

MACHINE LEARNING IN JAVA

DAVID STRÖM

CADEC 2019.01.24 & 2019.01.30 | [CALLISTAENTERPRISE.SE](https://callistaenterprise.se)

CALLISTA

— ENTERPRISE —

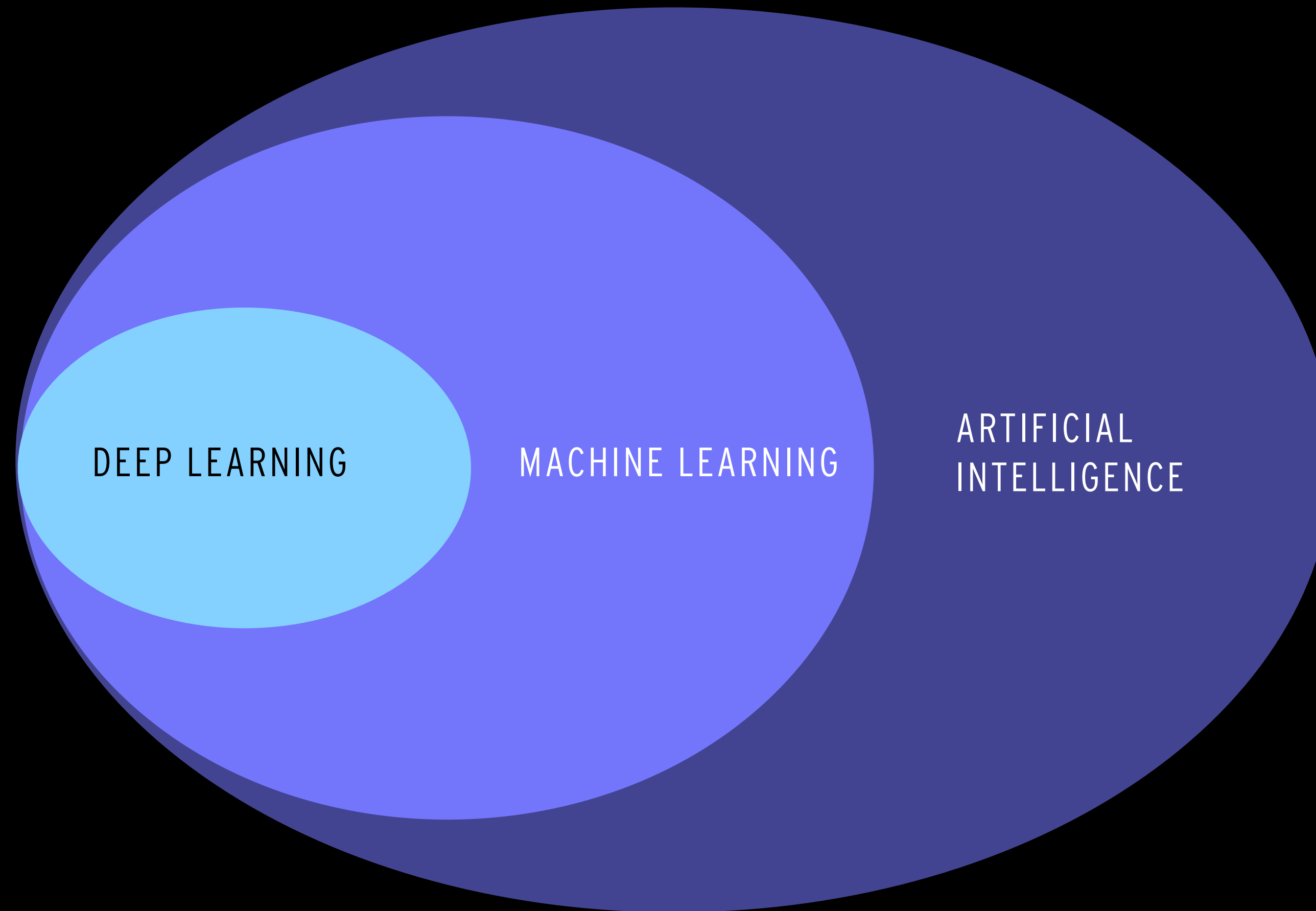
| TODAY'S QUESTIONS

- What is it?
- Why use machine learning, and why use Java?
- How do we do it?



