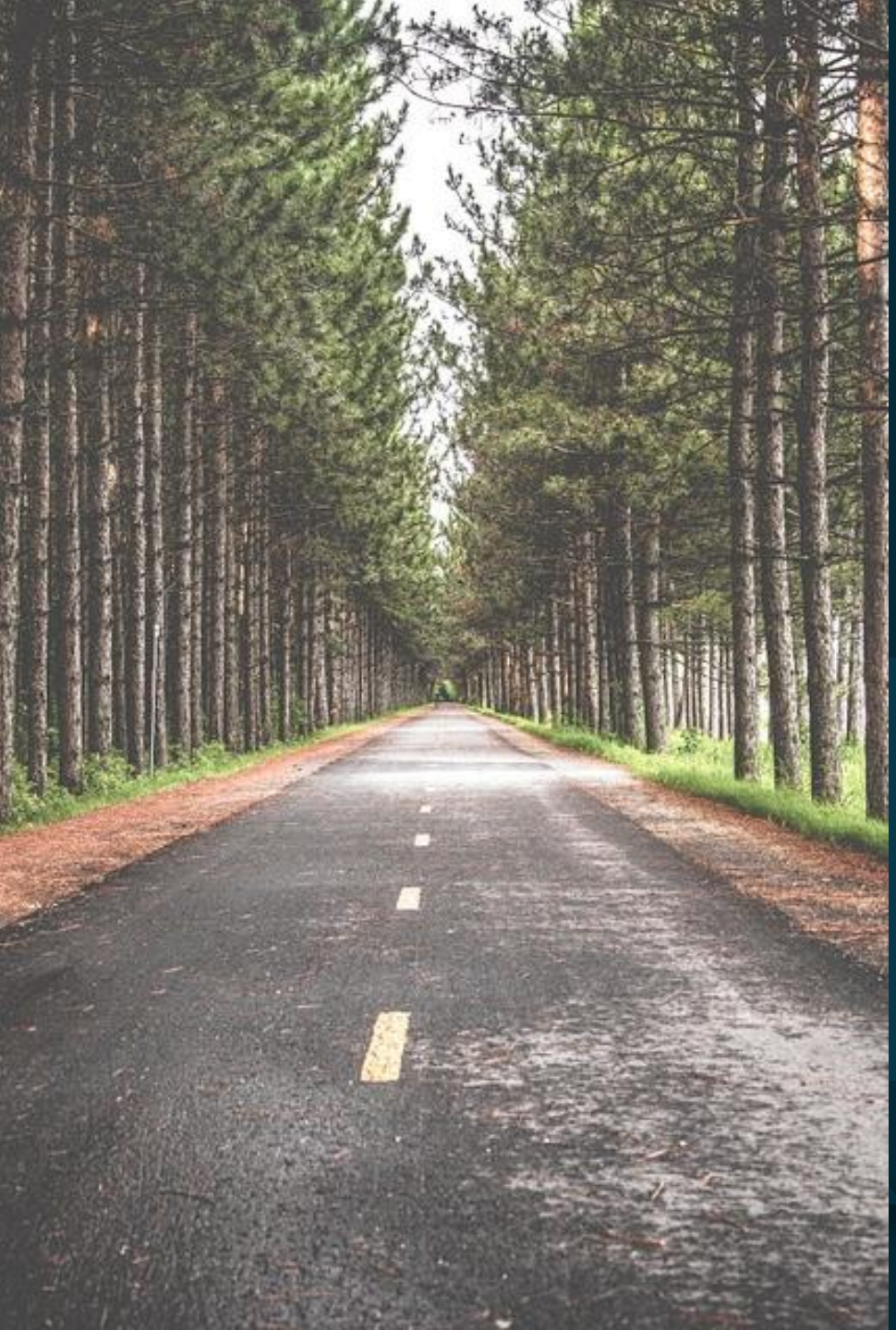
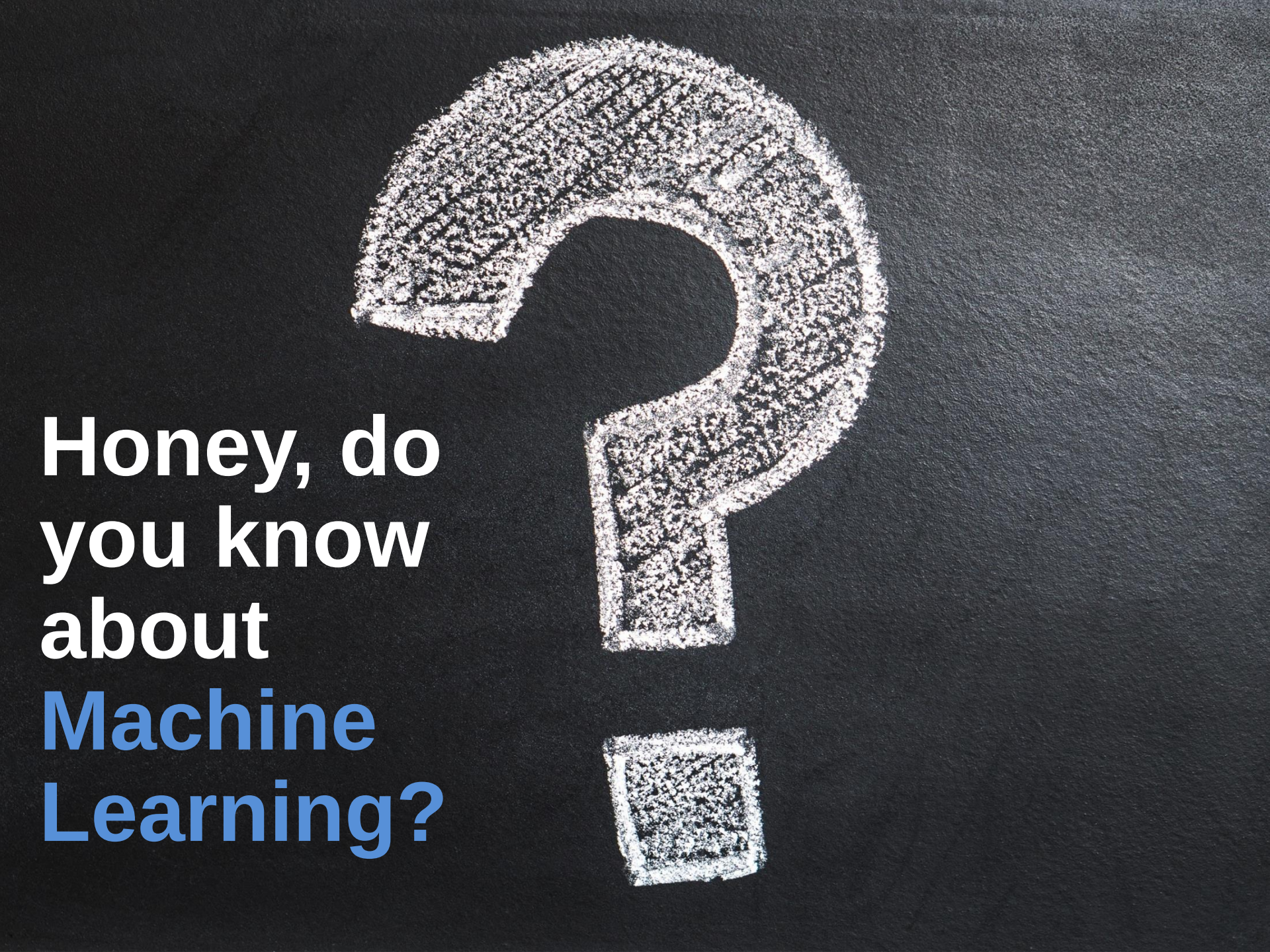
Machine Learning for SAS Programmers

Kevin Lee
Director of Data Science



The Agenda

- Introduction of Machine Learning
- Supervised and Unsupervised Machine Learning
- Deep Neural Network
- Machine Learning implementation
- Questions and Discussion

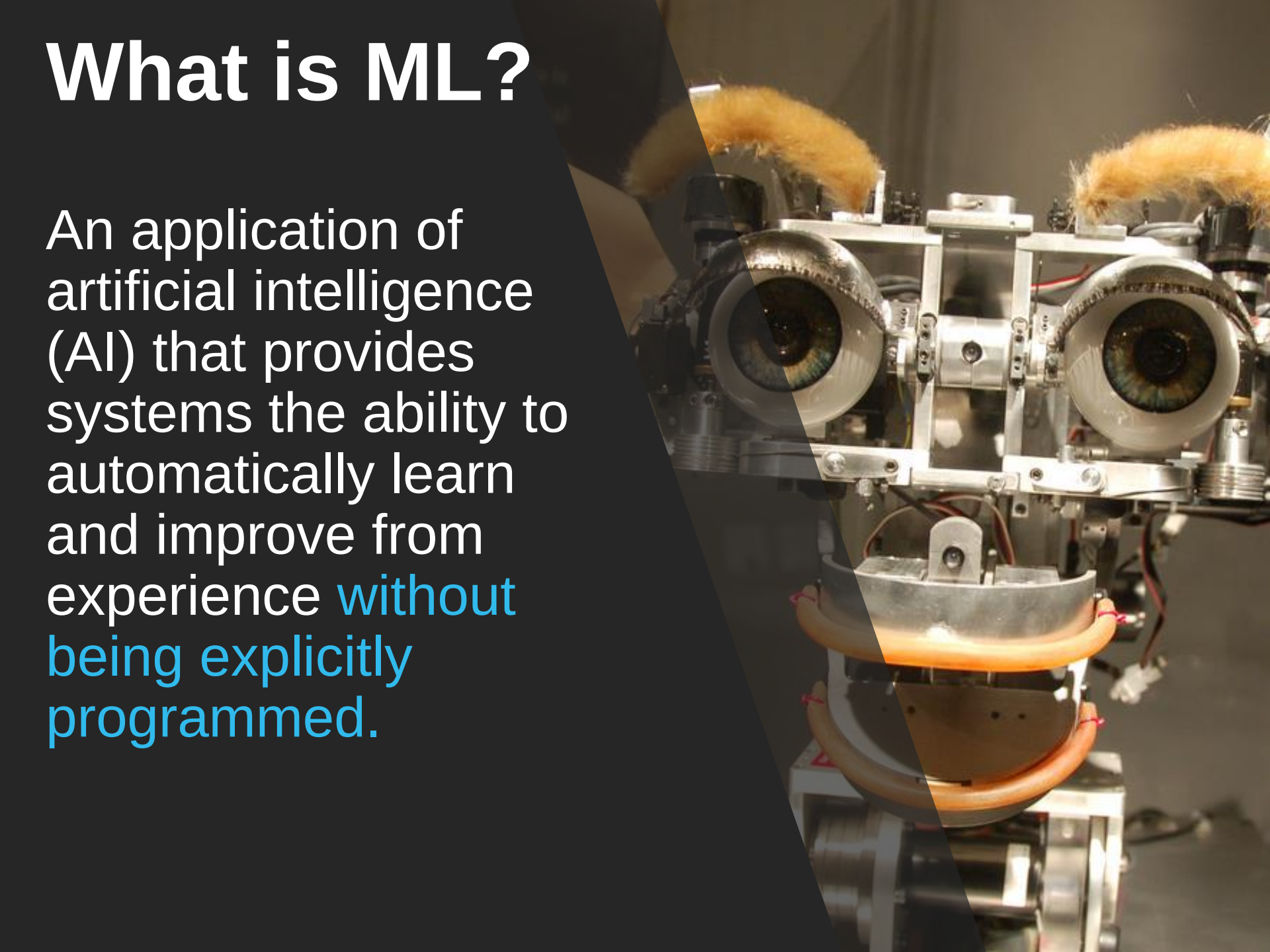

Honey, do you know about Machine Learning?

Why did people ask / expect me to know about Machine Learning?

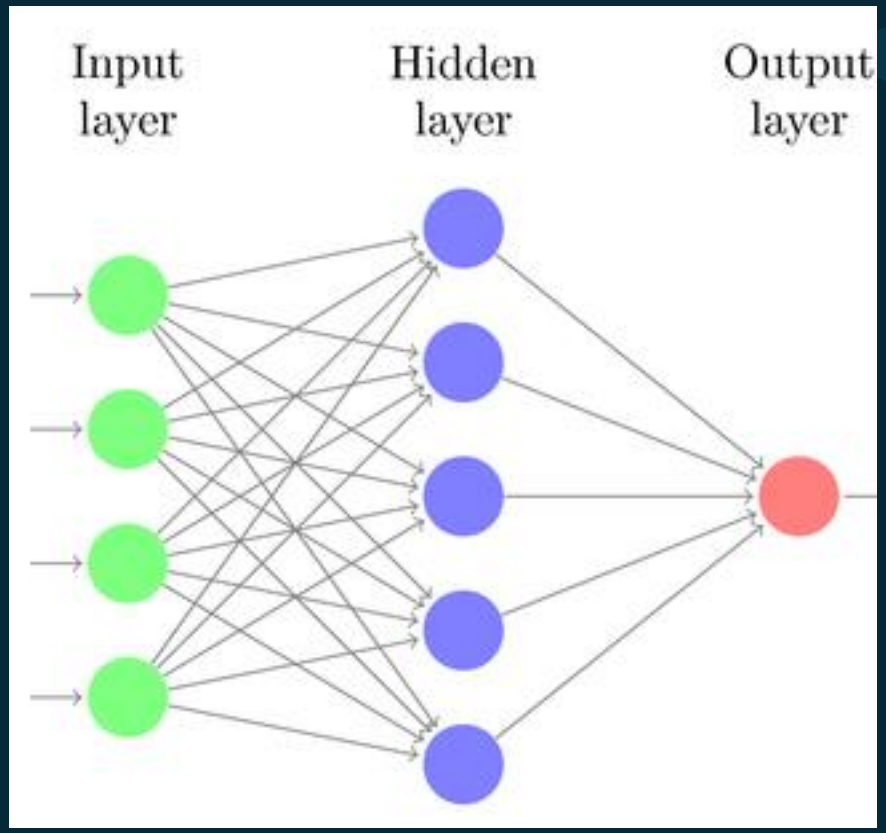
- Programming
- Statistics / modeling
- Working with data all the times

What is ML?

An application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.



```
** Derivation of 1st Age group **;
if age < 65 then do;
   agegr1n=1; agegr1="<65";
end;
if 65<=age<=69 then do;
   agegr1n=2; agegr1="65 - 69";
end;
else if 70<=age<=74 then do;
   agegr1n=3; agegr1="70 - 74";
end;
else if 75<=age<=79 then do;
   agegr1n=4; agegr1="75 - 79";
end;
else if 80<=age<=84 then do;
   agegr1n=5; agegr="80 - 84";
end;
else if age ge 85 then do;
   agegr1n=6; agegr1=">=85";
end;
```



Explicit programing

Machine Learning

How does Human Learn?
- Experience

How does Machine learn?

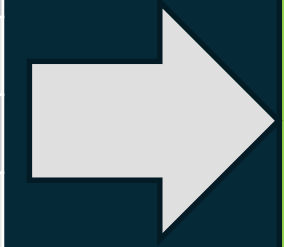
Input Data

Algorithm

How does Machine Learning work?

