

# Introduction to deep probabilistic models

Mauricio A. Álvarez PhD,  
H.F. Garcia C. Guarnizo (TA)

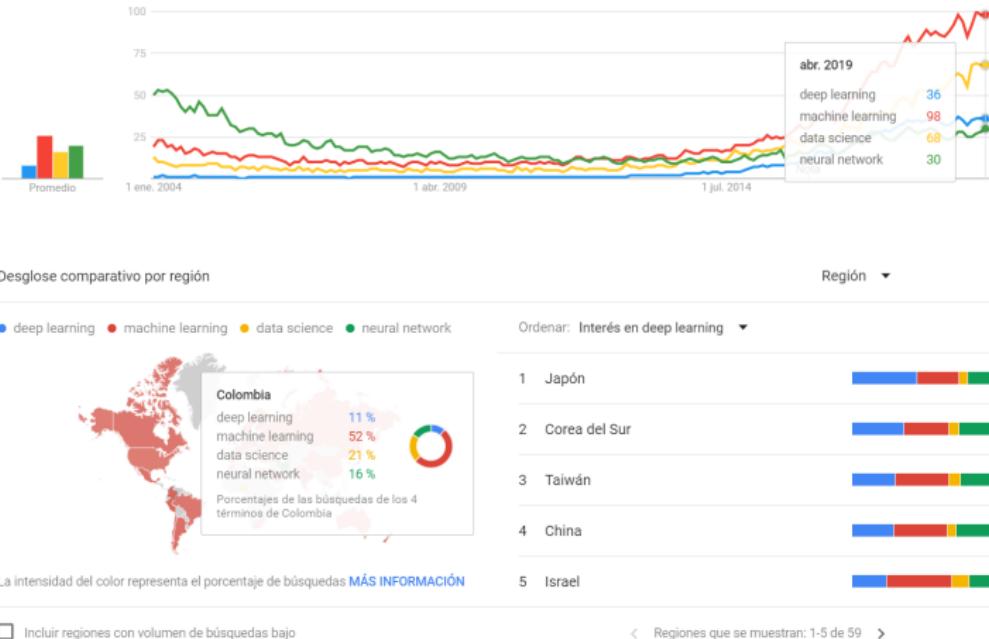


Universidad Tecnológica de Pereira, Pereira, Colombia

## 1 Introducción

## 1 Introducción

# Comenzando con un cliché ...



**Figure:** Tendencias de búsqueda en Google.

<https://trends.google.com/>



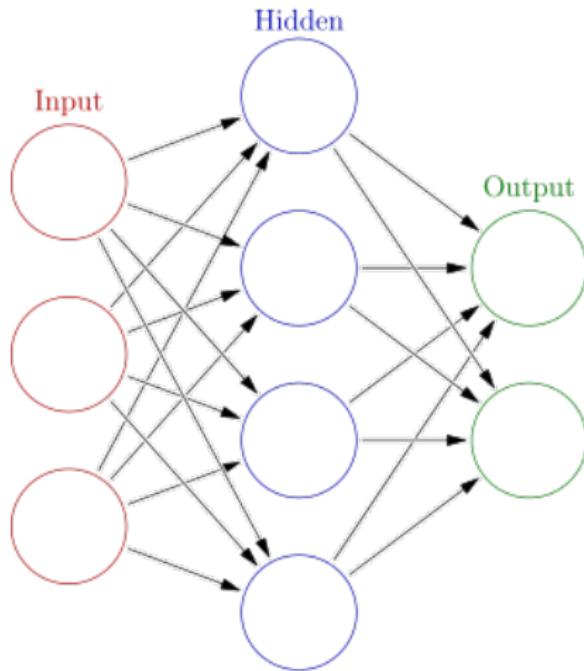
- **Fundamentals of Deep Probabilistic Model:** Introduction, Deep models, Bayesian inference.
- **Bayesian Neural Networks:** Approximate inference, Laplace approximation, Variational and Monte Carlo methods. Applications to CNNs and RNNs.
- **Deep Gaussian Processes:** Gaussian processes, deep Gaussian processes, inference in deep Gaussian processes, extensions.
- **Variational Autoencoders:** Autoencoders, variational autoencoders, normalizing flows, re-parameterization tricks, extensions.

