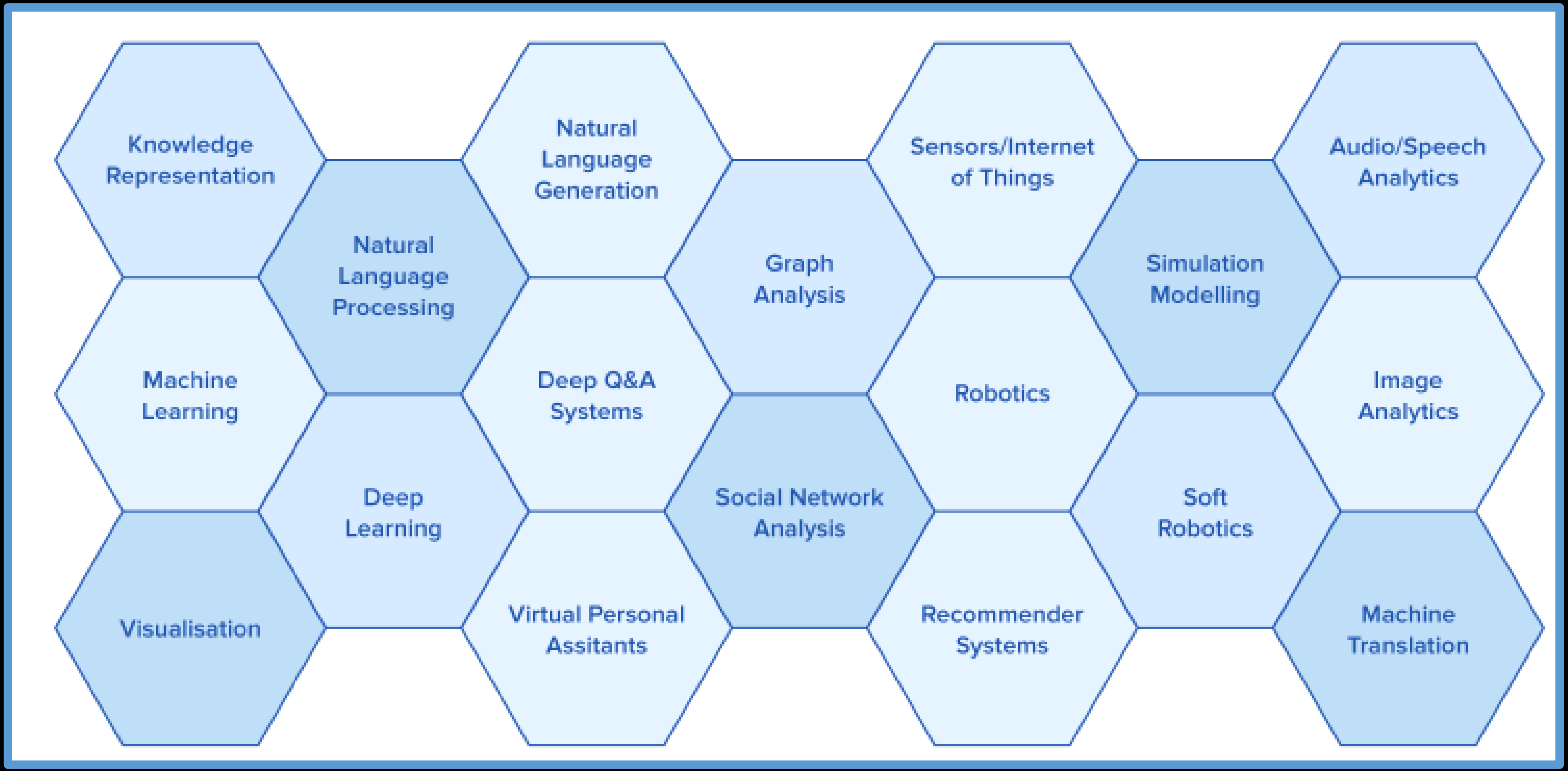
Advanced Machine Learning (AML) Overview

Palacode Narayana Iyer Anantharaman

7 Aug 2018

Copyright 2016 JNResearch, All Rights Reserved

AI Landscape (Source: PWC)



Copyright 2016 JNResearch, All Rights Reserved

AML course in perspective

- Artificial Intelligence is about building machines that have human like intelligence
  - How do we define intelligence in the context of AI?
- Deep Learning is the primary technique used to build AI systems
- Deep learning is a part of machine learning which relies on "learning from data" as opposed to algorithmic coding
- AML mostly is centred around deep learning but the scope includes other traditional ML approaches as well

Copyright 2016 JNResearch, All Rights Reserved

Our approach

- A number of MOOC courses are available on the subject of Deep Learning
- Our course includes the core content covered in these courses
- In addition, we address the following:
    - Theoretical rigor combined with adequate hours of hands on component
    - State of the art techniques that are relevant for product development
    - Traditional methods that are vital for product development
    - Techniques to develop practical, industry quality products

Copyright 2016 JNResearch, All Rights Reserved

Course Outline

- Nuts and Bolts
- Handling Spatially structured inputs
- Handling time series and sequence models
- Generative models and other topics

Copyright 2016 JNResearch, All Rights Reserved

Assumptions

- We assume the students are already familiar with core ML concepts, such as different types of learning, bias/variance, linear models, softmax distribution, clustering techniques, PCA, basic neural networks
- From the lab perspective, we assume that students are good at Python 3.x and can code with Keras/TensorFlow/Pandas/OpenCV/Matplotlib (Overview of these tools will be covered in the next lecture)
- We will emphasize on learning the theory and building systems from the first principles as opposed to mastering a specific framework like Keras. However we will use these frameworks to get hands on quickly to develop certain applications

Copyright 2016 JNResearch, All Rights Reserved

Unit 1

- Overview and Maths background
- Bayesian Approaches
- Review of Neural Networks and Autoencoders
- Keras, TensorFlow, Other tools

Copyright 2016 JNResearch, All Rights Reserved

Unit 2 : Reinforcement Learning

- Review of SVM and other core models

- Deep Reinforcement Learning and applications

Copyright 2016 JNResearch, All Rights Reserved

Unit 3 : Sequence Models

- Recurrent Neural Networks
- LSTM, GRU
- Sequence to Sequence Models, attention networks
- Memory Networks
- Applications to text data, image captioning

Copyright 2016 JNResearch, All Rights Reserved

Unit 4 : Computer Vision

- Core tasks of Computer Vision
- Convolutional Neural Networks
- Modern architectures for core tasks
- Transfer Learning
- Application of Computer Vision for autonomous vehicles and other cutting edge applications

Copyright 2016 JNResearch, All Rights Reserved

Unit 5 : Synthesis with Deep Learning

- Generative Adversarial Networks and its variants

- Data augmentation for images

- Techniques for generating text data

Copyright 2016 JNResearch, All Rights Reserved

Course Evaluation

In Semester Assessment (ISA) Max : 40

Activity	**Marks**	**Remarks**
Unit Evaluation 1	10	Hands On based
Unit Evaluation 2	**10**	Hands On based
T2	**20**	Regular paper (theory)
Unit Evaluation 3	**10**	Optional lab hands on
Total	**40**	Best 2 out of 3 labs + T2

