Deep Learning

Giri Iyengar

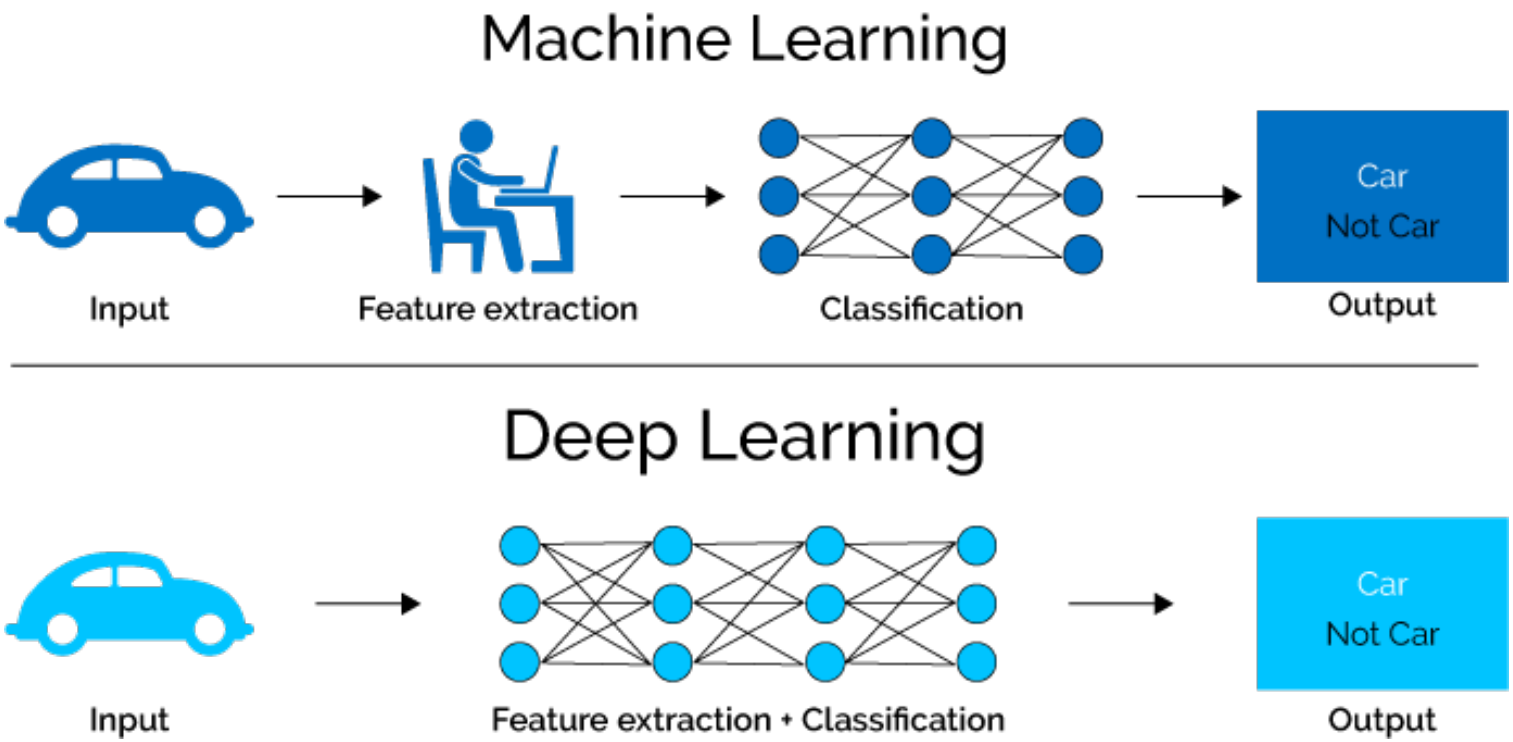
Cornell University

gi43@cornell.edu

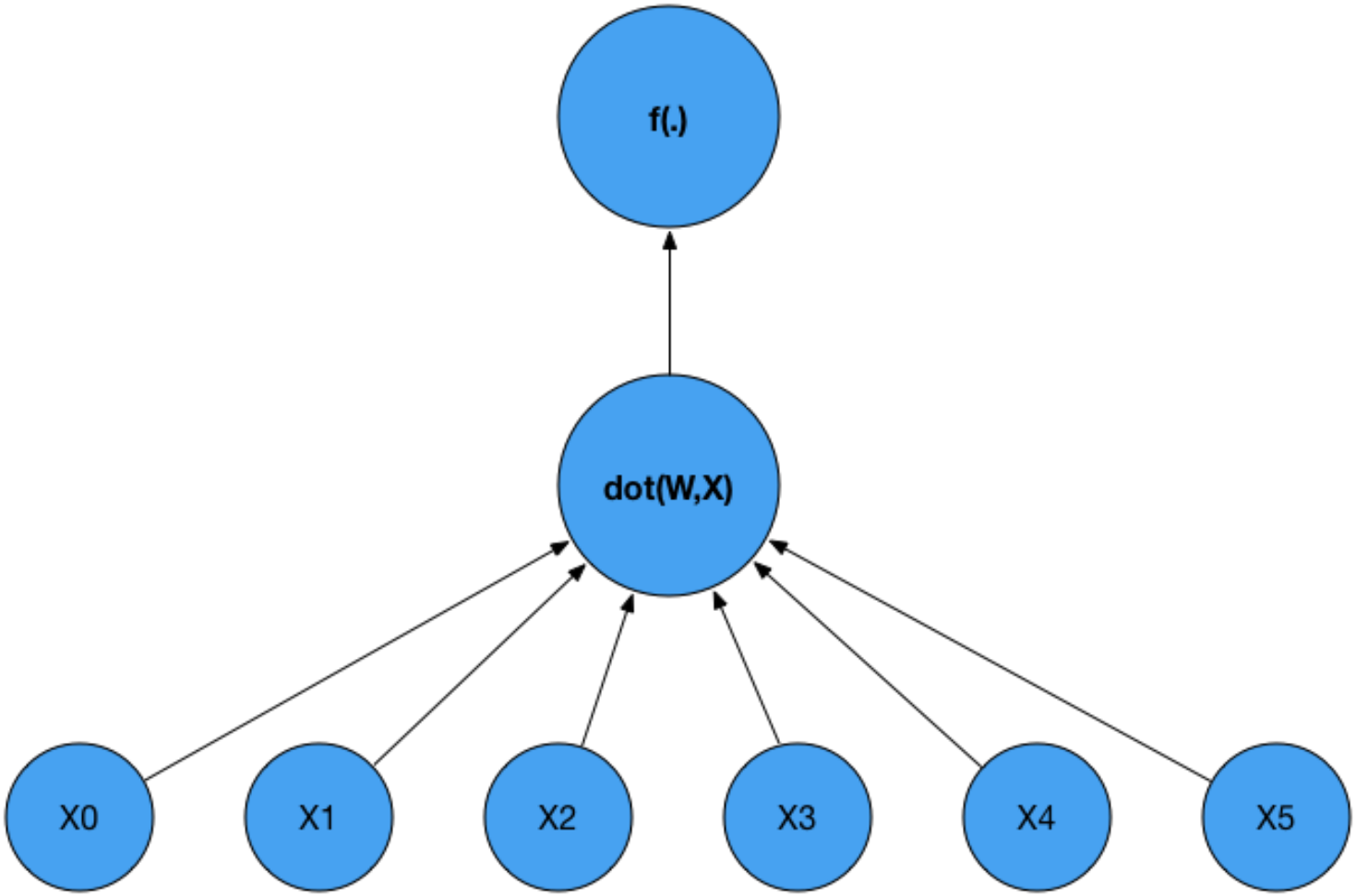
Feb 12, 2018

Overview

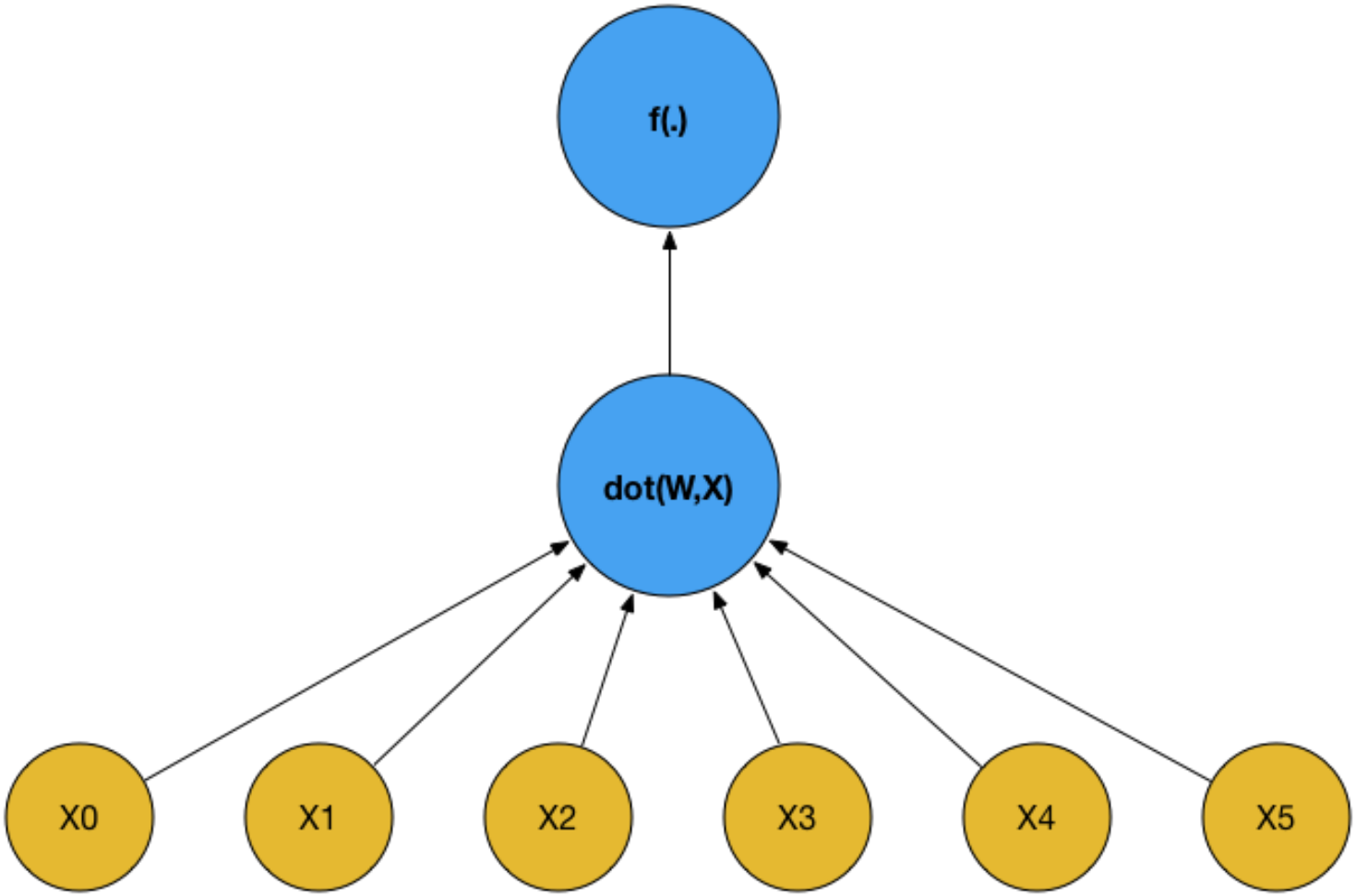
Deep Learning vs ML



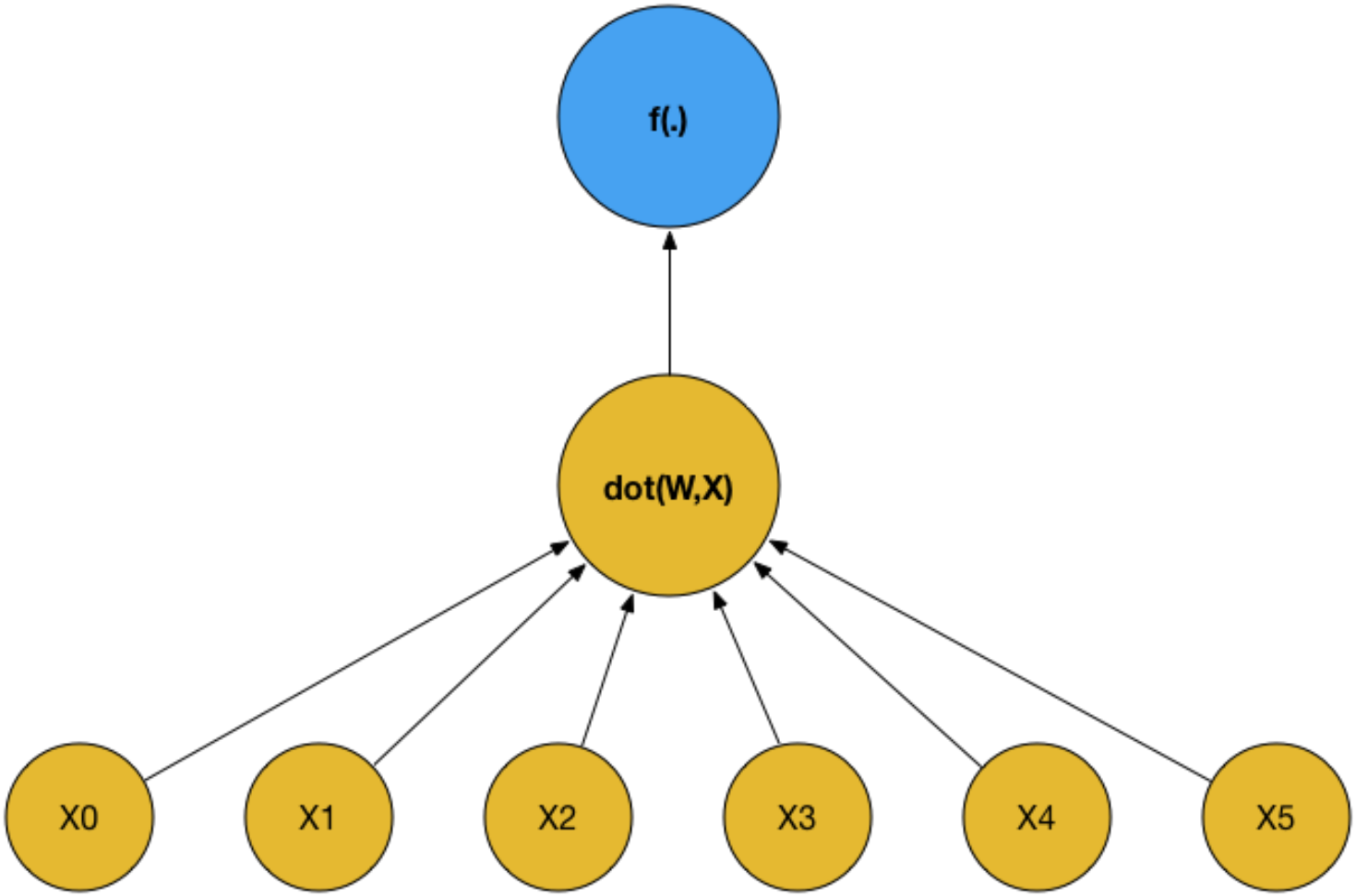
Basic Computation of the Neuron



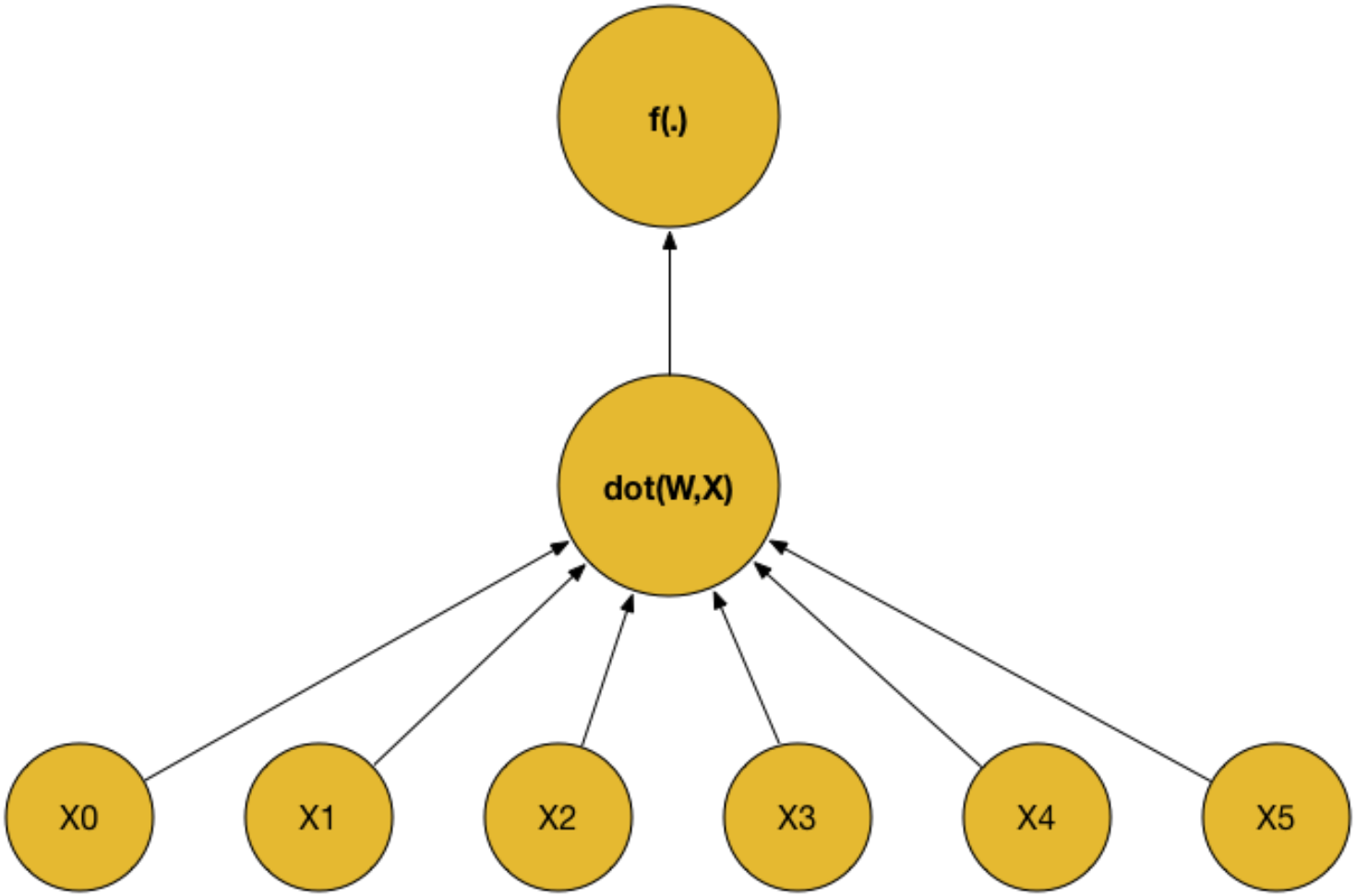
Basic Computation of the Neuron



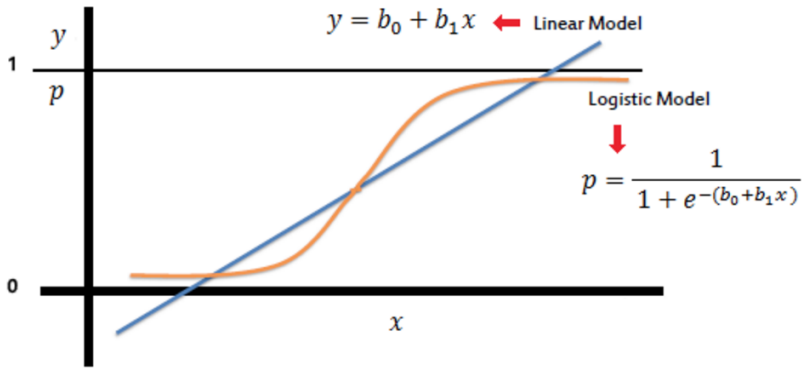
Basic Computation of the Neuron

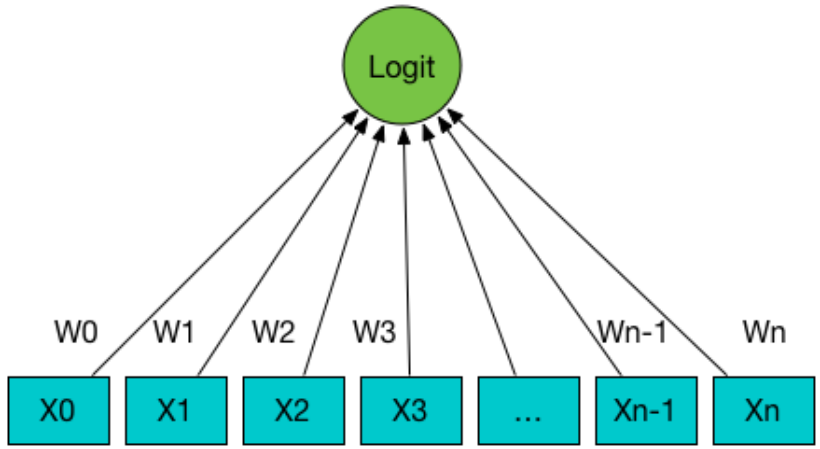


Basic Computation of the Neuron

