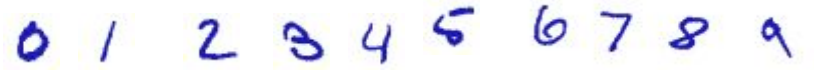
Machine Learning

Jaya Sil
Indian Institute of Engineering Science & Technology,
Shibpur

How to Learn?

•Human learns from **Experience** i.e. knowledge but they do not know how to describe the steps they take to reach to the answer.



- The machine's responsibility is to learn a way to go from an input to a label or output
- Machine learns from data. Data is cheap whereas knowledge is expensive
- Statistics used to make inference for a larger population of data with the help of statistical parameters.
- On a smaller sample data, i.e. LINEAR REGRESSION learns **relation** between the variables.

Introduction

- Machine learning is a subfield of Artificial Intelligence (AI).
- AI is intelligence demonstrated by machines (MI)
- AI research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions to maximize chance of successfully achieving the goals.
- The term AI is applied when a machine mimics "cognitive" functions of human brain, such as "learning“, “Reasoning” and "problem solving".
- The goal of machine learning (ML) is to understand the structure of data and fit that data into models and utilized by people.

Introduction

- Machine learning differs from traditional computational approaches.
- In traditional computing, algorithms are sets of explicitly programmed instructions used by computers to calculate or problem solve.
- Machine learning algorithms allow computers to learn by training the huge observed data inputs.
- Machine learning facilitates computers in building models from sample data in order to automate decision-making processes based on unknown data inputs.

Definition: Arthur Samuel (1959)

- Machine Learning is the field of study that gives the computer the ability to learn without being explicitly programmed.

- ML has the ability to automatically obtain deep insights, recognize unknown *patterns*, and create high performing *predictive models* from *data*.

Definition Tom Mitchell (1998)

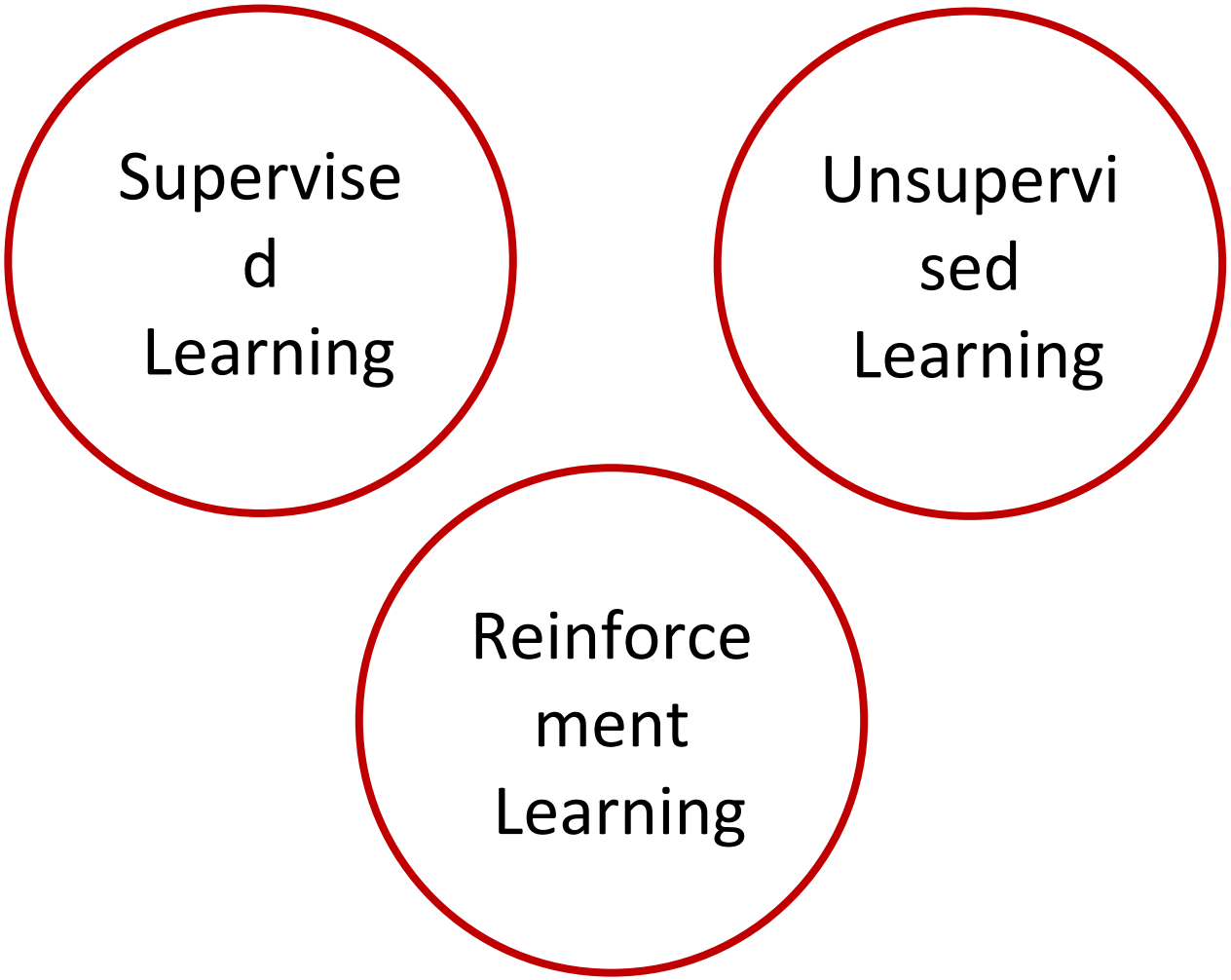
- *A computer is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."*

