



CORNELL
TECH

Deep Learning Clinic (DLC)

Lecture 11
Real World Tools

Jin Sun

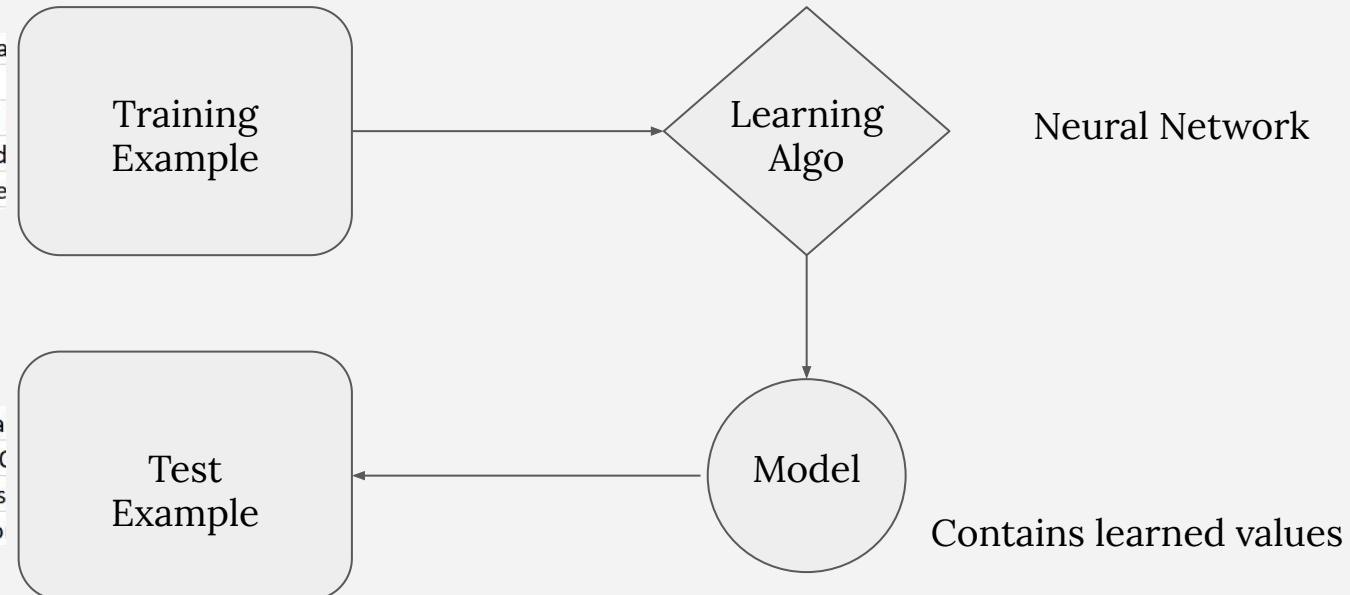
12/3/2019

Today - Real World Tools

- **What You Need for Real World Applications**
- Cloud Services
 - Cloud services comparison
 - Amazon SageMaker
 - Google AutoML
- Vision APIs
- Specialized Tools
 - Detectron by Facebook
 - Face++ for face analysis
 - Natural Language Toolkit (NLTK)
- Where to Find SOTA Models?

Recap: Machine Learning Paradigm

ham	Go until jurong point, crazy.. Availa
ham	Ok lar... Joking wif u oni...
spam	Free entry in 2 a wkly comp to win
ham	U dun say so early hor... U c alread
ham	Nah I don't think he goes to usf, he



What You Need for Real World Applications

ham	Go
ham	Ok
spam	Fre
ham	U d
ham	Nat

Data and labels

ham	Eve
ham	As
spam	Will
spam	Had

Models

Tools to Collect Data

Image data sources:

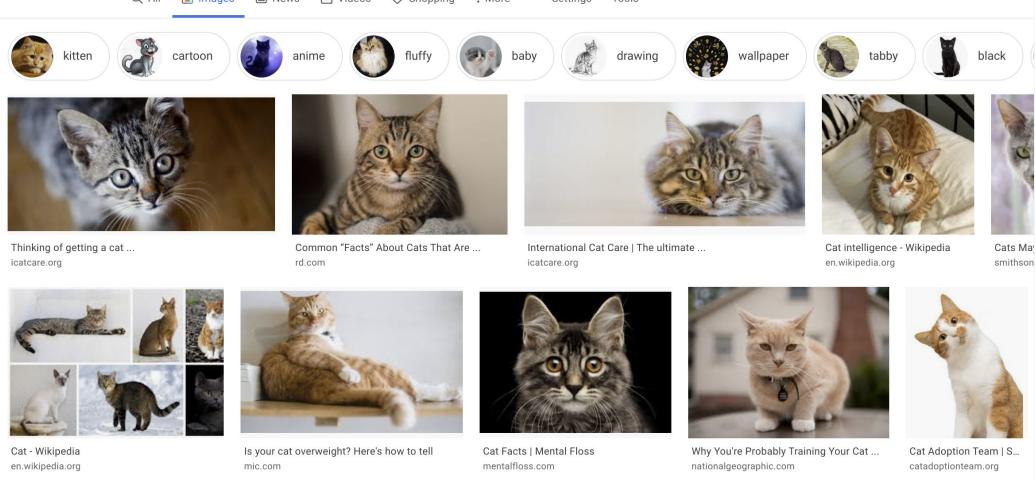
flickr Sign Up Explore Create Upload Photos, people, or groups Sign In

Explore



Google cat

All Images News Videos Shopping More Settings Tools



kitten cartoon anime fluffy baby drawing wallpaper tabby black

Thinking of getting a cat ...
icatcare.org

Common "Facts" About Cats That Are ...
rd.com

International Cat Care | The ultimate ...
icatcare.org

Cat Intelligence - Wikipedia
en.wikipedia.org

Cats Ma... smithson

Cat - Wikipedia
en.wikipedia.org

Is your cat overweight? Here's how to tell
mic.com

Cat Facts | Mental Floss
mentalfloss.com

Why You're Probably Training Your Cat ...
nationalgeographic.com

Cat Adoption Team | S...
catadoptionteam.org

Tools to Collect Data

Google Street View and Satellite



Data



New Datasets

View the most recently published datasets on the data catalog.