# Motivación para el aprendizaje de modelos profundos

- Descomponer la tarea de aprendizaje
- Aprende conceptos simples. Usa esto para construir conocimiento de conceptos más complejos.

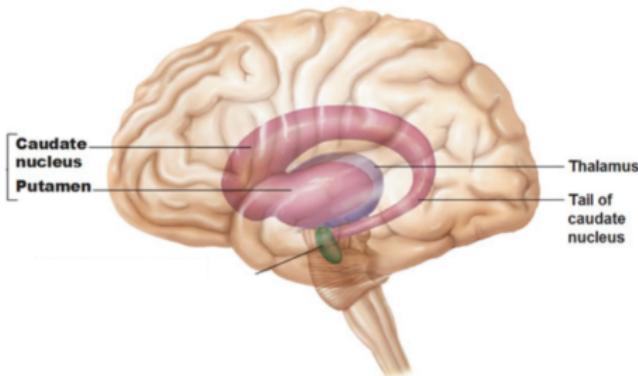


# Porque una red neuronal? I



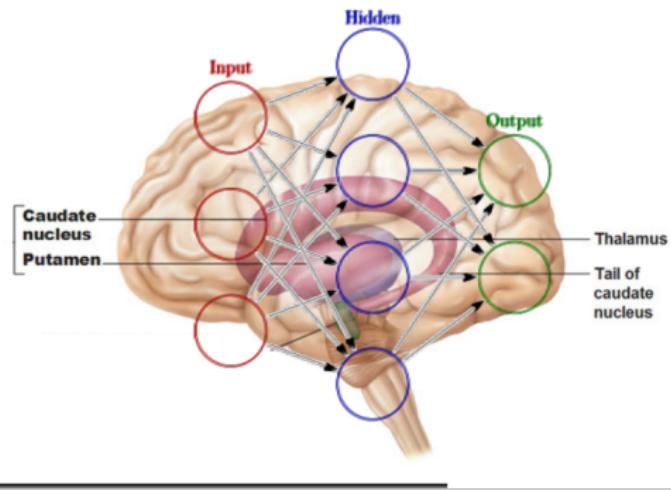
# Porque una red neuronal? II

## *Connectionism*



# Porque una red neuronal? III

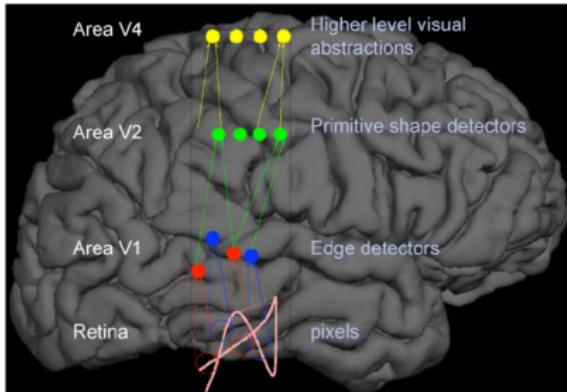
## Connectionism Neural Network



# Porque una red neuronal? IV



## Deep Architecture in the Brain



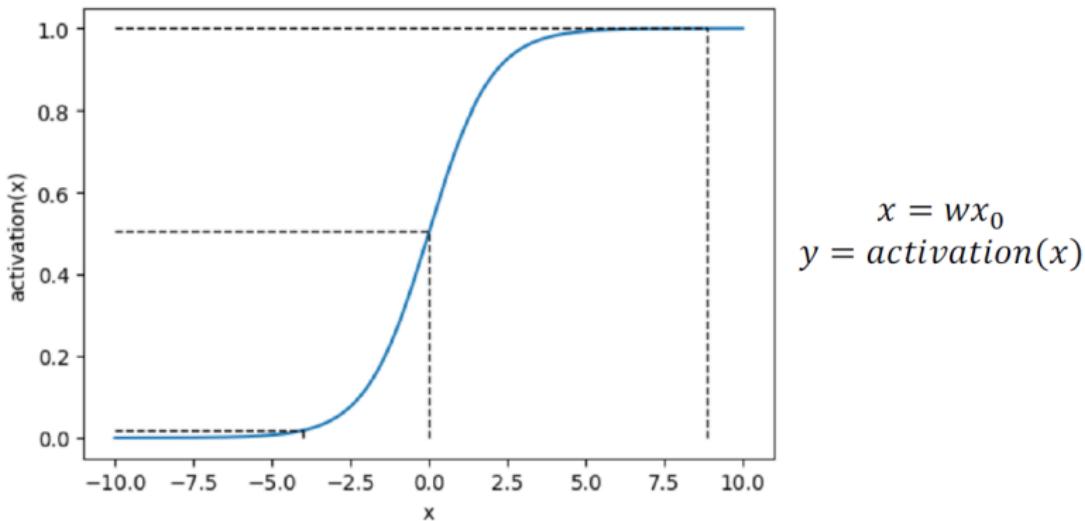
Ref: antranik.org

## Some applications . . .



# Porque una red neuronal? V

*Starting simple: Logistic Regression*



# Porque una red neuronal? VI

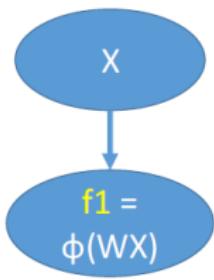
X (inputs)			y(outputs)
0	0	1	0
0	1	1	0
1	0	1	1
1	1	1	1

*This is a linear problem!*