Experience (data): games played by the program (with itself)

Performance measure: winning rate

- To predict the traffic patterns at a busy intersection (task T), write a machine learning algorithm with data about past traffic patterns (experience E) and, if it has successfully "learned", it will then do better at predicting future traffic patterns (performance measure P).

Types of Learning



Types of Learning

- **Supervised Learning -** based on example input - output data.
- Class label or target value is known
- Learn an Input to Output Map: Model
- Classification: Categorical output; Regression: Continuous output

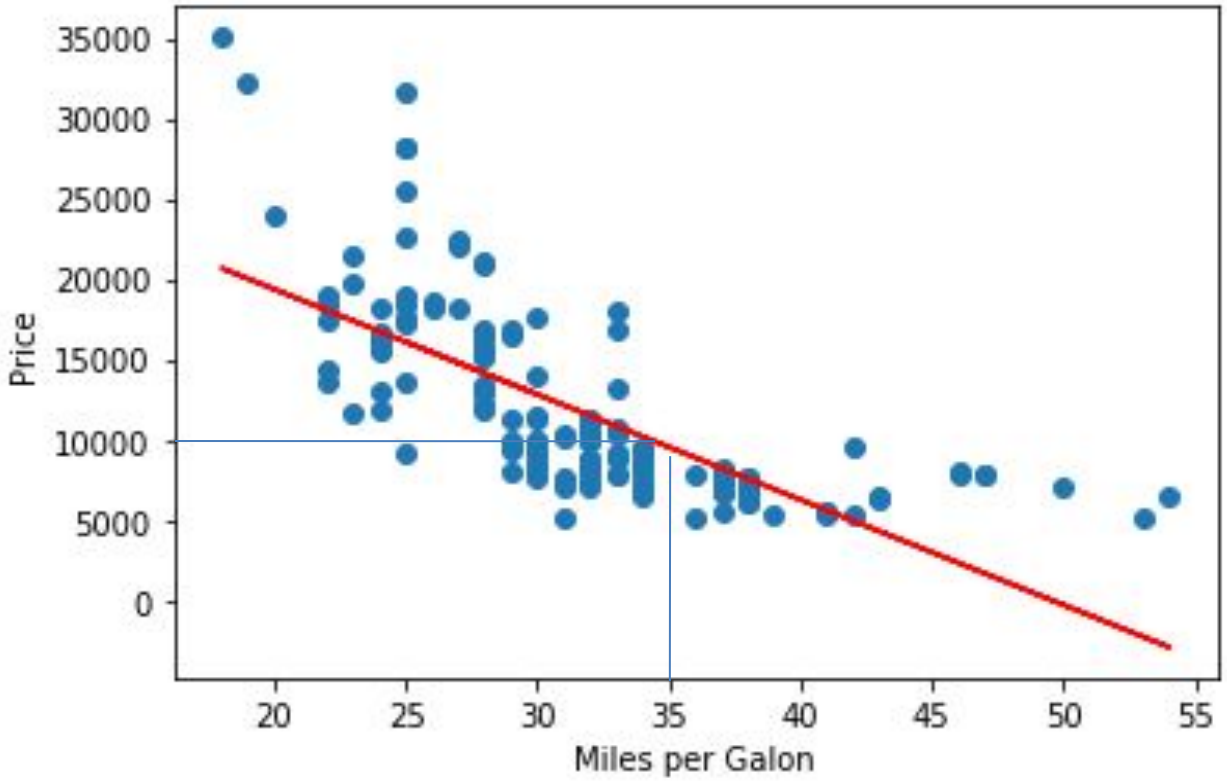
- **Unsupervised Learning -** the algorithm with no labeled data find structure within its input data by exploring similarity or commonality within the data.
- Discover Patterns in the data: Clustering, Association (frequent co-occurrence)