Popular Datasets

View some of the most popular datasets on the data catalog.

NYC 2018 Squirrel Census Map



Web Scraping with Python

The screenshot shows a web browser window with the Bloomberg Business website. On the left, the page displays the S&P 500 Index information, including the symbol SPX:IND, current price (2,022.19), change (+32.62), and percentage change (+1.64%). It also shows the open price (1,994.71), day range (1,994.71 - 2,022.37), volume (641,731,328), and previous close (1,989.57). On the right, the browser's developer tools are open, specifically the Elements tab, which shows the HTML structure of the page. The HTML code includes various divs and spans, with the current price (2,022.19) highlighted in the code.

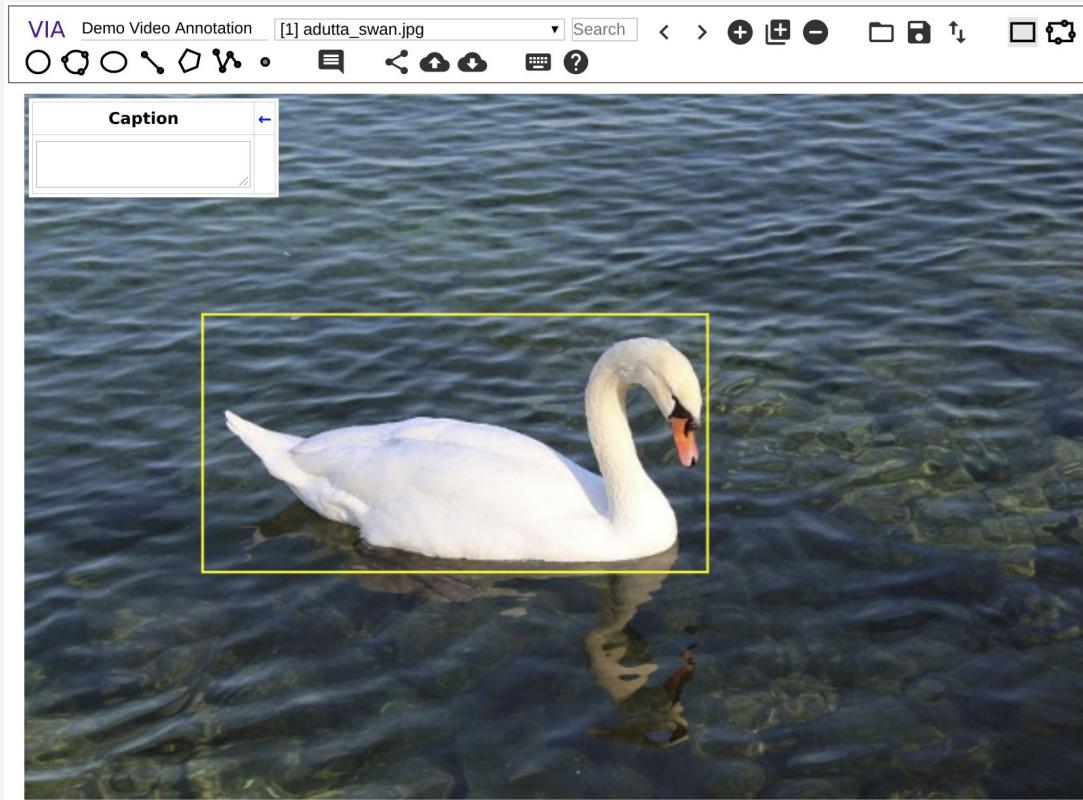
<https://www.freecodecamp.org/news/how-to-scrape-websites-with-python-and-beautifulsoup-5946935d93fe/>

Web page contents can be obtained by retrieving HTML tags.

Python library: BeautifulSoup, CSV

Be careful about copyrights and licensing issues!

Tools to Label Data



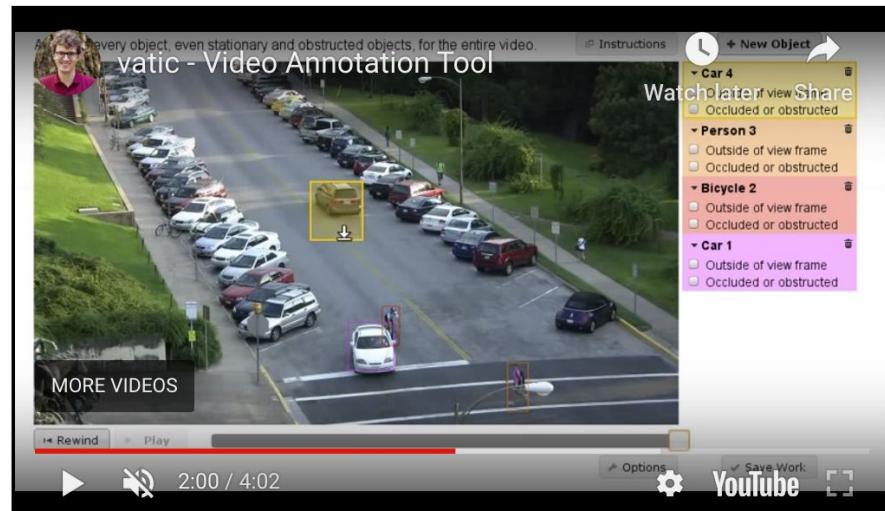
http://www.robots.ox.ac.uk/~vgg/software/via/demo/via_image_annotator.html

Tools to Label Data

Video Annotation Tool from Irvine, California

"These are some of the best HITS I've ever done!" — *Mechanical Turk worker*

vatic is a **free, online, interactive video annotation tool** for computer vision research that crowdsources work to Amazon's Mechanical Turk. Our tool makes it easy to build massive, affordable video data sets and can be deployed on a cloud. After three years of research, vatic is now used by labs around the world to annotate the next generation of data sets.