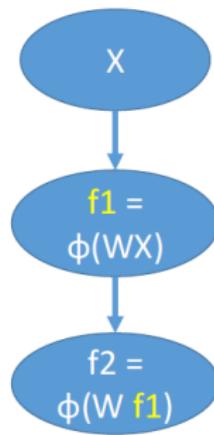


# Porque una red neuronal? VII

## Logistic Regression



## Deep Neural Network



# Porque una red neuronal? VIII



Get Started   Features   Ecosystem   Blog   Tutorials   Docs   Resources   GitHub   Q

## FROM RESEARCH TO PRODUCTION

An open source deep learning platform that provides a seamless path from research prototyping to production deployment.

Get Started >



# Porque una red neuronal? IX

 TensorFlow

Install Learn API Resources Community Why TensorFlow

Buscar Language GitHub Acceder

## An end-to-end open source machine learning platform

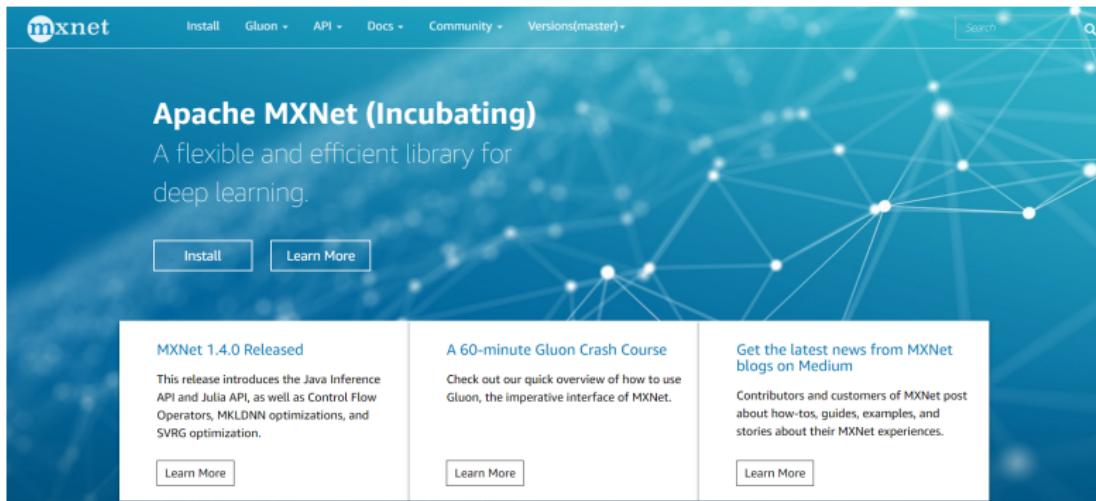
TensorFlow For JavaScript For Mobile & IoT For Production

The core open source library to help you develop and train ML models. Get started quickly by running Colab notebooks directly in your browser.

