Crashing Machine Learning

A Crash Course on Machine Learning

@joshuamabina

Who is this for?

What comes to mind, maybe...

- Artificial intelligence?
- Machine learning?
- Deep learning?
- Data-science?
- Big-data?

@joshuamabina

This is for...

- Artificial intelligence?
- Machine learning?
- Deep learning?
- Data-science?
- Big-data?

The aim is to raise the comfort-bar around machine learning concepts.

@joshuamabina

About me

Developer

Vim

Data-science

@joshuamabina

Worth mentioning...

1. Believer & preacher of the gospel of JavaScript
2. Assisting **@Andela** in Tanzania **#ALCinTanzania**
3. Community ninja: **@dardotjs @forLoopTanzania @GDGDarEsSalaam**
4. Marathoner! Next, Kili Full Marathon

What is machine learning?

Machine learning...

Machine learning (ML) is the **study of algorithms** and **mathematical models** that computer systems use to **progressively improve their performance on a specific task.**

~ Wikipedia

Machine learning...

Machine learning algorithms **build a mathematical model of sample data**, known as "training data", in order to **make predictions or decisions without being explicitly programmed** to perform the task.

~ Wikipedia

Simply put,

1. It is a subfield of Artificial intelligence (AI)
2. It is a study of algorithms that learn from examples and experience
3. It is a study of algorithms that help computers learn without being explicitly programmed to do so
4. It is a study of algorithms that help computers predict correct answers & execute instructions without your explicit directions

@joshuamabina

What is the connection between artificial intelligence, machine learning, deep learning, data-science and big-data?

#foodForThought

 @joshuamabina

Head first into machine learning...

Problem: Is it a Mouse or Rat?! 🤔

@joshuamabina

The steps

1. Import your data set
2. Train your model
3. Predict the correct answer

The rodents data...

Weight (units)	Location	Label
140	Town (0)	Mouse (0)
120	Town (0)	Mouse (0)
700	Country (1)	Rat (1)
500	Country (1)	Rat (1)
900	Country (1)	Rat(1)
100	Town (0)	Mouse (0)

@joshuamabina

Importing the data:

features = [[140, 0], [120, 0], ..., [500, 1] [900, 1]]

labels = [0, 0, 1, 1, 1, 0]

@joshuamabina

Training the classifier:

```
classifier = tree.DecisionTreeClassifier()

classifier = classifier.fit(features,label)
```

Predict the rodent:

```
print classifier.predict([[300, 0]])
```

@joshuamabina

```
from sklearn import tree

# 1. import the data
features = [[140, 0], [120, 0], [700, 1], [500, 1], [900, 1], [100,
0]]
label = [0, 0, 1, 1, 1, 0]

# 2. train the classifier
classifier = tree.DecisionTreeClassifier()
classifier = classifier.fit(features,label)

# 3. predict
print classifier.predict([[300, 0]])
```

@joshuamabina

How accurate will the classifier be when used on rodents not in the training data?

@joshuamabina

One way to evaluate an algorithm is by partitioning it into two: (1) the training set, on which we learn some properties; (2) the testing set, on which we test the learned properties.

```
from sklearn import tree

# 1. import the data
testing_features = [[140, 0], [700, 1]]
testing_labels = [0, 1]

training_features = [[120, 0], [500, 1], [900, 1], [100, 0]]
training_labels = [0, 1, 1, 0]

# 2. train the classifier
classifier = tree.DecisionTreeClassifier()
classifier = classifier.fit(training_features, training_labels)

# 3. predict
print classifier.predict(testing_features) # try to compare with below
print testing_labels
```

What are the common machine learning problems?

What comes to mind?

- **Classification**
- Regression
- Clustering
- Association
- Structured output
- Ranking

@joshuamabina

Classification

A common machine learning problem

@joshuamabina

Supervised learning, a type of machine learning, can be approached through classification.

Classification is a process in which **a computer learns to identify to which class the given input belongs to** by **learning from labelled data**.

@joshuamabina

Other classification use-cases:

Spam filtering: Is this email spam or not spam?

Google photos: Given all photos, which ones are selfies?

Recap

Keep calm. Machine learning is fun!

@joshuamabina

- Machine learning is a subfield of AI, focused on training computers to learn from examples and experience without human intervention.
- Supervised learning is a type of ML where by computers learn from labeled data.
- Classification is an approach to supervised learning, in which a computer predicts the category of given data based from training data used.
- To code a machine learning classifier, simply follow the steps: (1) import your data, (2) train your model, (3) predict the correct answer.
- Evaluate the accuracy of a classifier by using some of the input data to train the model, and the rest, to compare with the prediction results.