Input data

X0	X1	X2	…	Xn	Y



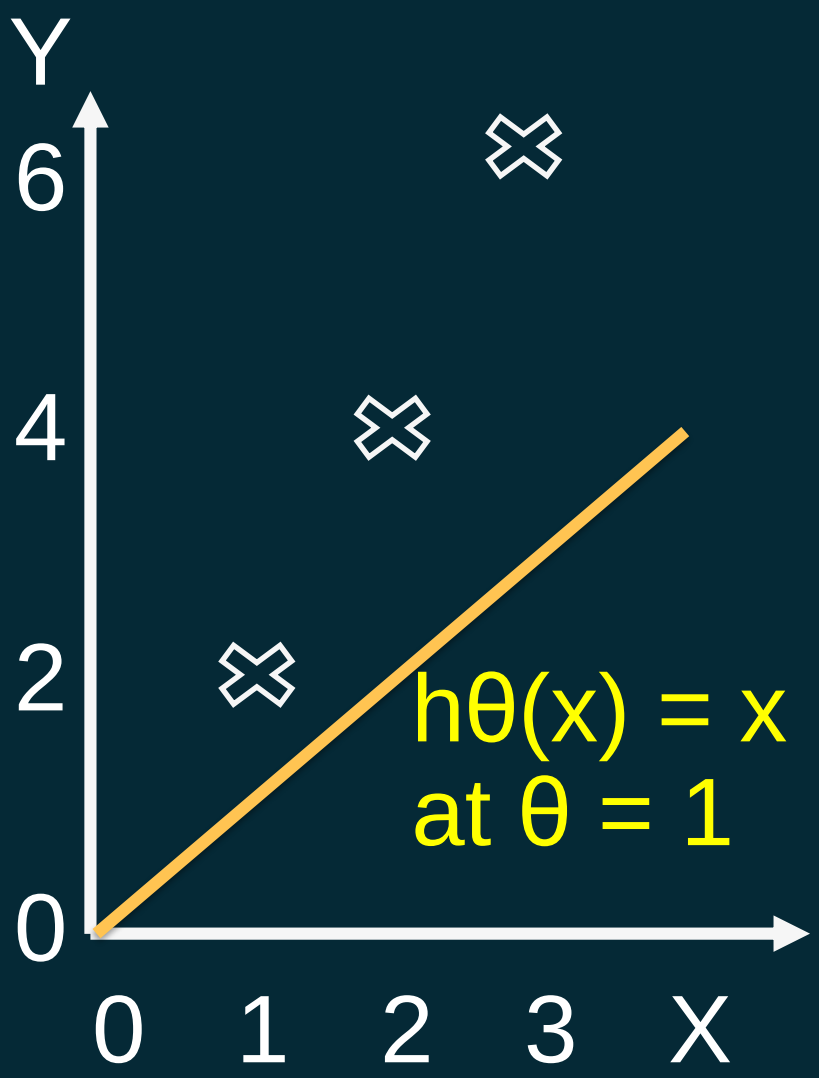
Algorithm

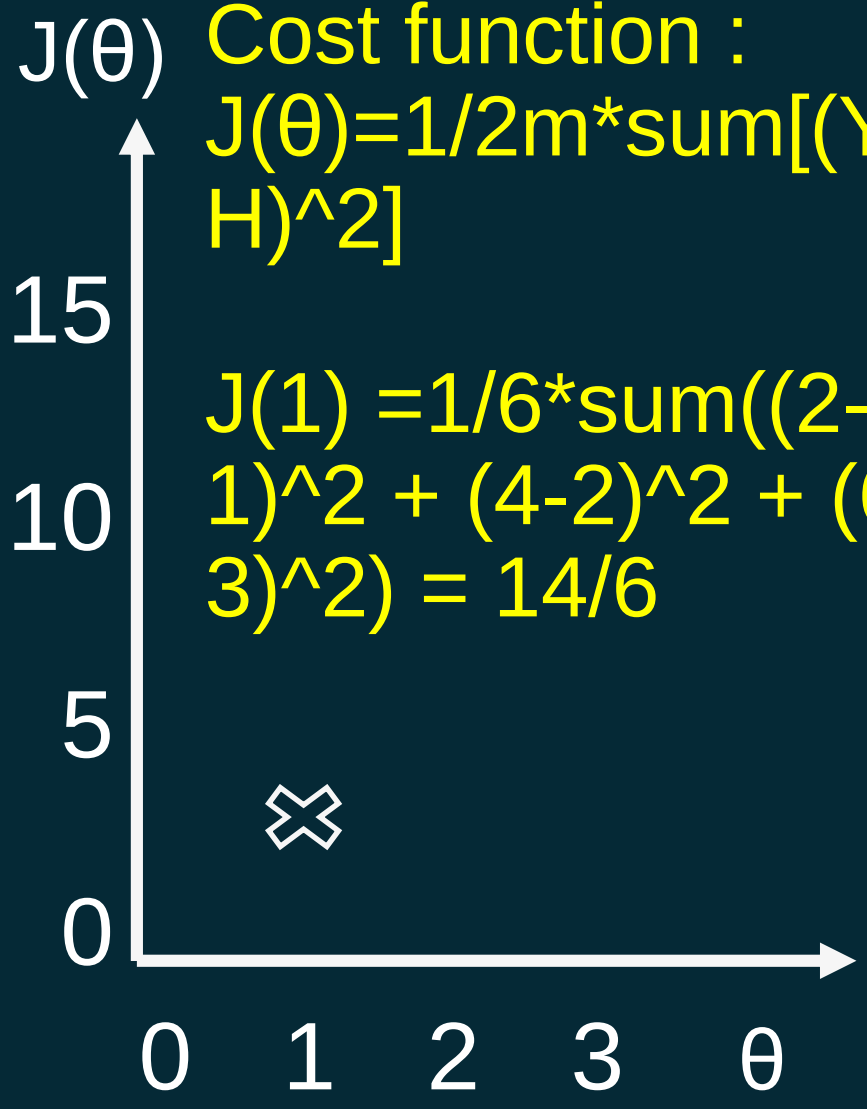
- Hypothesis Function - $h\theta(x) = \theta x + b$
- Minimize Cost Function, $J(\theta) = h\theta(x) – Y$, using Gradient Descent

Machine Learning Algorithms

- Hypothesis function
  - Model for data
  - $H_\theta(x) = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \ldots + \theta_n x_n$ (e.g., $Y = 2X + 30$)
- Cost function
  - measures how well hypothesis function fits into data.
  - Difference between actual data point and hypothesized data point. (e.g., $Y - H_\theta(x)$)
- Gradient Descent
  - Engine that minimizes cost function

Cost function with Gradient Descent

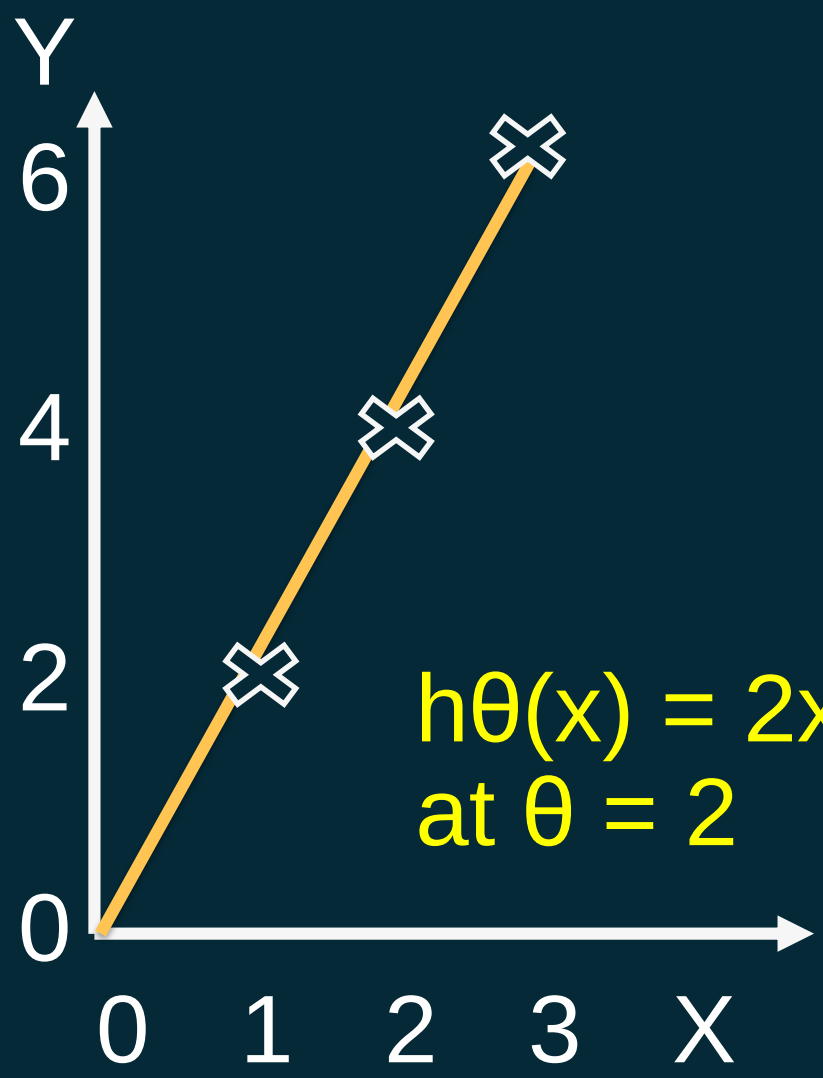


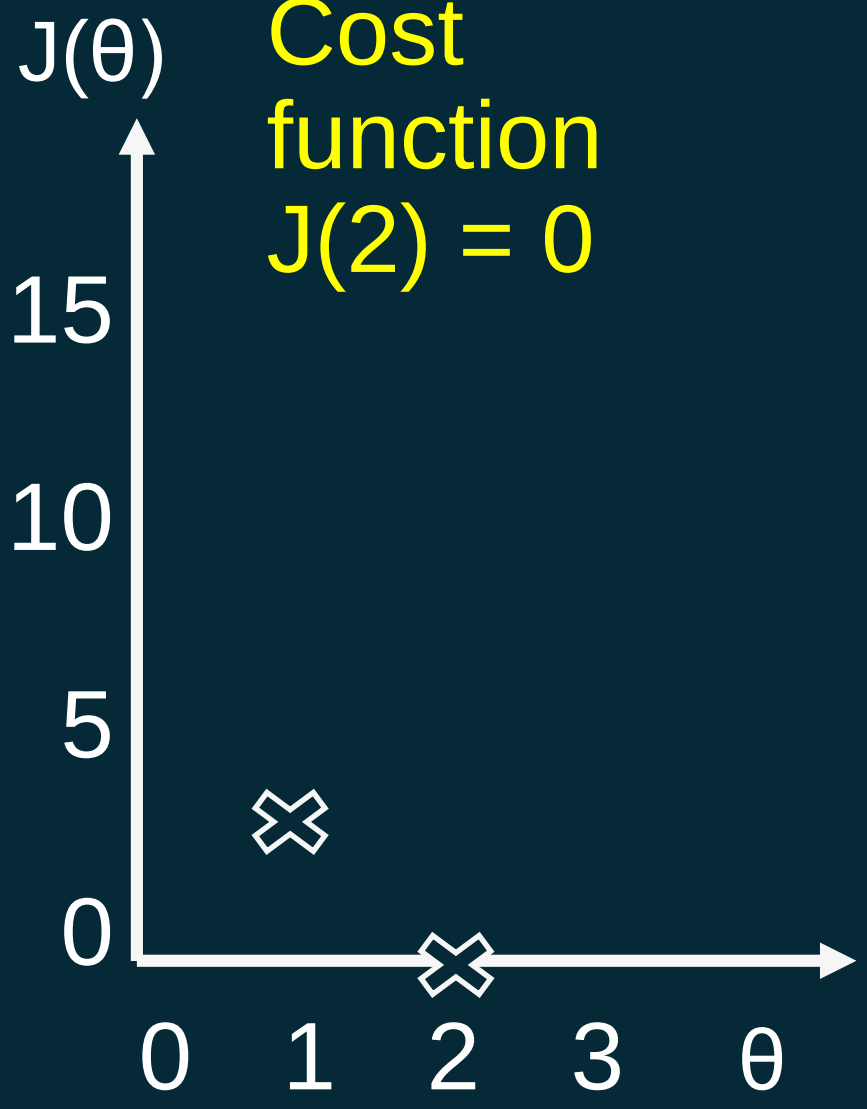


Cost function :
J(θ)=1/2m*sum[(Y - H)^2]

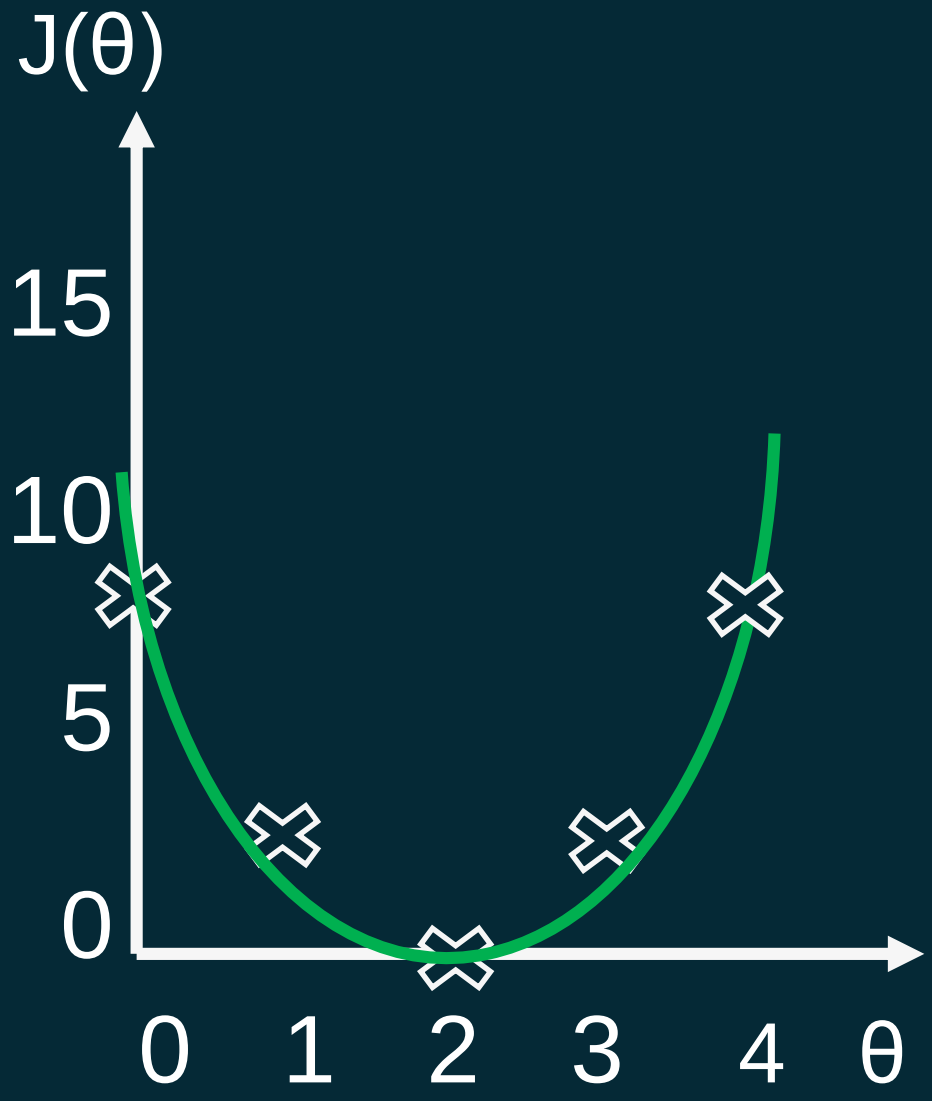
J(1) =1/6*sum((2-1)^2 + (4-2)^2 + (6-3)^2) = 14/6