Get started with TensorFlow



# Porque una red neuronal? X



The screenshot shows the Apache MXNet website. At the top, there is a navigation bar with links for 'Install', 'Gluon', 'API', 'Docs', 'Community', and 'Versions[master]'. A search bar is located in the top right corner. The main header features the text 'Apache MXNet (Incubating)' and a subtext 'A flexible and efficient library for deep learning.' Below the header are two buttons: 'Install' and 'Learn More'. The background of the page is a blue gradient with a white network graph overlay. The central content area is divided into three columns: 'MXNet 1.4.0 Released', 'A 60-minute Gluon Crash Course', and 'Get the latest news from MXNet blogs on Medium'. Each column contains a brief description and a 'Learn More' button.

**Apache MXNet (Incubating)**

A flexible and efficient library for deep learning.

[Install](#) [Learn More](#)

**MXNet 1.4.0 Released**

This release introduces the Java Inference API and Julia API, as well as Control Flow Operators, MKLDNN optimizations, and SVRG optimization.

[Learn More](#)

**A 60-minute Gluon Crash Course**

Check out our quick overview of how to use Gluon, the imperative interface of MXNet.

[Learn More](#)

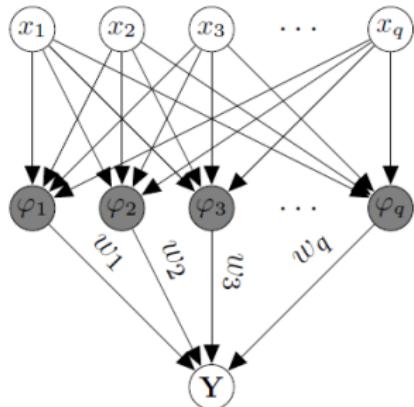
**Get the latest news from MXNet blogs on Medium**

Contributors and customers of MXNet post about how-tos, guides, examples, and stories about their MXNet experiences.

[Learn More](#)



# Una red neuronal estándar



- Definir  $\phi_j = \phi(x_j)$  y  $f_j = w_j\phi_j$  (ignorar el vías “por ahora”)
- Una vez que se han definido todos los  $w$ 's con *back-prop*, entonces  $f$  (y toda la red) se vuelve **determinista**
- Qué implica esto?

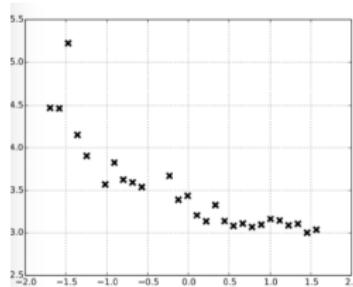
# La red neuronal entrenada es determinista. ¿Implicaciones? I

- **Generalización:** se produce sobre-entrenamiento.  
Necesidad de los regularizadores: abandono (drop-out), detención temprana (early stopping) ...
- **Generación de datos:** un modelo que generaliza bien, también debe comprender, o incluso poder crear ("imaginar") - variaciones.
- **Sin incertidumbre predictiva:** la incertidumbre debe propagarse en todo el modelo para que sea confiable.

