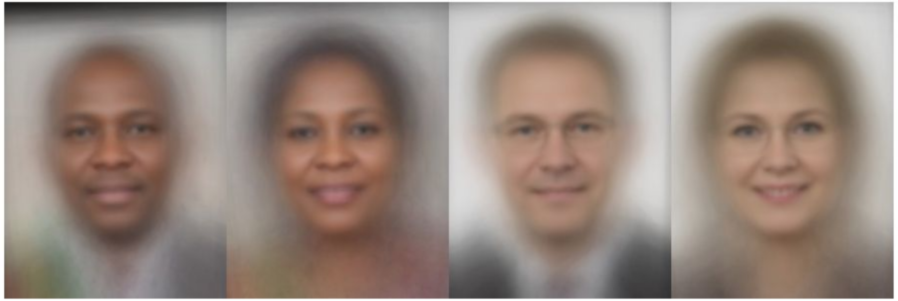
TALLER DE DEEP LEARNING

Lectura 6:
Contenidos adicionales

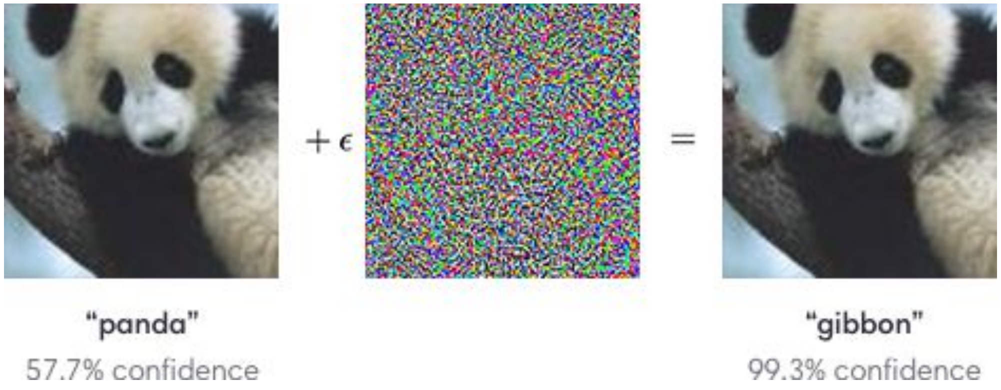
Dennis Núñez Fernández

Racial and Gender Bias in AI

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%



Attacking Machine Learning with Adversarial Examples



Attacking Machine Learning with Adversarial Examples

Glasses → Impersonation

[Sharif et al. 2016]

Banana + patch →Toaster

[Brown et al. 2017]

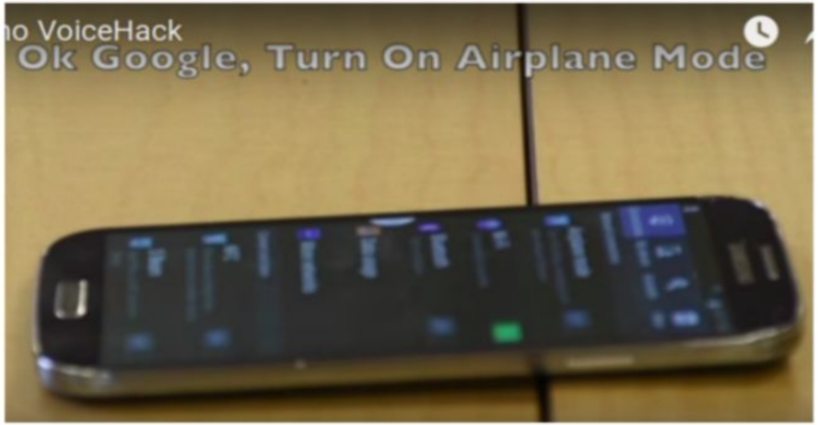
Stop + sticker →Yield

[Evtimov et al. 2017]