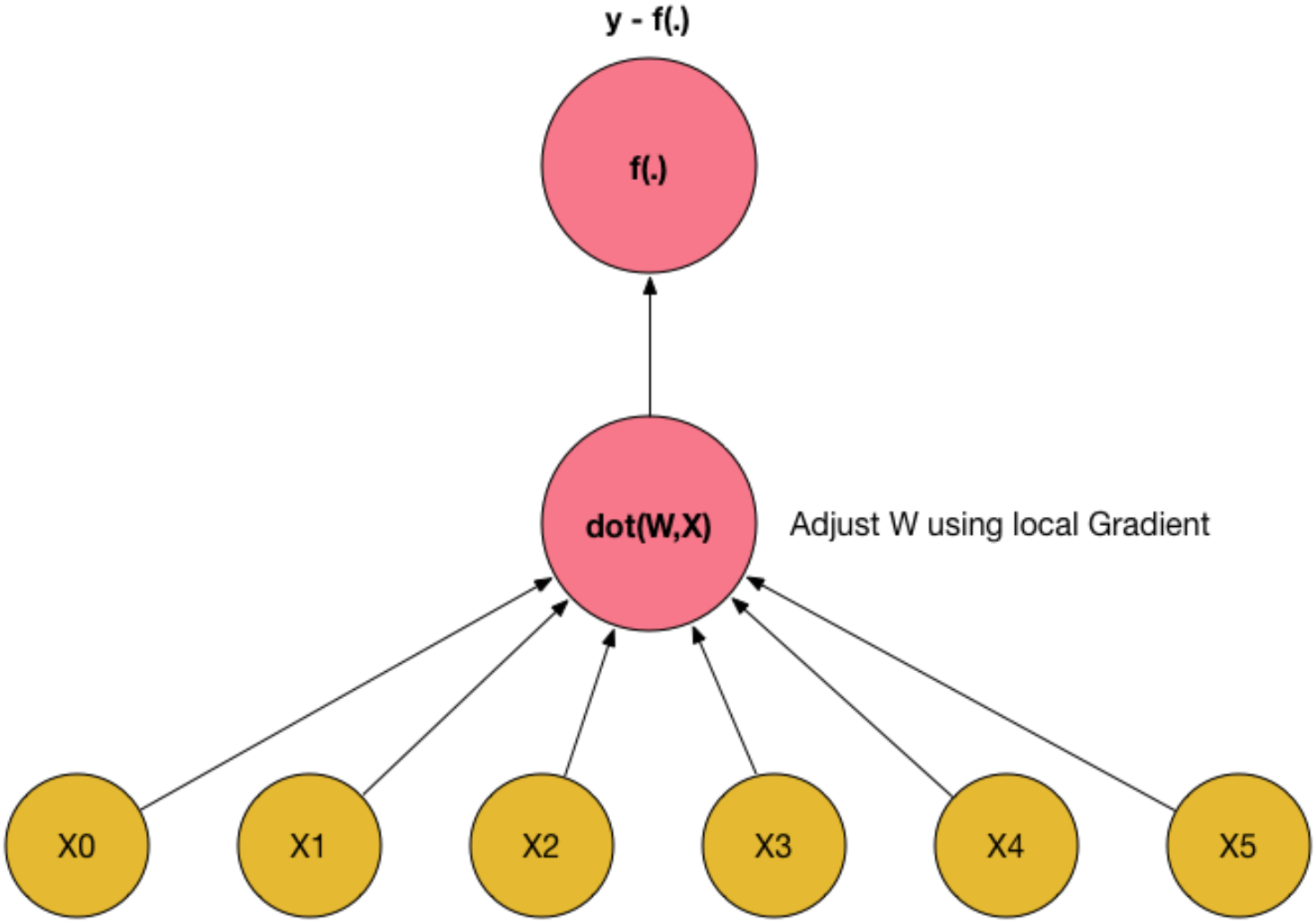


Logistic Regression is just a simple Neuron Computational Unit

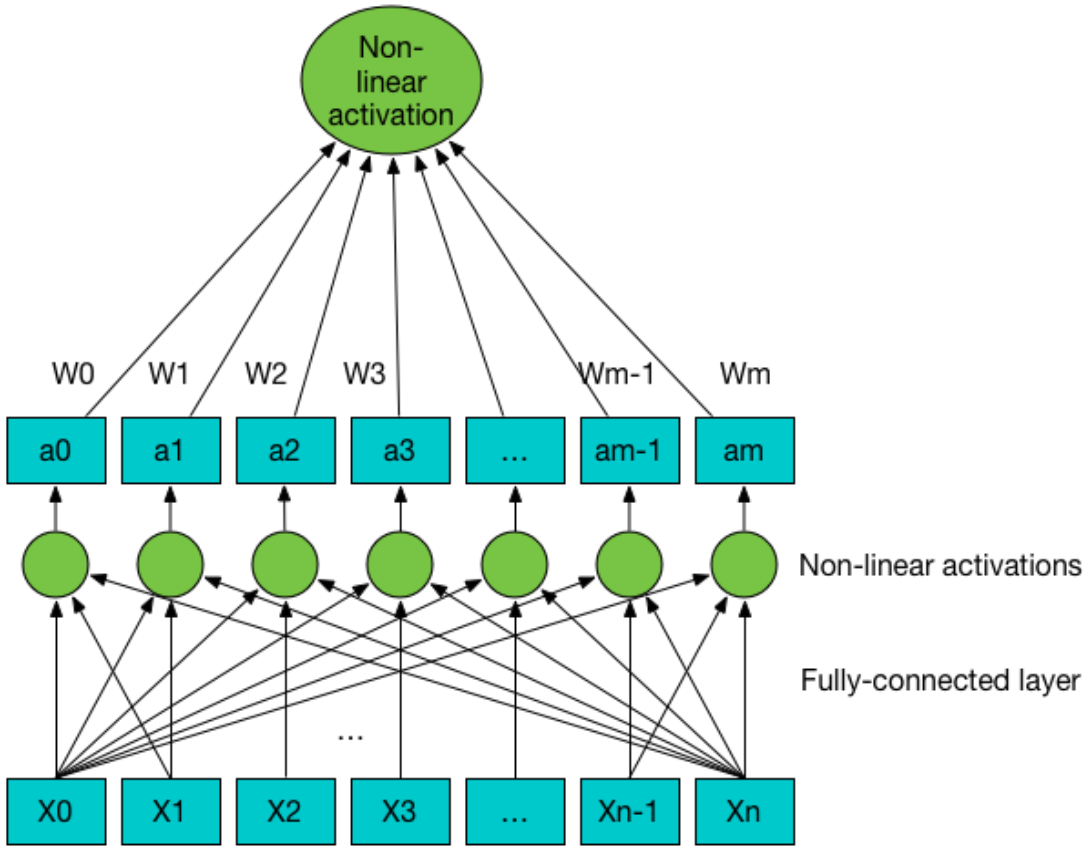




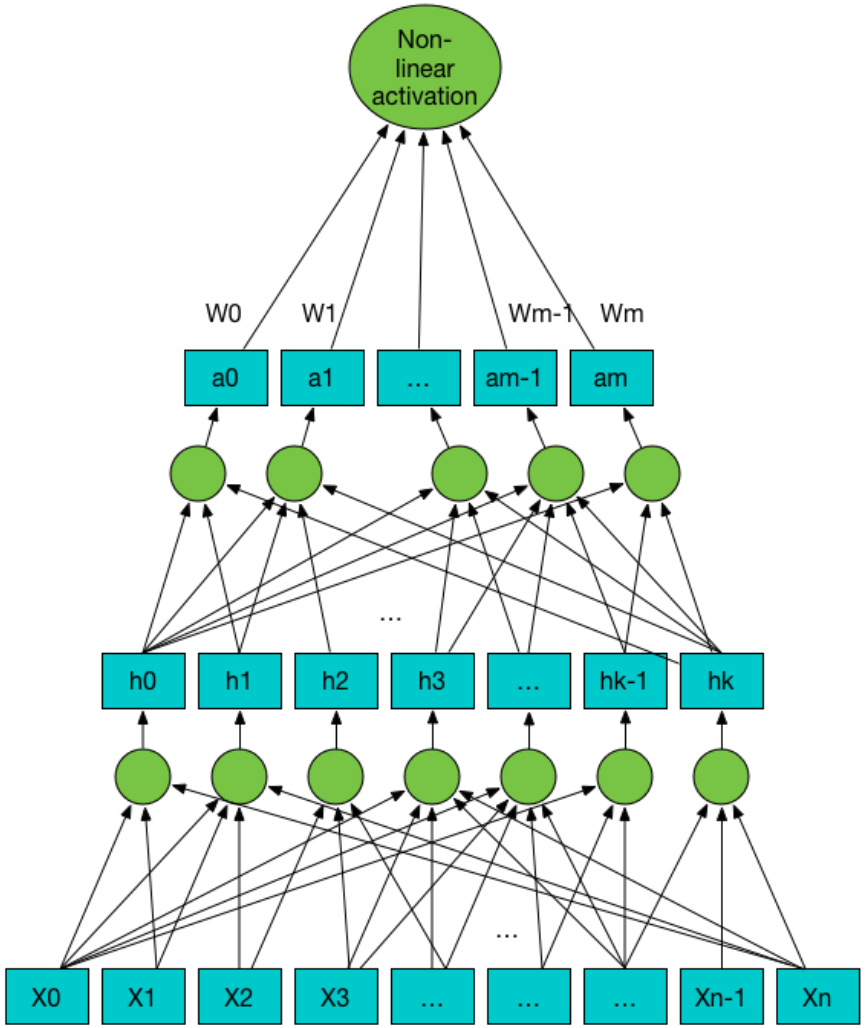
Training the Neuron



Multi Layer Perceptron



Multi Layer Perceptron



Why was this not done before?

- Too many model parameters
- Difficult to train
- Slow to train
- Prone to overfitting
- Beyond one hidden-layer, wasn't practical to fit
- Other algorithms such as SVMs had superior performance to MLP

What changed?

- Data set sizes grew exponentially