<http://www.cs.columbia.edu/~vondrick/vatic/>

Tools to Visualize Data

Tableau | Production | Content | Users | Groups | Schedules | Tasks | Status | Settings | Martin Rodriguez

Home > Manufacturing Indicators

Manufacturing Indicators PROJECT

Workbooks 14 Views 18 Data Sources 0 Permissions Details

Sort by Views: All (Most-Least)

General Filters

Owner Tag Modified on or after Modified on or before

Only my favorites Only my recently viewed Has an alert

Execution Overview Product Defects Key Metrics Systems Breakdown

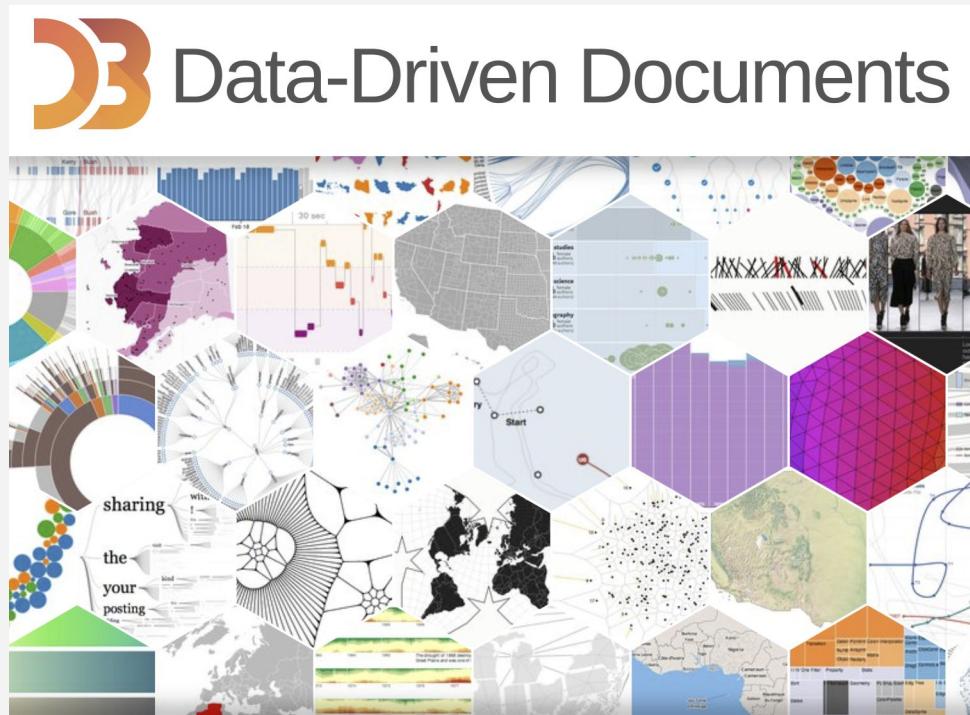
4,381 views 1 star 4,335 views 0 stars 4,199 views 1 star 4,039 views 0 stars

Cluster Analysis Base Sales Opportunities Active Trends Backend Site Performance

3,669 views 1 star 3,110 views 0 stars 2,440 views 0 stars 1,331 views 0 stars

<https://www.tableau.com/>

Tools to Visualize Data



<https://d3js.org/>

Data Collection and/or Labeling Services

The screenshot shows the Amazon Mechanical Turk (MTurk) Requester interface. The top navigation bar includes the MTurk logo, 'REQUESTER', and tabs for 'Home', 'Create', 'Manage', 'Developer', and 'Help'. Below the navigation is a sub-navigation bar with 'New Project' and 'New Batch with an Existing Project' on the left, and 'Create HITs individually' on the right.

The main content area is titled 'Start a New Project' and features a sidebar on the left with various project types: Categorization, Data Collection, Moderation of an Image, **Sentiment** (which is highlighted with an orange arrow), Survey, Survey Link, Tagging of an Image, Transcription from A/V, Transcription from an Image, Writing, and Other.

The central panel is titled 'Example of Sentiment' and contains the following instructions:

Instructions:
Pick the best sentiment based on the following criterion.

Strongly positive	Select this if the item embodies emotion that was extremely happy or excited toward the topic. For example, "Their customer service is the best that I've seen!!!"
Positive	Select this if the item embodies emotion that was generally happy or satisfied, but the emotion wasn't extreme. For example, "Sure I'll shop there again."
Neutral	Select this if the item does not embody much of positive or negative

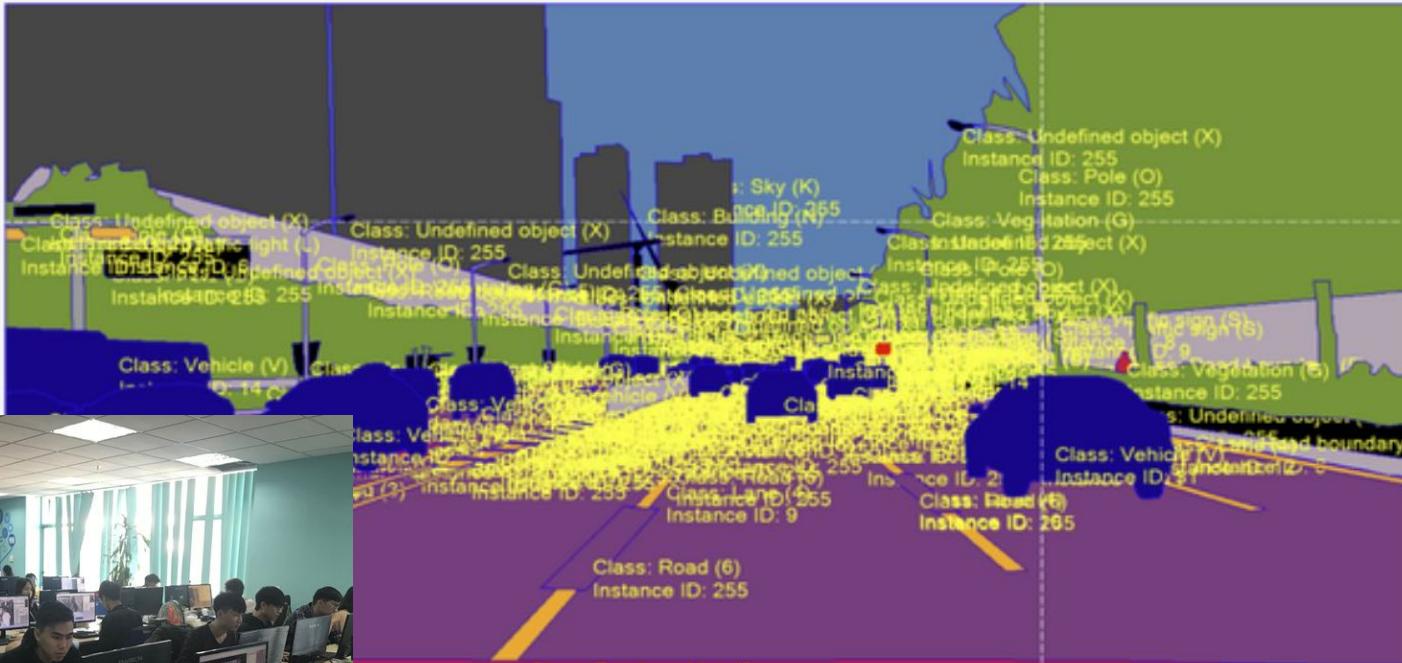
Judge the sentiment expressed by the following item toward: Amazon