WHAT?

| WHAT



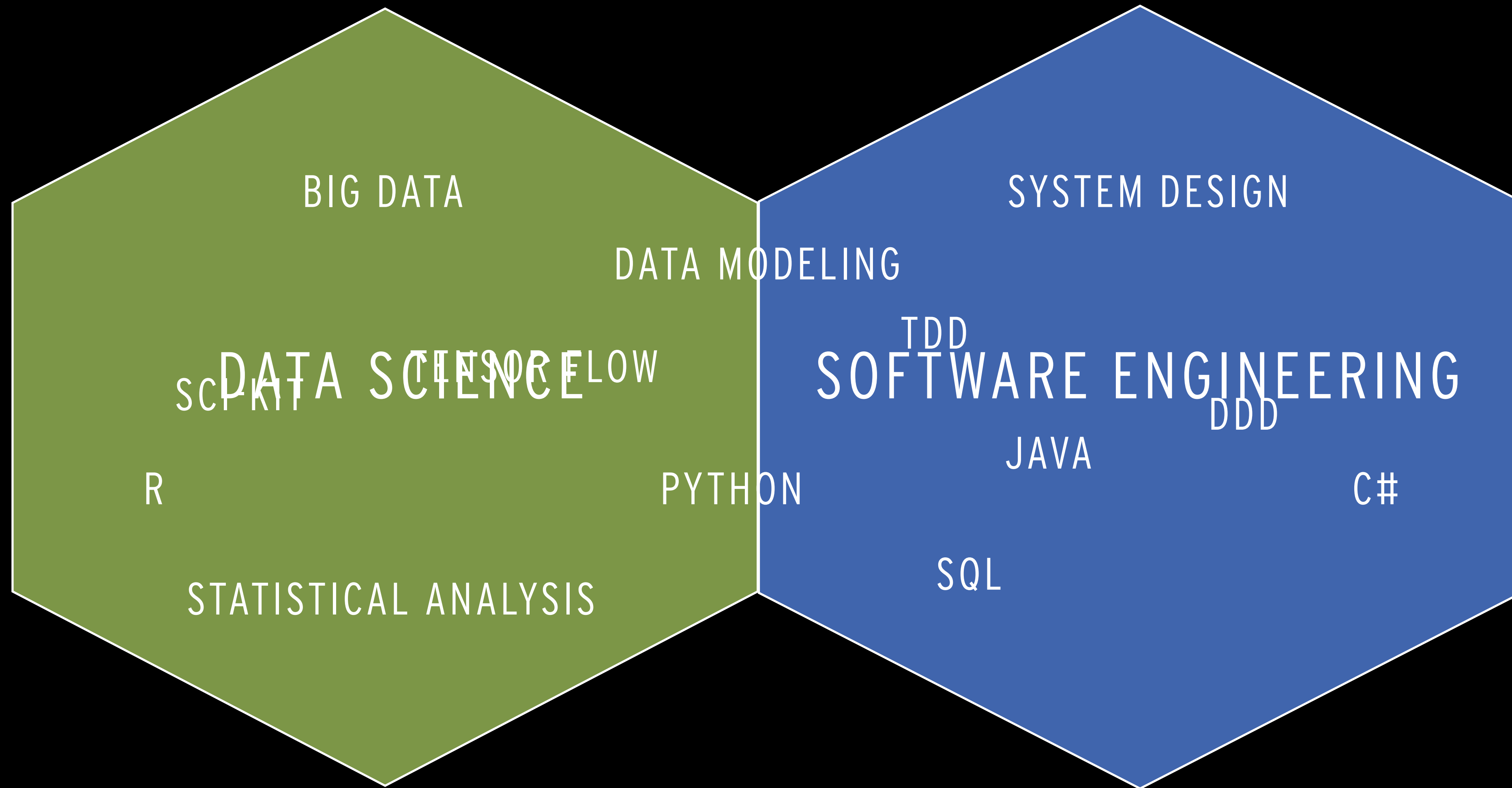
WHY?

WHY

- Some things are almost impossible to solve without, e.g.
 - Image recognition
- While other things get (a lot) better
 - Natural language processing
 - Recommendations
 - Robotic process automation
 - Anomaly detection
- Patterns are everywhere
 - Ask what your data can tell you?



CHALLENGE: NEW SKILLS



DEEPLARNING4J

- Open source
- Includes various tools for ML
 - ND4J
 - DataVec
 - Arbiter
 - Some visualization tools
- Import Keras models
- Supports dataprocessing on CUDA* enabled GPUs (Nvidia)



*CUDA: COMPUTE UNIFIED DEVICE ARCHITECTURE

WHY JAVA

- Python is *by far* more common machine learning language, but...
 - Java is versatile with huge ecosystem of tools
 - Great number of systems are built in Java
 - Great number of software engineers use Java primarily



HOW?