Cost function with Gradient Descent





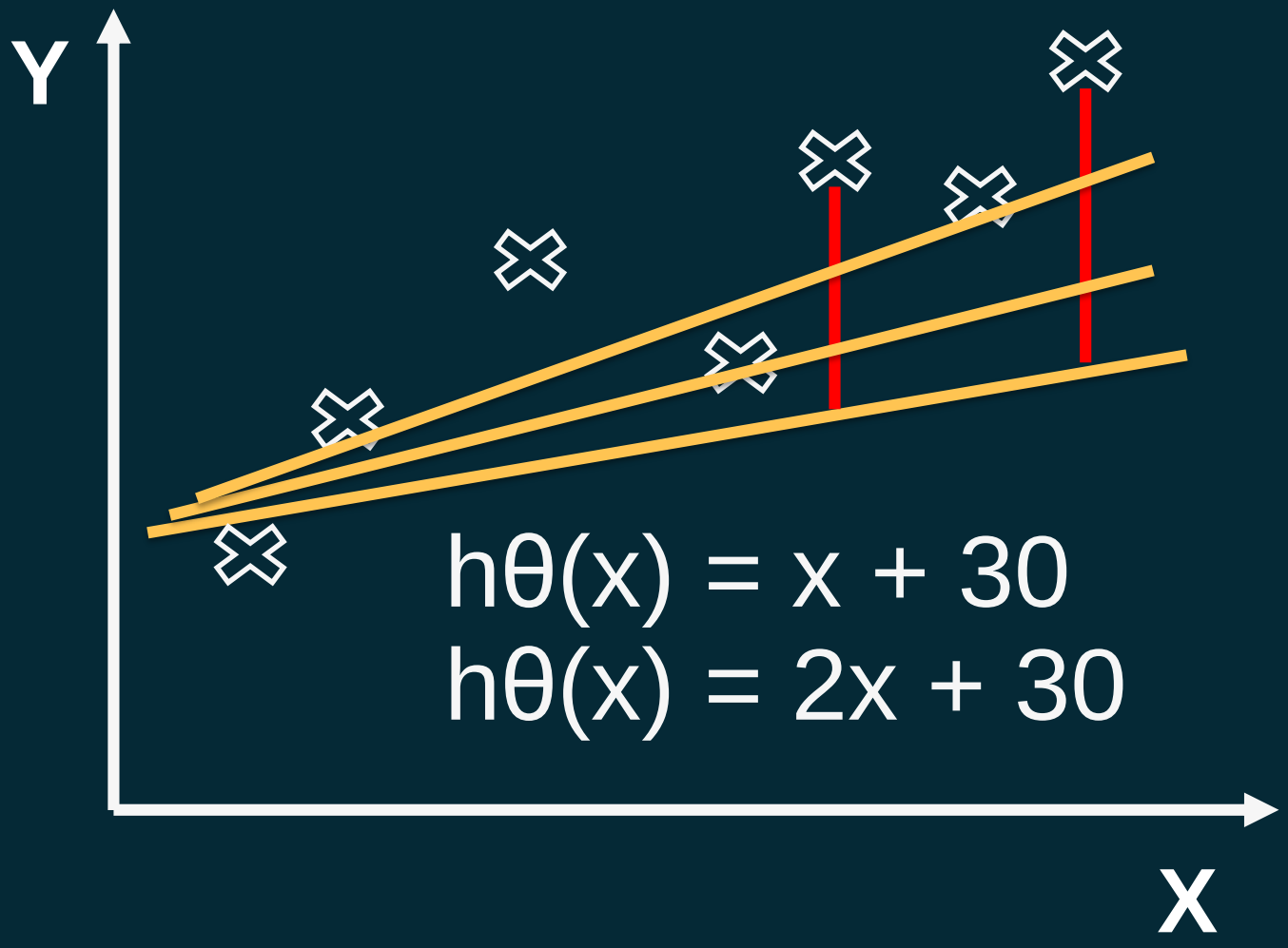
Cost function with Gradient Descent



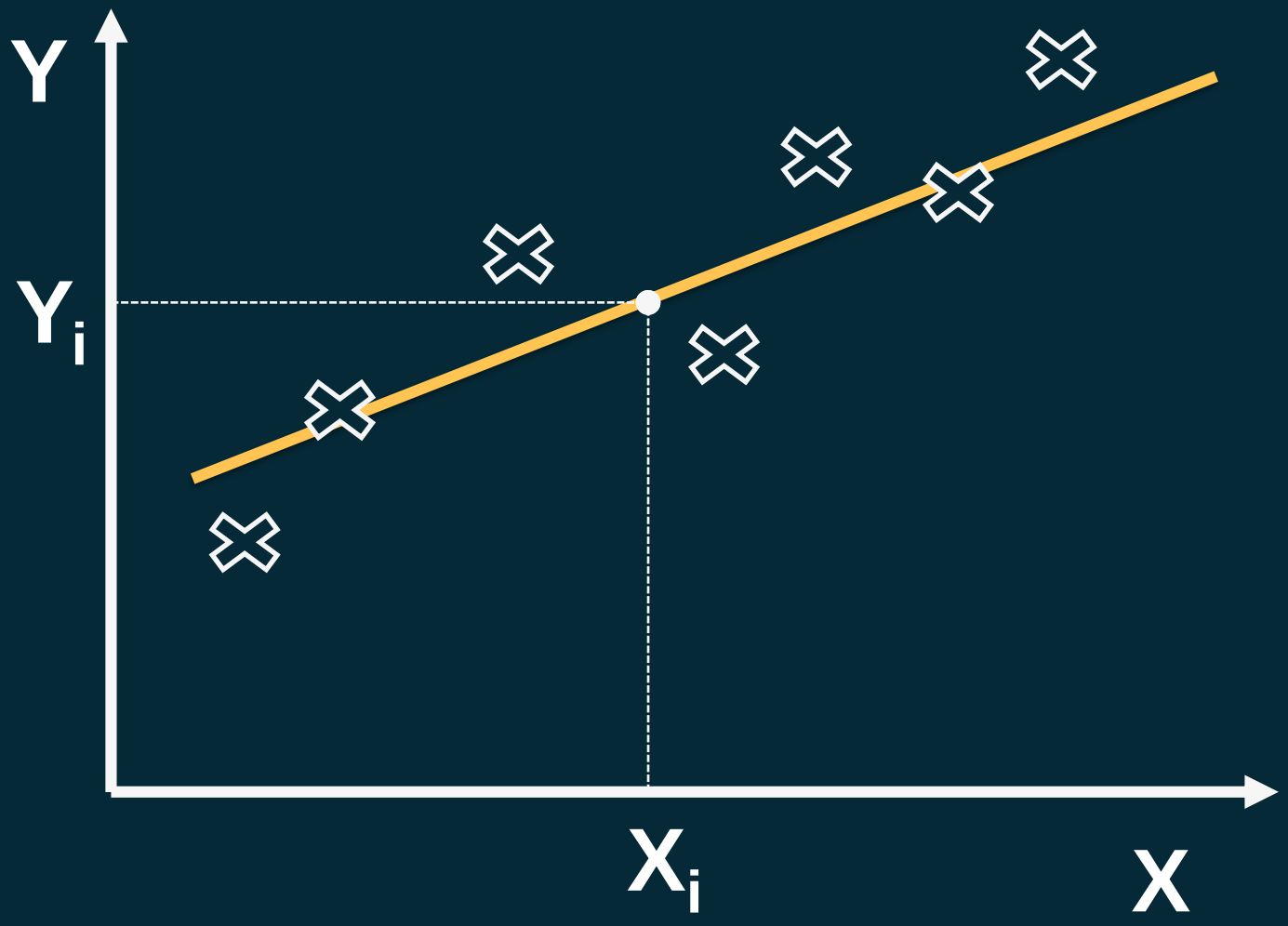
- J(0) =49/6 = 8.167
- J(1) =14/6 = 2.333
- J(2) =0/6 = 0
- J(3) =14/6 = 2.333
- J(4) =49/6 = 8.167

Optimum θ is 2 – minimize the cost function, best fitted model is h = 2X.

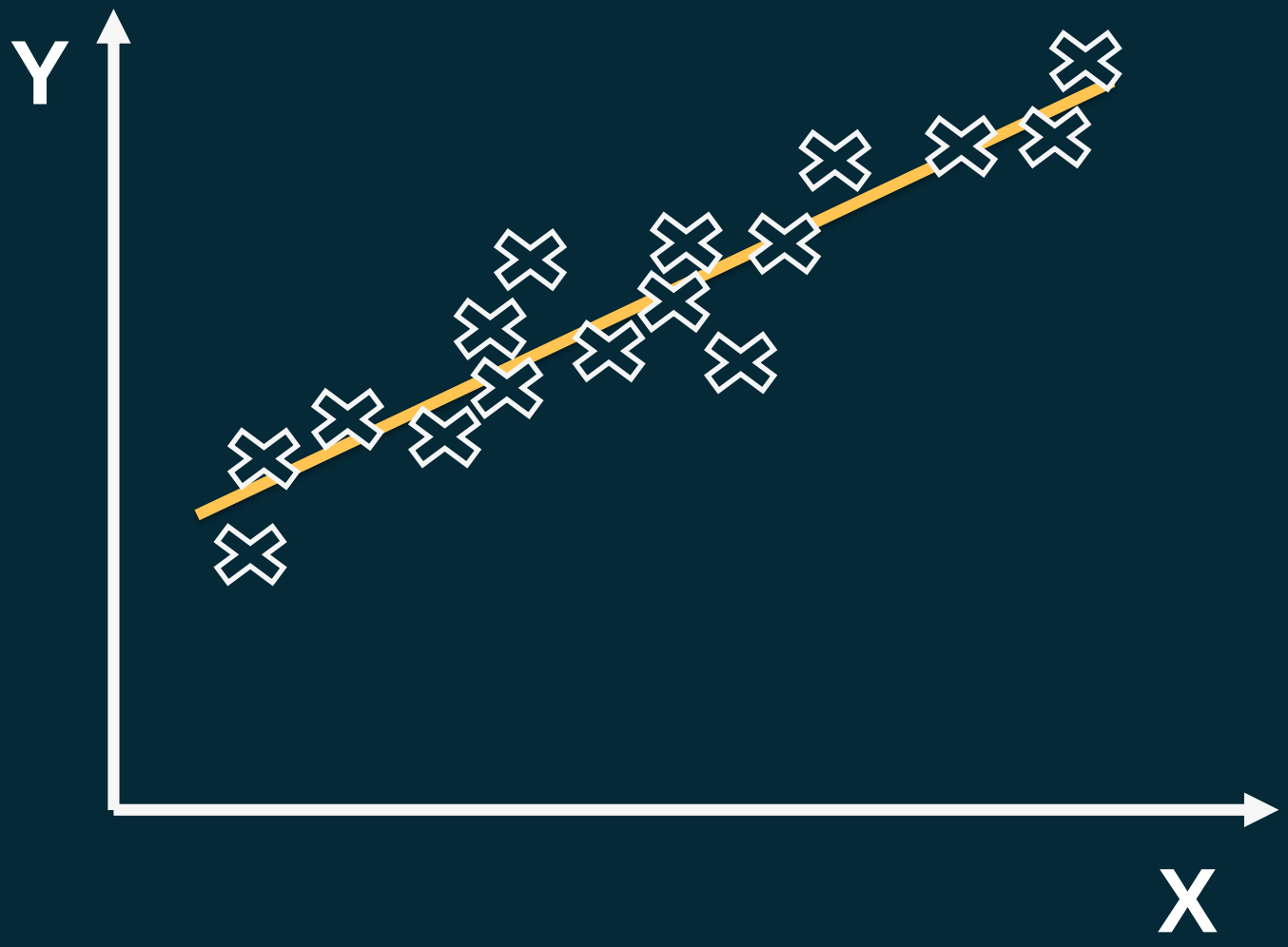
Machine finds best model using input data



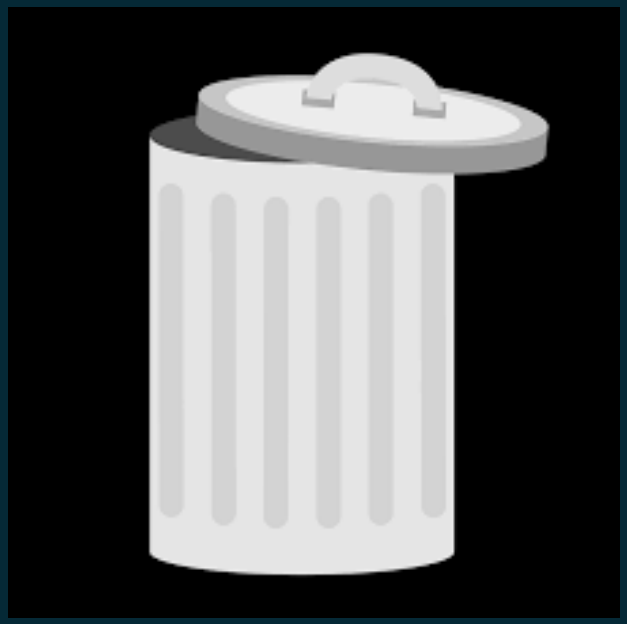
Best model can provide best predicted value.