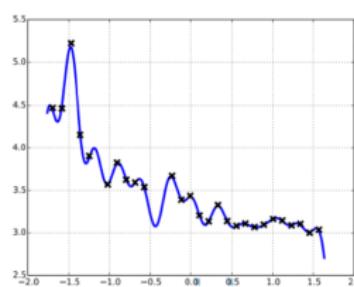


# La red neuronal entrenada es determinista. ¿Implicaciones? II

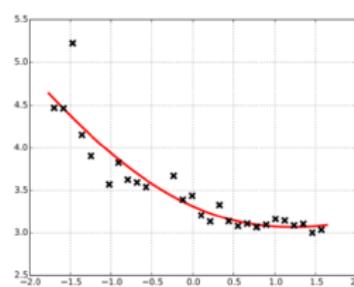
## Overfitting



(a)



(b)



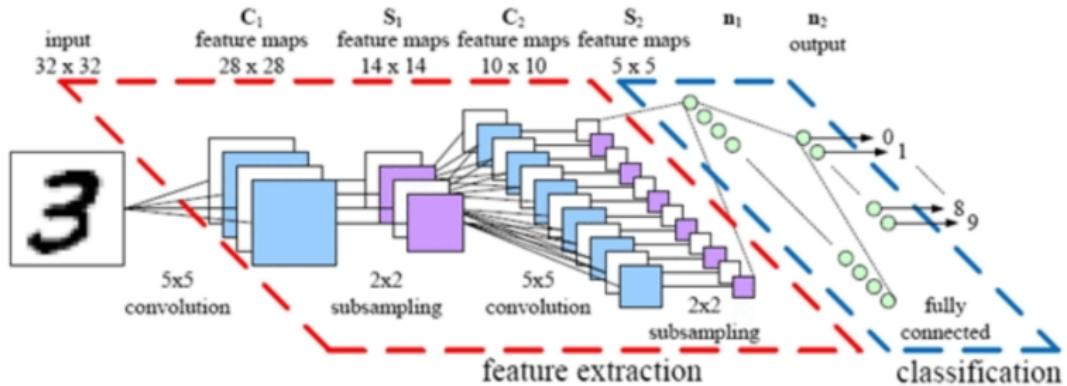
(c)

Which of the two curves (b) or (c) are better models for training data shown in (a)?

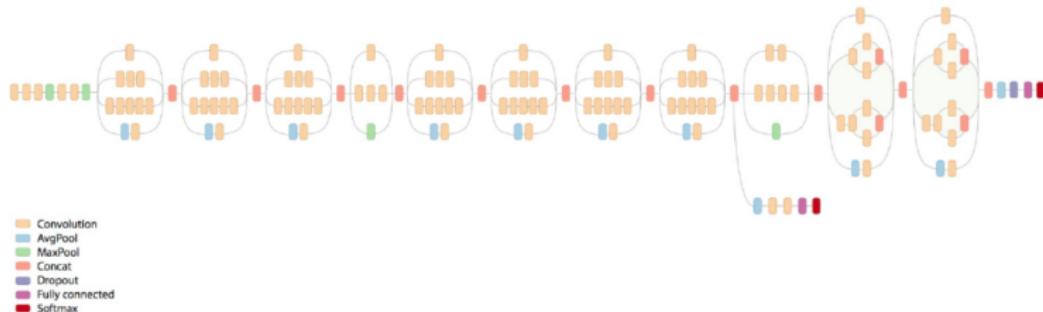
También podemos considerar diferentes activaciones y *wiring* de la red. Podemos combinar redes. Y más. Mientras todo siga siendo diferenciable. <http://playground.tensorflow.org/>