- **Reinforcement Learning:** Trade-off between *exploration and Exploitation Principle.*
- *The system tries* out new kinds of actions to see how effective they are, and *exploitation,* the system makes use of actions that are known to yield a high reward.

Machine Learning Methods

- Supervised learning is to use historical data to predict statistically likely future events. Like historical stock market information to anticipate upcoming fluctuations.
- As unlabeled data are more abundant than labeled data, unsupervised learning are particularly valuable for discovering hidden patterns within a dataset.
- Unsupervised learning is commonly used for transactional data.
- As a field, machine learning is closely related to computational statistics, so having a background knowledge in statistics is useful for understanding and leveraging machine learning algorithms.

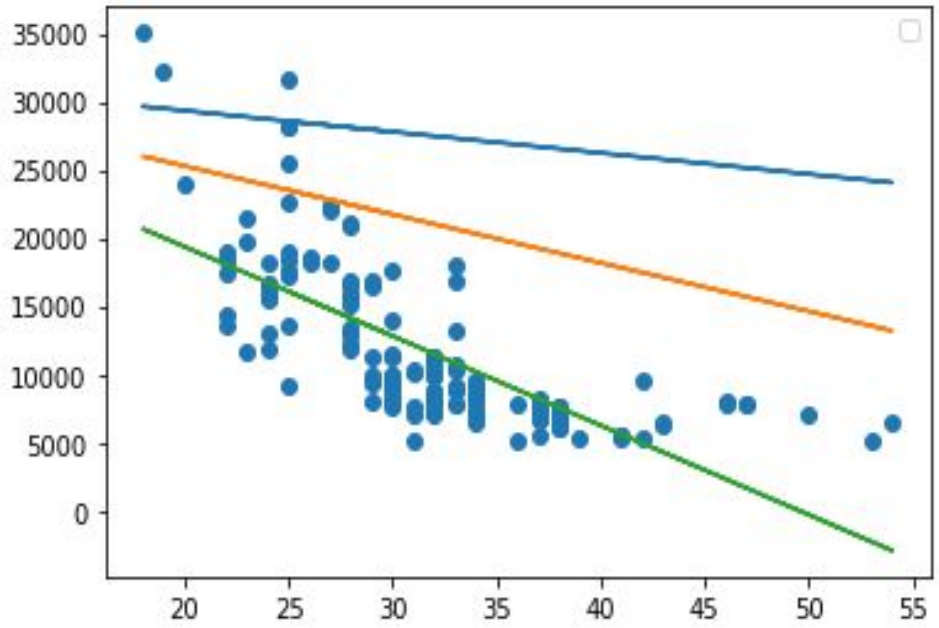
Linear Relationship between two Variables



How to find this straight line?

How to find which line is the best?

- *price = a + b * mpg = f (a, b, mpg) ; f is the model*
- Different values of ***a*** and ***b*** generate different straight lines.