Attacking Machine Learning with Adversarial Examples

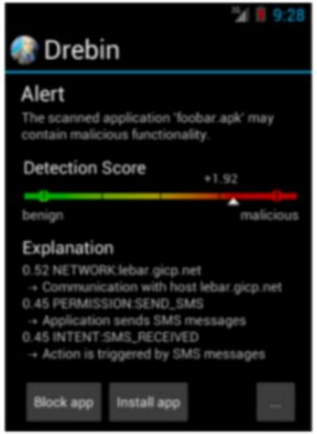
3D Turtle → Rifle

[Athalye et al. 2017]



Noise → "Ok Google"

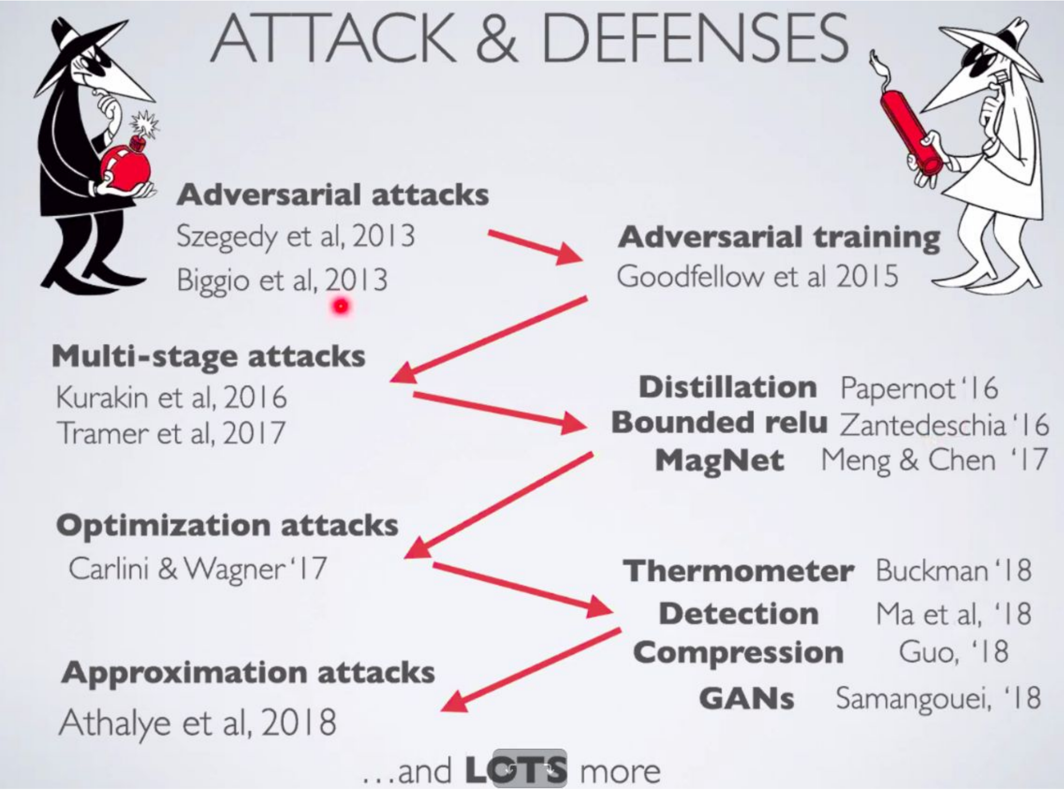
[Carlini et al. 2017]



