CLASS SESSION 7 (plus AI Lab session)

WHAT IS SIZE OF AI BUSINESS?

$100+ Trillion? Or $14 Trillion by 2025?

(Type your answer in the chat)

Algorithms: Model

Model Fitting in the Datascience sprint cycle

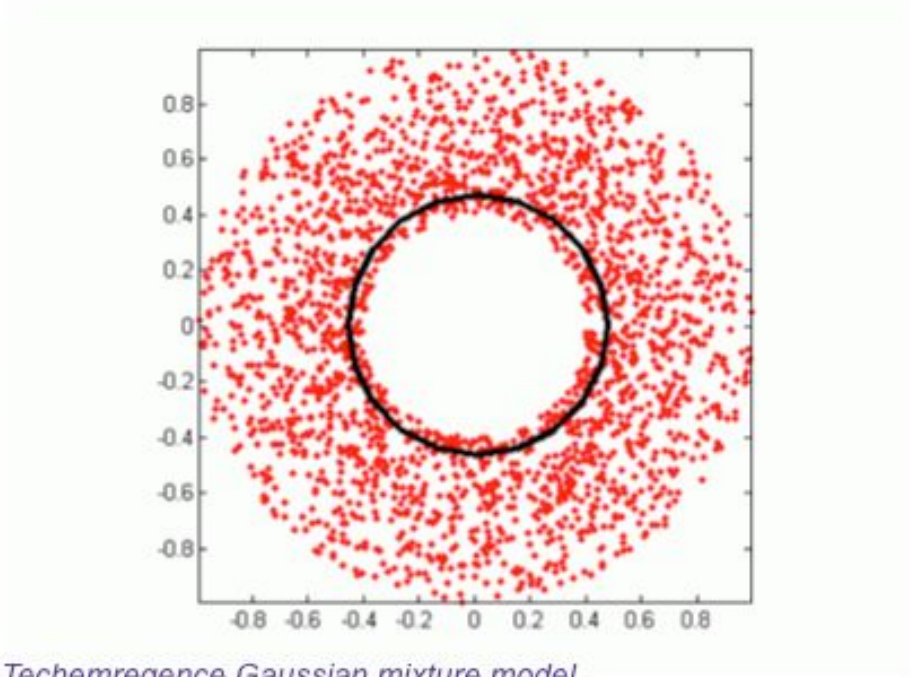
4 WAYS BY WHICH ALGORITHMS HELP CUSTOMERS

1. CLASSIFY AND FIND PATTERNS
2. RECOGNIZE OBJECTS
3. MAKE PREDICTIONS AND RECOMMENDATIONS
4. ANOMALY DETECTION

MACHINE LEARNING

Machine Learning takes in large volume of data and makes a decision by building Data Models (a geometric shape of data*)

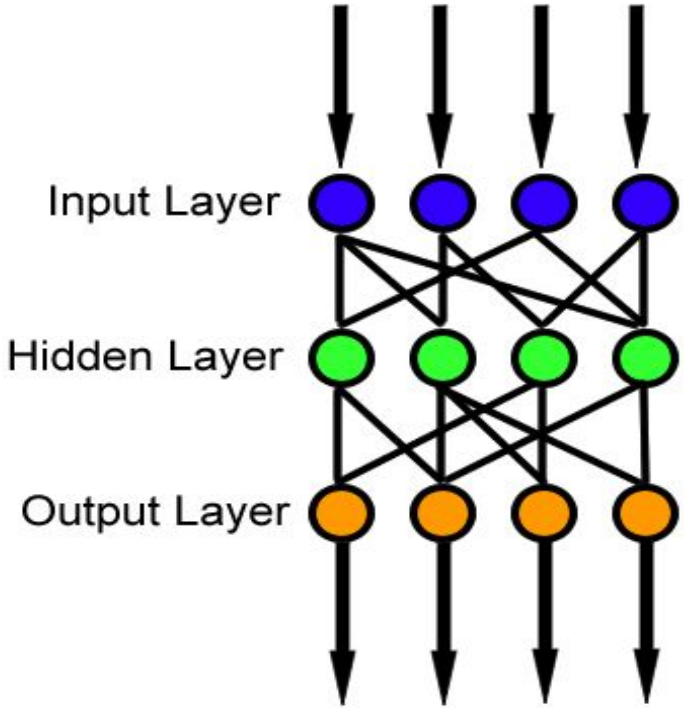
Training Data & Factors



Predict Outcomes to make decisions

image: Techemregence Gaussian mixture model

Neural Network



Deep Learning (Neural Networks)