If you loved Firefly TV show, amazing Amazon price for entire series: about \$27 BlueRay & \$17 DVD.

A horizontal scale at the bottom ranges from 'Strongly negative' to 'Strongly positive' with intermediate points 'Negative', 'Neutral', and 'Positive'.

Create Project »

Data Collection and/or Labeling Services



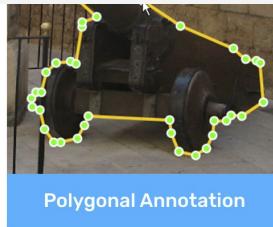
“... annotation company in Vietnam with over 500 plus permanent annotators and with a workforce of over 1000 plus ... the rate can be as low as USD 0.01 per image.”

<https://www.quora.com/Which-are-the-top-image-annotation-companies-with-the-lowest-hour-rate>

Data Collection and/or Labeling Services



Bounding Box



Polygonal Annotation



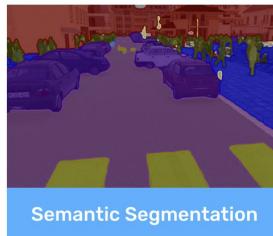
Line Annotation



Geospatial Annotation



Cuboidal Annotation



Semantic Segmentation



Clipping Path



Data Labelling



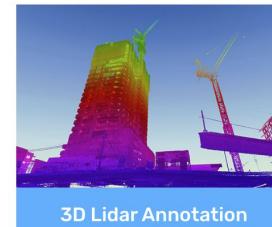
Image Tagging



Autonomous Vehicle



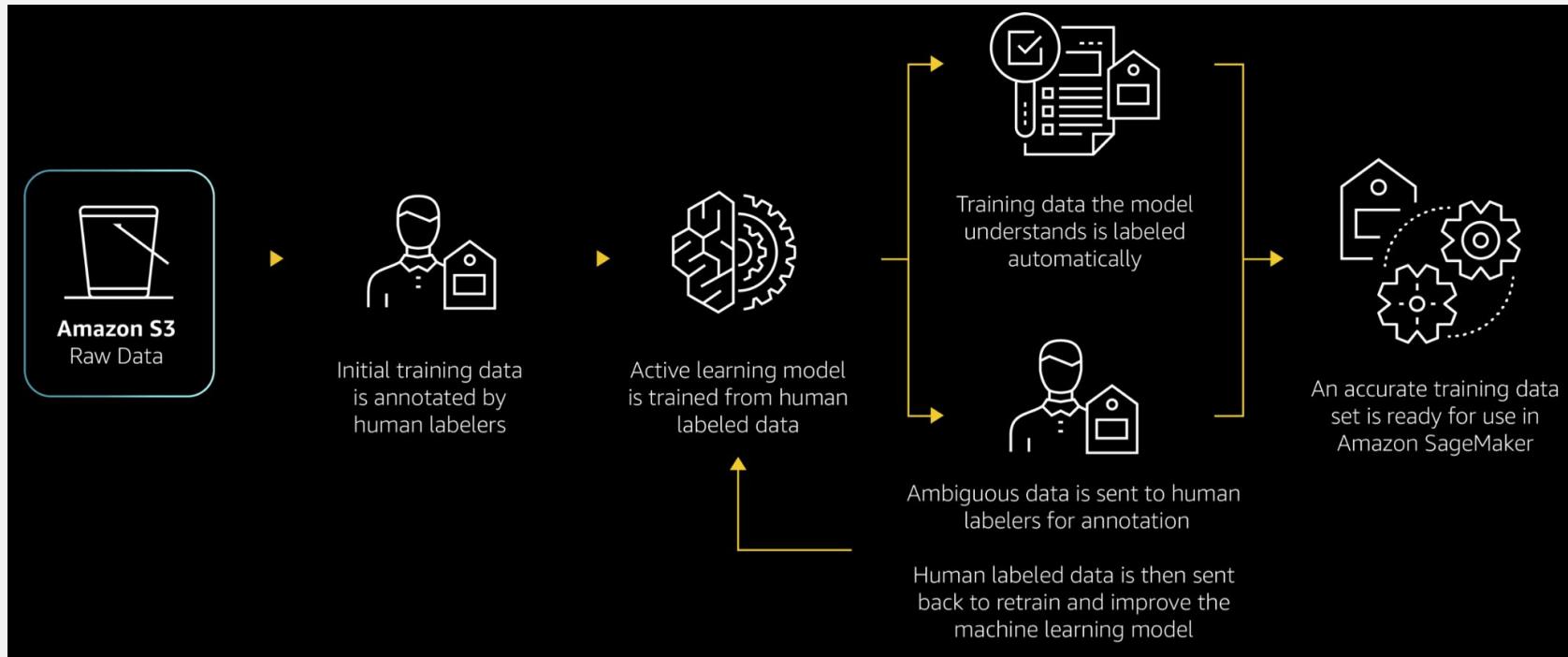
Image Masking



3D Lidar Annotation

Image source: <http://www.greettech.com/>

Data Collection and/or Labeling Services



Data Storage

Amazon S3		Overview	Features	Storage classes	Pricing	Getting started	Resources ▾	FAQs
S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data								
First 50 TB / Month								\$0.023 per GB
Next 450 TB / Month								\$0.022 per GB
Over 500 TB / Month								\$0.021 per GB
S3 Intelligent - Tiering * - Automatic cost savings for data with unknown or changing access patterns								
Frequent Access Tier, First 50 TB / Month								\$0.023 per GB
Frequent Access Tier, Next 450 TB / Month								\$0.022 per GB
Frequent Access Tier, Over 500 TB / Month								\$0.021 per GB
Infrequent Access Tier, All Storage / Month								\$0.0125 per GB
Monitoring and Automation, All Storage / Month								\$0.0025 per 1,000 objects
S3 Standard - Infrequent Access * - For long lived but infrequently accessed data that needs millisecond access								
All Storage / Month								\$0.0125 per GB
S3 One Zone - Infrequent Access * - For re-createable infrequently accessed data that needs millisecond access								
All Storage / Month								\$0.01 per GB
S3 Glacier ** - For long-term backups and archives with retrieval option from 1 minute to 12 hours								
All Storage / Month								\$0.004 per GB

Cloud service providers also have online storage plans. You need those for: cloud based training or outsourcing labeling.

Local Solution to Neural Network Training

What you need to consider when building a PC for deep learning?

GPU Nvidia only, newer models have ‘tensor cores’, need large ram

Hard drive NVMe M.2 SSD > SATA SSD > SATA HDD

If you want to be cautious: server-grade **CPU+memory (ECC)**

Double check **PCIe lanes** on your CPU and motherboard if you plan to use multiple GPUs and/or hard drives

Local Solution to Neural Network Training

Pre-built deep learning machines

Lambda

Servers Workstations TensorBook GPU Cloud Resources +1 (866) 711-2025 [Talk to an engineer](#)

TENSORBOOK GPU Laptop	LAMBDA DUAL 2x GPU Workstation	LAMBDA QUAD 4x GPU Workstation	LAMBDA BLADE 8x GPU Server	LAMBDA HYPERPLANE Tesla V100 Server
				