ASSESSMENT POLICY

End Semester Assessment (ESA) Max: 60

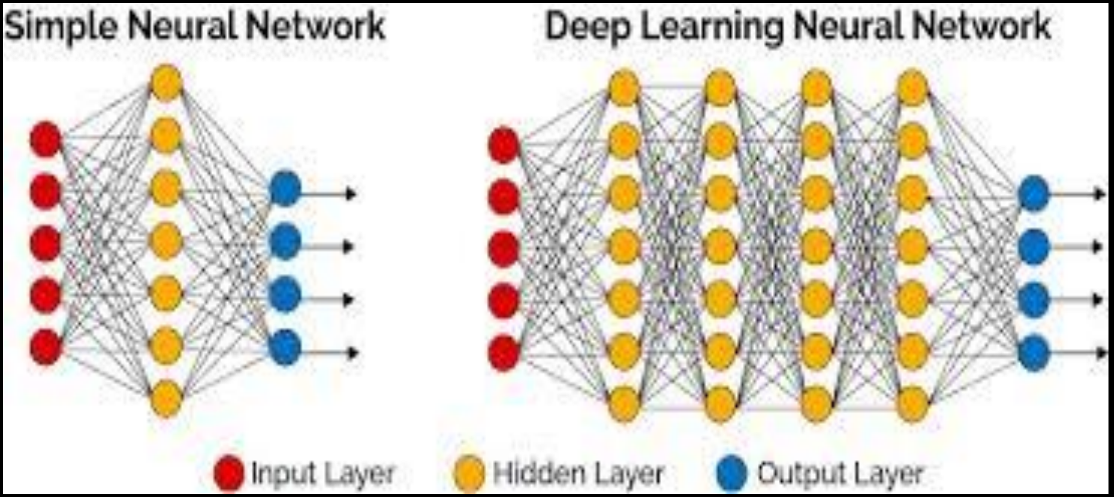
Activity	**Marks**	**Remarks**
Final Exam	25	Theory
3 day Hackathon	**35**	Hands On
Total	**60**	60

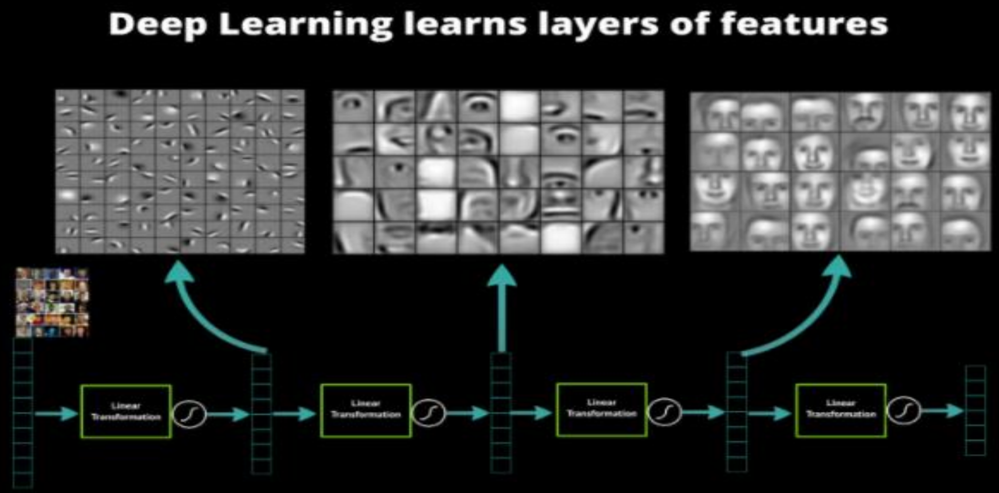
Copyright 2016 JNResearch, All Rights Reserved

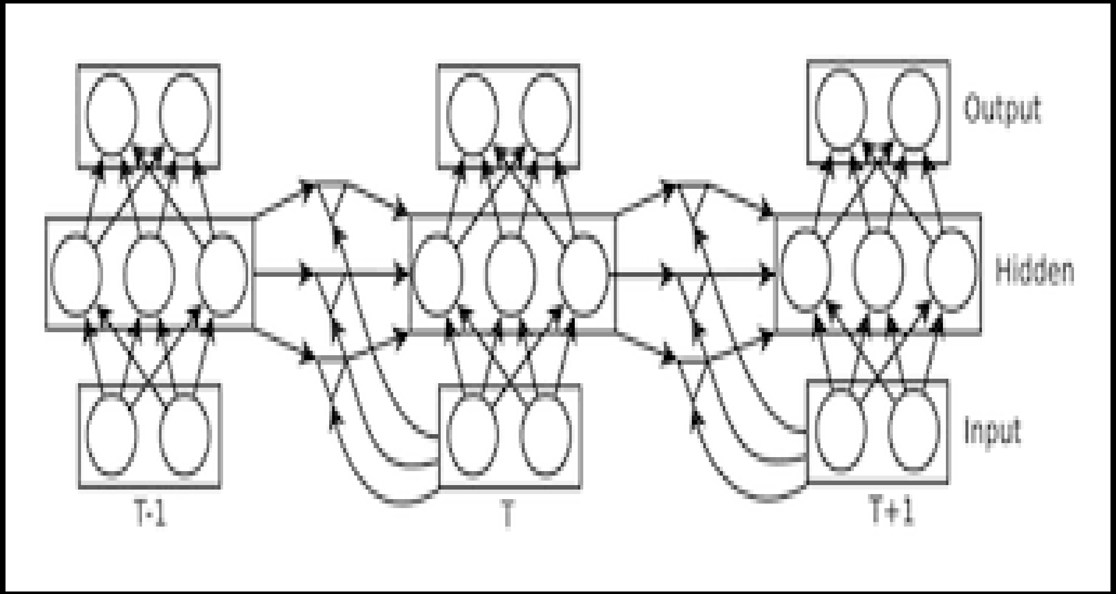
Deep Learning Overview

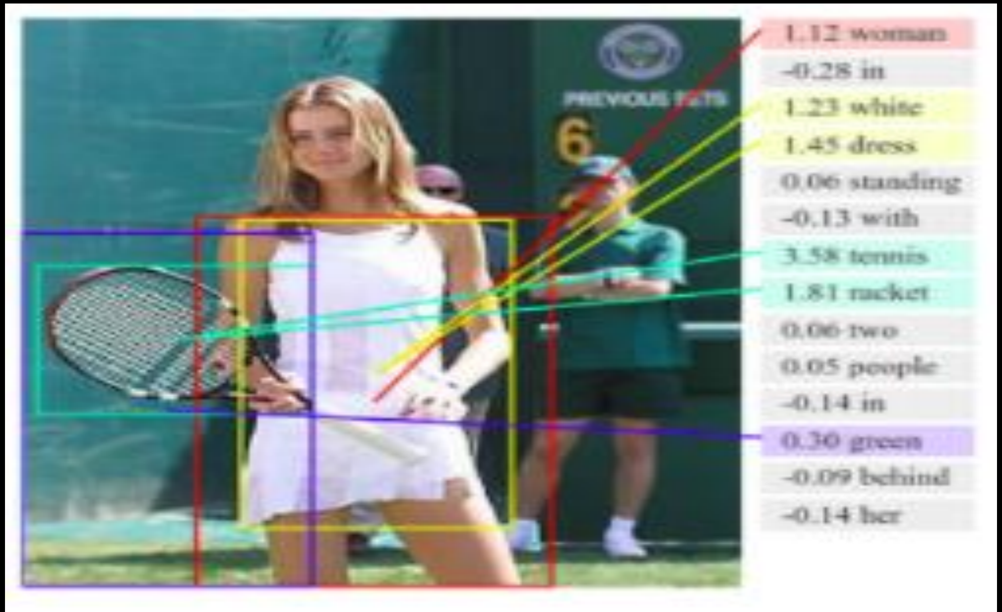
Copyright 2016 JNResearch, All Rights Reserved