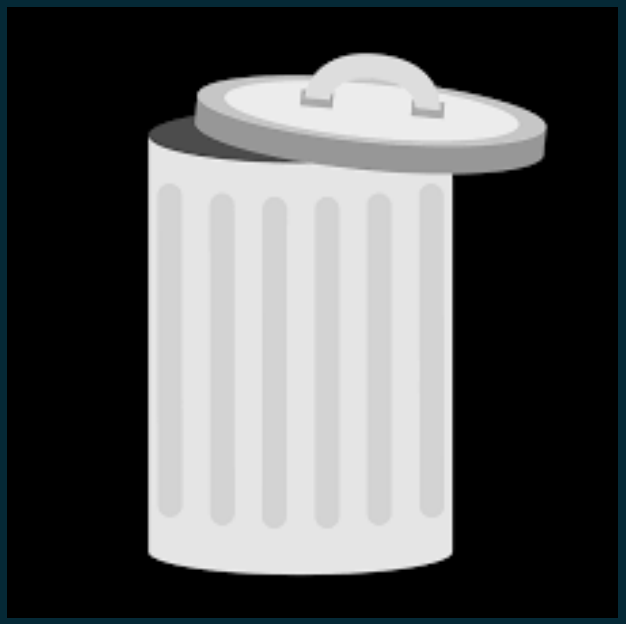
More data, the better model



Data Quality in Machine Learning



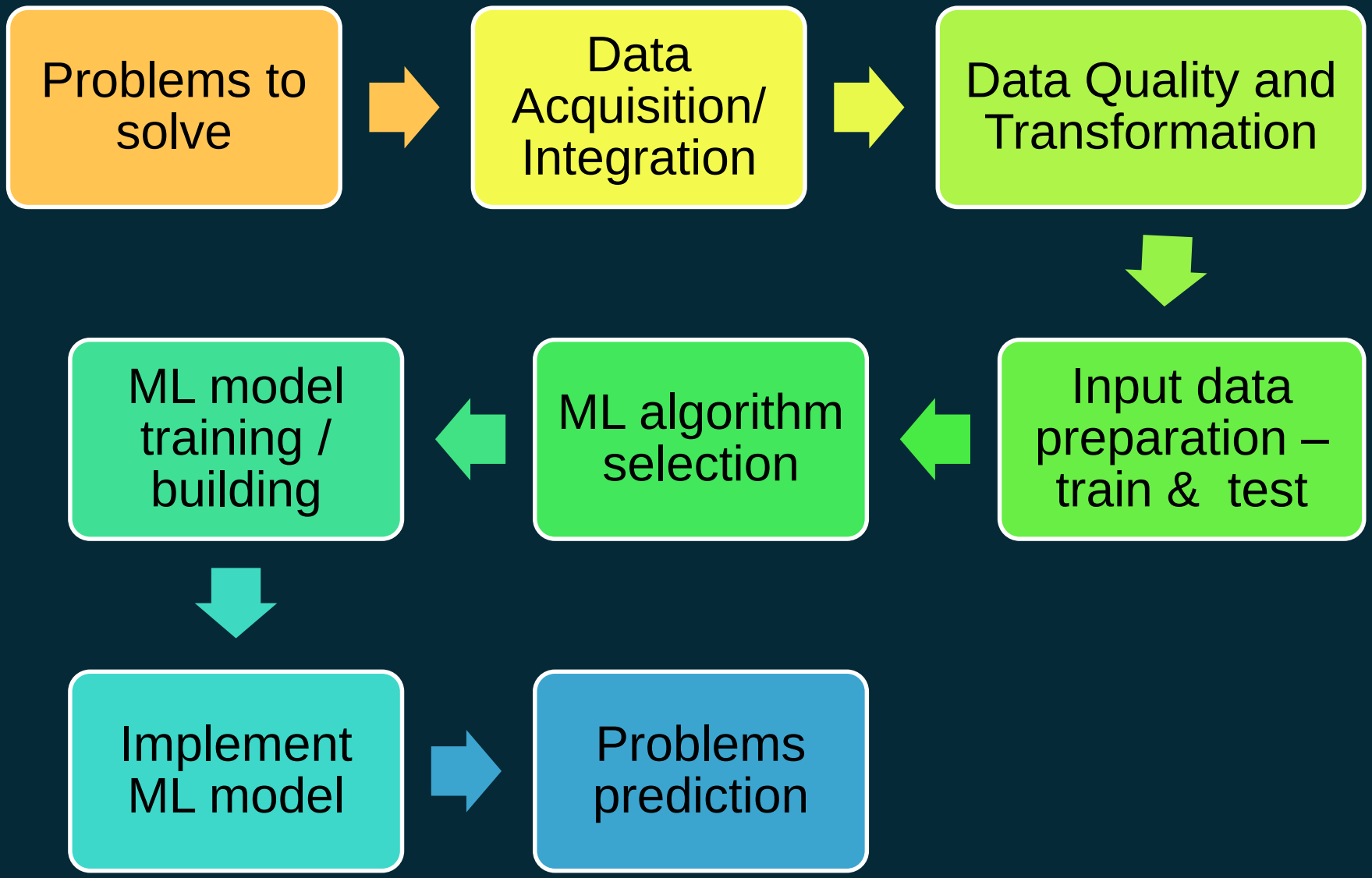
=



Garbage in

Garbage out

Typical Machine Learning Workflow



Machine Learning Type

- Supervised - we know the correct answers
- Unsupervised – no answers
- Artificial Neural Network – like human neural network

Supervised Machine Learning

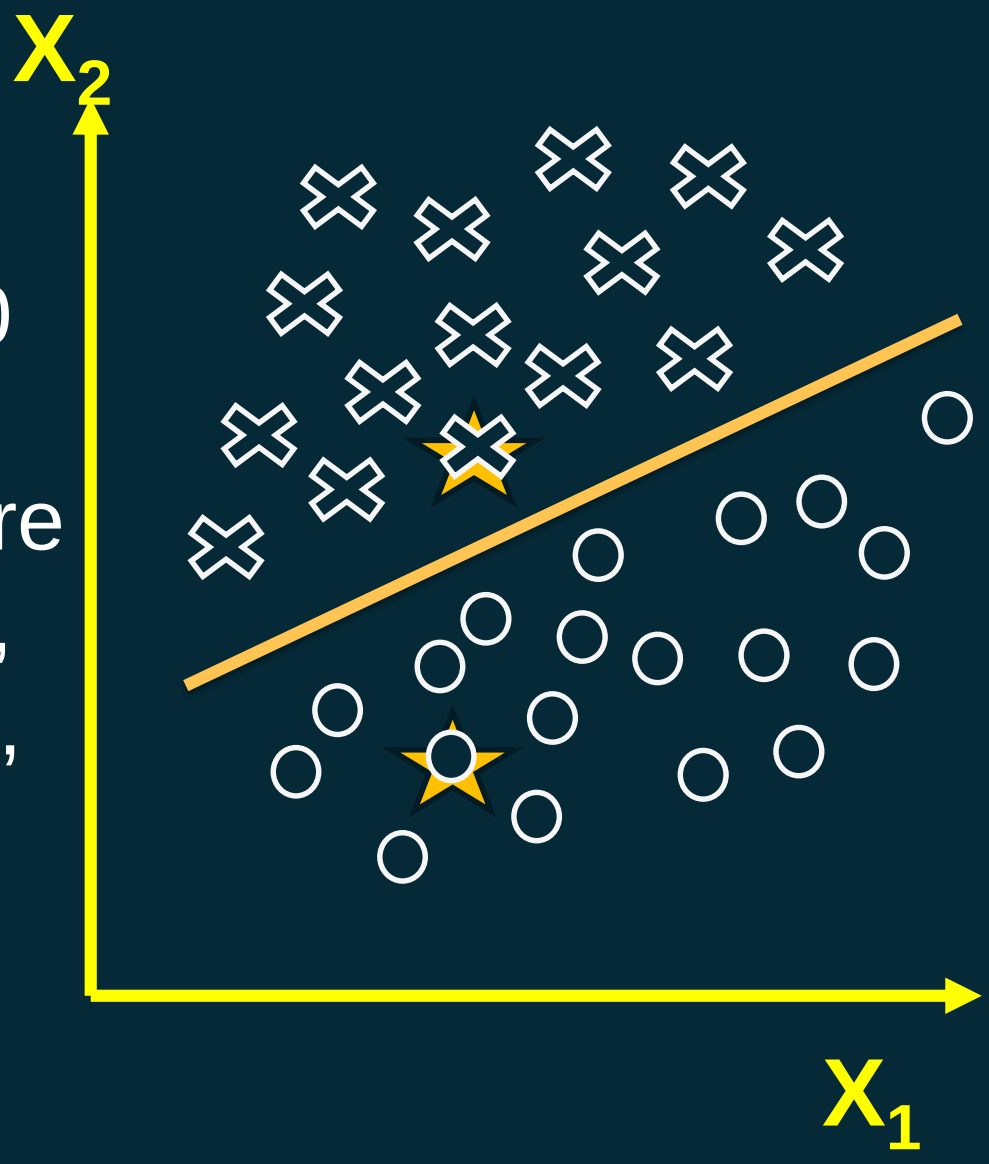
- Input data labeled – has correct answers

X0	X1	X2	…	Xn	Y

- Specific purpose
- Types
  - Classification for distinct output values
  - Regression for continuous output values

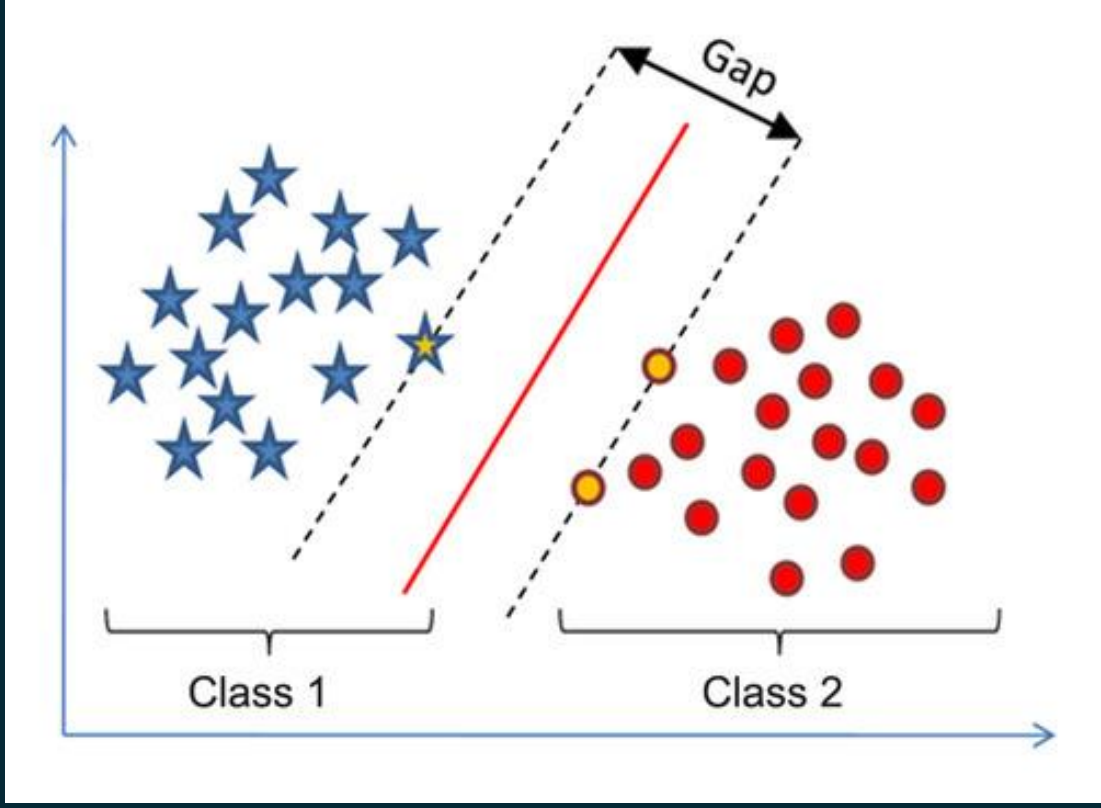
Classification

- Categorical target
- Often binary
- Example : Yes/No, 0 to 9, mild/moderate/severe
- Logistic Regression, SVM, Decision Tree, Forests



Support Vector Machine (SVM)

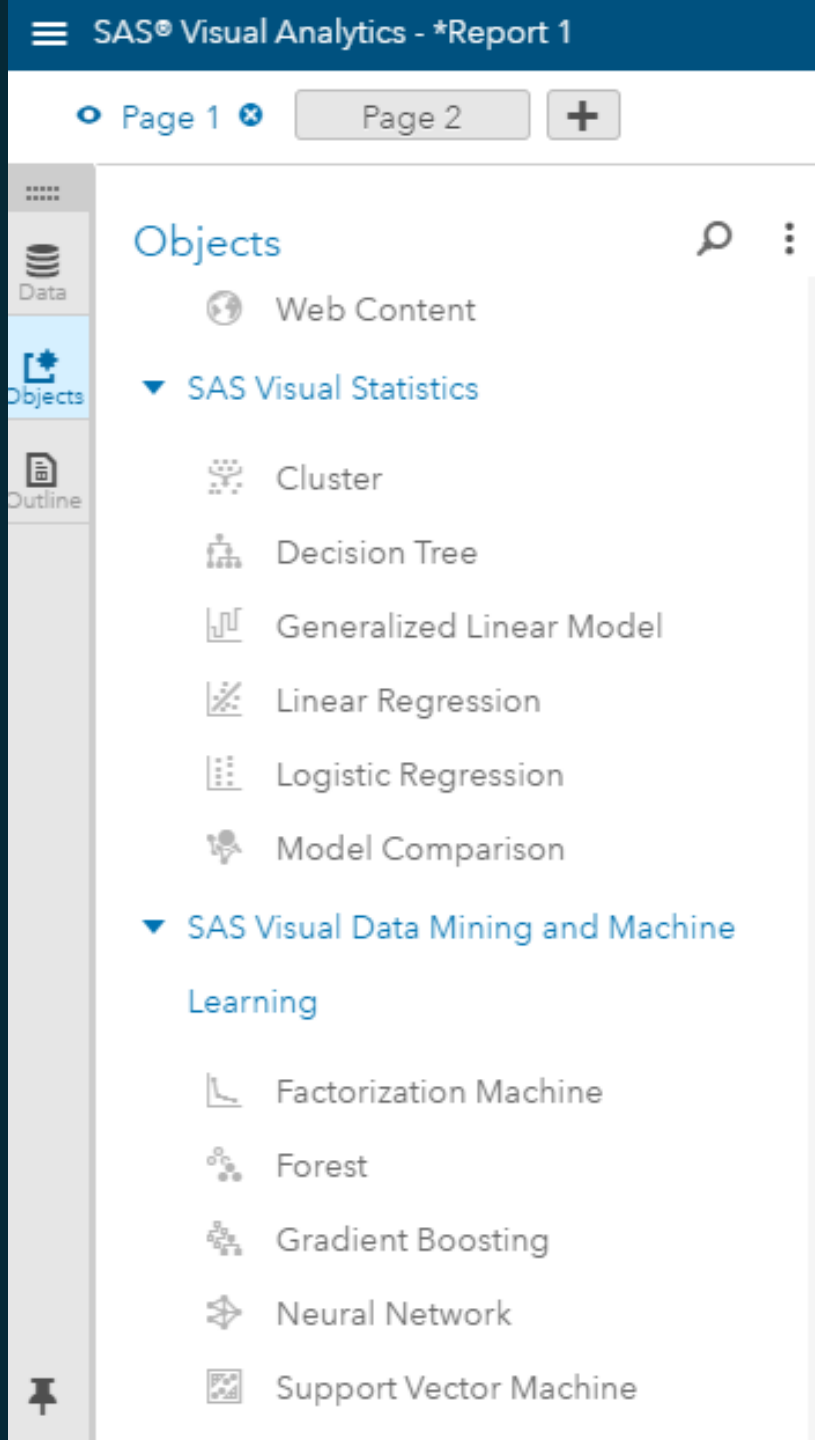
SVM is one of the most powerful classification model, especially for complex, but small/mid-sized datasets.