Deep Learning Laptop with RTX 2070 Max-Q or RTX 2080 Max-Q. Up to 32 GB Memory.	Deep Learning Workstation with RTX 2080 Ti, Titan RTX, RTX 8000, or Titan V GPUs. NVLink Available.	Deep Learning Workstation with RTX 2080 Ti, RTX 6000, RTX 8000, or Titan V GPUs. NVLink Available.	Deep Learning Server with up to 10x GPUs and NVLink. Quadro RTX 8000, RTX 6000, RTX 5000 available.	Tesla V100 Server with NVLink for Deep Learning. Up to 8 GPUs, 768 GB of Memory. Fully Customizable
Learn more	Learn more	Learn more	Learn more	Learn more

Today - Real World Tools

- What You Need for Real World Applications
- **Cloud Services**
 - Cloud services comparison
 - Amazon SageMaker
 - Google AutoML
- Vision APIs
- Specialized Tools
 - Detectron by Facebook
 - Face++ for face analysis
 - Natural Language Toolkit (NLTK)
- Where to Find SOTA Models?

Cloud Service for ML Comparison

Cloud Service	NVIDIA GPU	CPUs	GPU RAM	CPU RAM	Cost Per Hour	Wall Time	Cost to Train
Google Colab	K80	1	12	13	0.00	31.17	0.000
Google Cloud Compute Engine	P100	6	16	20	0.50	5.32	0.044
Google Cloud Compute Engine	K80	6	12	17	0.20	18.13	0.060
Google Cloud Compute Engine	V100	8	16	20	0.82	3.83	0.052
Google Cloud Compute Engine	P4	4	8	26	0.33	10.28	0.057
Google Cloud Compute Engine	V100 x 2	8	32	30	1.57	3.63	0.095
Google Cloud Compute Engine	V100 x 4	8	64	30	3.05	3.38	0.172
AWS EC2	K80 (p2.xlarge)	4	12	61	0.28	20.90	0.098
AWS EC2	K80 x 8 (p2.8xlarge)	32	96	488	2.35	16.12	0.631
AWS EC2	V100 (p3.2xlarge)	8	16	61	1.05	3.85	0.067
AWS EC2	V100 x 4 (p3.8xlarge)	64	128	488	4.05	2.97	0.200
vast.ai	GTX 1070 Ti	4	8.1	16	0.06	7.23	0.008
Paperspace	Quadro M4000	8	8	30	0.51	8.30	0.071

CLOUD MACHINE LEARNING SERVICES COMPARISON

	Amazon	Microsoft	Google	IBM
Automated and semi-automated ML services				
	Amazon ML	Microsoft Azure ML Studio	Cloud AutoML	IBM Watson ML Model Builder
Classification	✓	✓	✓	✓
Regression	✓	✓	✓	✓
Clustering	✓	✓	✗	✗
Anomaly detection	✗	✓	✗	✗
Recommendation	✗	✓	✓	✗
Ranking	✗	✓	✗	✗
Platforms for custom modeling				
	Amazon SageMaker	Azure ML Services	Google ML Engine	IBM Watson ML Studio
Built-in algorithms	✓	✗	✓	✓
Supported frameworks	TensorFlow, MXNet, Keras, Gluon, Pytorch, Caffe2, Chainer, Torch	TensorFlow, scikit-learn, Microsoft Cognitive Toolkit, Spark ML	TensorFlow, scikit-learn, XGBoost, Keras	TensorFlow, Spark MLlib, scikit-learn, XGBoost, PyTorch, IBM SPSS, PMML

Running your own PyTorch code on the cloud



AWS



Google Cloud
Platform



https://cloud.google.com/deep-learning-vm/docs/pytorch_start_instance



Microsoft Azure



Amazon SageMaker

Build, train, and deploy machine learning models fast

Amazon SageMaker enables developers and data scientists to quickly and easily build, train, and deploy machine learning models at any scale. It removes the complexity that gets in the way of successfully implementing machine learning across use cases and industries—from running models for real-time fraud detection, to virtually analyzing biological impacts of potential drugs, to predicting stolen-base success in baseball.