# La red neuronal entrenada es determinista. ¿Implicaciones? III



# La red neuronal entrenada es determinista. ¿Implicaciones? IV



Inception-v3



# Algunas aplicaciones I

Labradoodle or fried chicken



Image from: Yangyan Li



# Algunas aplicaciones II

Sheepdog or mop



Image from: Yangyan Li



# Algunas aplicaciones III

Chihuahua or muffin



Image from: Yangyan Li

@teenybiscuit



# Algunas aplicaciones IV

Barn owl or apple



Image from: Yangyan Li

@teenybiscuit



# Algunas aplicaciones V

Puppy or bagel



Image from: Yangyan Li



# Algunas aplicaciones VI

Parrot or guacamole



Image from: Yangyan Li

teenybiscuit



# Algunas aplicaciones VII

Raw chicken or Donald Trump



Image from: Yangyan Li



# Algunas aplicaciones VIII

## *Image Recognition: beyond binary classification*



The screenshot shows the Amazon Rekognition interface. At the top, there's a navigation bar with 'Menu', the Amazon logo, 'Products', 'Solutions', 'Pricing', 'Software', 'More', language 'English', 'My Account', and a 'Sign In to the Console' button. Below the navigation bar, the main content area has a heading 'Object and Scene Detection'. A descriptive text block explains that Rekognition identifies thousands of objects and scenes with confidence scores. To the right is a photograph of a person riding a mountain bike on a rocky trail. Overlaid on the image are several text labels with confidence percentages: 'PERSON 99.3%', 'OUTDOORS 83.1%', 'CREST 83.0%', 'MOUNTAIN BIKE 99.1%', and 'ROCK 82.8%'. The background of the main content area is white.



## *Convolutional Neural Network*

# Algunas aplicaciones X

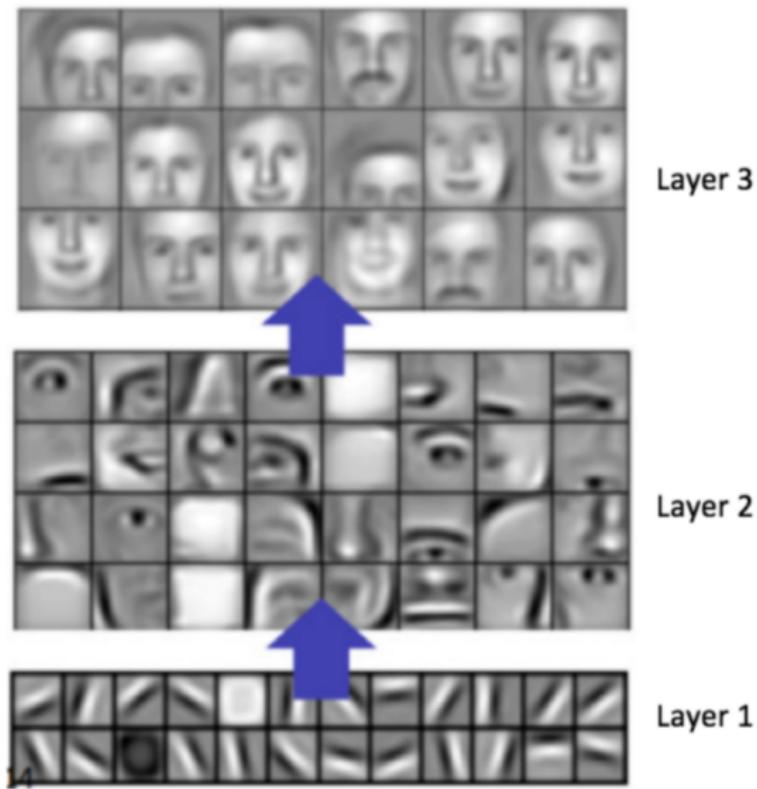


Input

[ujjwalkarn.me](http://ujjwalkarn.me)



# Algunas aplicaciones XI



[ujjwalkarn.me](http://ujjwalkarn.me)



# Necesidad de incertidumbre

- Aprendizaje por refuerzo
- Sistemas predictivos críticos
- Aprendizaje activo
- Sistemas semi-automaticos
- Escenarios de datos escasos.
- ...