- Choose the best values of ***a*** and ***b*** to get the best fitting line.
- The best line is the one that better fits the data points.
- **What a machine has to learn, i.e. *a* and *b***

Machine Learning

- Machine learning is referred - *predictive analytics*, or *predictive modeling*.
- Goal is to build new and/or leverage existing *algorithms* to *learn* from data, in order to *build* ***generalizable*** *models* that give *accurate predictions*, particularly with *new and unseen similar data*.
- Problem of ML is to learn or infer a functional relationship between a set of **attribute variables** and associated response or **target variables**.
- Popularity of this field in recent times: Neural network, deep learning, GPU, cloud enable and big data

Machine Learning: The Problem

- The process of learning begins with observations or data, such as examples, direct experience, or instruction.
- The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.
- Dataset as a table, rows are *observations* or samples (measurement, data point, etc), and columns for each observation or sample represent its property.
- If each sample has a multivariate or multi-dimensional entry, it is said to have attributes or **features**.
- A dataset is usually split into two subsets. subsets are the *training* and *test* datasets.

Feature Space

- For most practical applications, the original input variables are typically *preprocessed* to transform them into some new space.

- It is also called *feature extraction*.

- A form of dimensionality reduction.

Training set and Testing set

- A predictive model or classifier is trained using the training data, and then the model's predictive accuracy is determined using the test data.
- Machine learning leverages algorithms to automatically model and find patterns in data, usually with the goal of predicting some target output or *response*.
- These algorithms are heavily based on statistics and mathematical optimization.

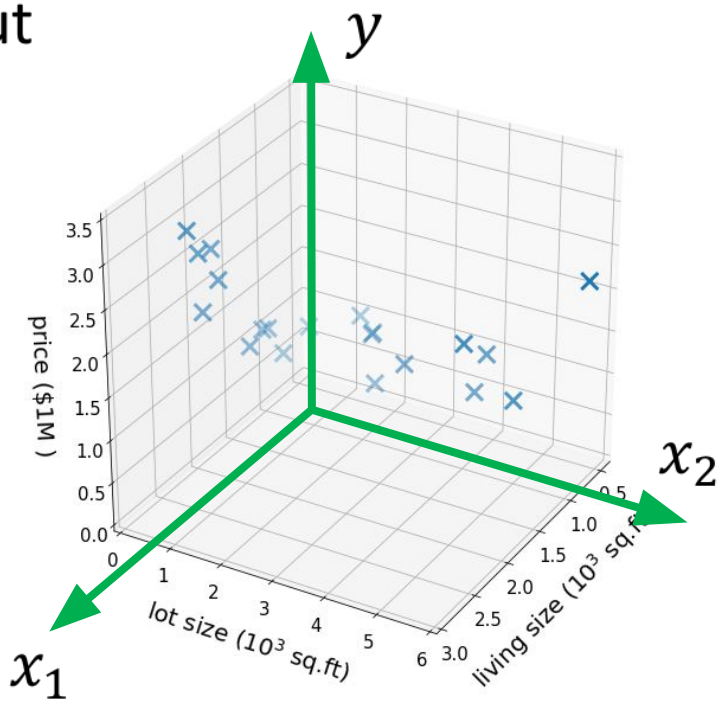
More Features

Task: find a function that maps

(size, lot size) $\rightarrow$ price

features/input
$x \in \mathbb{R}^2$

label/output
$y \in \mathbb{R}$



- Dataset: $\left(x^{(1)}, y^{(1)}\right), \ldots, \left(x^{(n)}, y^{(n)}\right)$