Format

Top Tech Companies that Rule the World

	Symbol	Company	Entry Price	Last Price	% Change	Market Cap (B)
1	NVDA	NVIDIA Corporation	$154.26	$191.17	19.31%	$116.33
2	AMD	Advanced Micro Devices, Inc.	$23.53	$28.46	17.32%	$30.78
3	AAPL	Apple Inc.	$173.15	$207.16	16.42%	$976.82
4	AMZN	Amazon.com, Inc.	$1,639.83	$1,901.75	13.77%	$935.76
5	F	Facebook, Inc.	$161.45	$182.58	11.57%	$521.38
6	GOOG	Alphabet Inc.	$1,119.92	$1,256.00	10.83%	$874.70
7	MSFT	Microsoft Corporation	$112.03	$125.01	10.38%	$959.10
8	MTCH	Match Group, Inc.	$55.38	$61.65	10.17%	$17.17
9	ATVI	Activision Blizzard, Inc.	$42.14	$46.75	9.86%	$35.71
10	INTC	Intel Corporation	$52.96	$58.00	8.69%	$264.16
11	PYPL	PayPal Holdings, Inc.	$98.07	$107.22	8.53%	$125.68
12	BABA	Alibaba Group Holding Limited	$183.03	$185.67	1.42%	$481.29
13	EA	Electronic Arts Inc.	$95.78	$94.00	-1.89%	$28.21
14	CR	salesforce.com, inc.	$163.65	$159.56	-2.56%	$123.64
15	TSLA	Tesla, Inc.	$319.88	$258.00	-23.98%	$44.93

Based on "Tech Stocks That Move The Market" by Yahoo Finance.

body Spreadsheet

Google AI video explains CNN well:
https://www.youtube.com/watch?v=OcycT1Jwsns

Computer Vision of Fruits used in Smart Agriculture

FINANCE: FRAUD DETECTION

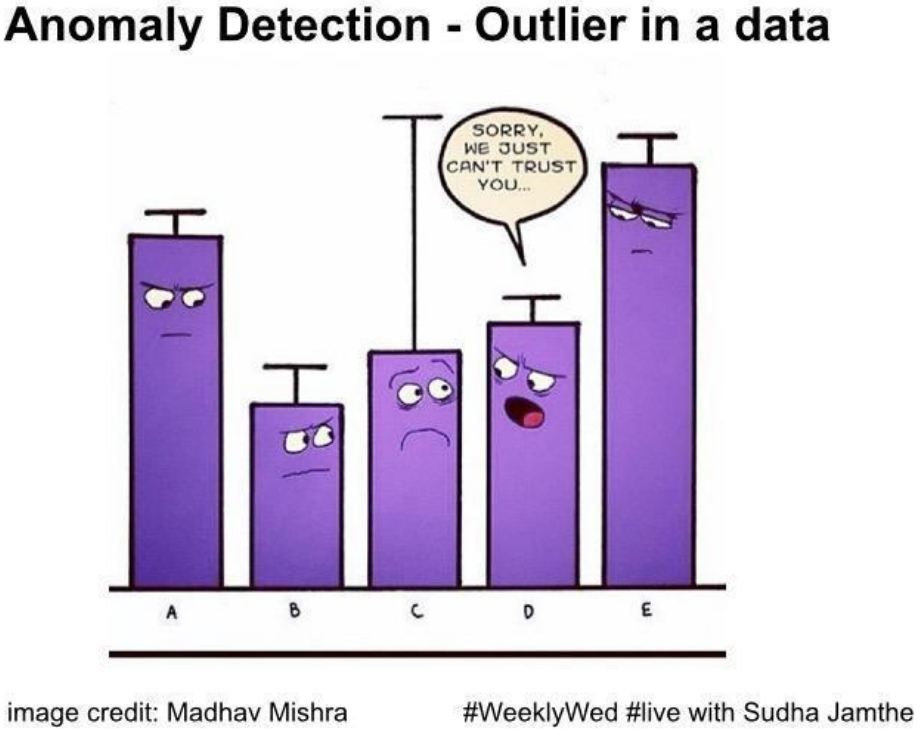
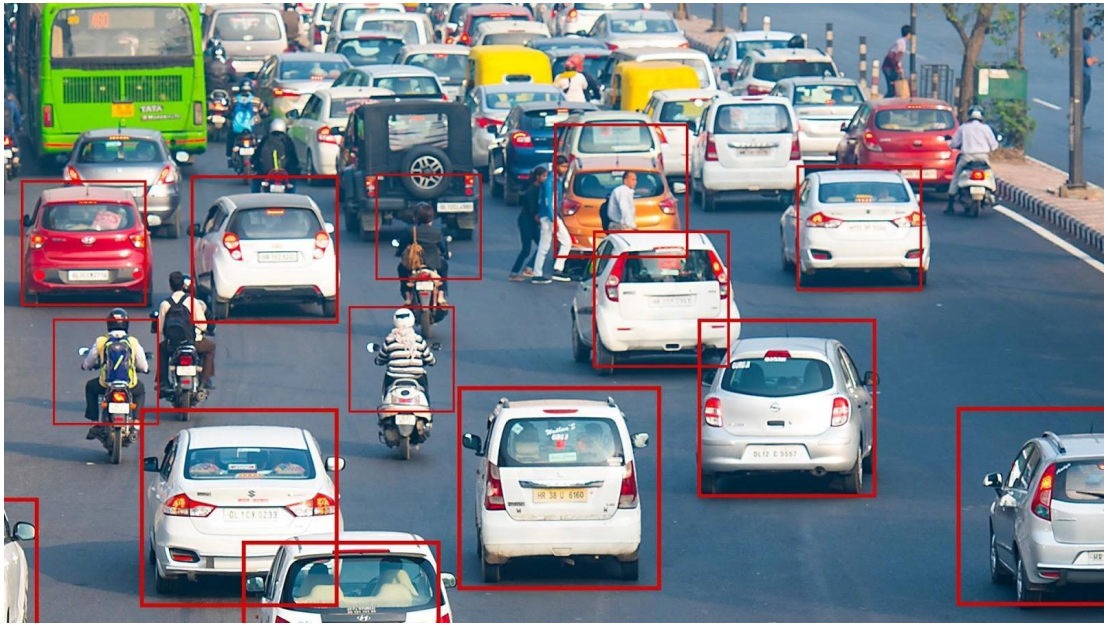