## Modelos conceptualmente sencillos

- **DATA:**  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ ,  $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$
- **Modelo:** dadas las matrices  $\mathbf{W}$  y funciones  $\sigma(\cdot)$ , la red multivariada se define como:

$$\tilde{\mathbf{y}}_i(\mathbf{x}_i) = \mathbf{W}_2 \cdot \sigma(\mathbf{W}_1 \mathbf{X}_i) \quad (1)$$

- **Objetivo:** encontrar  $\mathbf{W}$  para el cual  $\tilde{\mathbf{y}}_i(\mathbf{x}_i)$  es cercano a  $\mathbf{y}_i(\mathbf{x}_i)$  para todo  $i \leq N$

# Porque el Deep Learning? II

## Deep learning es cool!!

- Simple y modular
- Excelentes herramientas de software
- Escalable con los datos y el cómputo
- Existen aplicaciones de impacto real

... pero aún tiene muchos problemas (no uncertainty!)

- Qué no conoce mi modelo?
- No se puede interpretar fácilmente (black-boxes)
- Fácilmente engañable (AI safety)
- Carece de fundamentos matemáticos sólidos (mostly ad-hoc)
- Se basa principalmente en big data



# ¿Por qué debería preocuparme por la incertidumbre?

- Necesitamos una forma de saber qué sabe nuestro modelo y qué no.
  - Se crea un modelo para reconocer razas de perros.



# ¿Por qué debería preocuparme por la incertidumbre?

- Necesitamos una forma de saber qué sabe nuestro modelo y qué no.
  - Se crea un modelo para reconocer razas de perros.
  - Y se le da un gato para clasificar
  - ¿Qué te gustaría que hiciera tu modelo?



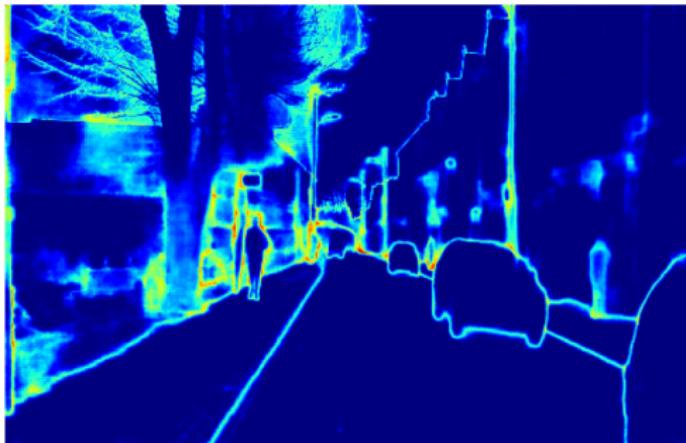
# ¿Por qué debería preocuparme por la incertidumbre?

- Necesitamos una forma de saber qué sabe nuestro modelo y qué no.
  - Se crea un modelo para reconocer razas de perros.
  - Y se le da un gato para clasificar
  - ¿Qué te gustaría que hiciera tu modelo?
  - Problemas similares en la toma de decisiones, física, ciencias de la vida, etc.



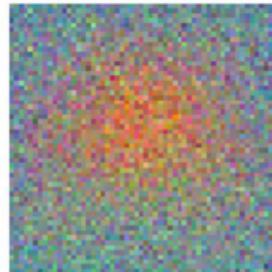
# ¿Por qué debería preocuparme por la incertidumbre?

- Necesitamos una forma de saber **qué sabe nuestro modelo** y qué no.
- La incertidumbre da una idea de la caja negra cuando falla. ¿Dónde no estoy seguro?



# ¿Por qué debería preocuparme por la incertidumbre?

- Necesitamos una forma de saber **qué sabe nuestro modelo** y qué no.
- La incertidumbre da una idea de la caja negra cuando falla. ¿Dónde no estoy seguro?
- ¡La incertidumbre incluso puede ser útil para identificar cuándo se ataca con ejemplos contradictorios (adversarial)!



# ¿Por qué debería preocuparme por la incertidumbre?

- Por último, necesita menos datos si la etiqueta solo es incierta cuando el modelo es incierto: desgaste en robótica, tiempo experto en análisis médico



# Conclusiones

- Los modelos clásicos de DL son matemáticamente simples; El reto está en cómo optimizarlos.
- ¿Eficiencia de datos? ¿Calibración de incertidumbre? Interpretabilidad? ¿La seguridad?