[Learn more »](#)

BUILD

10x better-performing algorithms: Use pre-installed, optimized algorithms for 10 times better performance with Amazon SageMaker.

Pre-built Jupyter notebooks: Fully managed instances for training data and preprocessing.

Algorithm marketplace: Choose from diverse algorithms and models available in the AWS Marketplace for Machine Learning.

70% cost reduction: Build highly accurate training datasets and reduce data-labeling costs by up to 70% with Amazon SageMaker Ground Truth.

TRAIN

One-click training and automatic model tuning with AutoML capabilities for the most accurate model predictions.

Train once, run anywhere with 2x performance: Amazon SageMaker Neo allows you to train your model once and run it on multiple hardware configurations with up 2 times performance increase.

Fully managed: Simplified model training through management all of the underlying infrastructure so you can easily scale models.

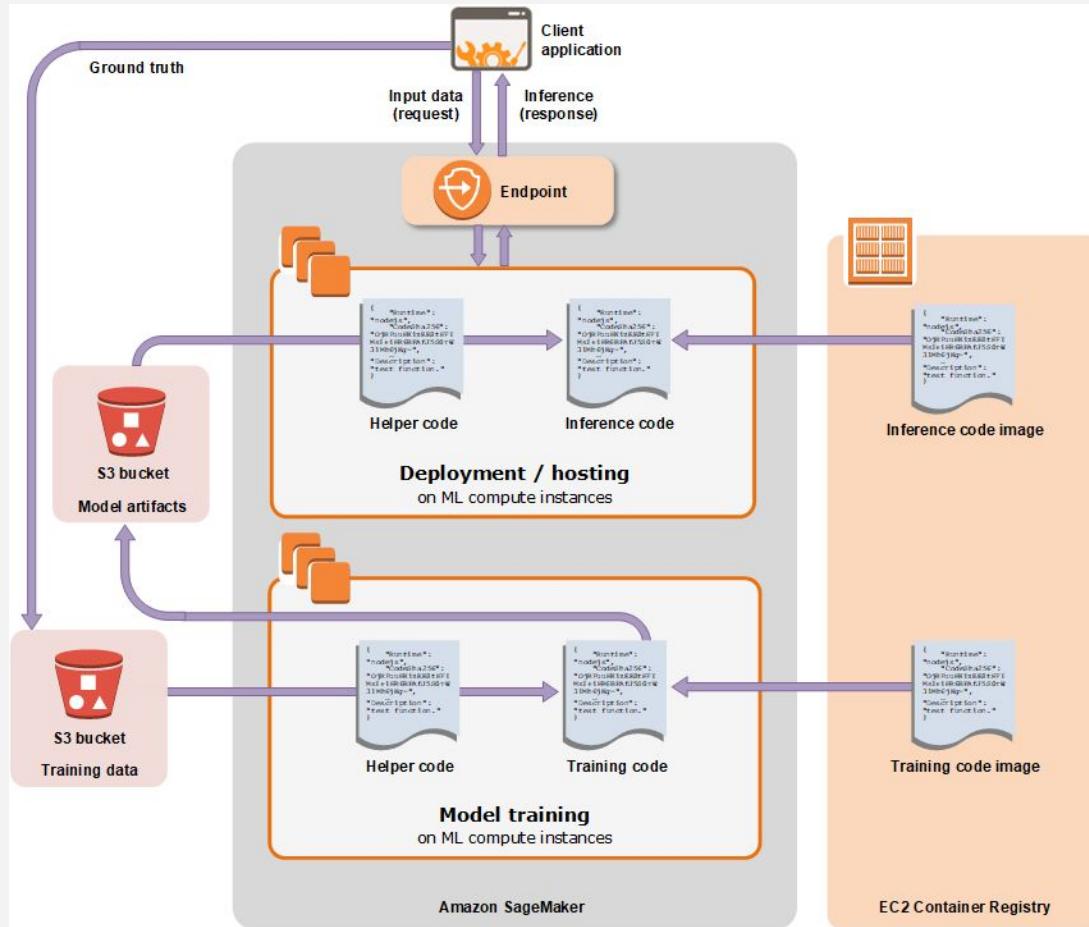
DEPLOY

One-click deployment: Deploy with auto-scaling clusters spread across multiple zones to deliver both high performance and high availability.

Reduce inference costs up to 75%: Amazon Elastic Inference allows you to provision just the right amount of GPU performance you need.

Automatic A/B Testing: Managed model A/B testing across as many as five different models giving you a lot of flexibility to run experiments.

Fully-managed hosting with auto scaling: Managed production infrastructure to perform health checks, apply security patches, and conduct routine maintenance.



<https://docs.aws.amazon.com/sagemaker/latest/dg/how-it-works-mlconcepts.html>



Categories

All Categories

Infrastructure Software (45)

DevOps (36)

Business Applications (6)

Machine Learning (321)

IoT (10)

Industries (96)

Filters

Clear all filters

Vendors

 Perception Health (31) Prosper Insights & Analytics (31) RocketML (23) Mphasis (22) Sensifai (18) Cloudwick (11) Figure Eight (11) Quantiphi (10) Persistent Systems (10) ClosedLoop.ai (9)

Show more

Pricing Plan

 Hourly (227) Free (94)

All Categories (321 results) showing 1 - 10

1 2 3 4 5 ... 33 ►



GluonCV YOLOv3 Object Detector

Version 1.1 | Sold by Amazon Web Services

1 AWS review

Given an input image, this will return object coordinates and category predictions. The format of coordinates is encoded as (left, top, right, bottom) of the absolute pixel locations. This model is trained on COCO dataset with 80 common object categories. It can be used as fast and reliable general...

Model Package - Fulfilled on Amazon SageMaker



Closing Price of Cryptocurrencies

Version 1.0 | Sold by ThingLogix Inc

1 AWS review

This model was created using data using over 785k records from the cryptocurrency marketing dating all the way back to 2013. The set used for the model uses a market data of over 1600 different cryptocurrencies to help make predictions. With only eight easily accessible and attributes making a...