image credit: Madhav Mishra

#WeeklyWed #live with Sudha Jamthe

Algorithms: Is your training data structured or unstructured?

Unstructured data



Unsupervised Learning (or)

Deep Learning

Structured data

ID › CUSTOMERID	CATEGORY › GENDER	CATEGORY › SENIORCITIZEN	CATEGORY › PARTNER	CATEGORY › DEPENDENTS	TE
9305-CDSKC	Female	0	No	No	8
1452-KIOVK	Male	0	No	Yes	22
6713-OKOMC	Female	0	No	No	10
7892-POOKP	Female	0	Yes	No	28
6388-TABGU	Male	0	No	Yes	62
9763-GRSKD	Male	0	Yes	Yes	13
7469-LKBCI	Male	0	No	No	16

Supervised Learning (or)

Machine Learning

Regression

- Supervised Learning
- Output is a continuous quantity
- Main aim is to forecast or predict
- Eg: Predict stock market price
- Algorithm: Linear Regression

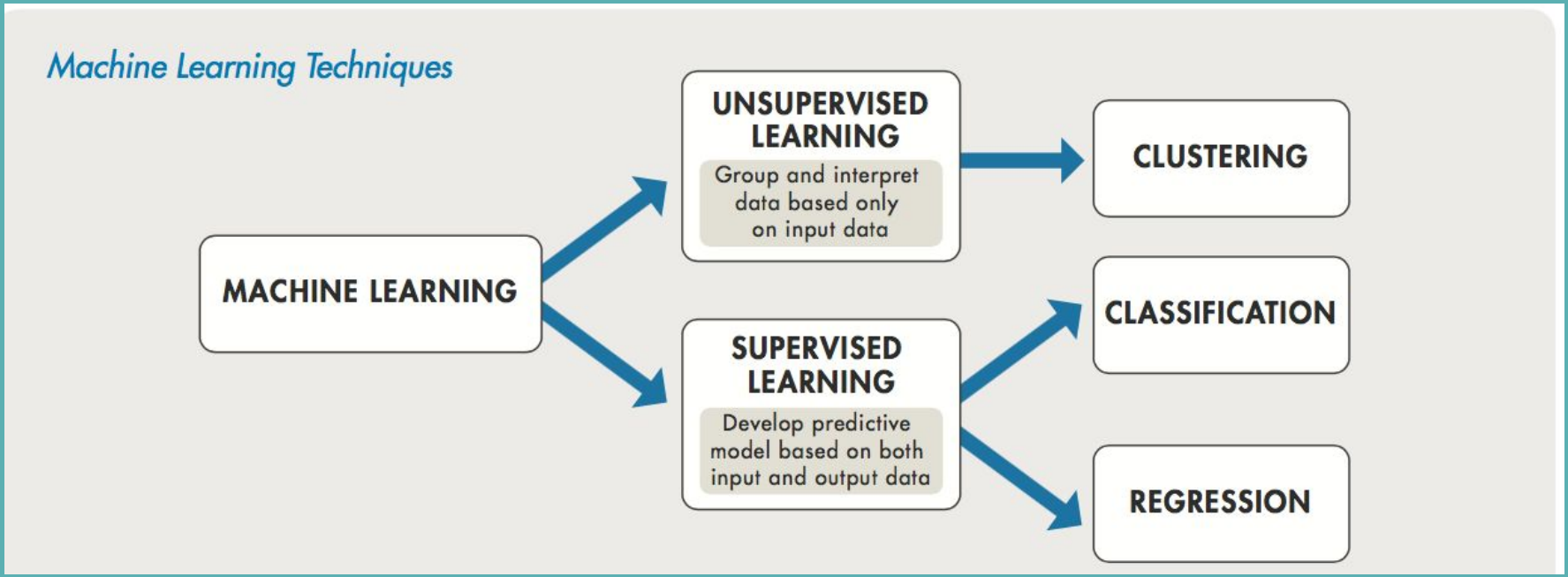
Classification

- Supervised Learning
- Output is a categorical quantity
- Main aim is to compute the category of the data
- Eg: Classify emails as spam or non-spam
- Algorithm: Logistic Regression