Deep Learning Architectures





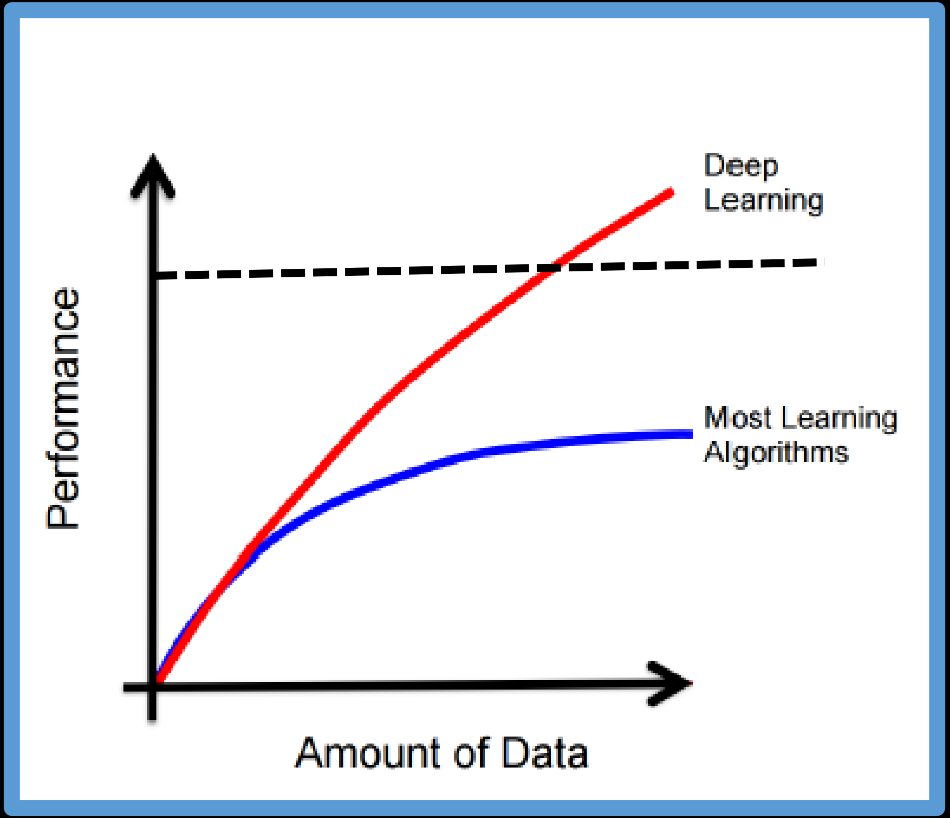




Copyright 2016 JNResearch, All Rights Reserved

Deep Learning

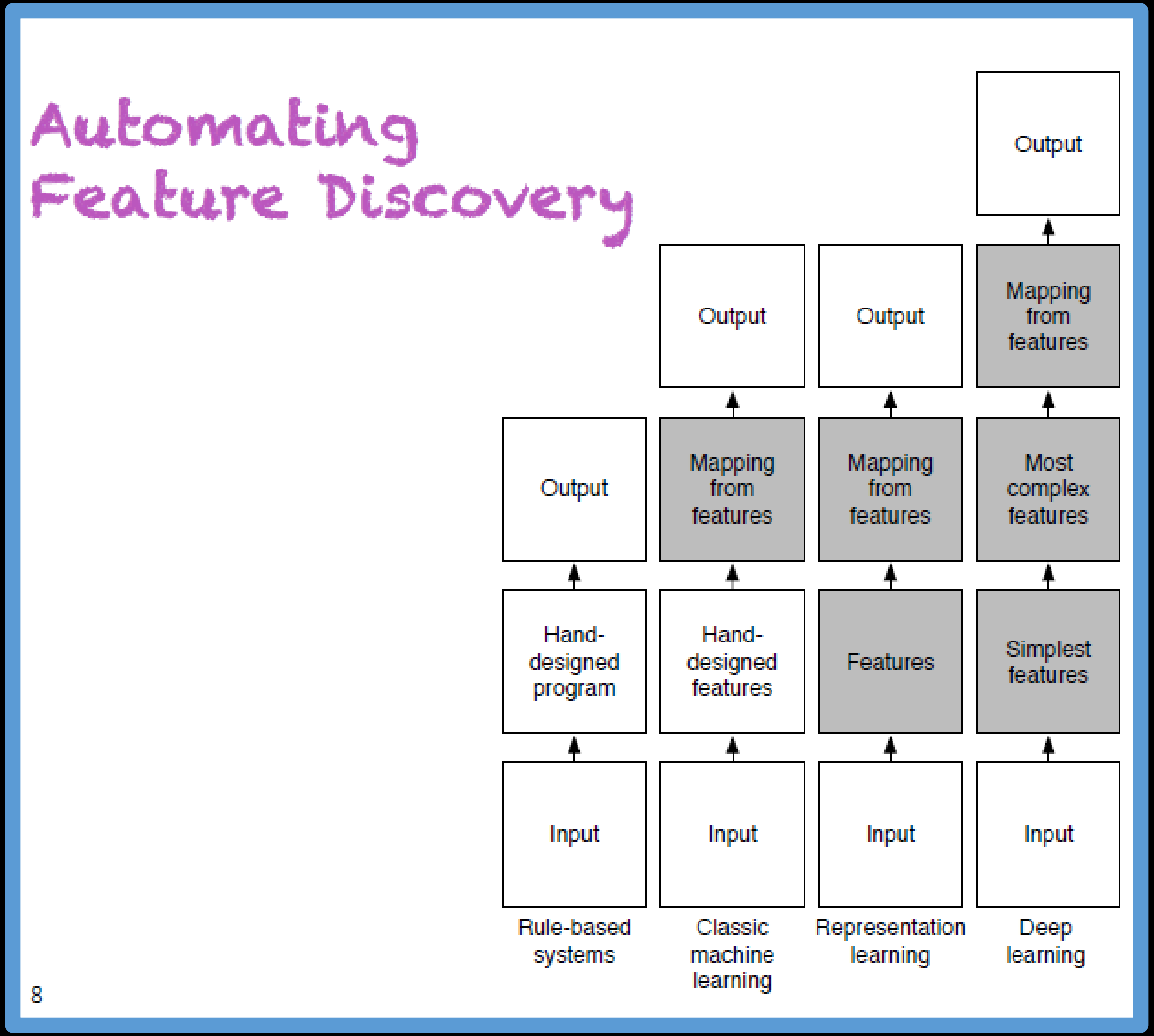
- Large number of layers forming a deep network
- The depth can be spatial or temporal
- More complex models but less dependency on human experts crafting the best features
- Due to the model's higher capacity, can leverage the data better – more the data you give, better can be the learning



Copyright 2016 JNResearch, All Rights Reserved

Feature Learning (fig from Y Bengio)

- Representation Learning
  - Automatically learn the "right" features at each hidden layer
  - Learn multiple levels of representations increasing in abstraction



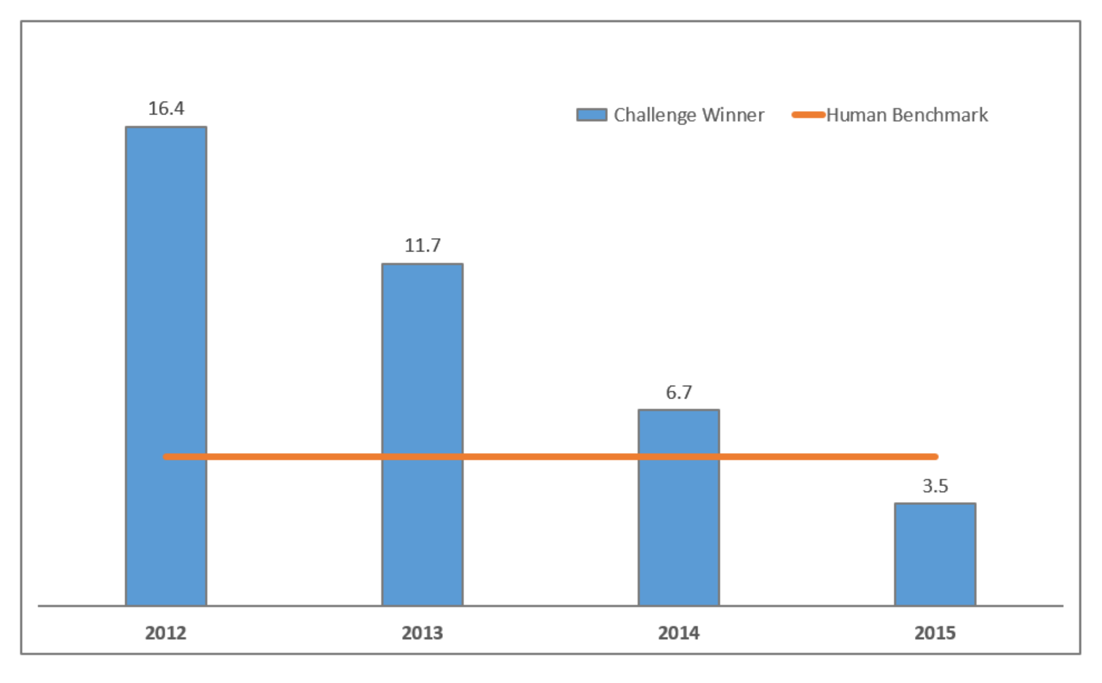
- Allow effective sharing of the learned parameters across different tasks: Multitask learning