Model Package - Fulfilled on Amazon SageMaker



H2O.ai H2O-3 Automl Algorithm

Version 0.2 | Sold by H2O.ai

H2O's AutoML can be used for automating the machine learning workflow, which includes automatic training and tuning of many models within a user-specified time-limit. Stacked Ensembles – one based on all previously trained models, another one on the best model of each family – will be automatically...

Algorithm - Fulfilled on Amazon SageMaker

link

Google AutoML

No need to find a model, tune hyperparameters, configure deep learning libraries

You provide data, AutoML will train it for you

You can even ask Google to label the data for you

<https://cloud.google.com/vision/automl/docs/beginners-guide>

Currently only support image classification and object detection for vision

Also support natural language processing, generic data classification/regression

Google AutoML

IMPORT IMAGES TRAIN EVALUATE TEST & USE Single-Label Classification

Model cloud_model Confidence threshold 0.5

Filter labels :

- All labels
- daisy
- dandelion
- roses
- sunflowers
- tulips

All labels

Total images 3,299
Test items 367
Precision 96.99%
Recall 96.46%

Use the slider to see which confidence threshold works best for your model on the precision-recall tradeoff curve.
[Learn more about these metrics and graphs.](#)

Precision



Recall

Confusion matrix

True Label	Predicted Label				
	tulips	dandelion	sunflowers	daisy	roses
tulips	98%	-	-	-	3%
dandelion	-	99%	-	1%	-
sunflowers	-	-	96%	1%	3%
daisy	-	-	-	100%	-
roses	6%	-	-	2%	92%

Recall

Precision



Confidence

— Recall — Precision

Today - Real World Tools

- What You Need for Real World Applications
- Cloud Services
 - Cloud services comparison
 - Amazon SageMaker
 - Google AutoML
- **Vision APIs**
- Specialized Tools
 - Detectron by Facebook
 - Face++ for face analysis
 - Natural Language Toolkit (NLTK)
- Where to Find SOTA Models?

Google Product

The screenshot shows the Google Cloud website with the 'Products' tab selected. The main content area is titled 'AI & Machine Learning Products' and features a section for 'VISION AI'. The 'VISION AI' section includes a large heading 'Vision AI', a description of its capabilities, and a 'GO TO CONSOLE' button.

Google Cloud Why Google Solutions **Products** Pricing Getting started

AI & Machine Learning Products

VISION AI

- Overview
- Vision API demo
- Benefits
- Features
- Customers
- Use cases
- Pricing
- Resources

Vision AI

Derive insights from your images in the cloud or at the edge with AutoML Vision or use pre-trained Vision API models to detect emotion, understand text, and more.

GO TO CONSOLE

AutoML vs Cloud API

Customized model vs
Pretrained model

<https://cloud.google.com/vision/>

Google Vision API

A set of services on existing models that provide result on your images

Image

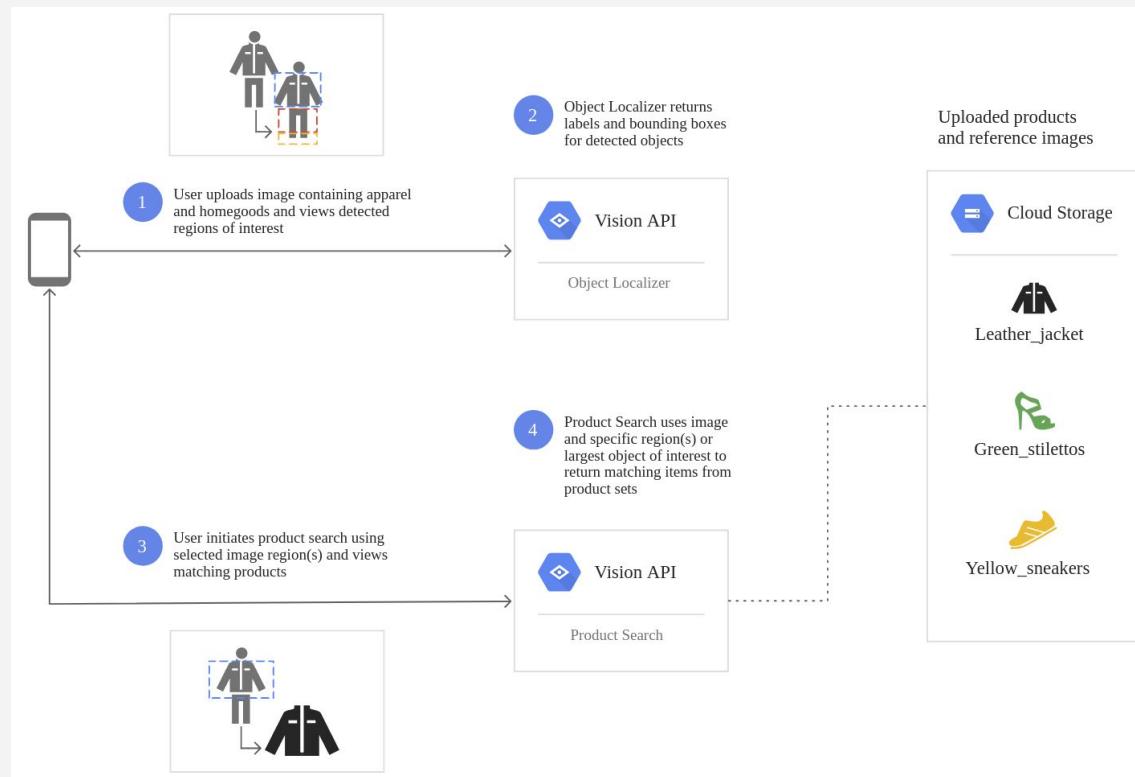
- Classification
- Object Detection
- OCR
- Product Search (Matching)
- Landmark Detection
- Face Detection
- Detect explicit content