Clustering

- Unsupervised Learning
- Assigns data points into clusters
- Main aim is to group similar items clusters
- Eg: Find all transactions which are fraudulent in nature
- Algorithm: K-means

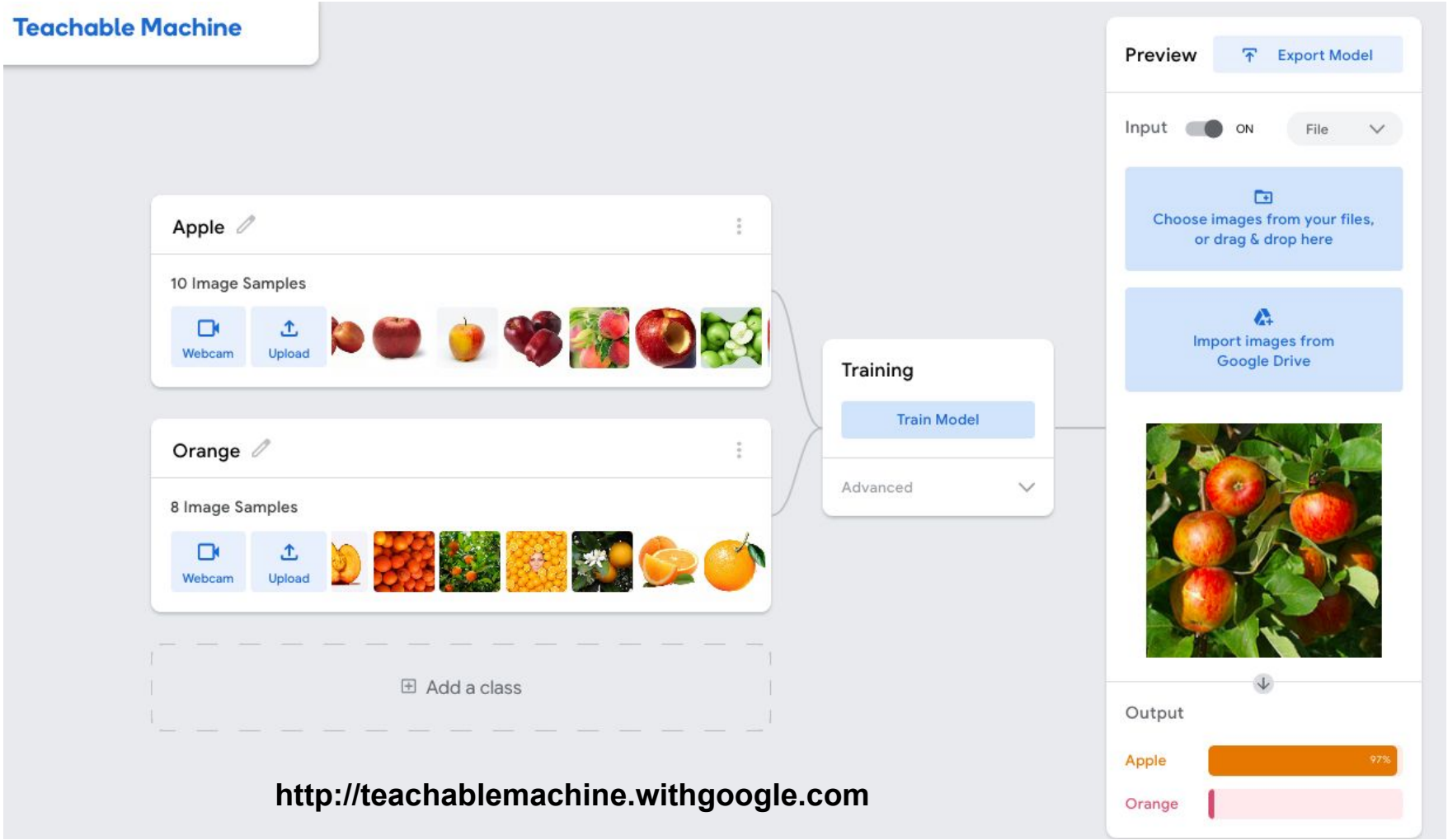
Type Of Problems Solved Using AI – Artificial Intelligence Algorithms – Edureka