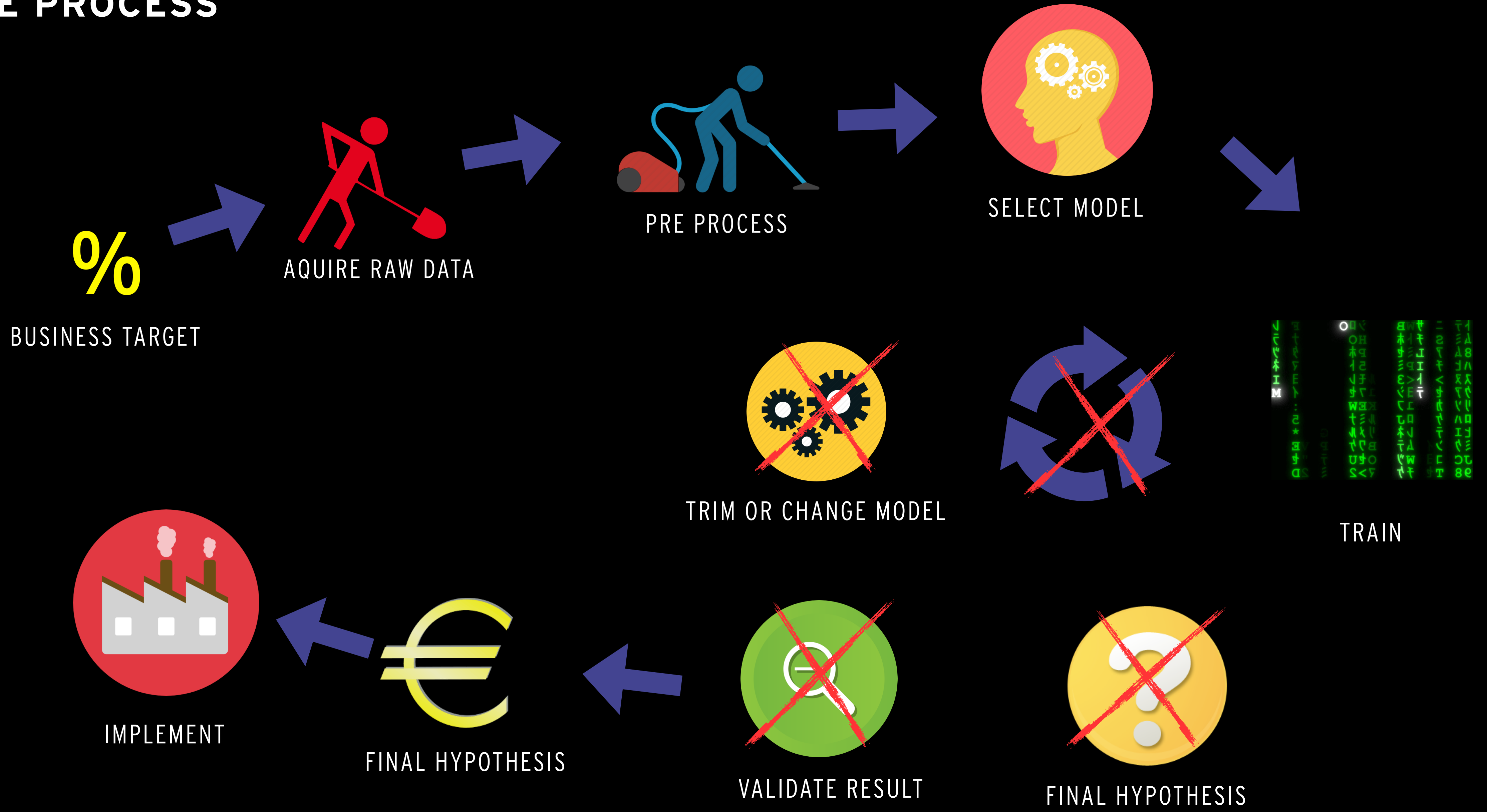
HOW BUILD - DEEPLARNING4J

```
dependencies {  
    compile("org.deeplearning4j:deeplearning4j-core:1.0.0-beta3")  
    compile("org.nd4j:nd4j-native-platform:1.0.0-beta3")  
}
```