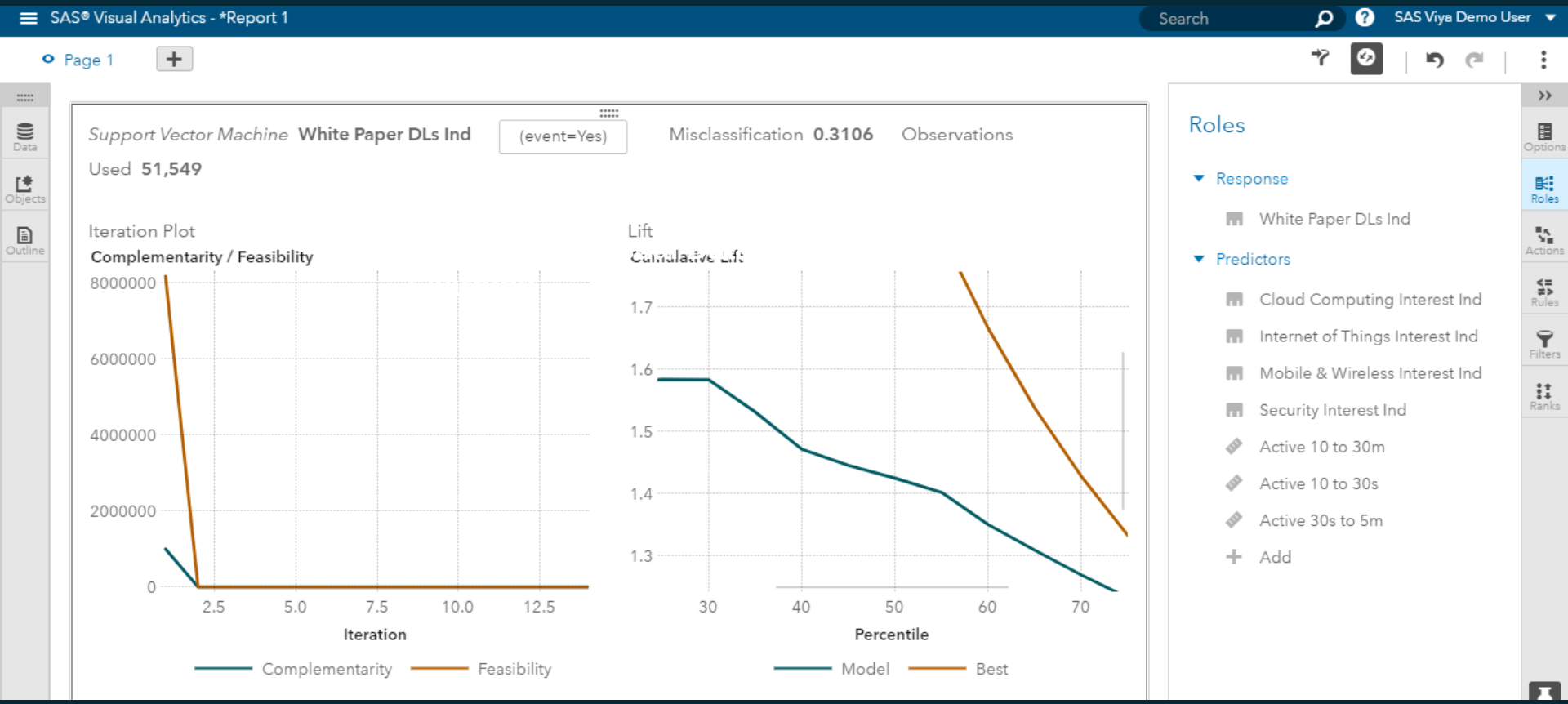

```
*** SVM;
proc svmachine data=x_train C=1.0;
    kernel linear;
    input x1 x2 x3 x4 / level=interval;
    target  y;
run;
```

SVM in SAS Visual Data Mining and Machine Learning

SAS Machine Learning portal can provide an interactive modeling.



SVM in SAS Visual Data Mining and Machine Learning

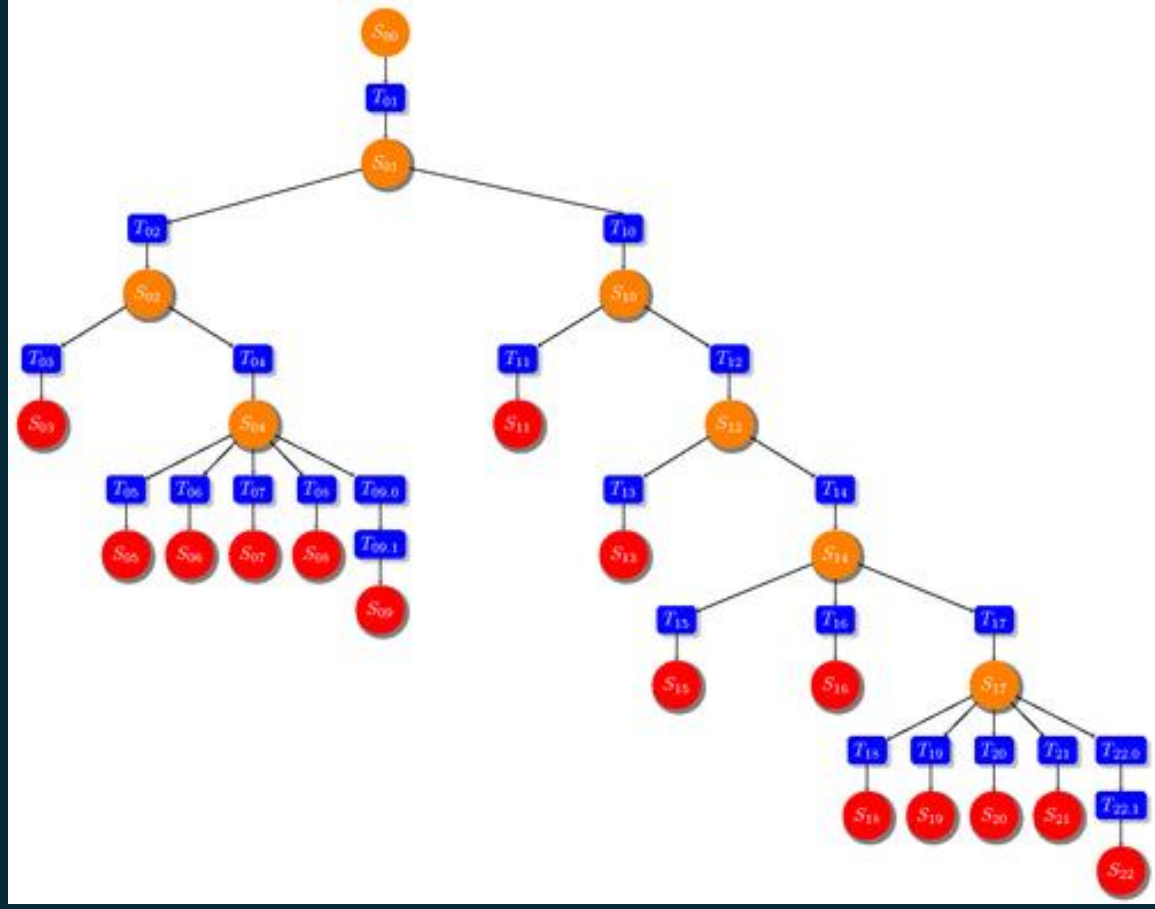


Python codes for SVM


```
#import ML algorithm
from sklearn.svm import SVC

#prepare train and test datasets
x_train = …
y_train = ….
x_test = ….

#select and train model
svm = SVC(kernel='linear', C=1.0, random_state=1)
svm.fit(x_train, y_train)

#predict output
predicted = svm.predict(x_test)
```

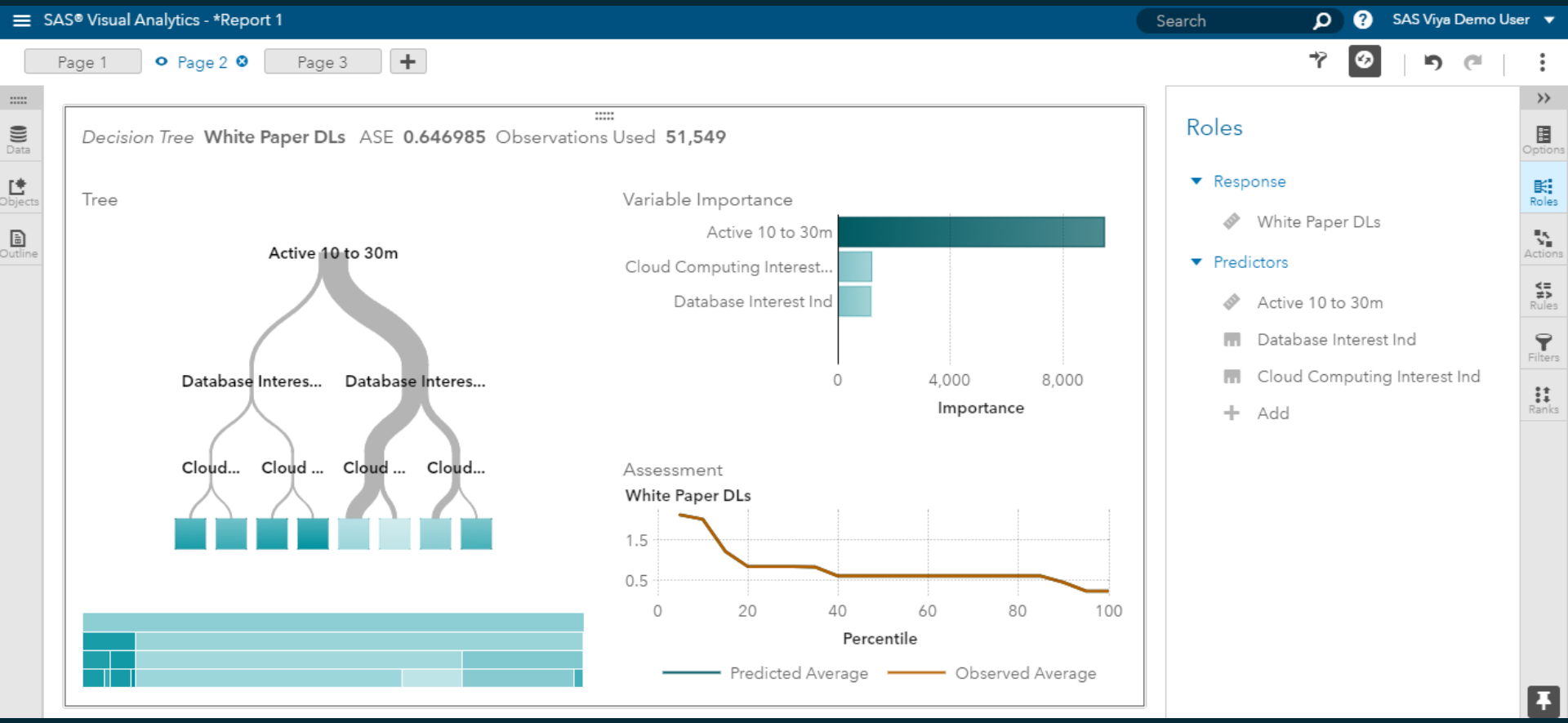
Decision Trees

- identify various ways of splitting a data set into branch-like segments.
- Example : predicting the conditions for death



```
PROC HPSPLIT data = ADAE maxleaves=100
    maxbranch = 4 leafsize=1 ;
    model Y(event='y') = x1 x2 x3 x4;
Run;
```

Decision Tree in SAS Visual Data Mining and Machine Learning

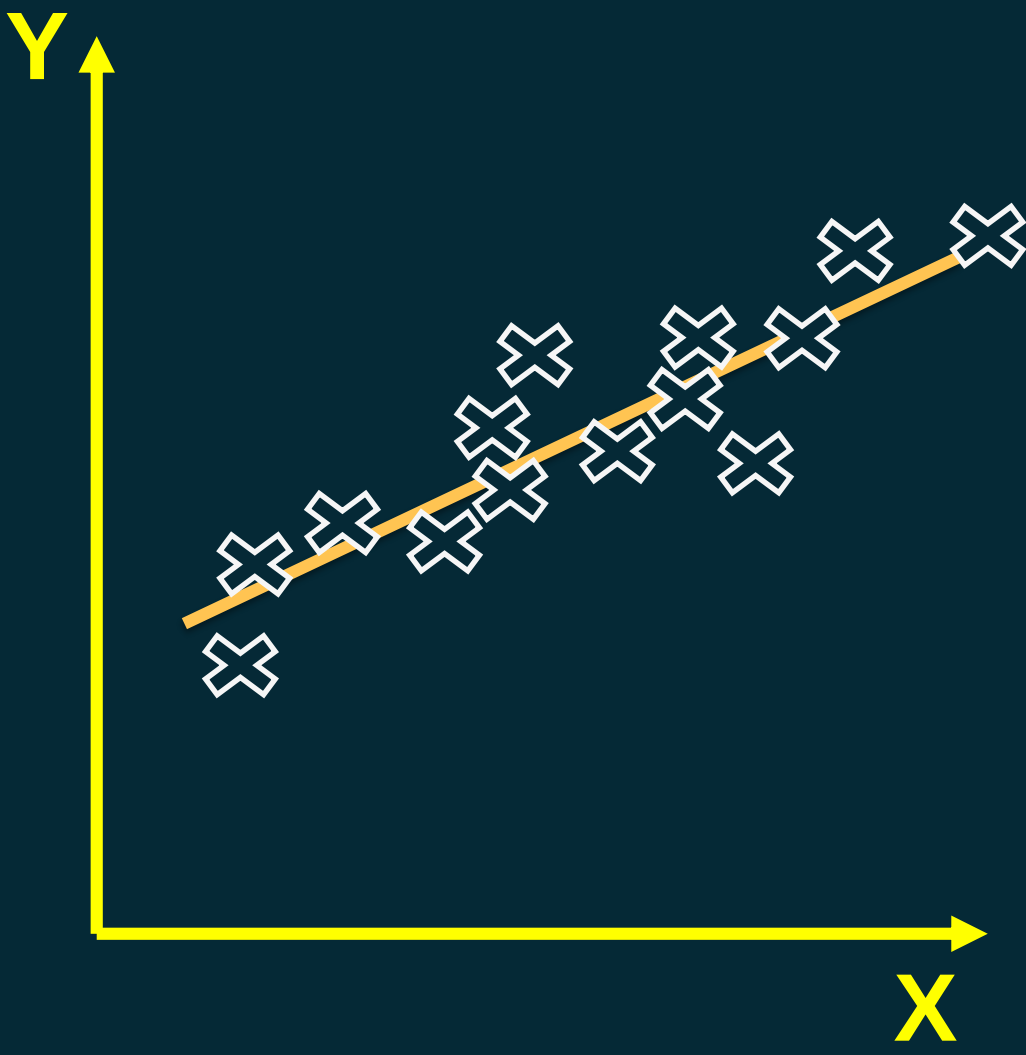


Python codes for Decision Tree


```
#import ML algorithm
from sklearn.tree import DecisionClassifier

#prepare train and test datasets
x_train = …
y_train = ….
x_test = ….

#select and train model
d_tree = DecisionClassifier(max_depth=4)
d_tree.fit(x_train, y_train)

#predict output
predicted = d_tree.predict(x_test)
```

Regression

- Numeric target
- Continuous variables
- Example : predicting house price per sqft
- Linear Regression, Polynomial Regression

