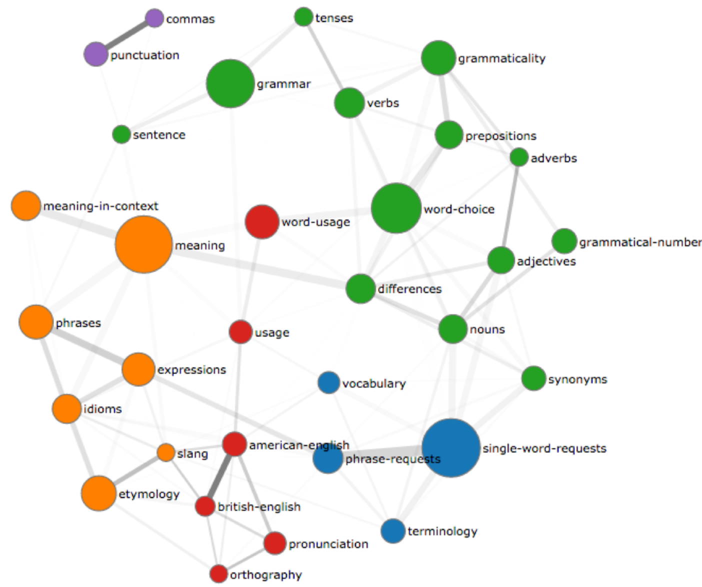


Deep Learning for NLP Workshop

Jay Urbain, PhD
Professor, Department of Electrical Engineering and Computer Science
Milwaukee School of Engineering

Deep Learning for NLP Workshop

- Deep Learning
- NLP, why NLP is hard,
- Lab 1 – Sentiment Classification
  - Text classification
  - NLP vocabulary
  - Embeddings
  - Network architectures: MLP, CNN, RNN
  - Model development and evaluation, review
- Lab 2 – Language/Date Translation
  - Encoder-decoder
  - Attention
  - Network architecture: LSTM with attention
  - Model development and evaluation, review

WORKSHOP PERSPECTIVE

- What this workshop is:
  - An intro to Natural Language Processing (NLP) using Deep Learning
  - Lectures followed by step-by-step hands-on exercises using Jupyter Notebooks, TensorFlow and Keras
  - Emphasis is on the hands-on exercises
  - Downloadable Docker image

- This workshop is not:
  - Comprehensive review of NLP
  - Deep Learning from first principles
  - Extensive programming and knowledge of APIs

EXPECTATIONS

- Assumptions:
  - Basic Python
  - Basic understanding of Machine Learning
- Nice to have:
  - Some deep learning, any NLP experience

Consider working with a partner

HOPEFULL TAKE AWAYS

- Ability to design a deep learning workflow to conduct NLP based text tasks
- Have a template for creating your own sentiment analysis NLP research

Running the NLP workshop labs

1) Install Docker CE

https://www.docker.com/community-edition

2) Open a terminal and run Docker to launch the workshop image in a Docker container

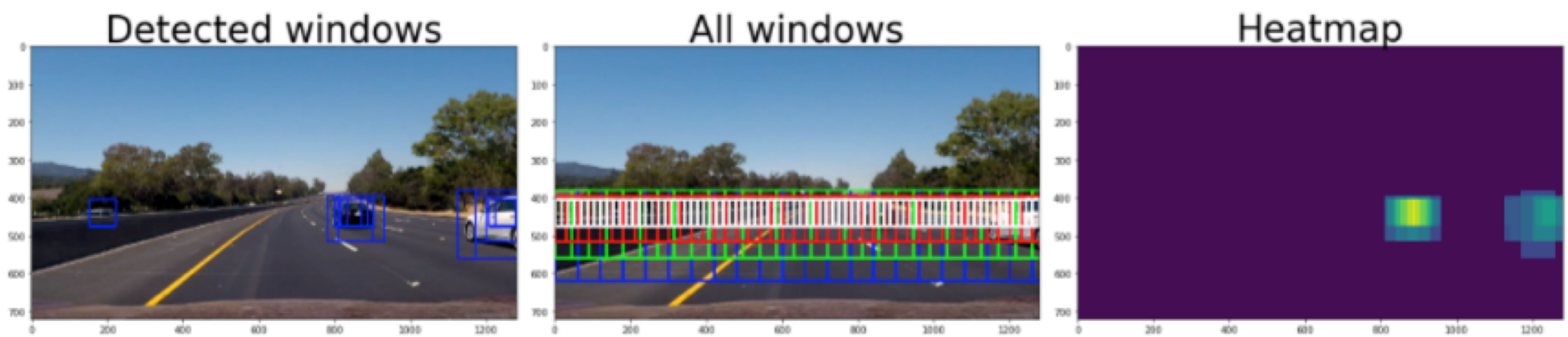
$ docker run -p 8889:8888 jayurbain/deepnlpintro:latest