Credit: MatLab

Classification Algorithms

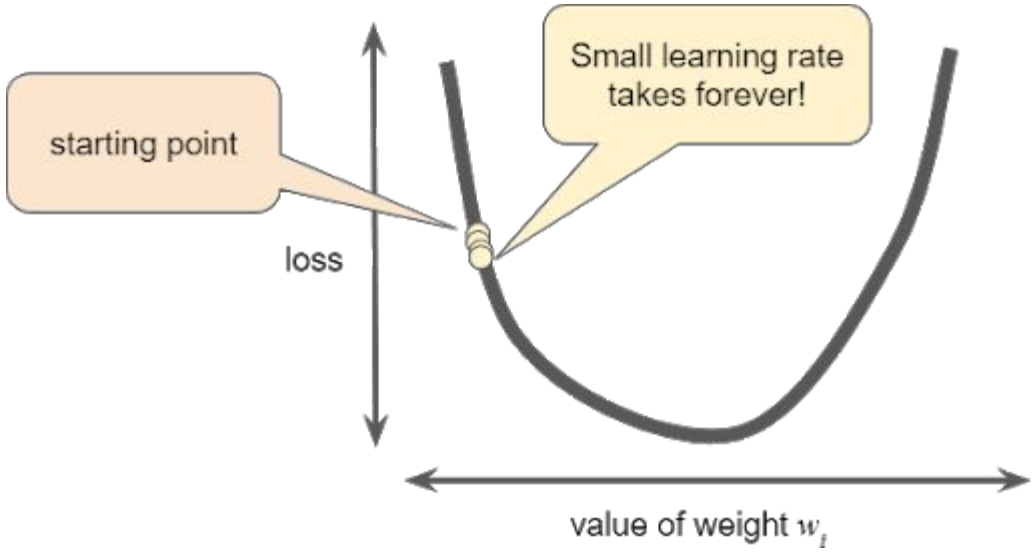
Binary Classification: Predict a Yes/No situation.
Give me an example


Teachable Machine
Apple
10 Image Samples
Webcam
Upload
Orange
8 Image Samples
Webcam
Upload
Add a class
Training
Train Model
Advanced
Preview
Export Model
Input
ON
File
Choose images from your files, or drag & drop here
Import images from Google Drive
Output
Apple
97%
Orange
http://teachablemachine.withgoogle.com

Classification Algorithms

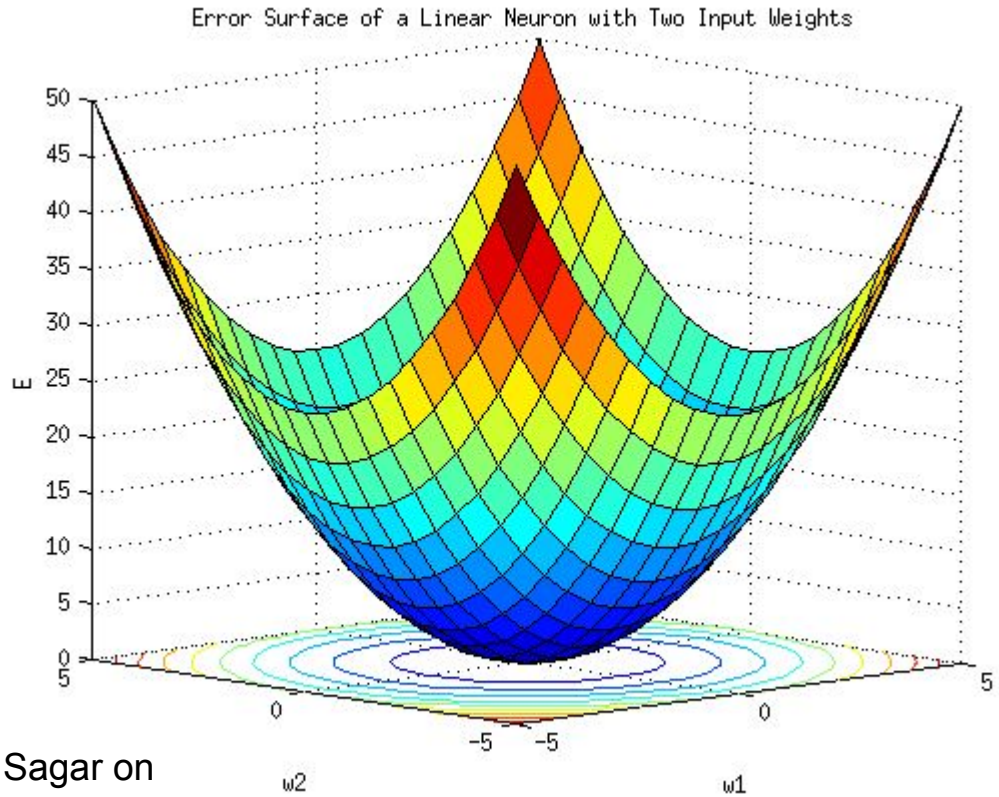
Gradient Descent: When you have large volume of data