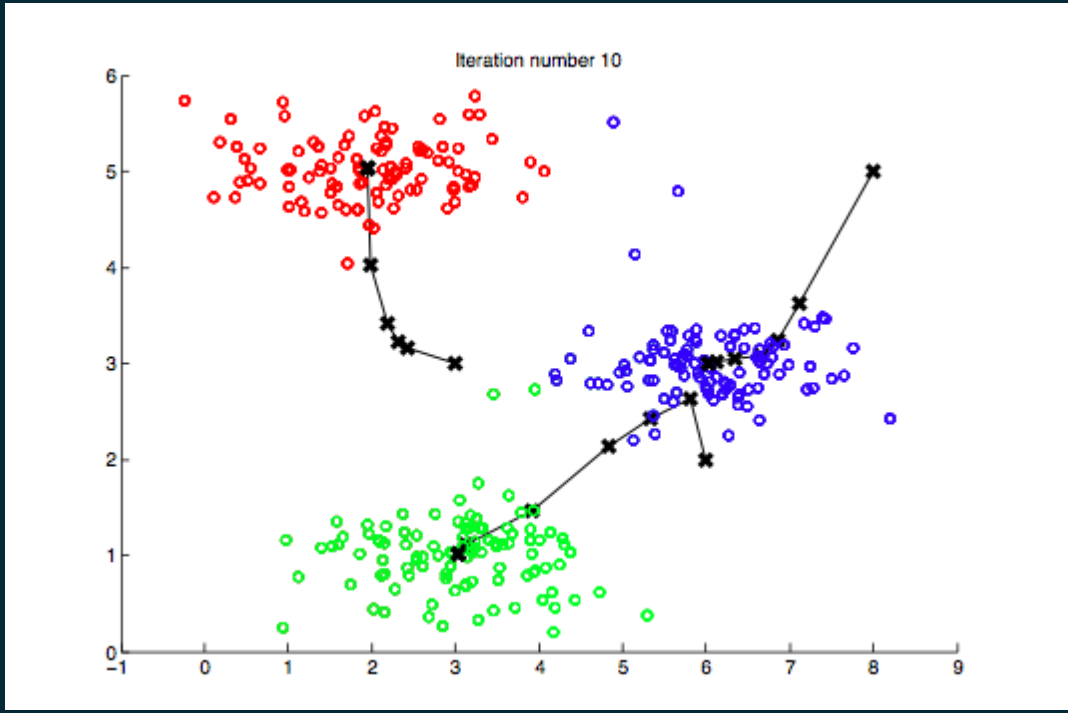


Python codes for ML Linear Regression

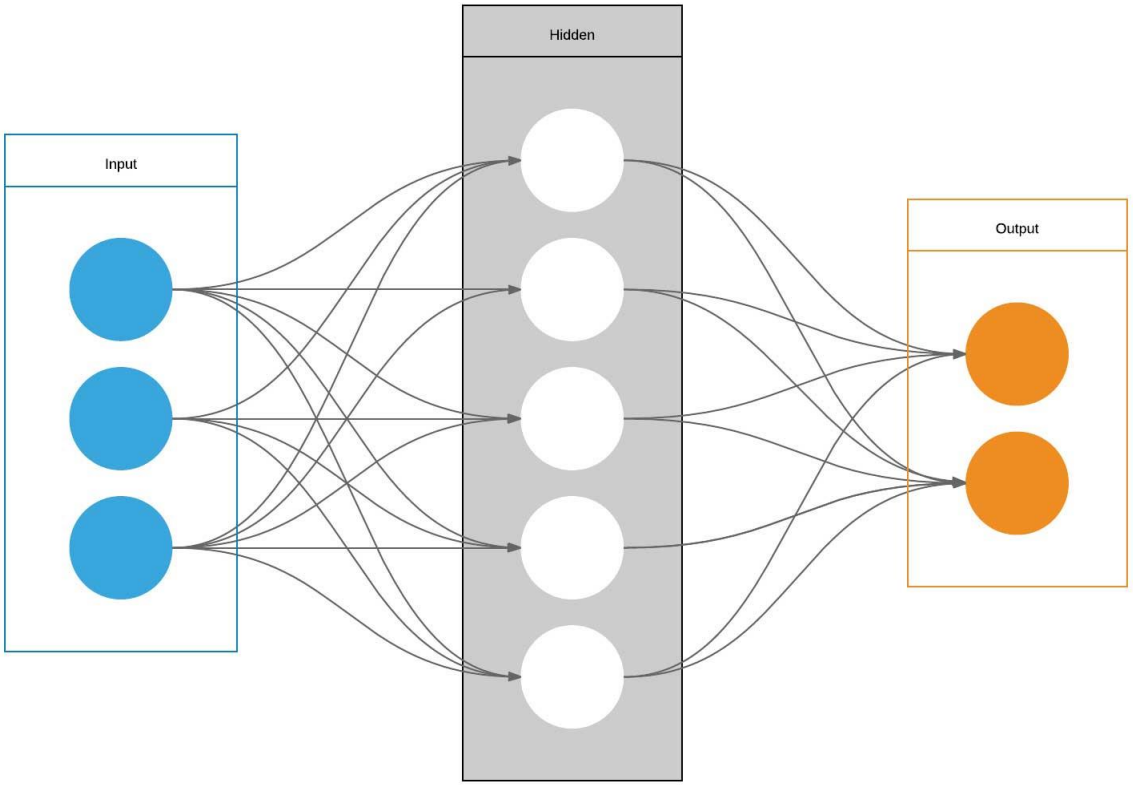

```
#import ML algorithm
from sklearn import linear_model

#prepare train and test datasets
x_train = …
y_train = ….
x_test = ….

#select and train model
linear = linear_model.LinearRegression()
linear.fit(x_train, y_train)

#predict output
predicted = linear.predict(x_test)
```

Unsupervised Machine Learning

- Input data not-labeled – no correct answers
- Exploratory
- Clustering – the assignment of set of observations into subsets (clusters)

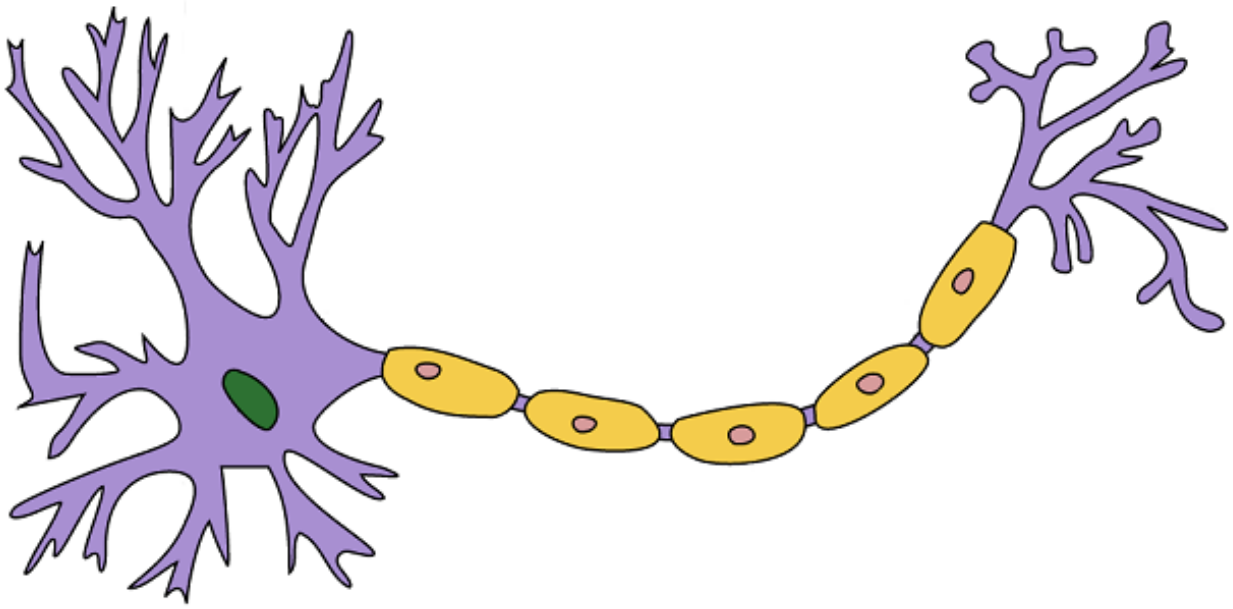


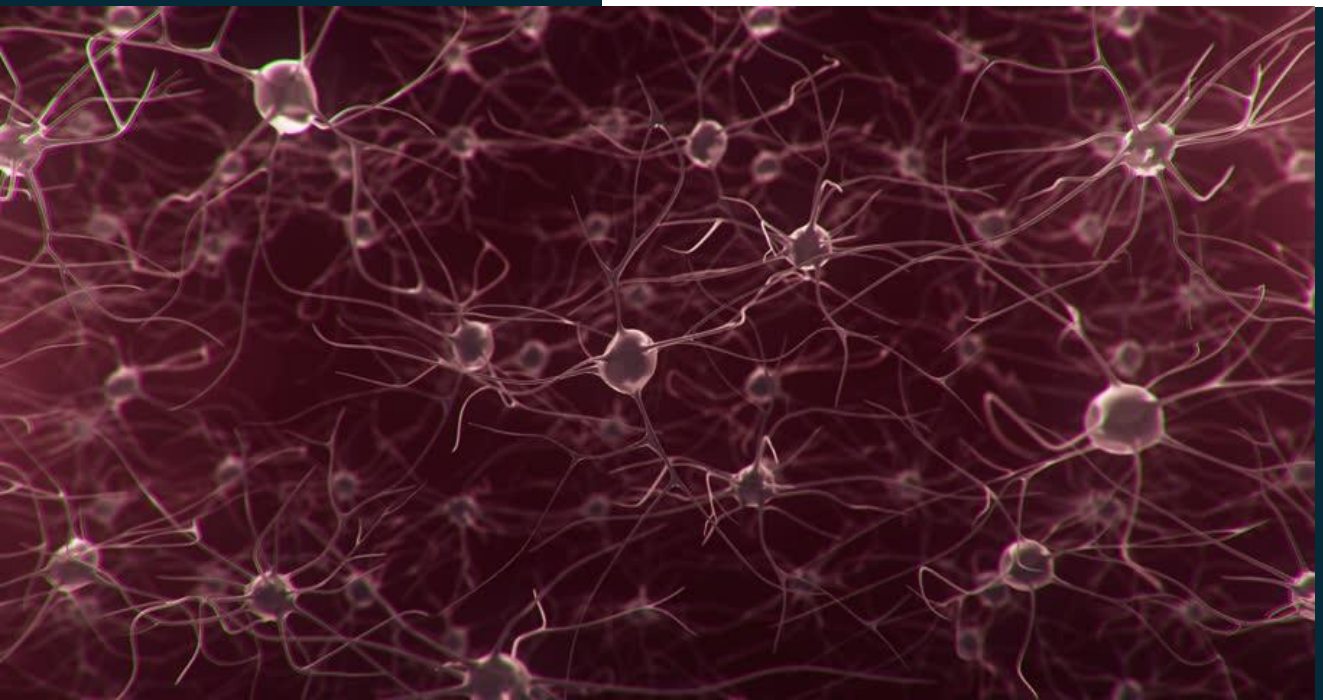
Artificial Neural Network (ANN)



- Most powerful ML algorithm
- Game Changer
- Works very much like human brain – Neural network