3) Open Chrome browser to localhost, port 8889, with the token

http://localhost:8889/?token=e1bc2fe905e1b0de7ec0820b03841c9d7d3bb43437754
0df

If you would prefer to set up your own environment, the lab exercises can be found at the following github repository:

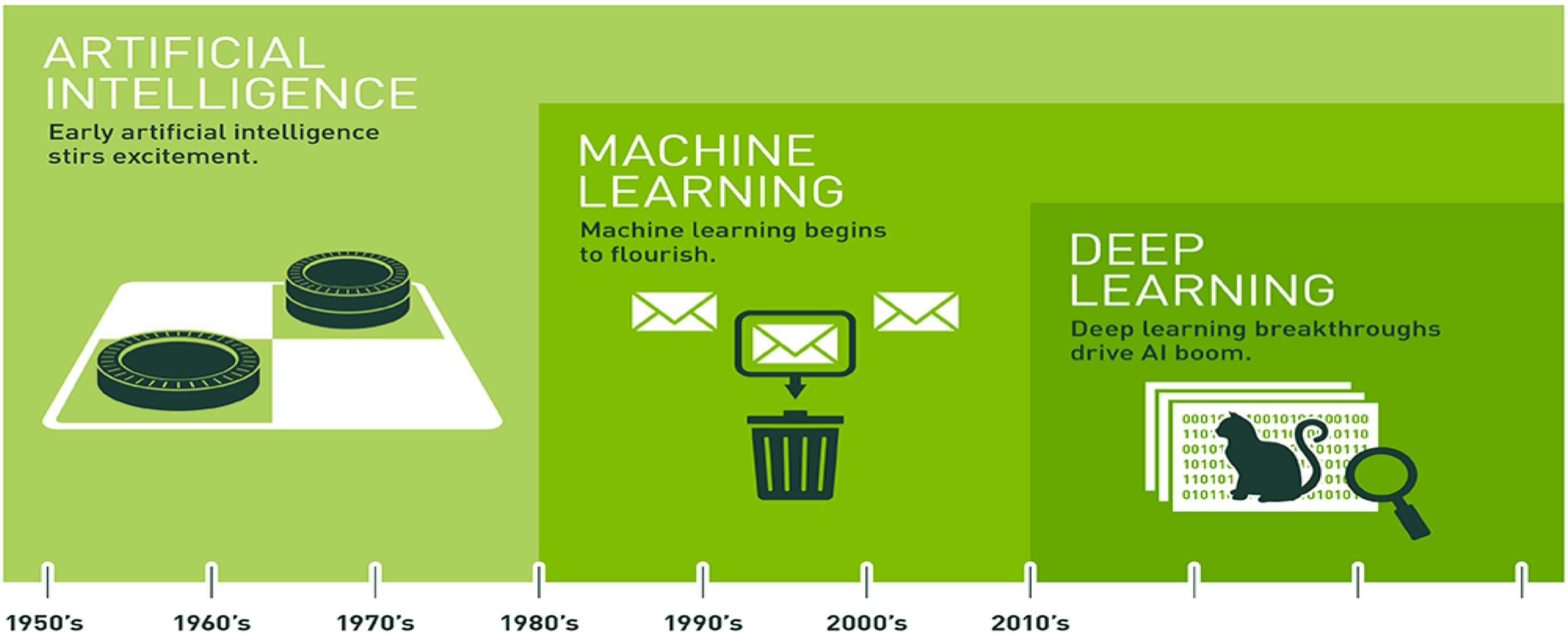
https://github.com/jayurbain/DeepNLPIntro



Intro to Deep Learning

Jay Urbain, PhD
Professor, Department of Electrical Engineering and Computer Science
Milwaukee School of Engineering