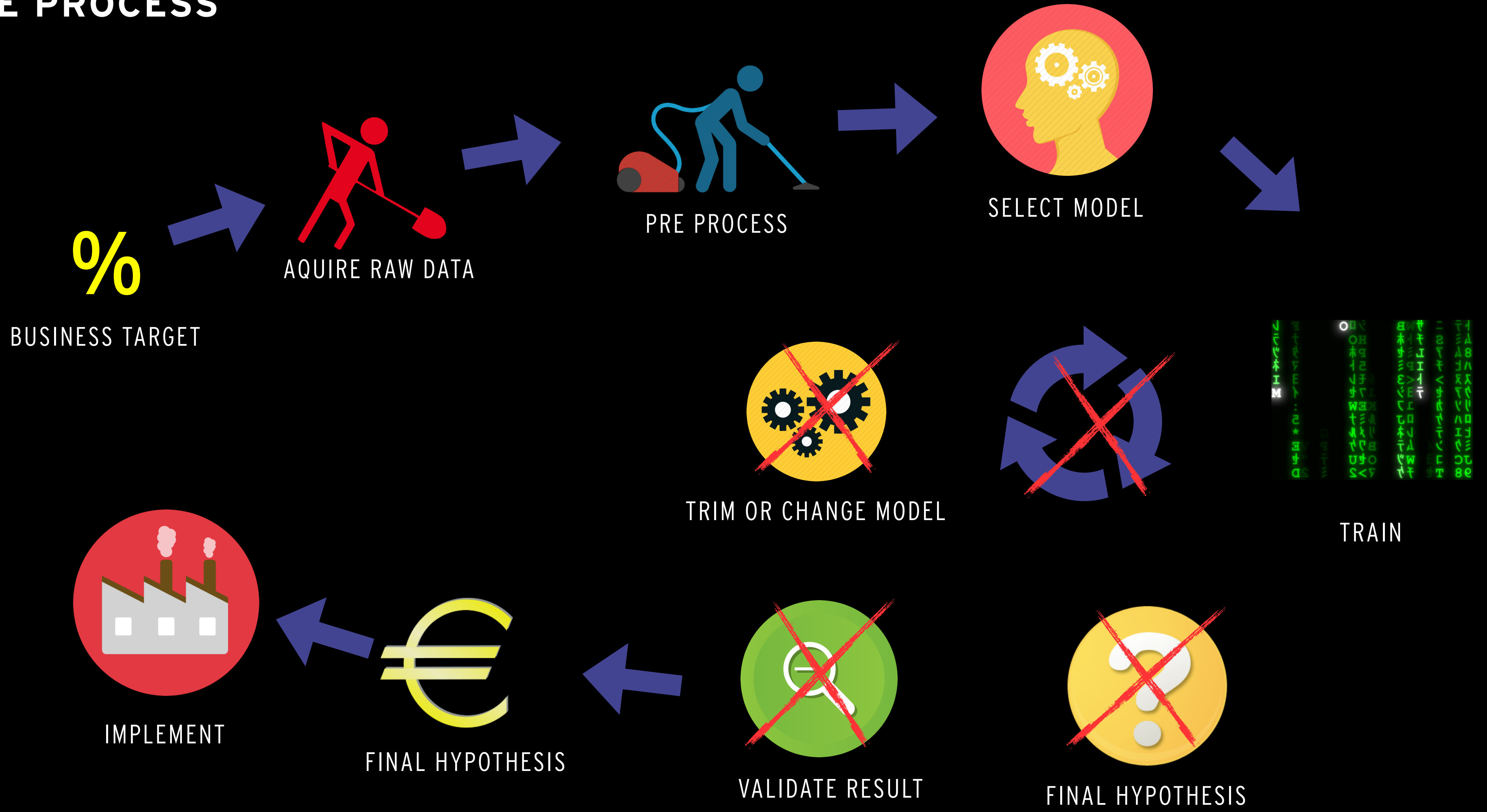
Alternatively:

```
dependencies {  
    compile("org.deeplearning4j:deeplearning4j-core:1.0.0-beta3")  
    compile("org.nd4j:nd4j-cuda-9.2:1.0.0-beta3")  
}
```

THE PROCESS



THE PROCESS



BUSINESS TARGET

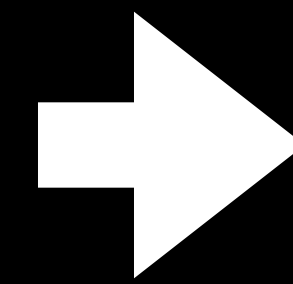
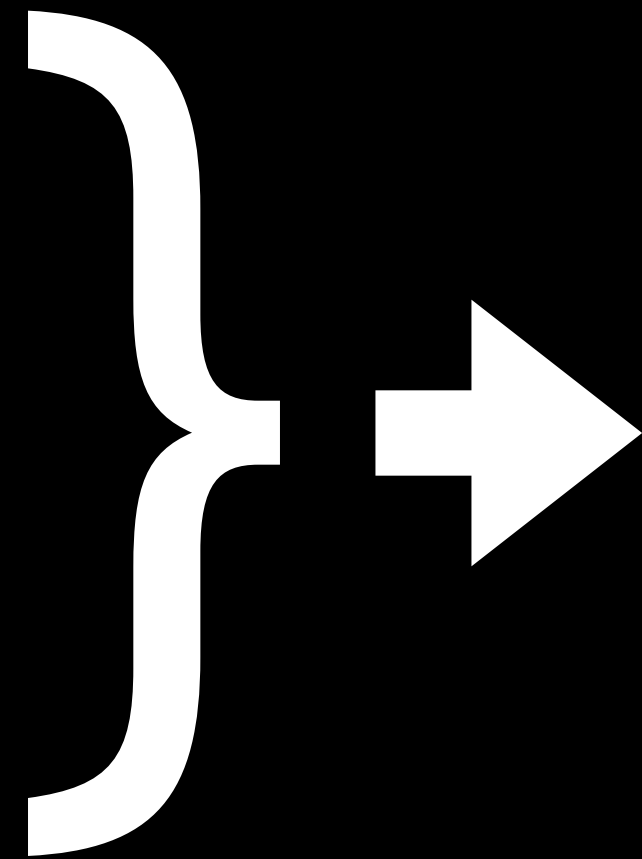
A

B

C

D

E



"A"