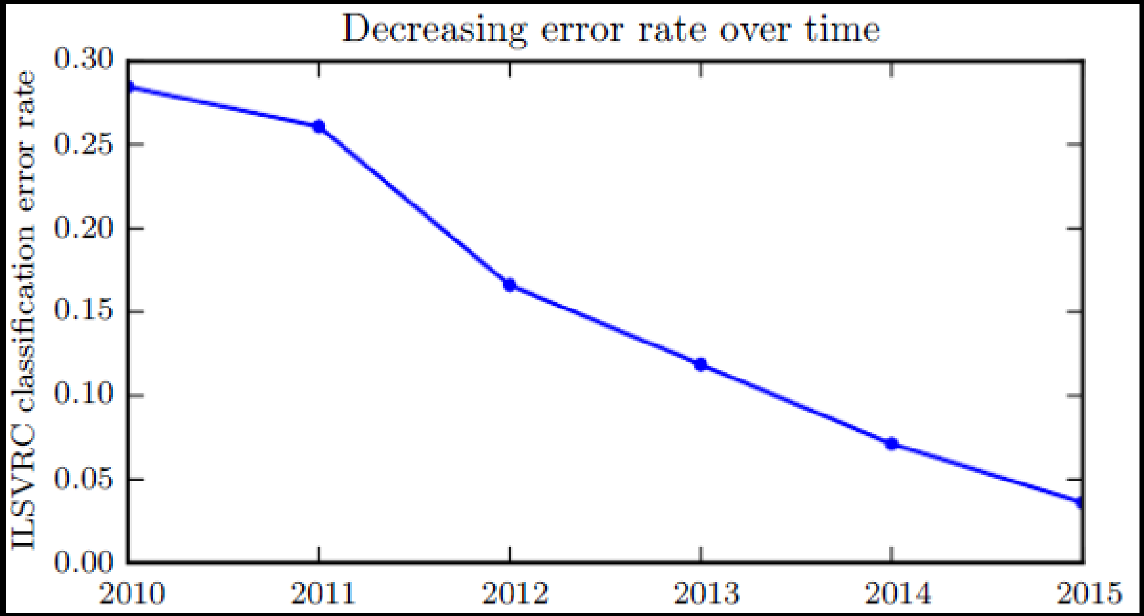
Three reasons to use deep learning

- **Performance**
  - The difference between 93% to 96% can make all the difference
  - Make cool technologies usable for a common man.
- **Broad Applicability (Domain independence)**
  - Not limited to a narrow set of problems
  - Minimize the need for domain specialized feature engineering
- **New class of applications**
  - Multimodal fusion
  - Generative models

Copyright 2016 JNResearch, All Rights Reserved

“**A dramatic moment in the meteoric rise of deep learning came when a convolutional network won this challenge for the first time and by a wide margin, bringing down the state-of-the-art top-5 error rate from 26.1% to 15.3%** (Krizhevsky *et al.*, 2012), meaning that the convolutional network produces a ranked list of possible categories for each image and the correct category appeared in the first five entries of this list for all but 15.3% of the test examples. **Since then, these competitions are consistently won by deep convolutional nets**, and as of this writing, advances in deep learning have brought the latest top-5 error rate in this contest down to 3.6%” – Ref: Deep Learning Book by Y Bengio et al

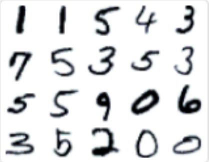




Starting with simple datasets: DL makes a difference

MNIST

who is the best in MNIST ?



MNIST 50 results collected

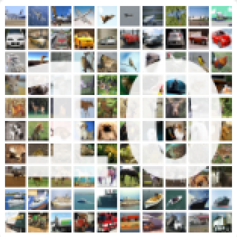
Units: error %

Classify handwriten digits. Some additional results are available on the original dataset page.

Result	Method	Venue	Details
0.21%	Regularization of Neural Networks using DropConnect	ICML 2013	
0.23%	Multi-column Deep Neural Networks for Image Classification	CVPR 2012	
0.23%	APAC: Augmented PAttern Classification with Neural Networks	arXiv 2015	
0.24%	Batch-normalized Maxout Network in Network	arXiv 2015	Details
0.29%	Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree	AISTATS 2016	Details

CIFAR-10

who is the best in CIFAR-10 ?



CIFAR-10 49 results collected

Units: accuracy %

Classify 32x32 colour images.

Result	Method	Venue	Details
96.53%	Fractional Max-Pooling	arXiv 2015	Details
95.59%	Striving for Simplicity: The All Convolutional Net	ICLR 2015	Details
94.16%	All you need is a good init	ICLR 2016	Details

Reference: http://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html

Copyright 2016 JNResearch, All Rights Reserved

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Au

Archive > Volume 529 > Issue 7587 > News > Article

NATURE | NEWS

عربي

Google AI algorithm masters ancient game of Go

Deep-learning software defeats human professional for first time.

Elizabeth Gibney

27 January 2016

Intelligent Machines

Baidu's Deep-Learning System Rivals People at Speech Recognition

China's dominant Internet company, Baidu, is developing powerful speech recognition for its voice interfaces.

by Will Knight December 16, 2015

Microsoft Store Products Support

Next The Official Microsoft Blog The Fire Hose Microsoft On the Issues Transform

Historic Achievement: Microsoft researchers reach human parity in conversational speech recognition

DEEP LEARNING IS MACHINE PERCEPTION FOR...

IMAGES • FACES • SELF-DRIVING VEHICLES	**TEXT** • CRM • SEARCH + • ADS
SOUND • VOICE SEARCH • MUSIC GEN. • TRANSLATION	**TIME SERIES** • HEALTH DATA • SENSORS • FINANCE