ACCOMPLISHING COMPLEX GOALS

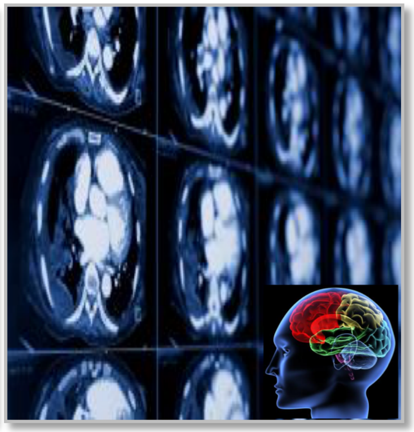


SWEEPING ACROSS INDUSTRIES

Internet Services

- Image/Video classification
- Speech recognition
- Natural language processing

Medicine



- Cancer cell detection
- Diabetic grading
- Drug discovery

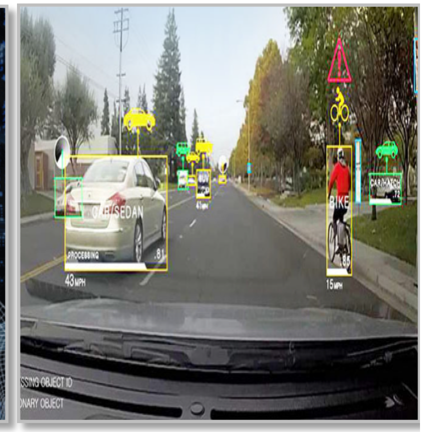
Media & Entertainment

- Video captioning
- Content based search
- Real time translation

Security & Defense

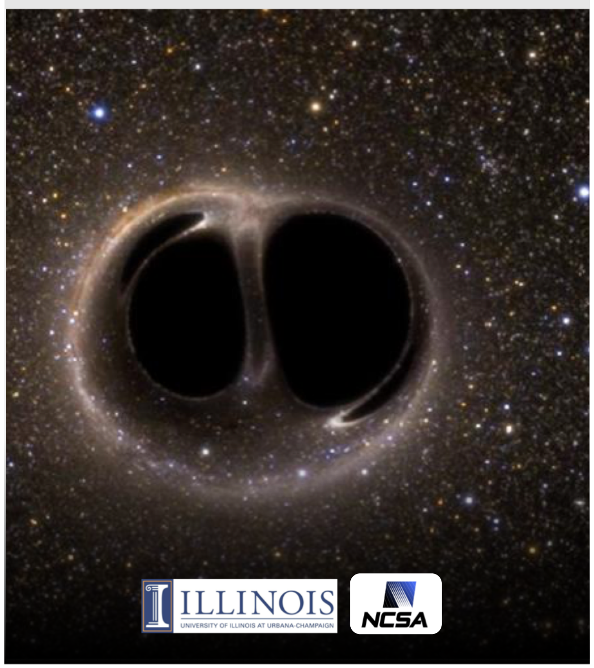
- Face recognition
- Video surveillance
- Cyber security