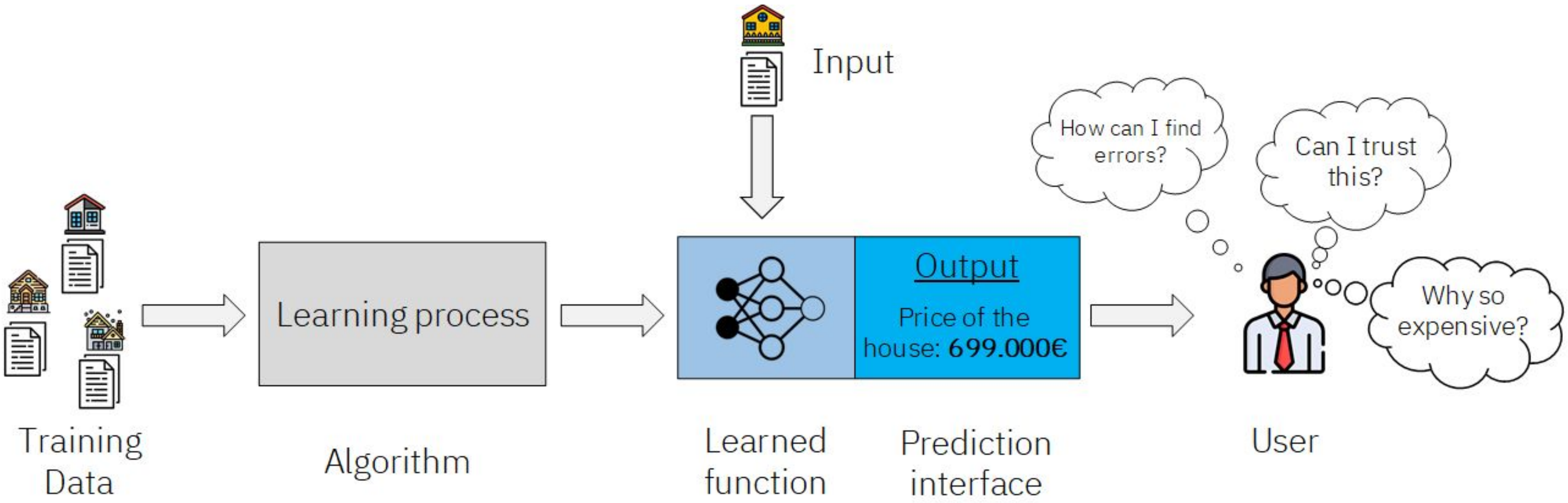
Malware → Benign

[Grosse et al. 2017]

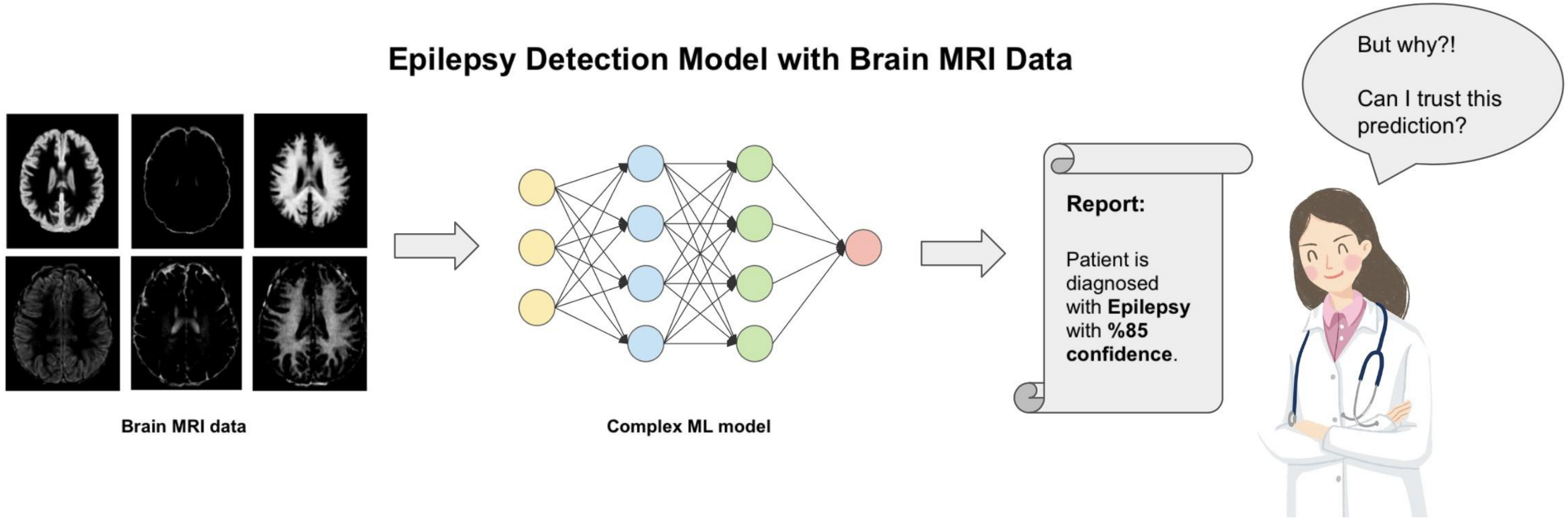
Attacking Machine Learning with Adversarial Examples