RECORD-BREAKING ACCURACY

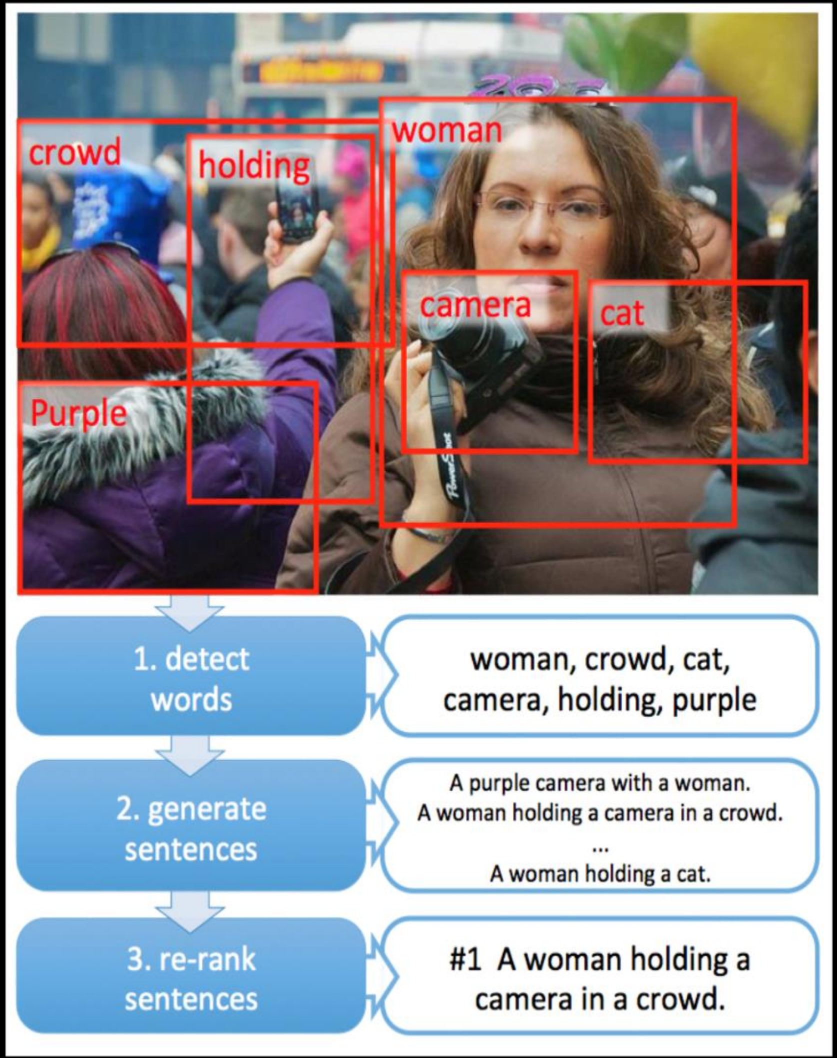
- FACIAL RECOGNITION = 97% accuracy
- GENERAL IMAGE RECOG. = 93%
- SPEECH RECOGNITION = 81%
- VIDEO ACTIVITY RECOG. = 52% - 94% (Varies by dataset)
- TEXT CLASSIFICATION = 94%

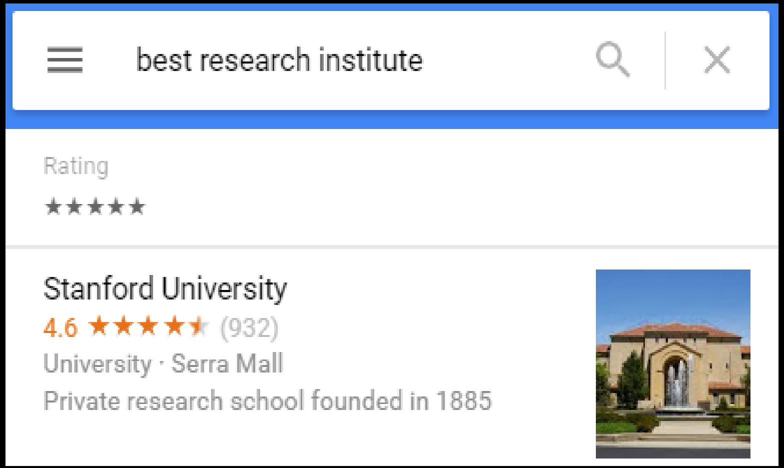
Image Reference Domino Data Lab

ML Everywhere!: Text, Speech, Image, Video

100 funny things to ask S Voice

Google Now won't entertain your humour in the same way as Siri and Cortana, but if you own a Samsung phone or tablet you can chat away with the S Voice personal assistant. Here are 100 funny questions to ask Hi Galaxy.





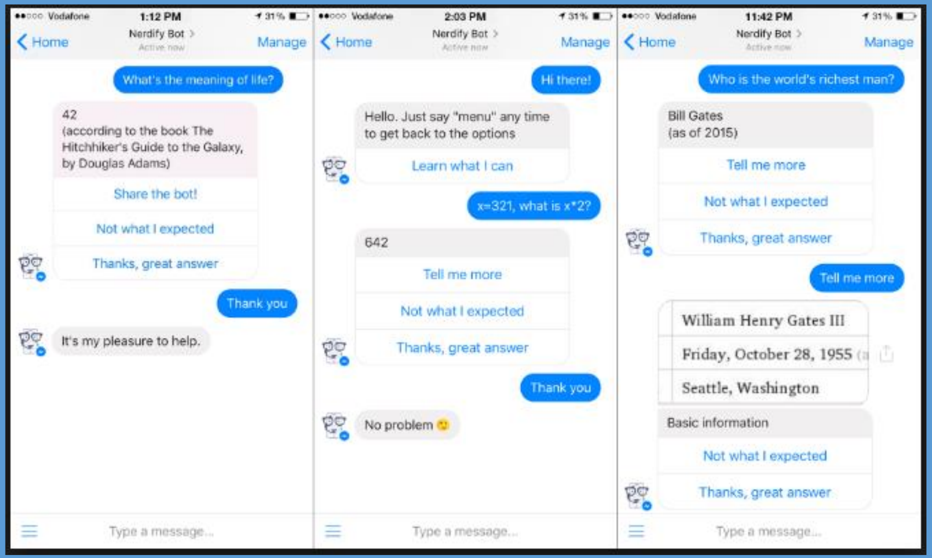
Copyright 2016 JNResearch, All Rights Reserved

AI First

From mobile-first to AI-first

Sundar believes the last 10 years were about building a mobile-first world, turning smartphones into remote controls for our lives. **But in the next 10 years, the shift will be towards a world that is AI-first, a world where computing becomes universally available** — be it at home, at work, in the car, or on the go — and interacting with all of these surfaces becomes much more natural, intuitive, and intelligent. Sundar said,

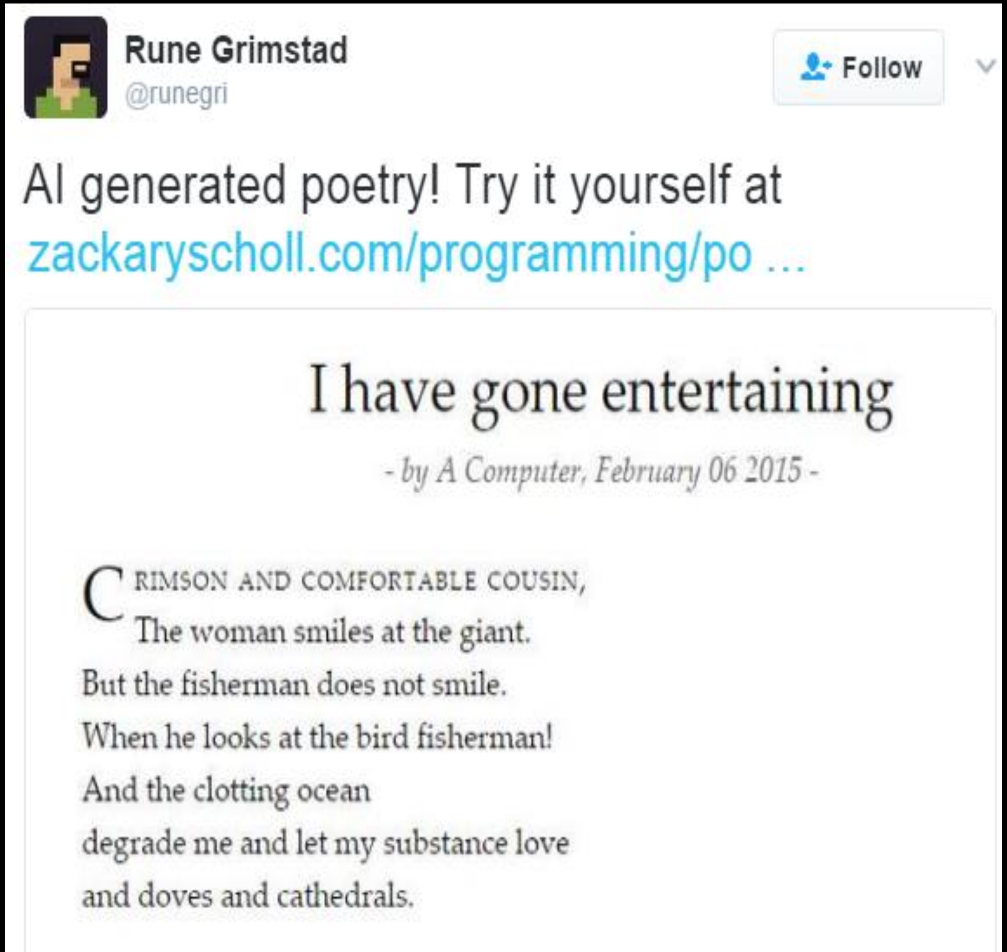
> **This is why we built the Google Assistant, which allows you to have a natural conversation between you and Google.**