Autonomous Machines

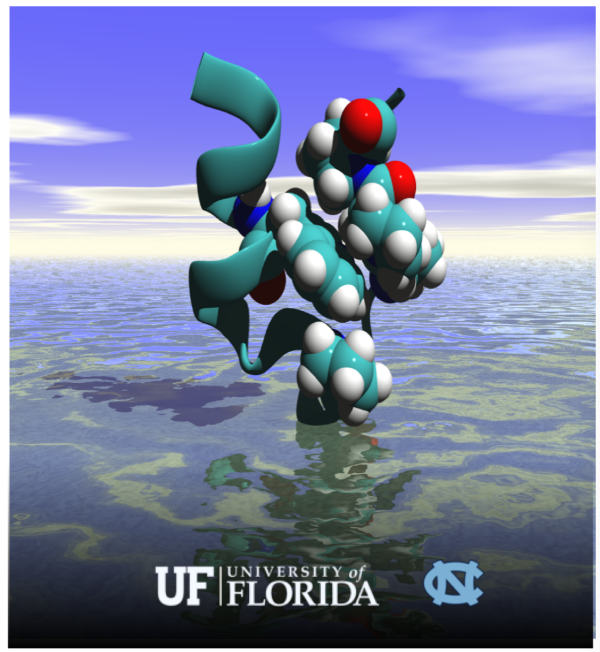


- Pedestrian detection
- Lane tracking
- Recognize traffic signs

TRANSFORMING RESEARCH



"Seeing" Gravity In Real Time



Accelerating Drug Discovery

92%

believe AI will impact their work

93%

using deep learning seeing positive results

insideHPC.com Survey
November 2016

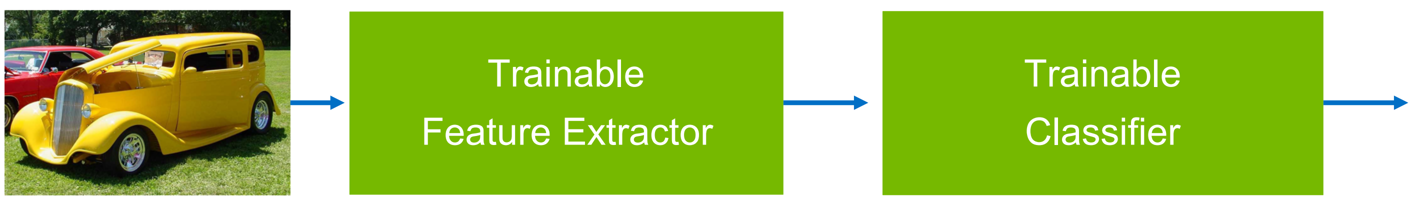
Deep learning =
Learning representations/features

– The traditional model of pattern recognition (since the late 50's)

  – Fixed/engineered features (or fixed kernel) + trainable classifier

– End-to-end learning / Feature learning / Deep learning

  – Trainable features (or kernel) + trainable classifier

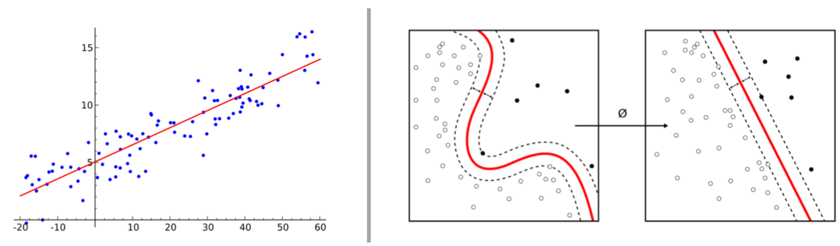


Difference in Workflow

Classic Machine Learning [1990 : now]

Input → Hand Designed **Features** → Model / Mapping → Output

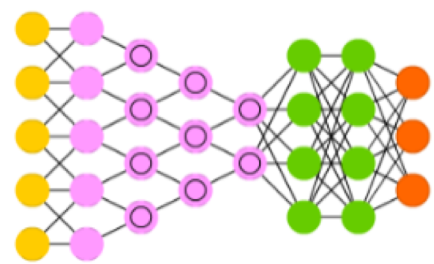
Examples [Regression and SVMs]



Deep/End-to-End Learning [2012 : now]

Input → Simple Features → Complex Features → Model/ Mapping → Output

Example [Conv Net]

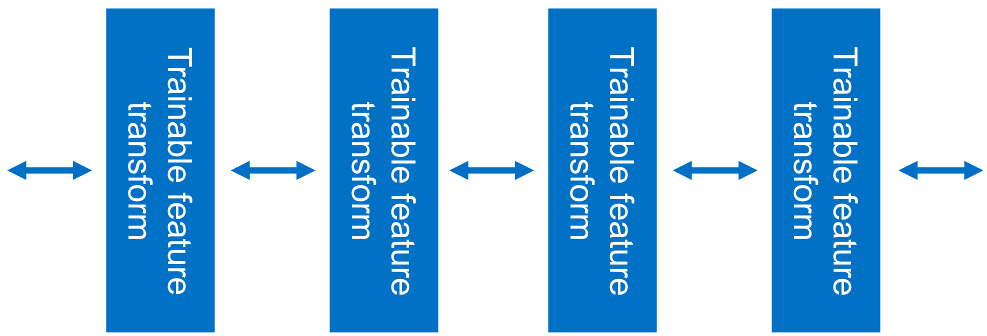


Machine learning workflow shifts from engineering features for “shallow” models
to architecting deep learning models with the ability to learn hierarchical representations of features