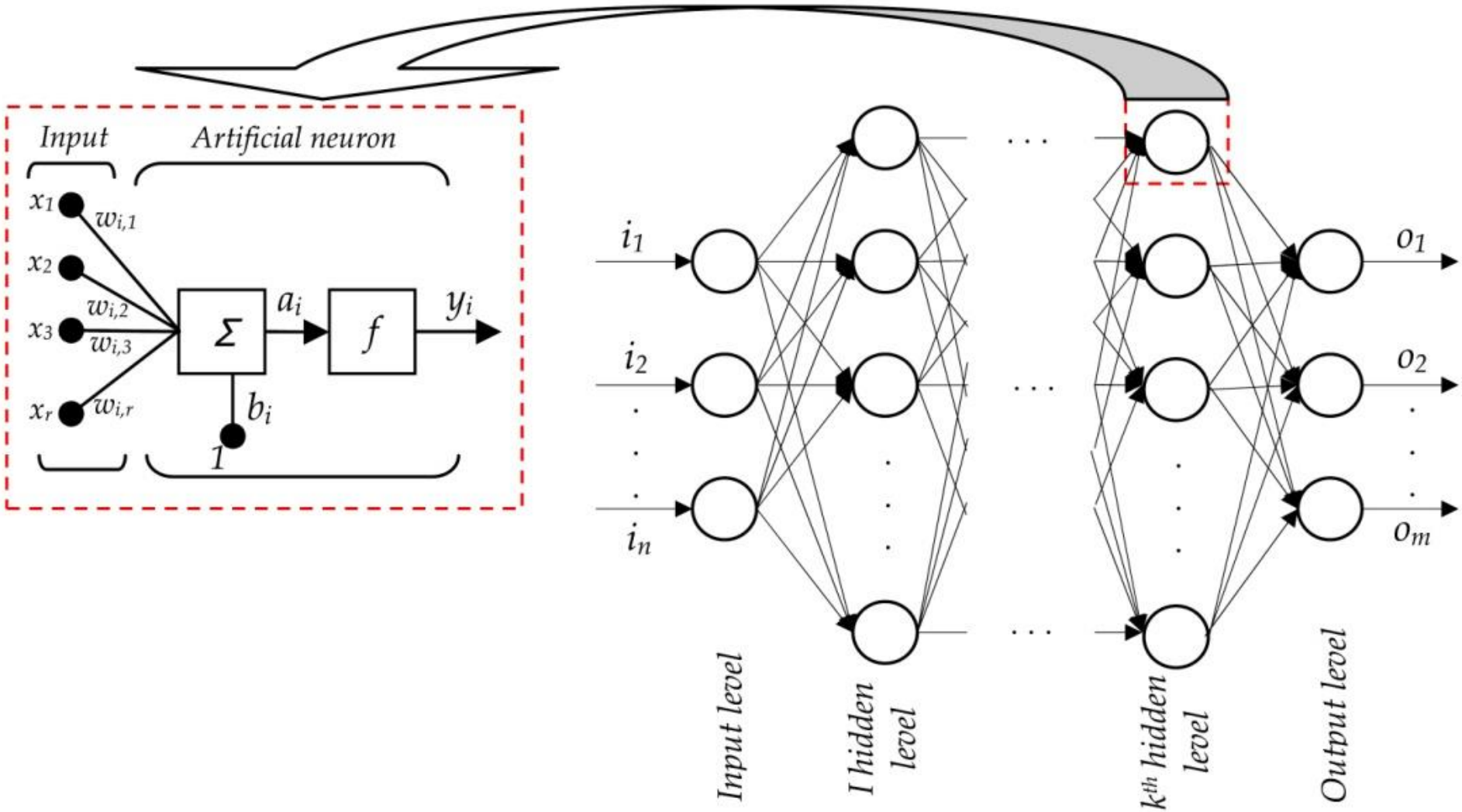
Human Neuron





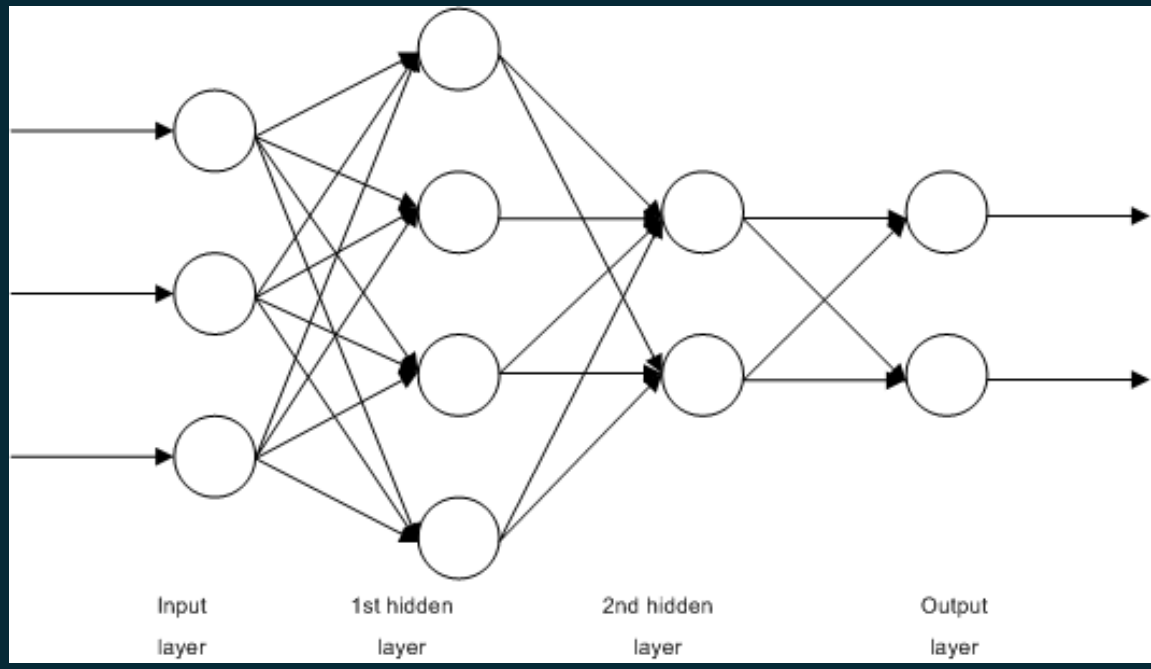
Neural Network – 100 billions

Artificial Neural Network (ANN) Introduction



ANN Architecture

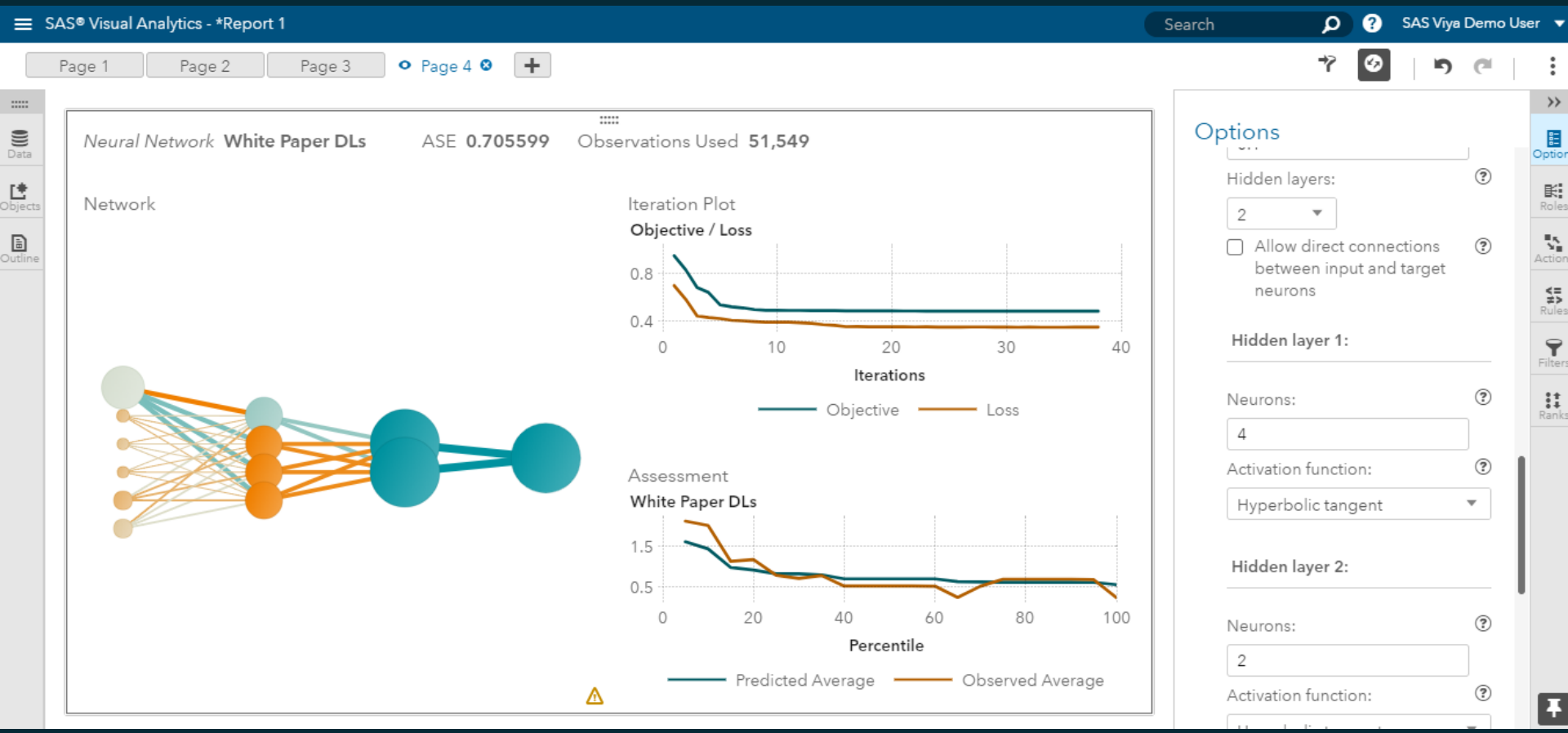
- Input layer
  - 3 features (variables)
- Hidden layer
  - Hidden layer1 - 4 neurons
  - Hidden layer2 - 2 neurons
  - Other parameters – weight, activation function, learning rate
- Output layer – 2 outputs



Neural Network in SAS using proc nnet


```
Proc nnet data=Train;
      architecture mlp;
      hidden 4;
      hidden 2;
      input x1 x2 x3 x4;
      target Y;
Run;
```

Neural Network in SAS Visual Data Mining and Machine Learning

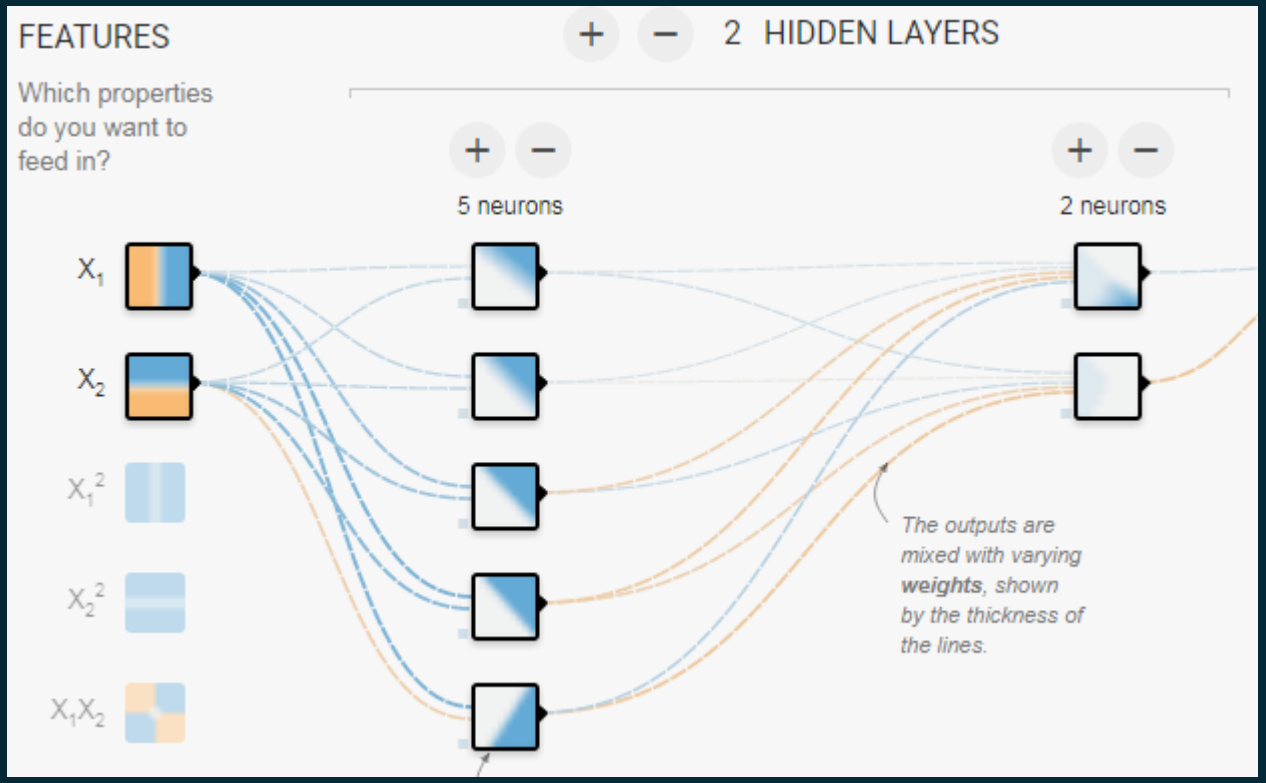


Python codes for DNN

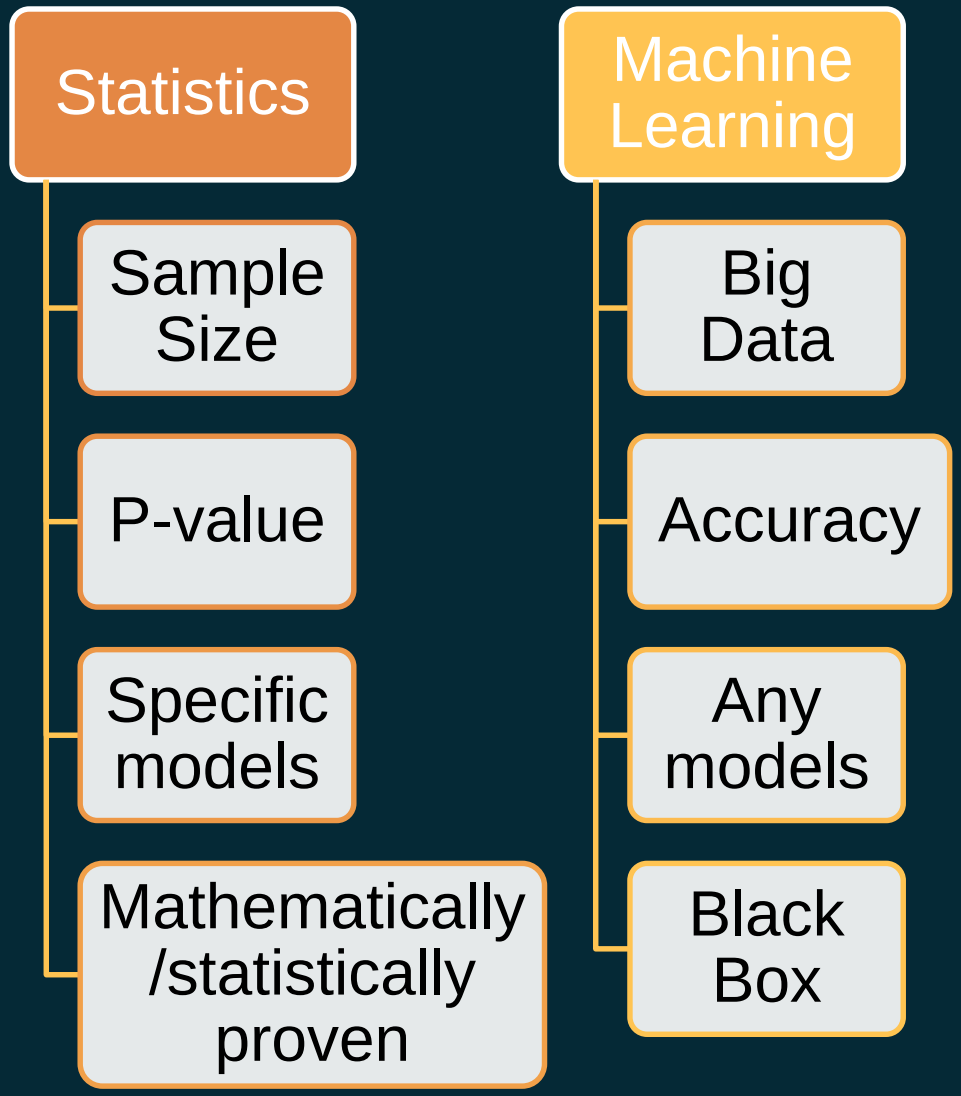

```
#import ANN - TensorFlow
Import tensorflow as tf

X = tf.placeholder(..)
Y = tf.placeholder(..)
hidden1 = tf.layer.dense(X, 4, activation=tf.nn.relu)
hidden2 = tf.layer.dense(hidden1, 2, activation=tf.nn.relu)
logits = neuron_layer(hidden2, 2)
….
loss = tf.reduce_mean(….)
optimizer = tf.train.GradientDescentOptimezer(0.1)
traing_op = optimizer.minimizer(loss)

tf.Session.run(training_op, feed_dict={X:x_train, Y:y_train})
```

Tensor Flow Demo

http://playground.tensorflow.org



Difference between Statistics and Machine Learning



Where is SAS in ML?

SAS Visual Data Mining and Machine Learning

- Linear Regression
- Logistic Regression
- Forest
- Support Vector Machine
- Neural Networks (limited layers)