DATA

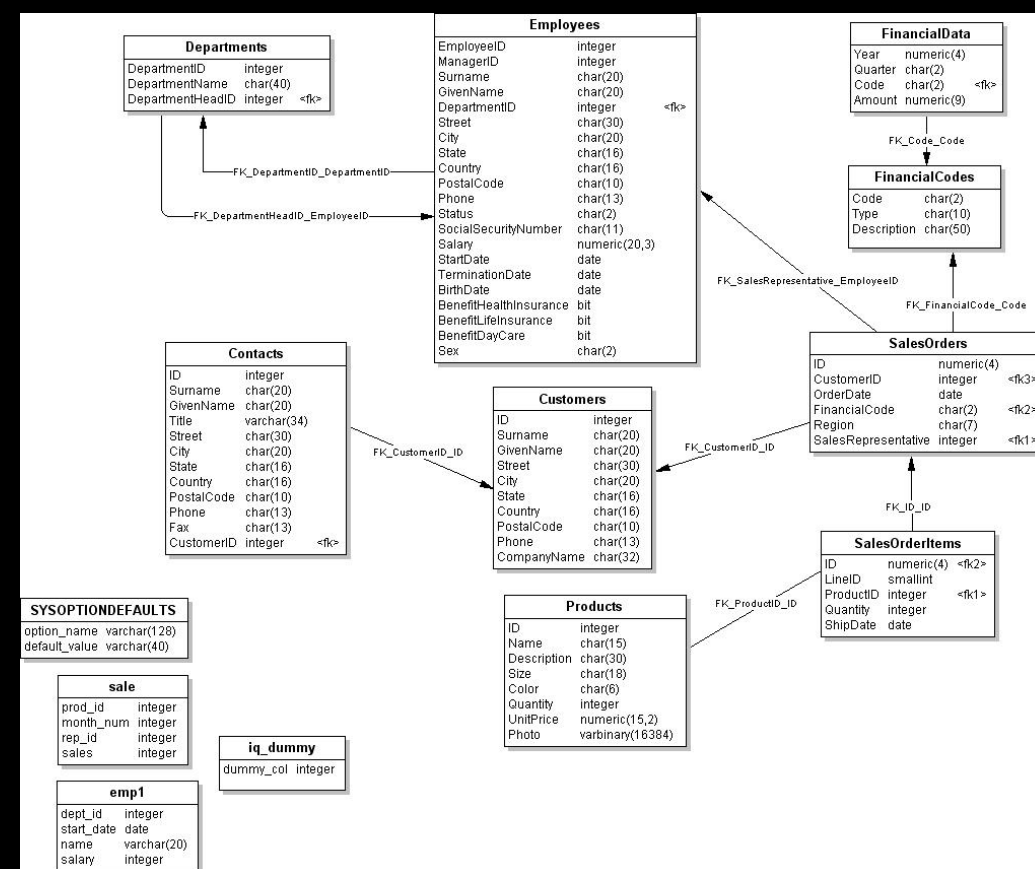
HOW - INPUT DATA



SOUND



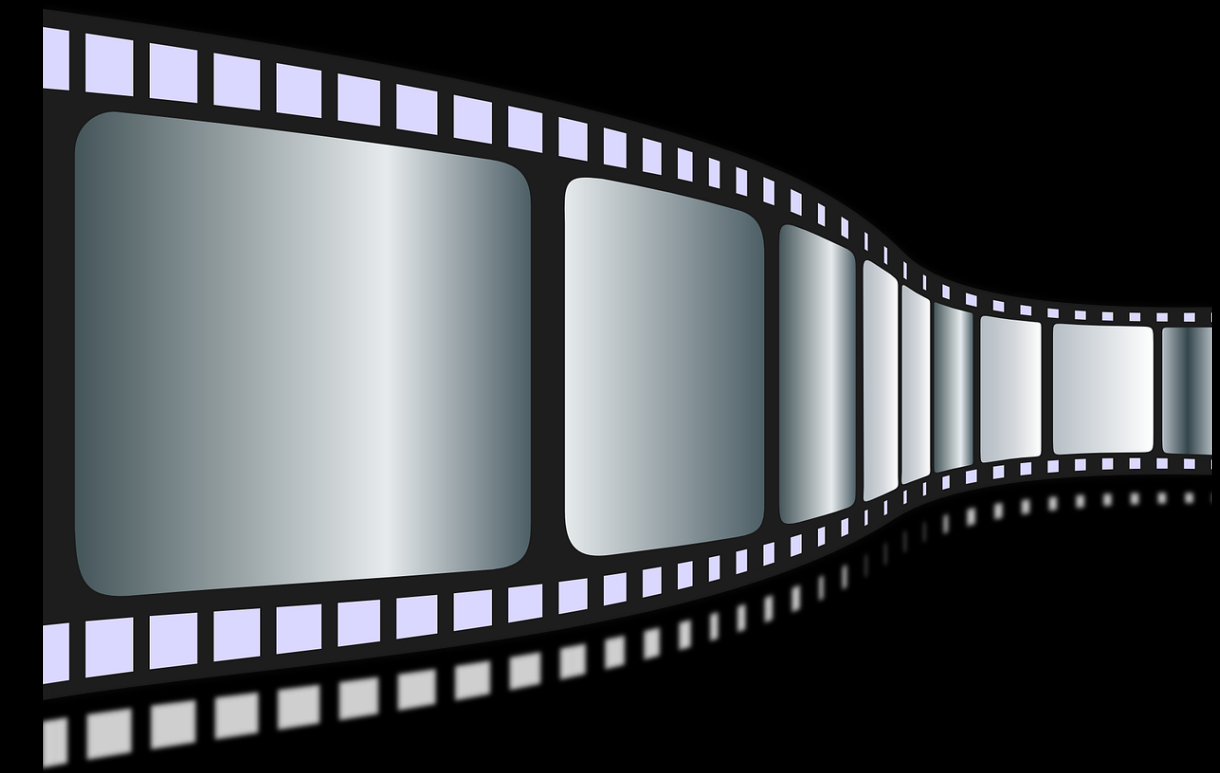
IMAGES



DATABASE/TABULAR DATA



DOCUMENTS/TEXT

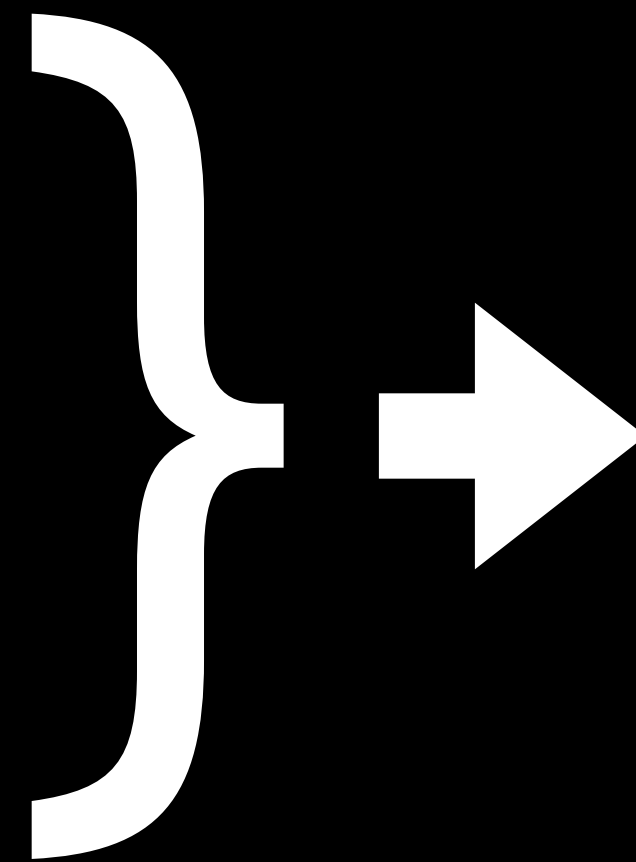


VIDEO

HOW - INPUT PREPROCESSING



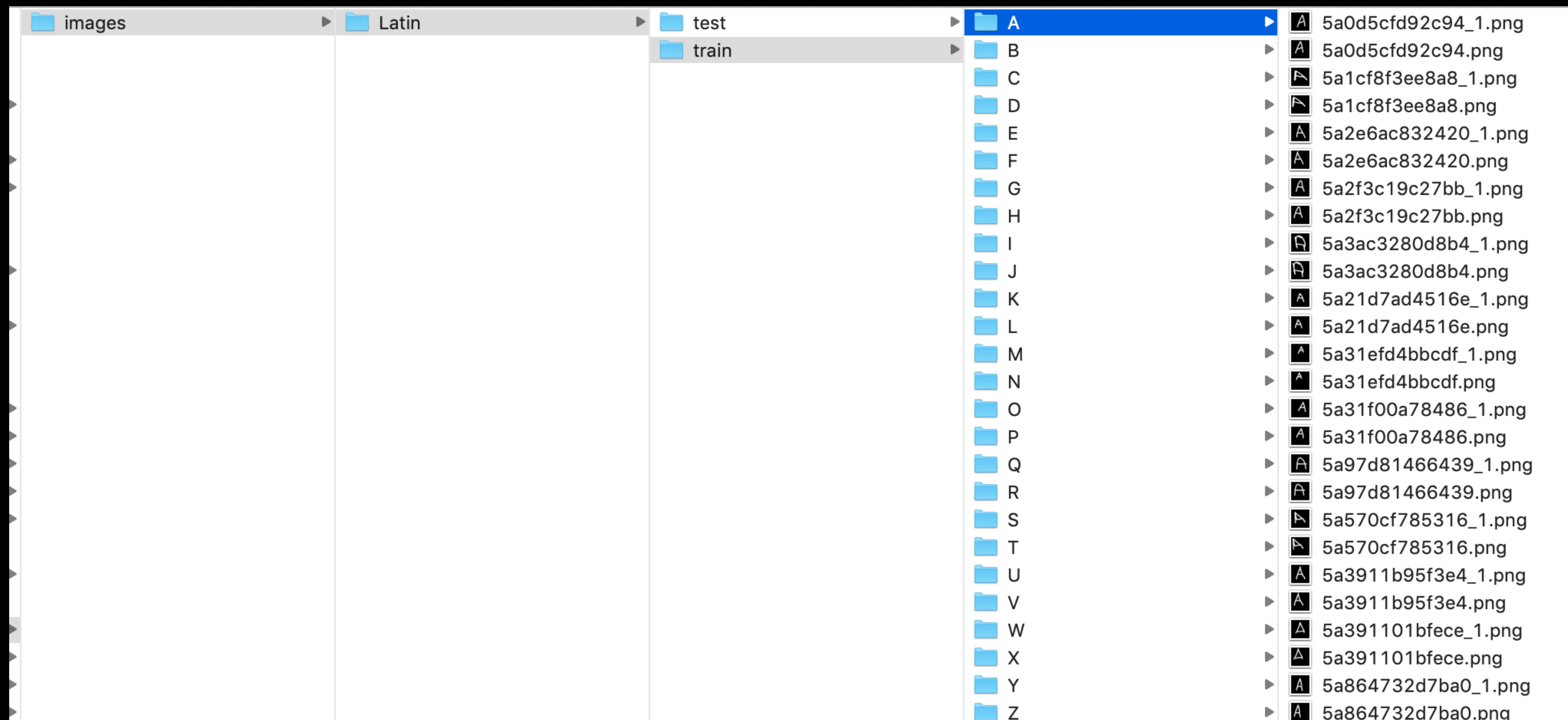