Trainable feature hierarchy

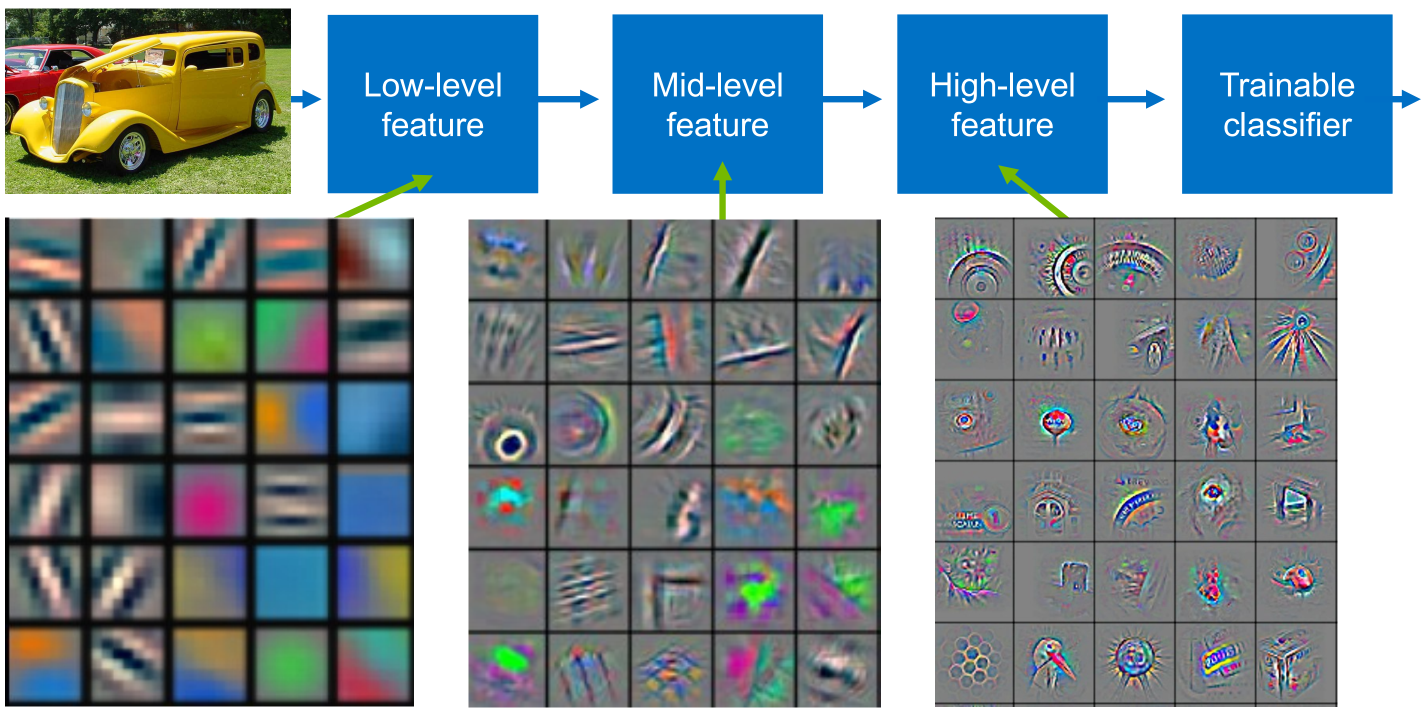
Hierarchy of representations with increasing level of abstraction. Each stage is a kind of trainable feature transform

- Image recognition
  - Pixel → edge → motif → part → object
- Text
  - Character → word → word group → clause → sentence → story/semantic understanding
- Speech
  - Sample → spectral band → sound → … → phone → phoneme → word