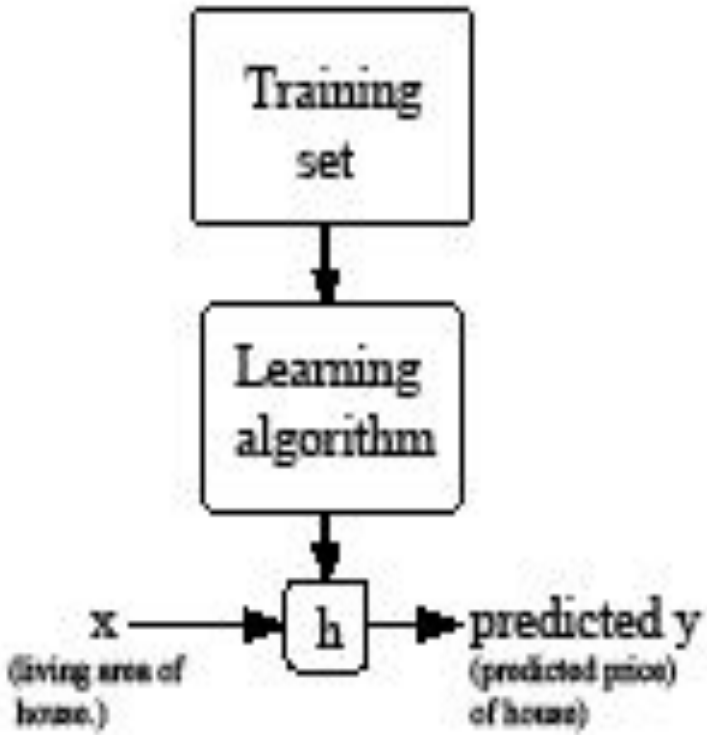
where $x^{(i)} = \left(x_1^{(i)}, x_2^{(i)}\right)$

- "Supervision" refers to $y^{(1)}, \ldots, y^{(n)}$

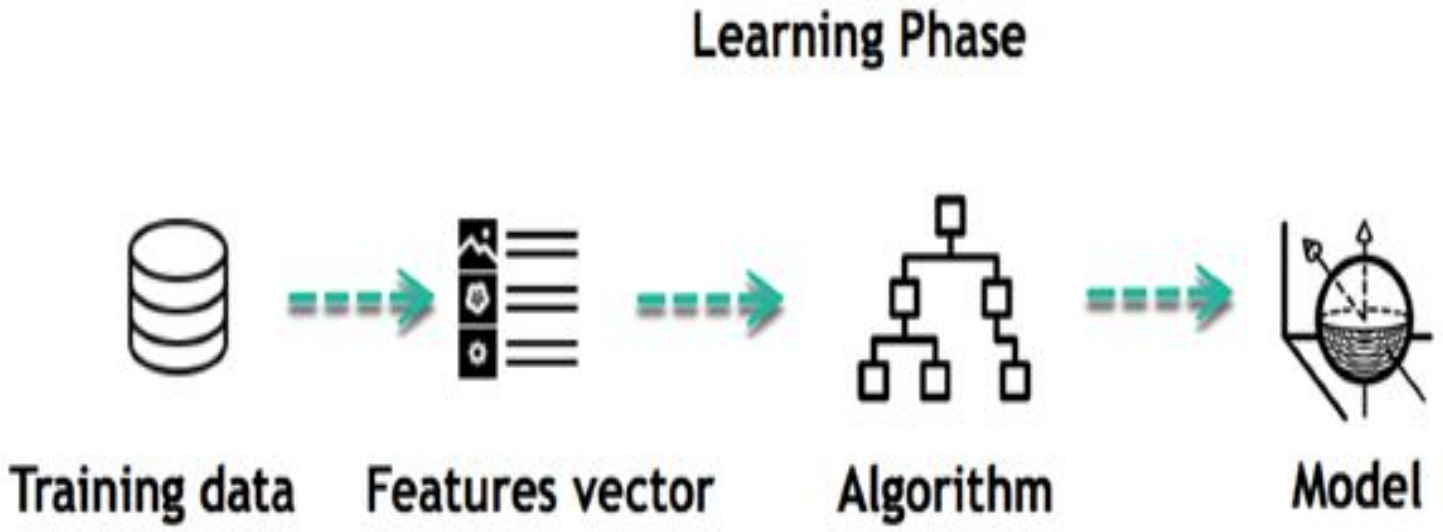
High-dimensional Features

- $x \in \mathbb{R}^d$ for large d
- E.g.,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ \vdots \\ x_d \end{bmatrix} \begin{matrix} \text{--- living size} \\ \text{--- lot size} \\ \text{--- \# floors} \\ \text{--- condition} \\ \text{--- zip code} \\ \vdots \end{matrix} \rightarrow y \text{ --- price}$$


Training set
Learning algorithm
x
(living area of house.)
h
predicted y
(predicted price) of house)