INPUTS



233, 123, 34...
212, 12, 43...
123, 243, 221...
2, 87, 123...
233, 123, 34...
233, 123, 34...
233, 123, 34...

HOW - INPUT PREPROCESSING

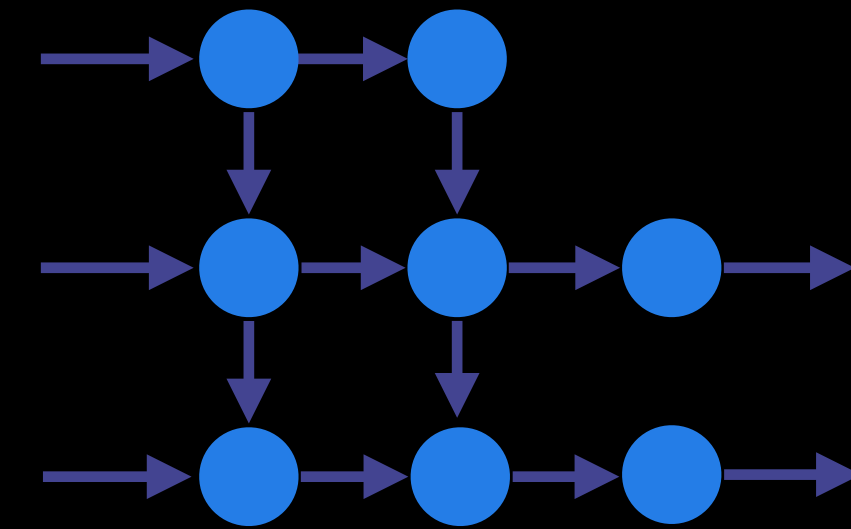
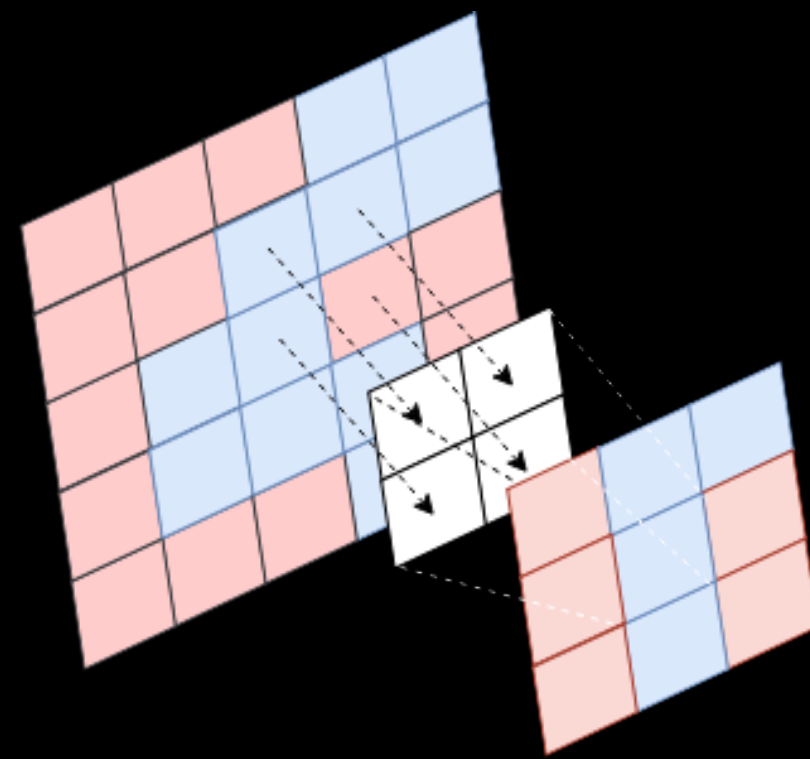
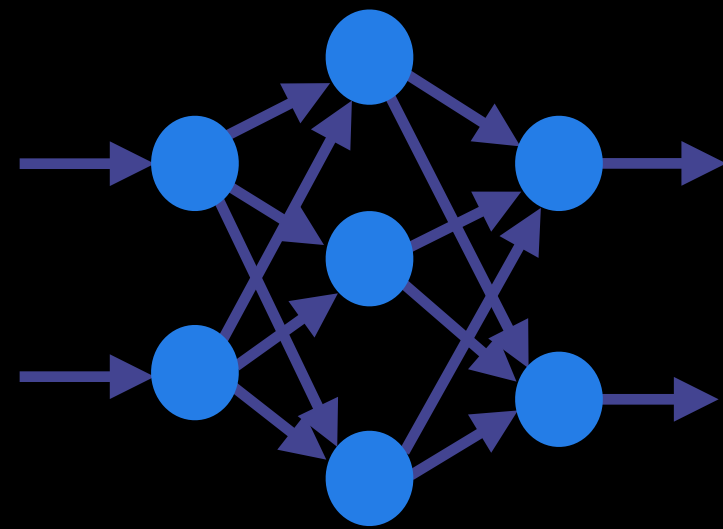
```
// ...  
FileSplit fileSplit = new FileSplit(directory, {".png"});  
ParentPathLabelGenerator labelMaker = new ParentPathLabelGenerator();  
ImageRecordReader recordReader = new ImageRecordReader(28, 28, 1, labelMaker);  
recordReader.initialize(fileSplit);  
// ...
```