Video

- Most of the image functions +
- Scene change detection
- Detect and track objects
- Automatic video transcription

https://cloud.google.com/vision/docs/?_ga=2.3742384.-1411060093.1557114861

Combine multiple cloud vision products together



Comparison of Cloud APIs for CV

	Google	Azure	Rekognition	clarifai	IBM	KAIROS
FACE DETECTION	✓	✓	✓	✓	✓	✓
FACE RECOGNITION	✗	✓	✓	✗	✗	✓
FACIAL LANDMARKS	✓	✓	✓	✗	✗	✓
FEATURE DETECTION	✓	✓	✓	✓	✗	✓
SIMILAR FACES	✗	✓	✓	✗	✗	✓
EMOTION	✓	✓	✓	✗	✗	✓
LABEL DETECTION	✓	✓	✓	✓	✓	✗
LANDMARKS	✓	✓	✗	✗	✓	✗
CELEBRITIES	✗	✓	✓	✓	✗	✗
LOGO DETECTION	✓	✗	✗	✓	✗	✗
OCR	✓	✓	✓	✗	✓	✗
NSFW	✓	✓	✓	✓	✓	✗
IMAGE ANALYSIS	✓	✓	✓	✓	✓	✗
VIDEO ANALYSIS	✓	✓	✓	✗	✗	✓ for faces
CUSTOM MODEL CREATION	✗	✓	✗	✓	✓	✗

[Source](#)

IMAGE ANALYSIS APIs COMPARISON

	Amazon	Microsoft	Google	IBM
Object Detection	✓	✓	✓	✓
Scene Detection	✓	✓	✓	✗
Face Detection	✓	✓	✓	✓
Face Recognition (person face identification)	✓	✓	✗	✗
Facial Analysis	✓	✓	✓	✓
Inappropriate Content Detection	✓	✓	✓	✓
Celebrity Recognition	✓	✓	✓	✗
Text Recognition	✓	✓	✓	✓
Written Text Recognition	✗	✓	✓	✗
Search for Similar Images on Web	✗	✗	✓	✗
Logo Detection	✗	✗	✓	✗
Landmark Detection	✗	✓	✓	✗
Food Recognition	✗	✗	✗	✓
Dominant Colors Detection	✗	✓	✓	✗

VIDEO ANALYSIS APIs COMPARISON

	Amazon	Microsoft	Google
Object Detection	✓	✓	✓
Scene Detection	✓	✓	✓
Activity Detection	✓	✗	✗
Facial Recognition	✓	✓	✗
Facial and Sentiment Analysis	✓	✓	✗
Inappropriate Content Detection	✓	✓	✓
Celebrity Recognition	✓	✓	✗
Text Recognition	✓	✓	✗
Person Tracking on Videos	✓	✓	✗
Audio Transcription	✗	✓	✓
Speaker Indexing	✗	✓	✗
Keyframe Extraction	✗	✓	✗
Video Translation	✗	9 languages	✗
Keywords Extraction	✗	✓	✗
Brand Recognition	✗	✓	✗
Annotation	✗	✓	✗
Dominant Colors Detection	✗	✗	✗
Real-Time Analysis	✓	✗	✗

Today - Real World Tools

- What You Need for Real World Applications
- Cloud Services
 - Cloud services comparison
 - Amazon SageMaker
 - Google AutoML
- Vision APIs
- **Specialized Tools**
 - Detectron by Facebook
 - Face++ for face analysis
 - Natural Language Toolkit (NLTK)
- Where to Find SOTA Models?

Detectron by Facebook

Research oriented

Contains:

[Mask R-CNN](#) -- *Marr Prize at ICCV 2017*

[RetinaNet](#) -- *Best Student Paper Award at ICCV 2017*

[Faster R-CNN](#)

[RPN](#)

[Fast R-CNN](#)

[R-FCN](#)



Code: https://github.com/facebookresearch/Detectron/blob/master/GETTING_STARTED.md

Models: https://github.com/facebookresearch/Detectron/blob/master/MODEL_ZOO.md

Object detection API by Google

Based on Tensorflow

Model name	Speed	COCO mAP
ssd_mobilenet_v1_coco	fast	21
ssd_inception_v2_coco	fast	24
rfcn_resnet101_coco	medium	30
faster_rcnn_resnet101_coco	medium	32
faster_rcnn_inception_resnet_v2_atrous_coco	slow	37

Can finetune to custom categories

Face++ for face analysis





Age 24

Gender Female

Smiling value: 0.104; threshold: 50



Head Pose 

Emotion neutral

Ethnicity WHITE

Left Eye Status Open, no glass

NLTK for Natural Language Text Processing

Natural Language Toolkit

NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to [over 50 corpora and lexical resources](#) such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active [discussion forum](#).

Thanks to a hands-on guide introducing programming fundamentals alongside topics in computational linguistics, plus comprehensive API documentation, NLTK is suitable for linguists, engineers, students, educators, researchers, and industry users alike. NLTK is available for Windows, Mac OS X, and Linux. Best of all, NLTK is a free, open source, community-driven project.

NLTK has been called “a wonderful tool for teaching, and working in, computational linguistics using Python,” and “an amazing library to play with natural language.”

TABLE OF CONTENTS

NLTK News
Installing NLTK
Installing NLTK Data
Contribute to NLTK
FAQ
Wiki
API
HOWTO

Today - Real World Tools

- What You Need for Real World Applications
- Cloud Services
 - Cloud services comparison
 - Amazon SageMaker
 - Google AutoML
- Vision APIs
- Specialized Tools
 - Detectron by Facebook
 - Face++ for face analysis
 - Natural Language Toolkit (NLTK)
- **Where to Find SOTA Models?**

ML papers with code

Browse state-of-the-art

820 leaderboards • 1099 tasks • 898 datasets • 11656 papers with code

Follow on  Twitter for updates

Computer Vision



Semantic Segmentation

16 leaderboards

409 papers with code



Image Classification

38 leaderboards

340 papers with code



Object Detection

33 leaderboards

294 papers with code



Image Generation

24 leaderboards

153 papers with code



Pose Estimation

26 leaderboards

140 papers with code

▶ See all 598 tasks

<https://paperswithcode.com/>

Today - Real World Tools

- What You Need for Real World Applications
- Cloud Services
 - Cloud services comparison
 - Amazon SageMaker
 - Google AutoML
- Vision APIs
- Specialized Tools
 - Detectron by Facebook
 - Face++ for face analysis
 - Natural Language Toolkit (NLTK)
- Where to Find SOTA Models?