How ML is being used in our daily life

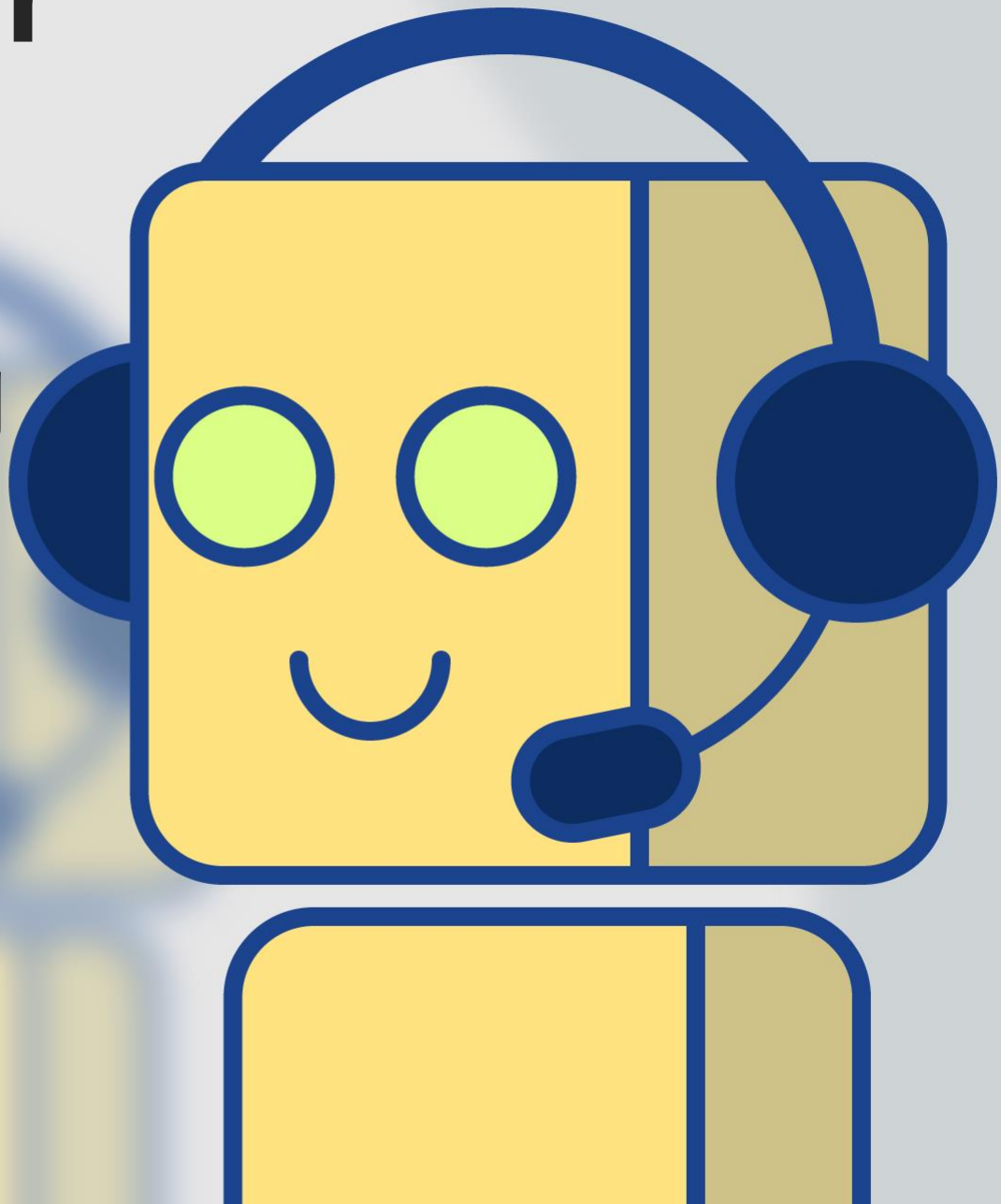

HEY
SIRI

Recommendation

- Amazon
- Netflix
- Spotify

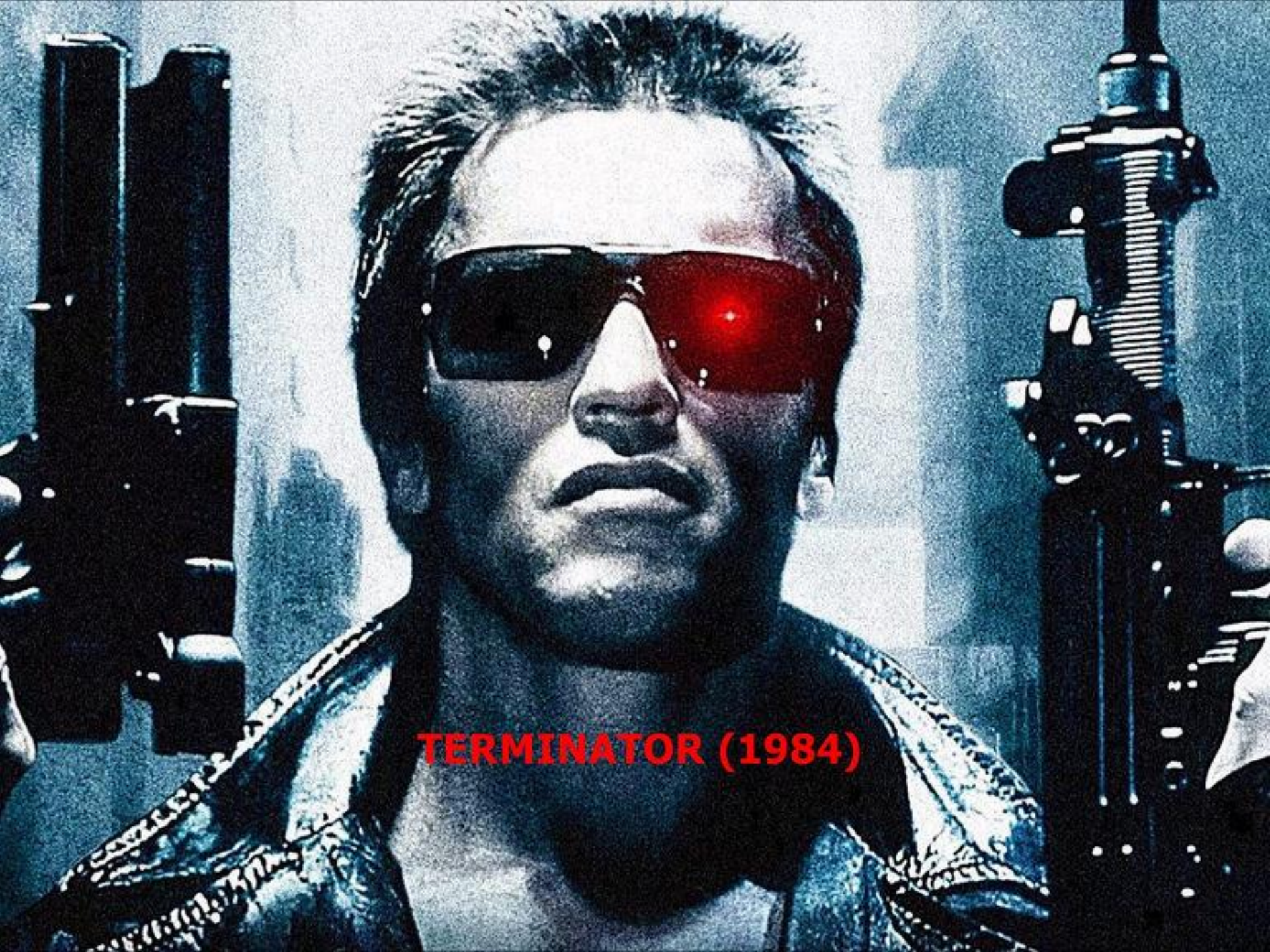
Customer Service

- Online Chatting
- Call



AlphaGO


A B C D E F G H I J K L M N O P Q R S
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
ALPHAGO
01:27:15
LEE SEDOL
00:45:18


TERMINATOR (1984)

Why is AI(ML) so popular now?

- Cost effective
  - Automate a lot of works
  - Can replace or enhance human labors
  - “Pretty much anything that a normal person can do in <1 sec, we can now automate with AI” Andrew Ng
- Accurate
  - Better than humans
- Can solve a lot of complex business problems

Now, how Pharma goes into AI/ML market

- GSK sign $43 million contract with Exscientia to speed drug discovery
  - aiming ¼ time and ¼ cost
  - identifying a target for disease intervention to a molecule from 5.5 to 1 year

- J&J
  - Surgical Robotics – partners with Google. Leverage AI/ML to help surgeons by interpreting what they see or predict during surgery

Now, how Pharma goes into AI/ML market

- Roche
  - With GNS Healthcare, use ML to find novel targets for cancer therapy using cancer patient data
- Pfizer
  - With IBM, utilize Watson for drug discovery
  - Watson has accumulated data from 25 million articles compared to 200 articles a human researcher can read in a year.

Now, how Pharma goes into AI/ML market

- Novartis
  - With IBM Watson, developing a cognitive solution using real-time data to gain better insights on the expected outcomes.
  - With Cota Healthcare, aiming to accelerate clinical development of new therapies for breast cancer.

ML application in Pharma R&D

- Drug discovery
- Drug candidate selection
- Clinical system optimization
- Medical image recognition
- Medical diagnosis
- Optimum site selection / recruitment
- Data anomality detection
- Personalized medicine

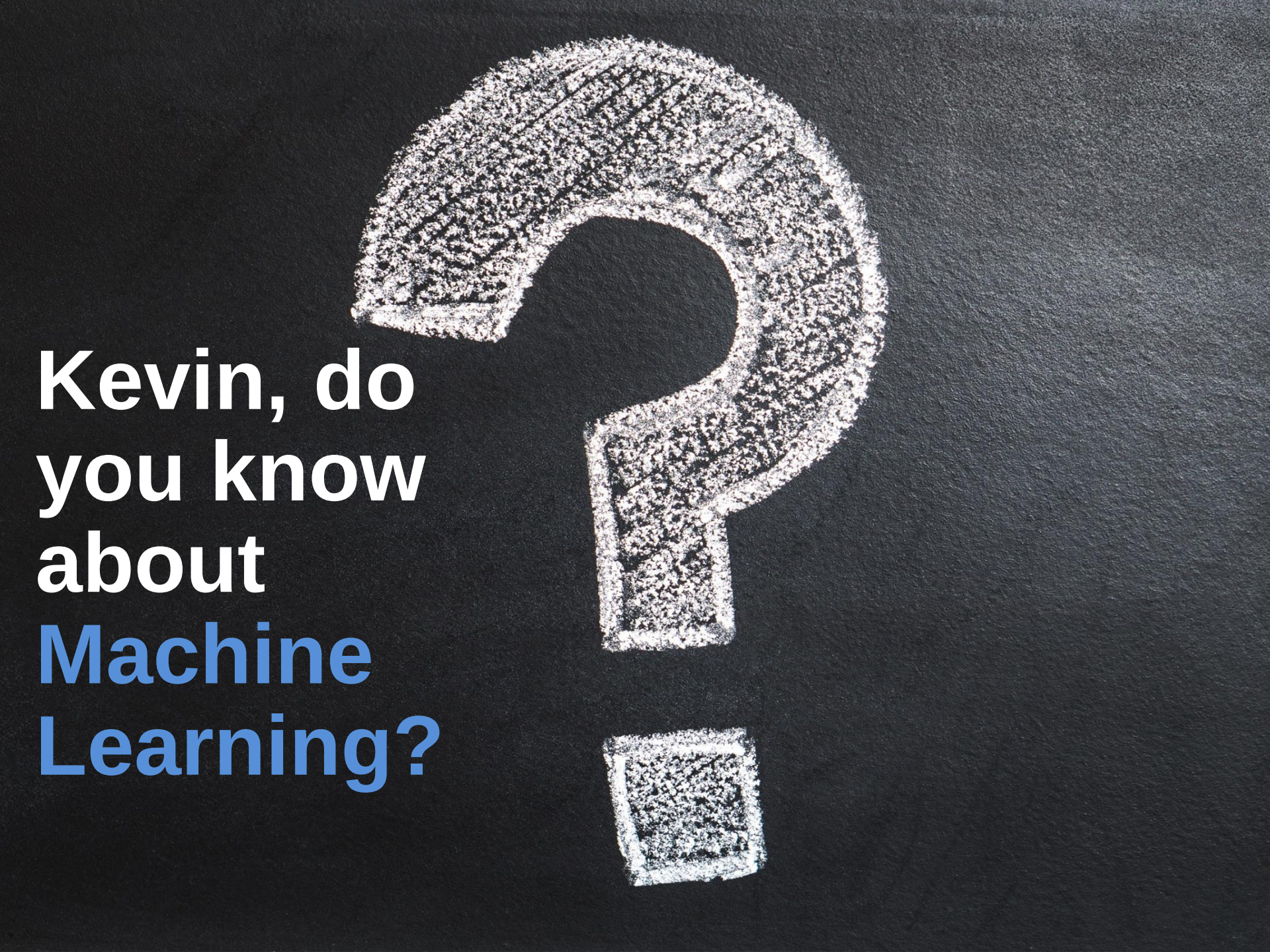
Adoption of AI/ML in Pharma

- Slow
- Regulatory restriction
- Machine Learning Black Box challenge – need to build ML models, statistically or mathematically proven and validated, to explain final results.
- Big investment in Healthcare and a lot of AI Start up aiming Pharma.

Healthcare AI/ML market

- US - 320 million in 2016
- Europe – 270 million in 2016
- 40% annual rate
- 10 billion in 2024
- Short in talents


At the Center of Innovation


Kevin, do you know about Machine Learning?

Contact Us!

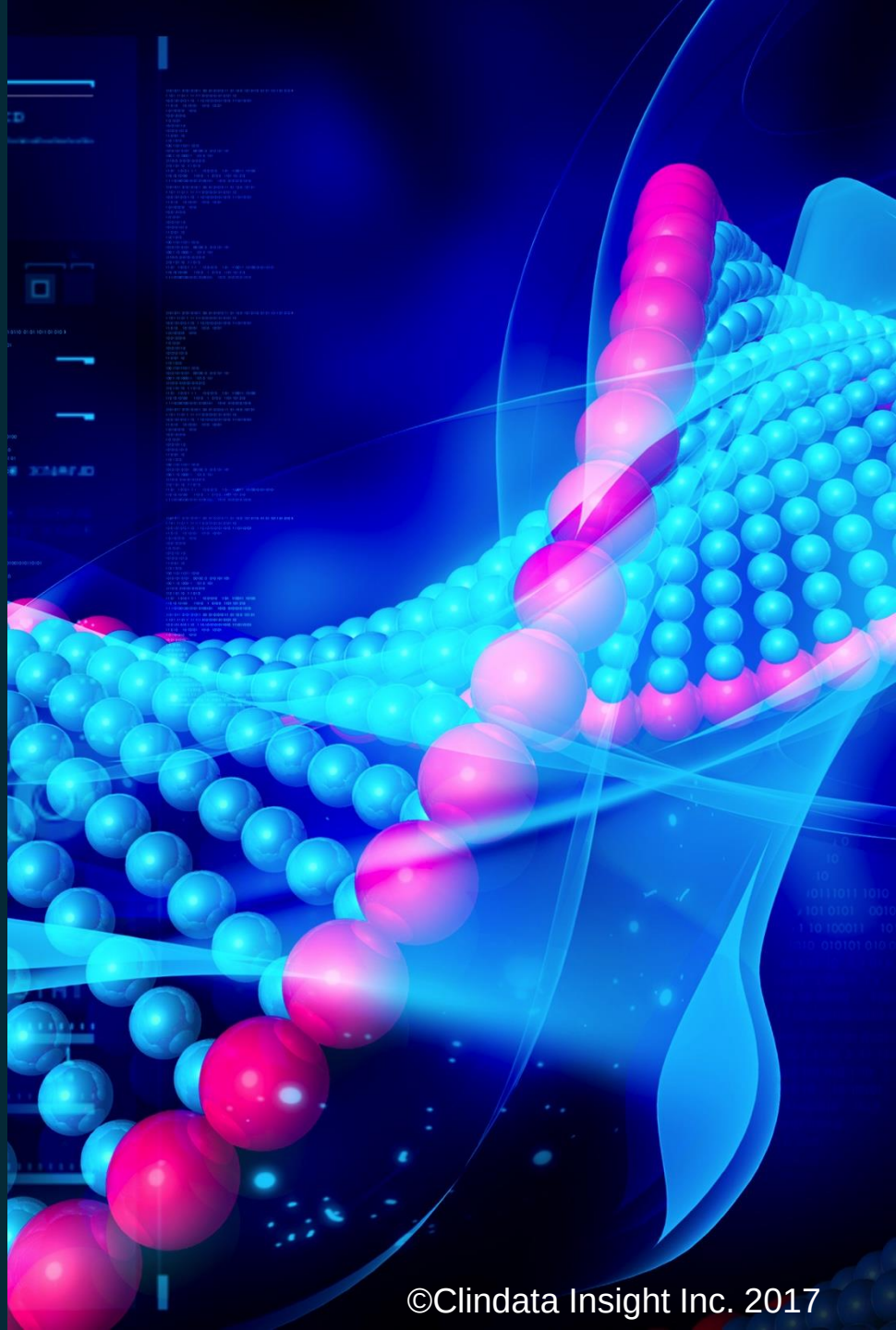
Contact Clindata Insight to learn more about Big Data and Machine Learning.

Email us at
klee@clindatainsight.com
consulting@clindatainsight.com
http://www.clindatainsight.com/

Like us on Facebook @
Facebook.com/clindatainsight

Twitter @clindatainsight

WeChat @clindatainsight



THAT'S ALL

THANKS

Kevin Lee
klee@clindatainsight.com