Learning rate and loss function



Source: analyticsindiamag.com

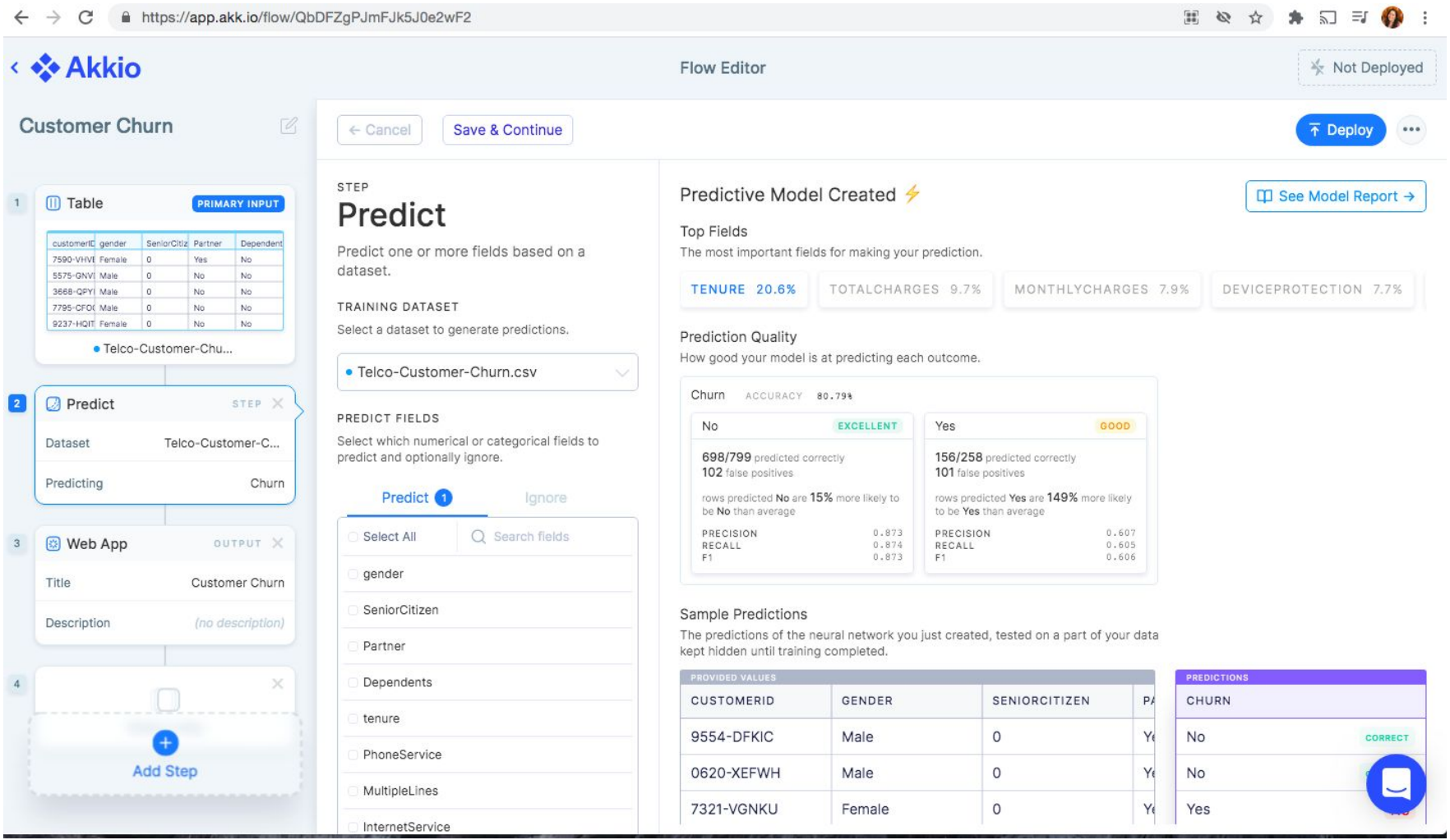
Gradient Descent



Credit: Ram Sagar on https://analyticsindiamag.com/a-lowdown-on-alternatives-to-gradient-descent-optimization-algorithms/

Live Lab: IBM AutoAI and Akkio

Lets try a binary classification and compare the NOCODE AI AutoML model fitting experient