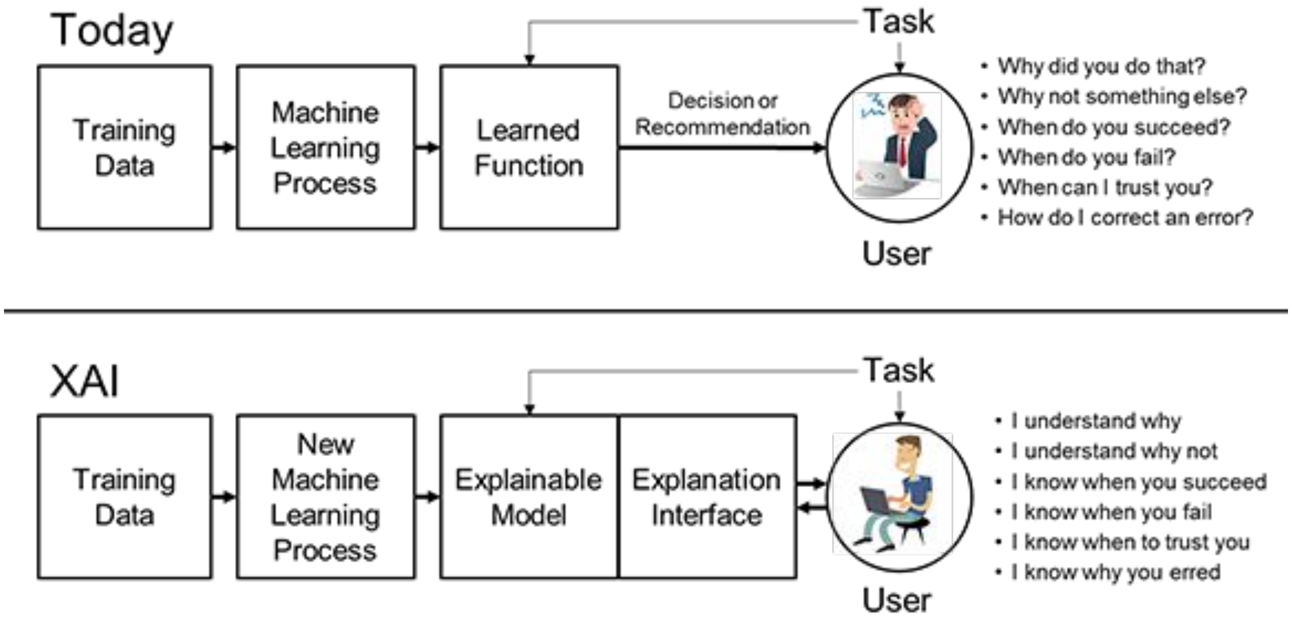
Explainable AI



Explainable AI



Explainable AI



Bayesian Neural Networks

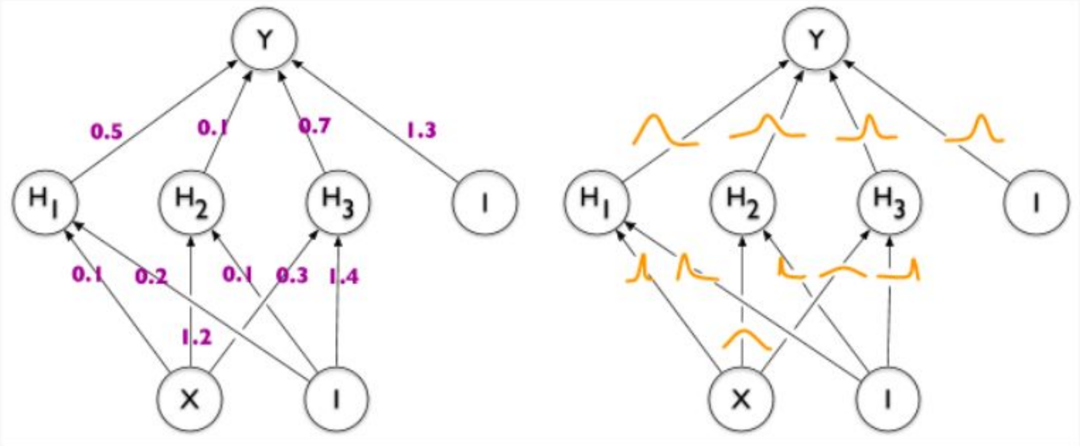
Dealing with Overconfidence in Neural Networks: Bayesian Approach