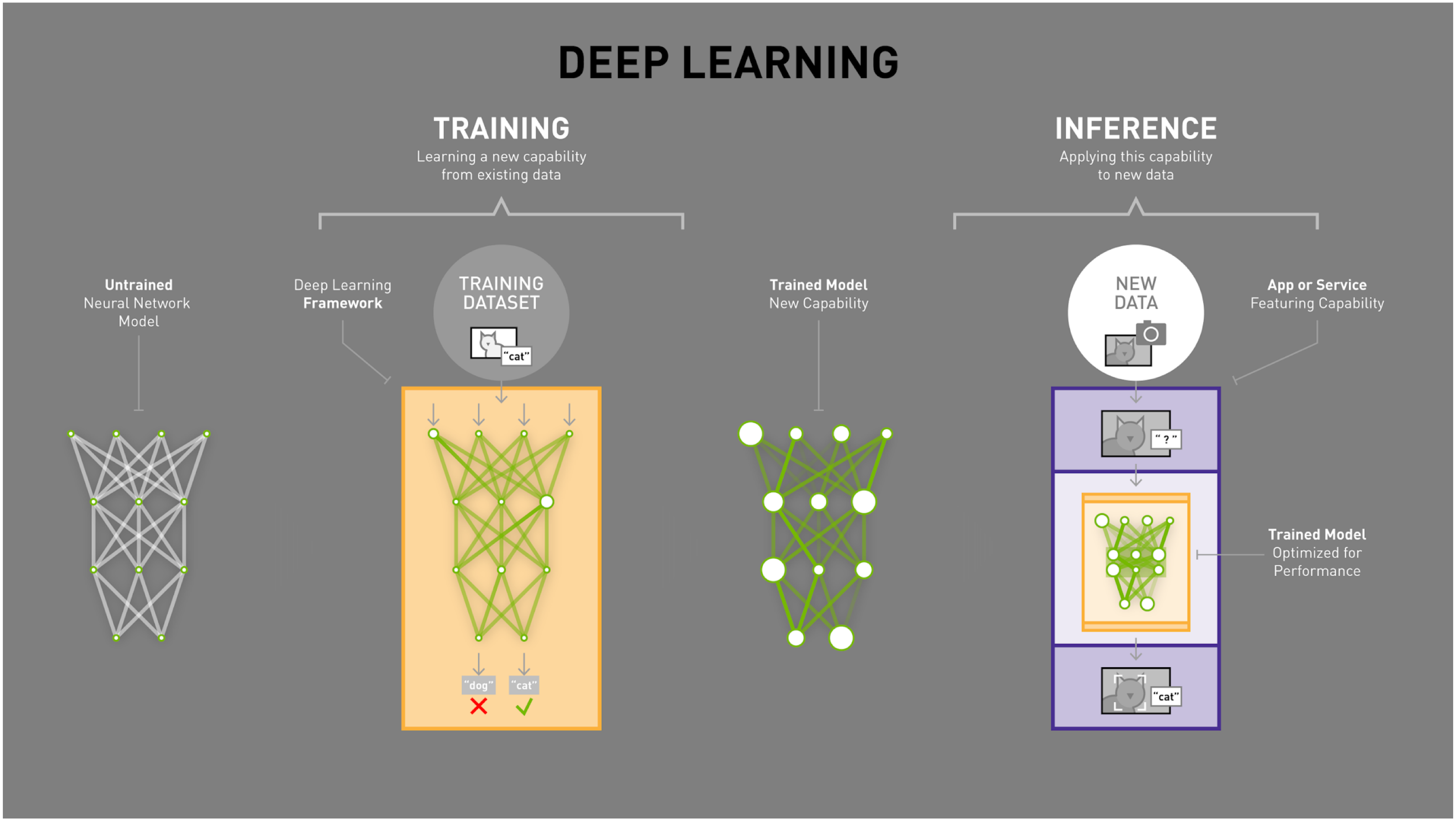


Deep learning = learning hierarchical representations

It's deep if it has more than one stage of non-linear feature transformation



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]


DEEP LEARNING
TRAINING
Learning a new capability from existing data
INFERENCE
Applying this capability to new data
Untrained Neural Network Model
Deep Learning Framework
TRAINING DATASET
"cat"
"dog"
"cat"
Trained Model New Capability
NEW DATA
App or Service Featuring Capability
" ? "
Trained Model Optimized for Performance
"cat"

Rationale for Deep Learning

Costs of acquiring and storing large quantities of heterogeneous data have dropped significantly.

The ability to extract actionable knowledge from such datasets has become critical for companies to maintain a competitive advantage.

Significant improvements in deep learning algorithms enabled by GPU processing provide new application opportunities across industries.

Deep Learning has set new performance standards in many machine learning applications.