https://app.akk.io/flow/QbDFZgPJmFJk5J0e2wF2
Akkio
Flow Editor
Not Deployed
Customer Churn
Cancel
Save & Continue
Deploy
1 Table PRIMARY INPUT
customerID gender SeniorCitiz Partner Dependent
7590-VHVE Female 0 Yes No
5575-GNVI Male 0 No No
3668-QPYI Male 0 No No
7795-CFOC Male 0 No No
9237-HQIT Female 0 No No
Telco-Customer-Chu...
2 Predict STEP
Dataset Telco-Customer-C...
Predicting Churn
3 Web App OUTPUT
Title Customer Churn
Description (no description)
4
Add Step
STEP
Predict
Predict one or more fields based on a dataset.
TRAINING DATASET
Select a dataset to generate predictions.
Telco-Customer-Churn.csv
PREDICT FIELDS
Select which numerical or categorical fields to predict and optionally ignore.
Predict 1
Ignore
Select All
Search fields
gender
SeniorCitizen
Partner
Dependents
tenure
PhoneService
MultipleLines
InternetService
Predictive Model Created
See Model Report
Top Fields
The most important fields for making your prediction.
TENURE 20.6%
TOTALCHARGES 9.7%
MONTHLYCHARGES 7.9%
DEVICEPROTECTION 7.7%
Prediction Quality
How good your model is at predicting each outcome.
Churn ACCURACY 80.79%
No EXCELLENT
698/799 predicted correctly
102 false positives
rows predicted No are 15% more likely to be No than average
PRECISION 0.873
RECALL 0.874
F1 0.873
Yes GOOD
156/258 predicted correctly
101 false positives
rows predicted Yes are 149% more likely to be Yes than average
PRECISION 0.607
RECALL 0.605
F1 0.606
Sample Predictions
The predictions of the neural network you just created, tested on a part of your data kept hidden until training completed.
PROVIDED VALUES
CUSTOMERID GENDER SENIORCITIZEN
9554-DFKIC Male 0
0620-XEFWH Male 0
7321-VGNKU Female 0
PREDICTIONS
CHURN
No CORRECT
No
Yes

Rank	Pipeline	Holdout Accuracy (Optimiz...	Algorithm	Enhancements	Build time	
2	Pipeline 3	0.813	Gradient Boosting Classi...	HPO-1 FE	00:01:48	Save as

Gradient Boosting Classifier

EVALUATION

- Model Evaluation
- Confusion Matrix
- Precision Recall Curve

MODEL VIEWER

- Model Information
- Feature Transformations
- Feature Importance

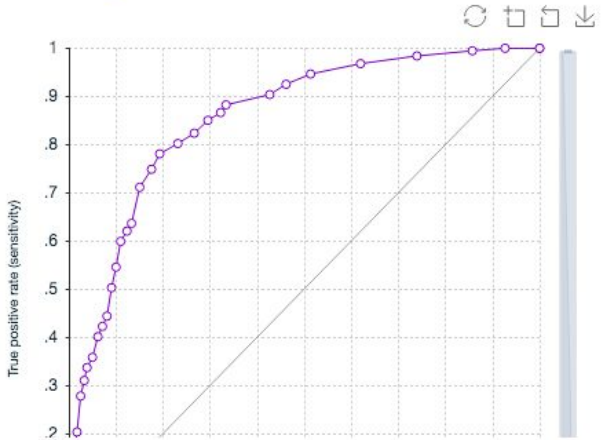
Model Evaluation

TARGET : CHURN

Model Accuracy

0.813

ROC Curve



Model Evaluation Measures

	Holdout Score	Cross Validation Score
Accuracy	0.813	0.805
Area Under ROC Curve	0.858	0.844
Precision	0.673	0.663
Recall	0.572	0.543

Rank
4

Pipeline 1

Holdout Accuracy *(Optimiz...*
0.809

Algorithm
Gradient Boosting Classi...

Gradient Boosting Classifier

EVALUATION

Model Evaluation

Confusion Matrix

Precision Recall Curve

MODEL VIEWER

Model Information

Feature Importance

Confusion Matrix

TARGET : CHURN

Observed	Predicted		
	Yes	No	Percent Correct
Yes	101	86	54.0%
No	49	469	90.5%
Percent Correct	67.3%	84.5%	80.9%

Less correct — More correct

IBM Cloud Pak for Data | All | Search | Upgrade

Rank	Pipeline 4	Holdout Accuracy (Optimiz...	Algorithm	Enhancements	Build time
1		0.814	Gradient Boosting Classi...	HPO-1 FE HPO-2	00:01:43

Gradient Boosting Classifier

EVALUATION

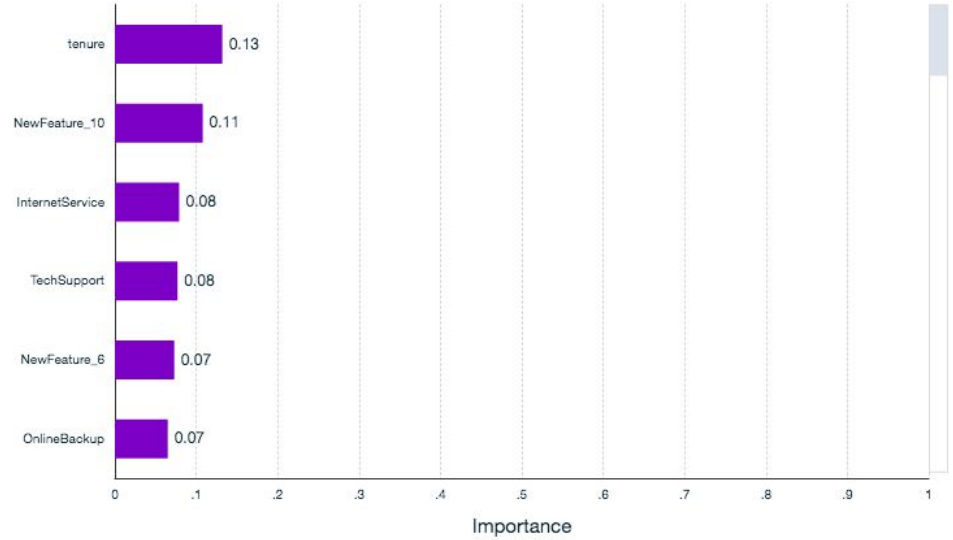
- Model Evaluation
- Confusion Matrix
- Precision Recall Curve