Measure of uncertainty in the prediction is missing from the current neural networks architectures, but Bayesian neural networks incorporate this

Prediction: dog
Probability: 0.98

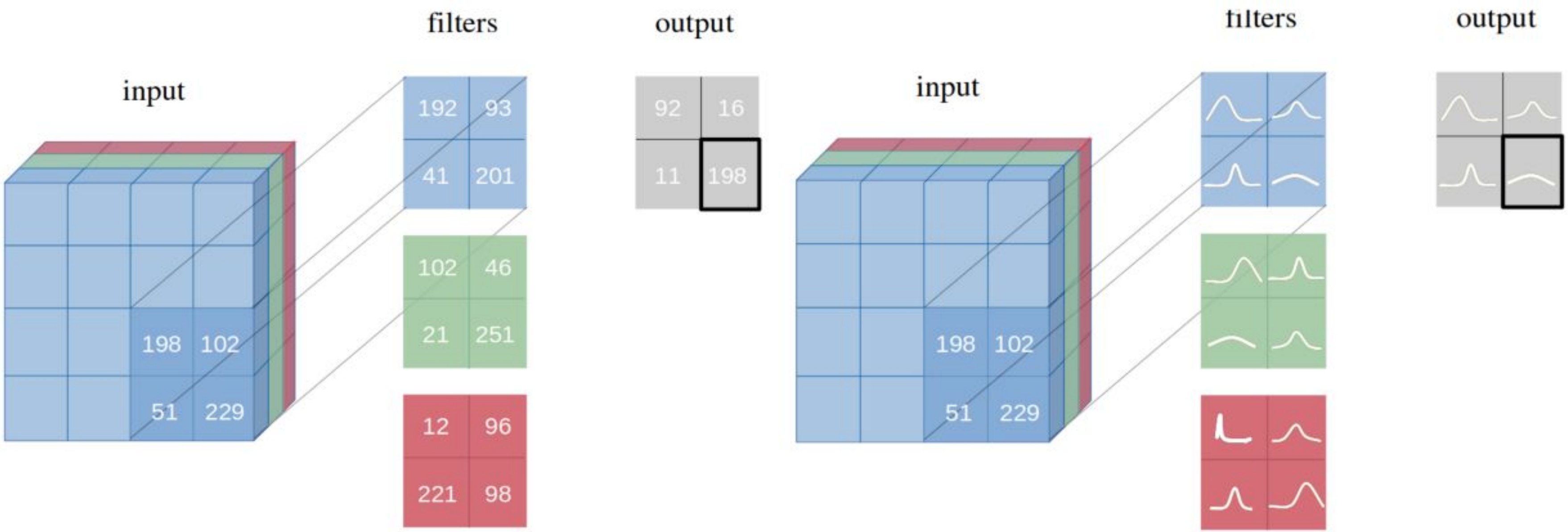
Prediction: dog
Probability: 0.95

Bayesian Neural Networks



Unlike deterministic Neural Networks(left) that have a fixed value of their parameters, Bayesian Neural Networks(right) has a distribution defined over them.