Goal of Learning

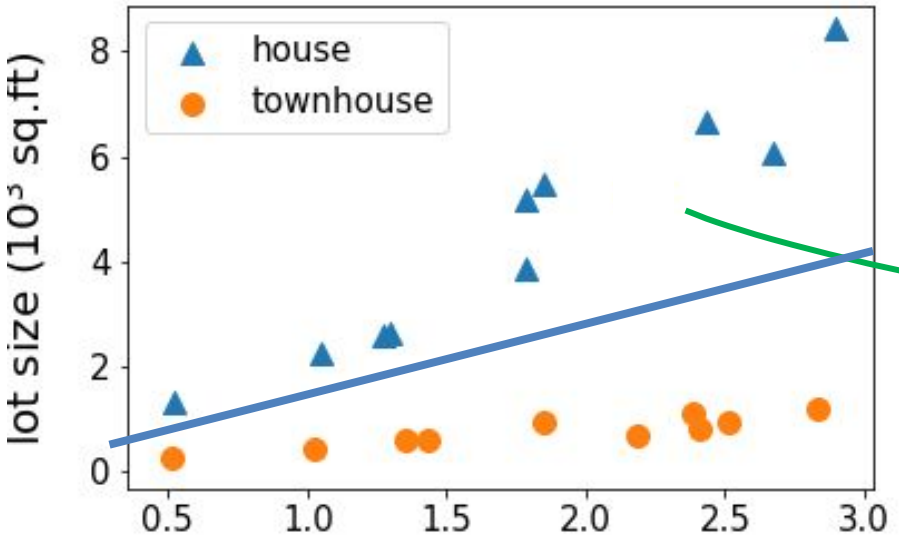


The ability to categorize correctly new examples that differ from those used for training is known as *generalization*.

Regression vs Classification

- regression: if $y \in \mathbb{R}$ is a continuous variable
  - e.g., price prediction
- classification: the label is a discrete variable
  - e.g., the task of predicting the types of residence

(size, lot size) → house or townhouse?



y = house or townhouse?

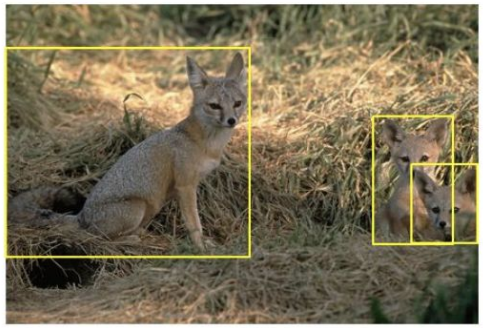
Inductive Learning

- Inductive Learning is where we are given examples as data (x) and the output of the function ($f(x)$). The goal of inductive learning is to learn the function for new data (x).

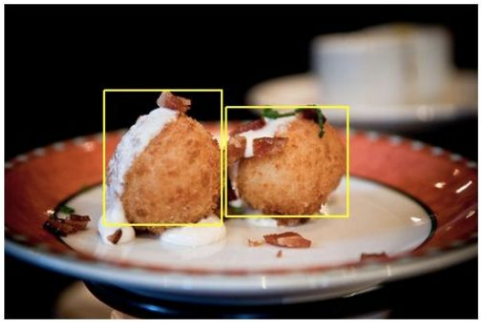
- **Classification**: when the function being learned is discrete.
- **Regression**: when the function being learned is continuous.
- **Probability Estimation**: when the output of the function is a probability.

Supervised Learning in Computer Vision

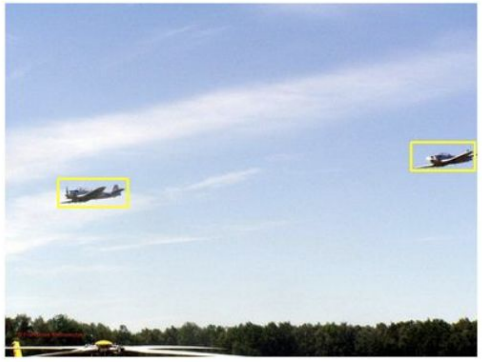
- Object localization and detection
  - x = raw pixels of the image, y = the bounding boxes



kit fox



croquette



airplane

frog

ImageNet Large Scale Visual Recognition Challenge. Russakovsky et al.'2015