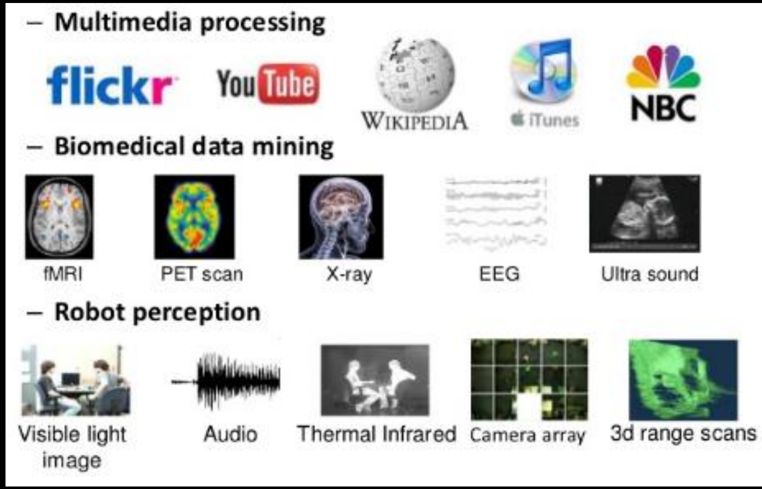


Copyright 2016 JNResearch, All Rights Reserved

Novel Applications

- Self driving cars
    - Multimodal Fusion
- AI generated art (Images, poetry, music etc)
- Video Activity Recognition
- Image Transcription





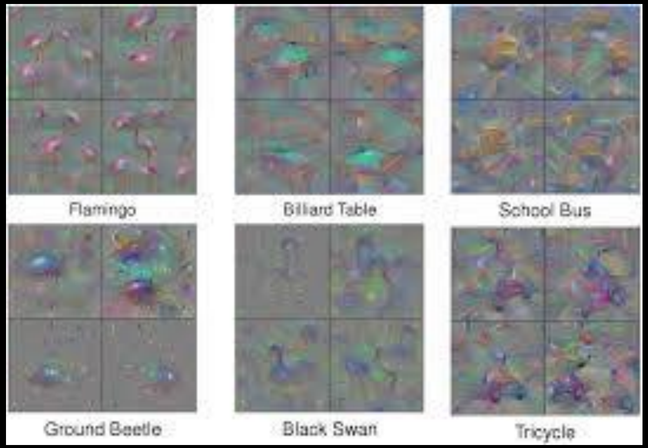
Copyright 2016 JNResearch, All Rights Reserved

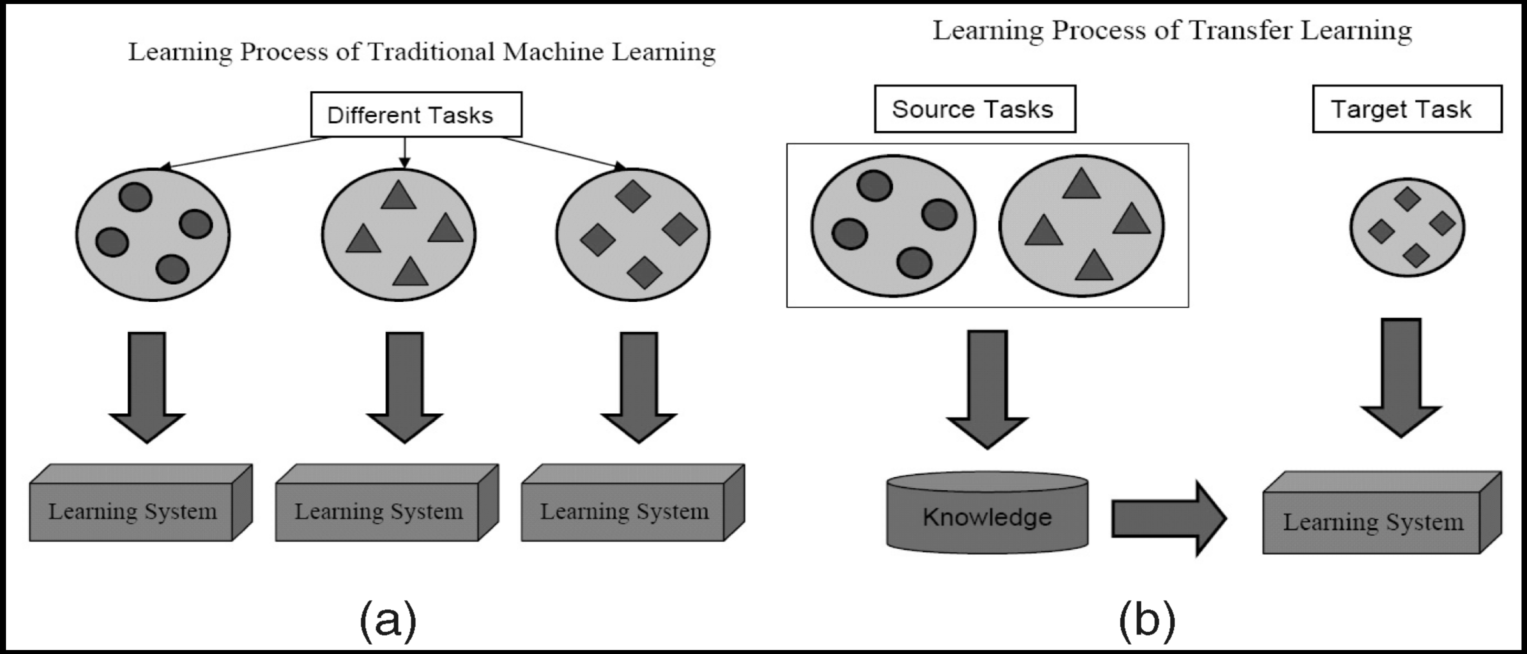
Drawbacks (might be addressed in future)

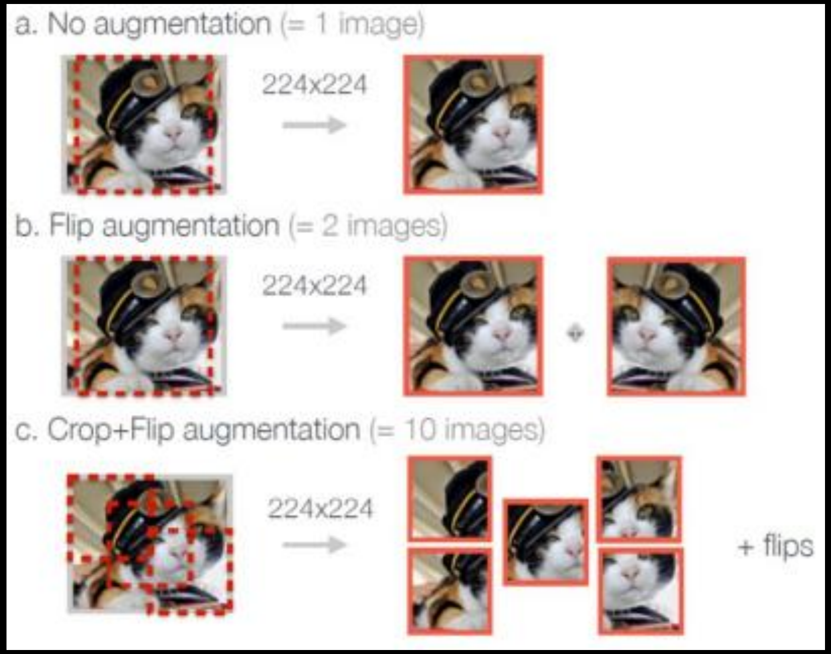
- Needs large amount of data – more the data, more can be the accuracy
  - Results in computer vision that report insane accuracies use insane number of images to train the deep network.
  - The number of labelled training images may be scarce for certain domain specialized applications such as medical imaging
- Needs high end GPU based hardware resources for many complex problems
  - Some of the CNN architectures discussed in this course take days/weeks to train
- Lack of theoretical foundations
  - You can explain the intuition but you can’t prove or guarantee!
  - High performance system design is still an art rather than science or engineering
- It is difficult to comprehend what the network learns

