	The Next Generation of Computing is here.	
EdgeQ	Hardware	Edge devices	A	53	2018	Bay Area		Raised 51M, out of stealth	https://www.edgeq.io/	Wave Computing is revolutionizing AI and deep learning with its dataflow-based systems and embedded solutions.	
LeapMind	Hardware	Edge devices	C	50	2012	Tokyo			https://leapmind.io/	GreenWaves' GAP8 is the industry's first ultra-low-power processor enabling battery-operated AI in IoT applications.	
Lightelligence	Hardware	Accelerator	A	36	2017	Boston		Raised 26M	https://www.lightelligence.ai/	The World's Top Performing AI Processor for Edge Devices Hailo offers a breakthrough microprocessor uniquely designed to accelerate embedded AI applications on edge devices. Breathe life into your edge AI product today with Hailo-8.	
SiMa.ai	Hardware	Edge devices	A	30	2018	Bay Area		Raised 30M	https://sima.ai/	Cambricon Technologies builds core processor chips for intelligent cloud servers, intelligent terminals, and intelligent robots.	
BrainChip	Hardware	Edge devices	IPO	27.8	2006	LA			https://brainchipinc.com/	Removing the Barrier to Custom Silicon	
GreenWaves Technologies	Hardware	Edge devices	A	12.5	2014	France			https://greenwaves-technologies.com/	Is your ML Green?TM We believe that the future of compute is high performance machine learning at the edge – and today, power is the limiter.	
Luminous Computing	Hardware	Accelerator	A	9	2018	Bay Area		Raised unknown	https://luminous.co/	Silicon design reimagined for a compute-intensive world.	
Zero ASIC	Hardware	Edge devices		0	2020	Boston		Changed name from Adapteva	https://www.zeroasic.com/	Graphcore has built a new type of processor for machine intelligence to accelerate machine learning and AI applications for a world of intelligent machines.	
Boulder AI	Hardware	Edge devices		0	2017	Boulder			https://boulderai.com/	Kneron develops an application-specific integrated circuit and software that offers artificial intelligence-based tools.	