Bayesian Neural Networks



Bayesian Neural Networks

$$P(\mathbf{w}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathbf{w}).P(\mathbf{w})}{P(\mathcal{D})}$$

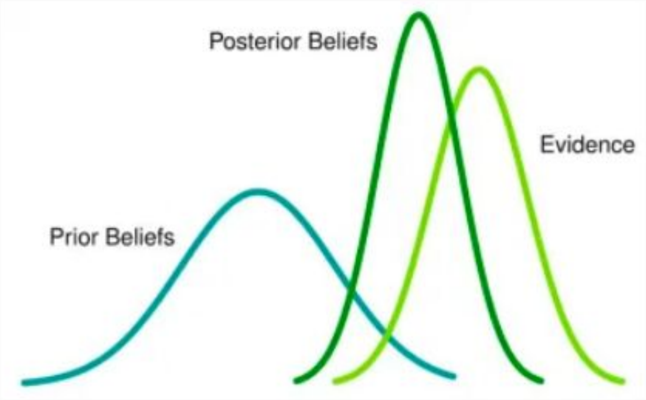
where,

$$P(\mathbf{w}|\mathcal{D}) \Longrightarrow \text{Posterior parameter distribution}$$

$$P(\mathcal{D}|\mathbf{w}) \Longrightarrow \text{Data likelihood}$$

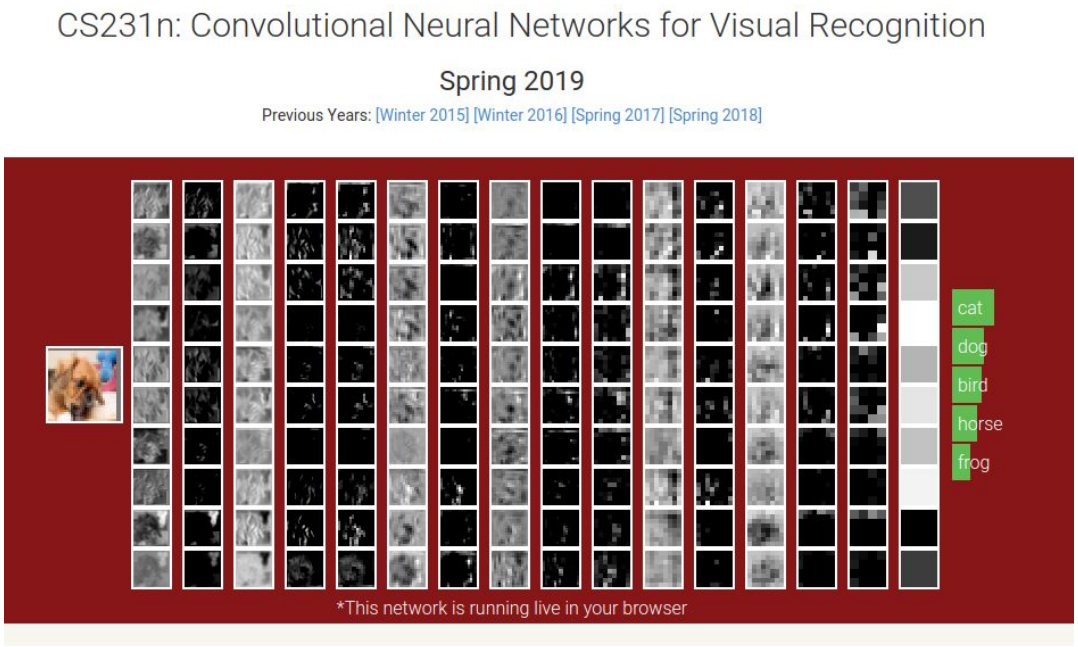
$$P(\mathbf{w}) \Longrightarrow \text{Prior parameter distribution}$$

$$P(\mathcal{D}) \Longrightarrow \text{Evidence}$$



Bayesian inference adjusts the beliefs about a distribution in the light of data or evidence

Dónde comenzar?



http://cs231n.stanford.edu/

Dónde aprender más?

coursera

UDACITY

edX®

MITx

fast.ai

deeplearning.ai

Herramientas para AI en Sistemas Embebidos

Jetson Nano

Coral TPU

Movius USB stick

Congresos relacionados a AI

- Perú

INTERCON 2019

SIMBig19

SHIRCON 2019
SCIENCES AND HUMANITIES INTERNATIONAL RESEARCH CONFERENCE

XPOTRON 2019

Congresos importantes

- Iberoamérica

Congresos importantes

- A nivel mundial

ICML
International Conference
On Machine Learning

ECCV'20

CVPR