Some techniques to the rescue

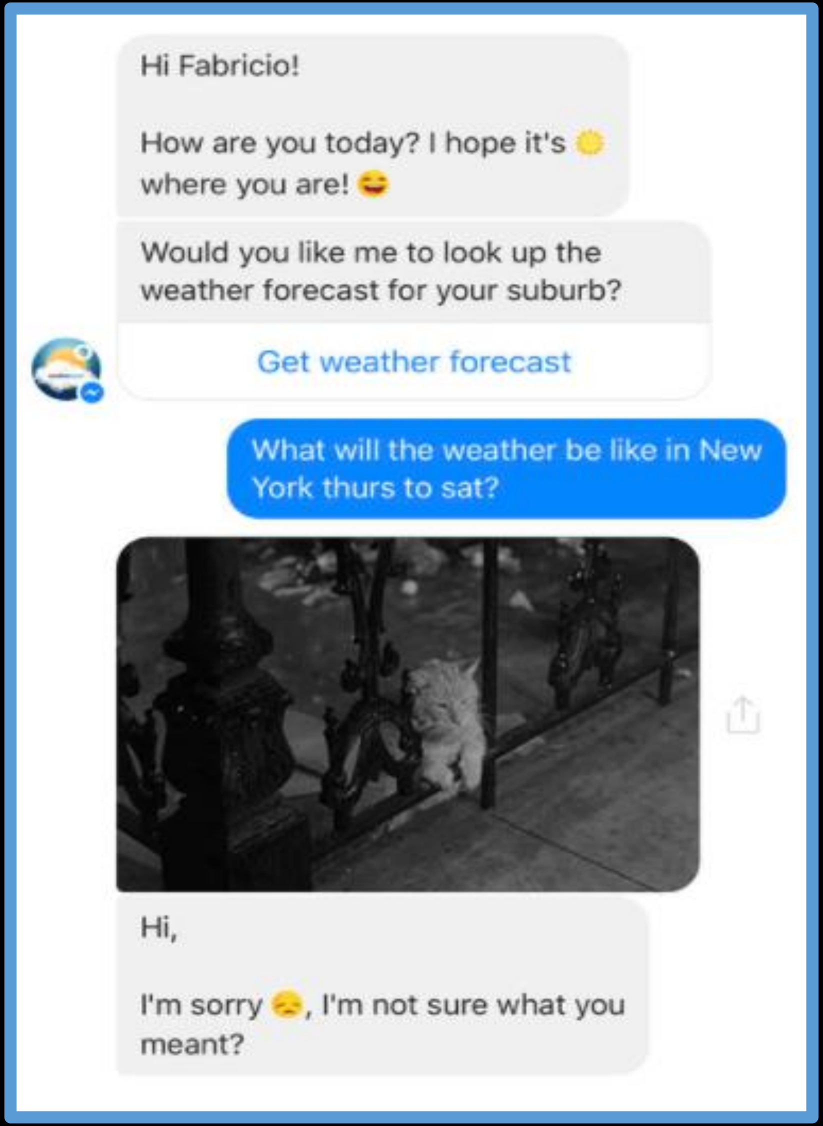
- Data augmentation
- Synthetic Data
- Transfer Learning
- Visualization techniques







Artificial General Intelligence is still hard despite the current progress



Copyright 2016 JNResearch, All Rights Reserved

Can we "generate" data?

- Often there is a need for "generating" synthetic data
  - Applications may require a synthesized data : e.g Speech synthesis (TTS)
  - We may need larger datasets to train classifiers that need labelling
- What kind of data we typically need?
  - Images
  - Speech
  - Handwriting
  - Text

Copyright 2016 JNResearch, All Rights Reserved

Summary and Key Insights

- Machine Learning and Deep Learning are powerful approaches to AI that have enabled novel applications
- Classical Approach: Occam's Razor principle: simple models, bias-variance trade offs
- Contemporary Approach: Don't be afraid to build complex models, complex problems require commensurately powerful models.
- Choose the right deep learning framework (such as TensorFlow/Keras) with other complementing domain specific libraries
- DL techniques aren't perfect, but they are powerful enough to be immensely useful

Copyright 2016 JNResearch, All Rights Reserved

Developing Products, Solutions : Examples

Copyright 2016 JNResearch, All Rights Reserved

Hands On: Top 10 tools for a ML Engineer

1. Numpy
2. Pandas
3. Matplotlib, Seaborn
4. Jupyter
5. Scikit-learn
6. TensorFlow
7. Keras
8. Nltk
9. Opencv
10. Big data processing tools

Copyright 2016 JNResearch, All Rights Reserved

Problem Statement

- How do we price the product, given its core features?

- How much something is really worth?

Sweater A:	Sweater B:
"Vince Long-Sleeve Turtleneck Pullover Sweater, Black, Women's, size L, great condition."	"St. John's Bay Long-Sleeve Turtleneck Pullover Sweater, size L, great condition"

Copyright 2016 JNResearch, All Rights Reserved

Dataset

- Training dataset size is Number of rows: 1482535, columns: 8
- train_id or test_id - the id of the listing
- name - the title of the listing.
- item_condition_id - the condition of the items provided by the seller
- category_name - category of the listing
- brand_name
- price - the price that the item was sold for. (This is the target variable for prediction)
- shipping - 1 if shipping fee is paid by seller and 0 by buyer
- item_description - the full description of the item.

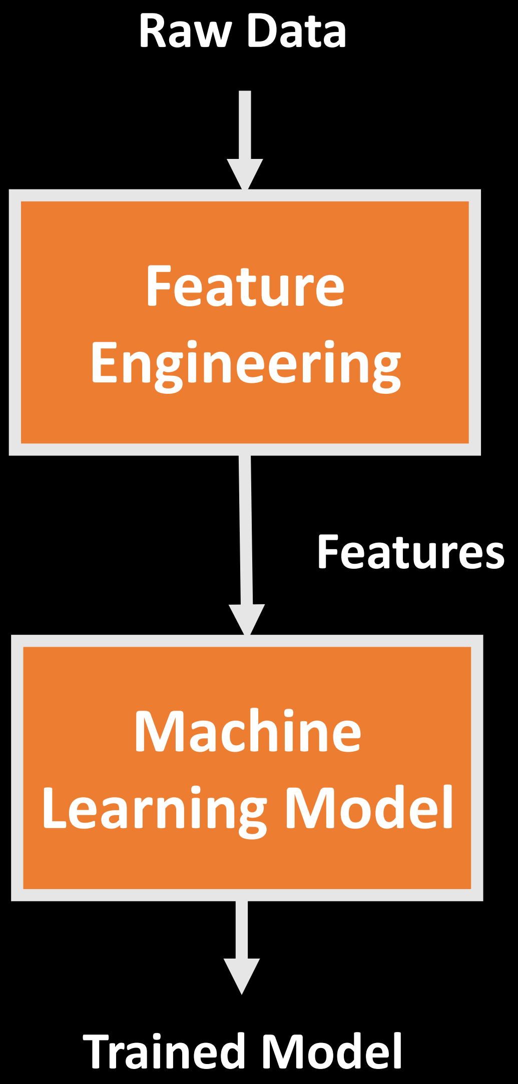
Copyright 2016 JNResearch, All Rights Reserved

Feature Representation

- A labelled dataset provides the inputs X and the corresponding labels Y

- Our first task is to turn these in to a good representation for X (features) and also the target variable Y

- Inspecting, visualizing and analysing the dataset gives us insights in to how to form the features