- New techniques for training were invented
  1. Stochastic Gradient Descent
  2. HogWild
  3. Stochastic L-BFGS
  4. Fast hyper-parameter search techniques
- Computational Capacity grew making training times manageable
- New algorithms together with techniques for managing model complexity (regularization) demonstrated superior performance – shot past all other models in terms of performance
- Finally delivering on **universal function approximator** promise

What Technologies make it work?

- Stochastic Gradient Descent
  - With SGD, training times of Deep Learning networks were dramatically reduced
  - Other techniques such as unsupervised initialization dramatically improved generalization ability of these networks to novel tasks
- Automatic Differentiation
  - Advances in algorithmic differentiation (not numerical differentiation or symbolic differentiation) greatly simplified writing gradient descent code for such models
- GPU
  - Availability of programmable Graphics Processing Units and compilers (CUDA) further sped up training times. ▸ Mark Harris interview

What Technologies make it work?

- Regularization
  - Advances in the understanding of regularization techniques to handle model complexity reduced overfitting issues
- Dataset sizes
  - Exponential Growth in dataset sizes further reduced overfitting issues and enabled these networks to show significantly better performance than traditional models
- Compute Capabilities
  - Improved compute capabilities make it feasible to deploy such models in production where they are able to score billions of rows a day

Some successes of Deep Learning

- Image/Object Recognition
- Speech Recognition
- Natural Language Processing
- Machine Translation

Object Recognition

Describes without errors	Describes with minor errors	Somewhat related to the image	Unrelated to the image
 A person riding a motorcycle on a dirt road.	 Two dogs play in the grass.	 A skateboarder does a trick on a ramp.	 A dog is jumping to catch a frisbee.
 A group of young people playing a game of frisbee.	 Two hockey players are fighting over the puck.	 A little girl in a pink hat is blowing bubbles.	 A refrigerator filled with lots of food and drinks.
 A herd of elephants walking across a dry grass field.	 A close up of a cat laying on a couch.	 A red motorcorcycle parked on the side of the road.	 A yellow school bus parked in a parking lot.

Object Recognition

MS-COCO Dataset

Common Objects in Context - a new Captioning and Detection challenge

- a woman is playing a frisbee with a dog.
- a woman is playing frisbee with her large dog.
- a girl holding a frisbee with a dog coming at her.
- a woman kneeling down holding a frisbee in front of a white dog.
- a young lady is playing frisbee with her dog.