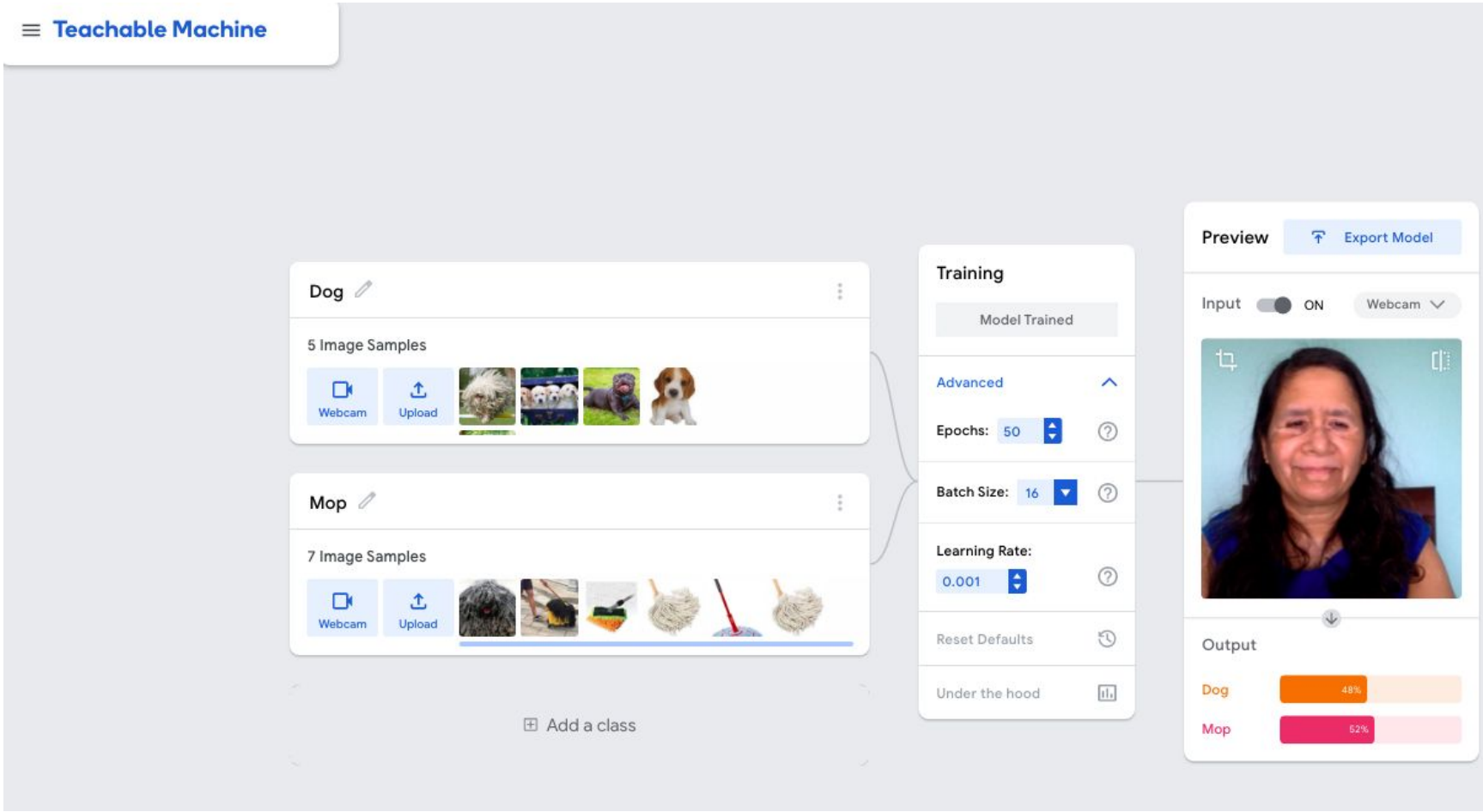
MODEL VIEWER

- Model Information
- Feature Transformations
- Feature Importance

Feature Importance

TARGET : CHURN




Teachable Machine
Dog
5 Image Samples
Webcam
Upload
Mop
7 Image Samples
Webcam
Upload
Add a class
Training
Model Trained
Advanced
Epochs: 50
Batch Size: 16
Learning Rate:
0.001
Reset Defaults
Under the hood
Preview
Export Model
Input
ON
Webcam
Output
Dog
48%
Mop
52%