Copyright 2016 JNResearch, All Rights Reserved

Data Analysis : Loading and Inspecting the dataset

Read the CSV file with 1 line of code using Pandas, create a dataframe

```
df1 = pd.read_csv(TRAIN_TSV_NAME, sep="\t")
```

A dataframe can hold a large number of rows - let us quickly visualize the first 5 rows

```
df1.head()
```

	train_id	name	item_condition_id	category_name	brand_name	price	shipping	item_description
0	0	MLB Cincinnati Reds T Shirt Size XL	3	Men/Tops/T-shirts	NaN	10.0	1	No description yet
1	1	Razer BlackWidow Chroma Keyboard	3	Electronics/Computers & Tablets/Components & P...	Razer	52.0	0	This keyboard is in great condition and works ...
2	2	AVA-VIV Blouse	1	Women/Tops & Blouses/Blouse	Target	10.0	1	Adorable top with a hint of lace and a key hol...
3	3	Leather Horse Statues	1	Home/Home Décor/Home Décor Accents	NaN	35.0	1	New with tags. Leather horses. Retail for [rm]...
4	4	24K GOLD plated rose	1	Women/Jewelry/Necklaces	NaN	44.0	0	Complete with certificate of authenticity

Copyright 2016 JNResearch, All Rights Reserved

Data Analysis Example

- Visualizing a sampling of rows from the dataframe in this example reveals the following:
  - There are missing data (See the NaN in the brand names)
  - The category_name field uses / as separator
  - The item_description field has some cells that indicate that there is “No description yet”
  - The price field is a positive number that is not normalized
  - The shipping field is a Boolean that seems to have a uniform distribution
- From the above, we can decide on how to handle missing values, how to represent the category_name, etc

Copyright 2016 JNResearch, All Rights Reserved

Analyzing the target variable

Analyzing the target variable gives us an idea about its representation. e.g. here we find that this is a positive, continuos variable that goes between 0 to 2009 with mean 26, sd 38.6

```
df1["price"].describe()
```

```
count    1.482535e+06
mean     2.673752e+01
std      3.858607e+01
min      0.000000e+00
25%      1.000000e+01
50%      1.700000e+01
75%      2.900000e+01
max      2.009000e+03
Name: price, dtype: float64
```

Mean Normalization using Pandas

```
price_df = df1["price"]
```

```
normalized_price_df = (price_df - df1["price"].mean())/df1["price"].std()
```

```
normalized_price_df.min(), normalized_price_df.max(), normalized_price_df.std()
```

```
(-0.69293189746856265, 51.372494613939061, 0.99999999999999989)
```

Copyright 2016 JNResearch, All Rights Reserved

Input Features – name field

```
from nltk import word_tokenize
remove_nonascii = lambda text: ''.join([i if ord(i) < 128 else '' for i in text])
```

```
name_toks = []
for item in name_list:
    words = word_tokenize(remove_nonascii(item))
    name_toks.append(words)
```

```
name_toks[:100]
```

```
[['MLB', 'Cincinnati', 'Reds', 'T', 'Shirt', 'Size', 'XL'],
 ['Razer', 'BlackWidow', 'Chroma', 'Keyboard'],
 ['AVA-VIV', 'Blouse'],
 ['Leather', 'Horse', 'Statues'],
 ['24K', 'GOLD', 'plated', 'rose'],
 ['Bundled', 'items', 'requested', 'for', 'Ruie'],
 ['Acacia', 'pacific', 'tides', 'santorini', 'top'],
 ['Girls', 'cheer', 'and', 'tumbling', 'bundle', 'of', '7'],
 ['Girls', 'Nike', 'Pro', 'shorts'],
 ['Porcelain', 'clown', 'doll', 'checker', 'pants', 'VTG'],
 ['Smashbox', 'primer'],
 ['New', 'vs', 'pi', 'k', 'body', 'mists'],
 ['Black', 'Skater', 'dress'],
 ['Sharpener', 'and', 'eraser'],
```

Copyright 2016 JNResearch, All Rights Reserved

Input Features – Item Condition Id

- All rows have an item_condition_id
- This is a discrete variable that can take values {1, 2, 3, 4, 5}
- The distribution indicates that the frequency decreases as the id increases

```
condition_df = df1["item_condition_id"]
```

```
condition_df.unique()
```

```
array([3, 1, 2, 4, 5])
```

```
condition_df.value_counts()
```

```
1    640549
3    432161
2    375479
4     31962
5      2384
Name: item_condition_id, dtype: int64
```

Copyright 2016 JNResearch, All Rights Reserved

Input Features – category_name field

- We observe that the category_name encodes a hierarchical structure
- We can represent this field as either:
  - A bag of words
  - A tree structured input
- The tree can have an arbitrary number of subtrees
- What is the right representation?

```
cat_df = df1["category_name"]
```

```
cat_df.describe()
```

```
count                                          1476208
unique                                            1287
top       Women/Athletic Apparel/Pants, Tights, Leggings
freq                                             60177
Name: category_name, dtype: object
```

```
cat_df[:10]
```

```
0                               Men/Tops/T-shirts
1    Electronics/Computers & Tablets/Components & P...
2                         Women/Tops & Blouses/Blouse
3                  Home/Home Décor/Home Décor Accents
4                            Women/Jewelry/Necklaces
5                                 Women/Other/Other
6                          Women/Swimwear/Two-Piece
7                   Sports & Outdoors/Apparel/Girls
8                   Sports & Outdoors/Apparel/Girls
9             Vintage & Collectibles/Collectibles/Doll
Name: category_name, dtype: object
```

Input Features – brand_name

```
brands_df = df1["brand_name"]
```

```
brands_df.value_counts()
```

```
PINK                   54088
Nike                   54043
Victoria's Secret      48036
LuLaRoe                31024
Apple                  17322
FOREVER 21             15186
Nintendo               15007
Lululemon              14558
Michael Kors           13928
American Eagle         13254
Rae Dunn               12305
Sephora                12172
Coach                  10463
Disney                 10360
Bath & Body Works      10354
Adidas                 10202
```

```
Clearasil                 1
Caruso                    1
Giorgio Brutini           1
Bananafish                1
Sansabelt                 1
Tech  Kids                1
CharmLeaks                1
Undercover Mama           1
Zanerobe                  1
Chequer                   1
Petit Tresor              1
Hill's Science Diet       1
Plackers                  1
Jones Jeans               1
Nurture by Lamaze         1
Finders Keepers           1
wallis                    1
Marks and Spencer         1
Youth To The People       1
China Glaze Co. Ltd.      1
Sergeants                 1
Name: brand_name, Length: 4809, dtype: int64
```

Copyright 2016 JNResearch, All Rights Reserved

Input Features – item_description

- We note that:
  - "No description yet" is the most occurring description (5.5%)
  - The descriptions that are "never used", "like new", "worn once", "excellent condition", etc indicate that the item is a used product

```
desc_df = df1["item_description"]
```

```
desc_df.describe()
```

```
count                    1482531
unique                   1281426
top          No description yet
freq                       82489
Name: item_description, dtype: object
```

```
desc_df.value_counts()
```

Copyright 2016 JNResearch, All Rights Reserved

Additional Complexity

- Product pricing gets even harder at scale, considering just how many products are sold online.

- Clothing has strong seasonal pricing trends and is heavily influenced by brand names, while electronics have fluctuating prices based on product specs.

Copyright 2016 JNResearch, All Rights Reserved

Sample Projects

Copyright 2016 JNResearch, All Rights Reserved

Sample Projects (most of them need GPU)

- Speech LSTM (recognition, generation)
- Deep Q Networks for practical applications (e.g chat quality)
- Object Recognition with Fast RCNN
- Image Denoiser
- Deep Q&A systems
- Smart Search with TensorFlow js
- Video Activity Recognition : tensorflow js
- Real time video/image analysis on RaspberryPi
- Classifying Toxic Comments
- Localize text in natural scenes

Copyright 2016 JNResearch, All Rights Reserved