Object Recognition

MS-COCO 2015

Some sample captions from 2015 challenge

The man at bat readies to swing at the pitch while the umpire looks on.

A large bus sitting next to a very tall building.

Speech Recognition, Translation, and Synthesis

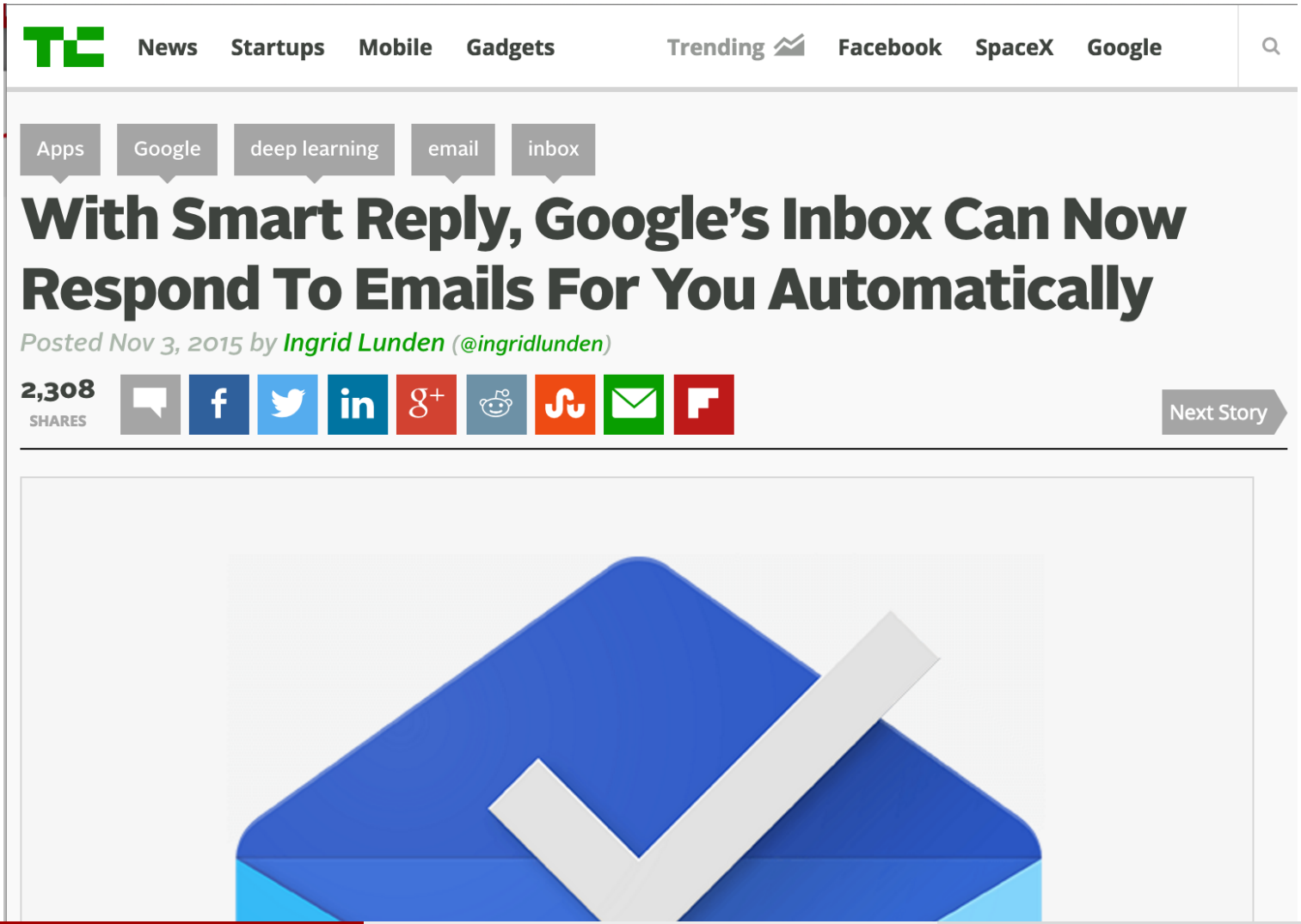
- Microsoft Research Speech Breakthrough ▸ Play

Named Entity Recognition

- Tutorial on NER ▸ Go

Natural Language Processing

NLP for Smart Email Replies



Some criticisms

- Theoretical Guarantees
- Models that are not well-understood
- Will these things last?
- ▸ Ali Rahimi blog


thank you
danke
謝謝
ngiyabonga
teşekkür ederim
gracias
спасибо
Баярлалаа
рахмат
bedankt
dziękuję
obrigado
hvala
sukriya
kop khun krap
go raibh maith agat
mochchakkeram
arigatō
dakujem
мерси
grazie
terima kasih
감사합니다
ありがとう
merci
ευχαριστώ
dank je
tapadh leat
sagolun
taiku
takk
xвала