MODEL

MODEL - ARTIFICIAL NEURAL NETWORK ARCHITECTURE

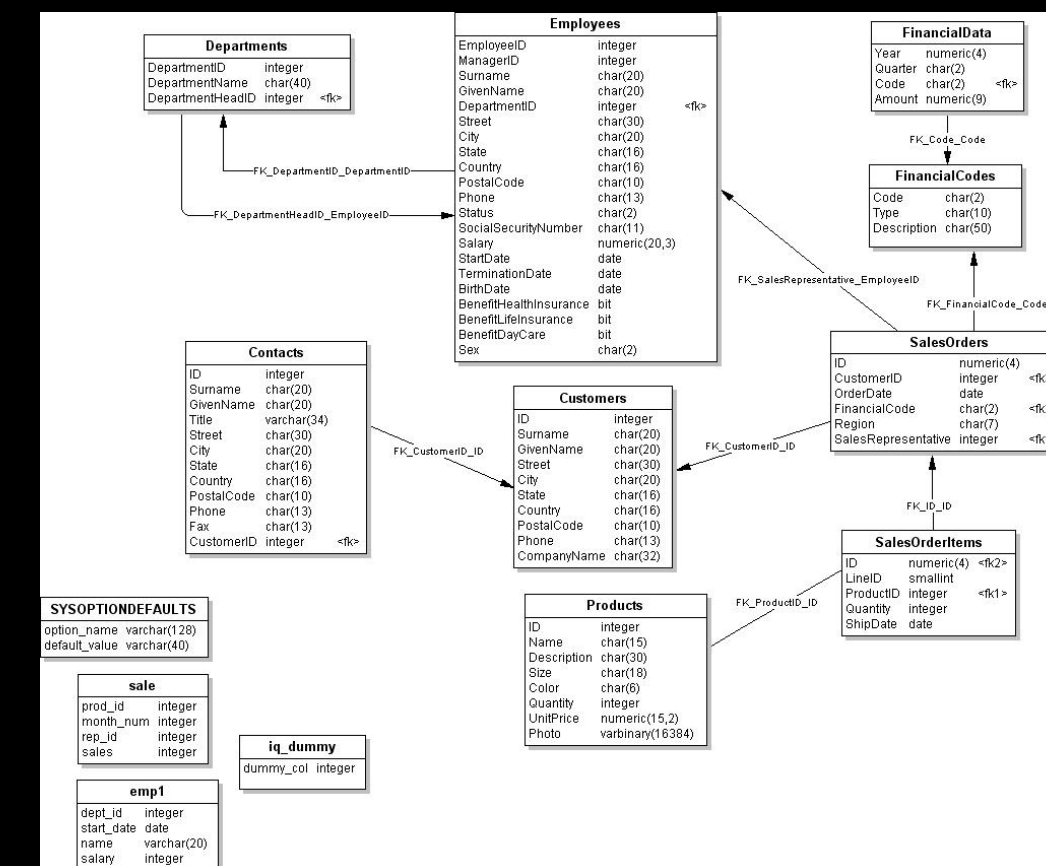
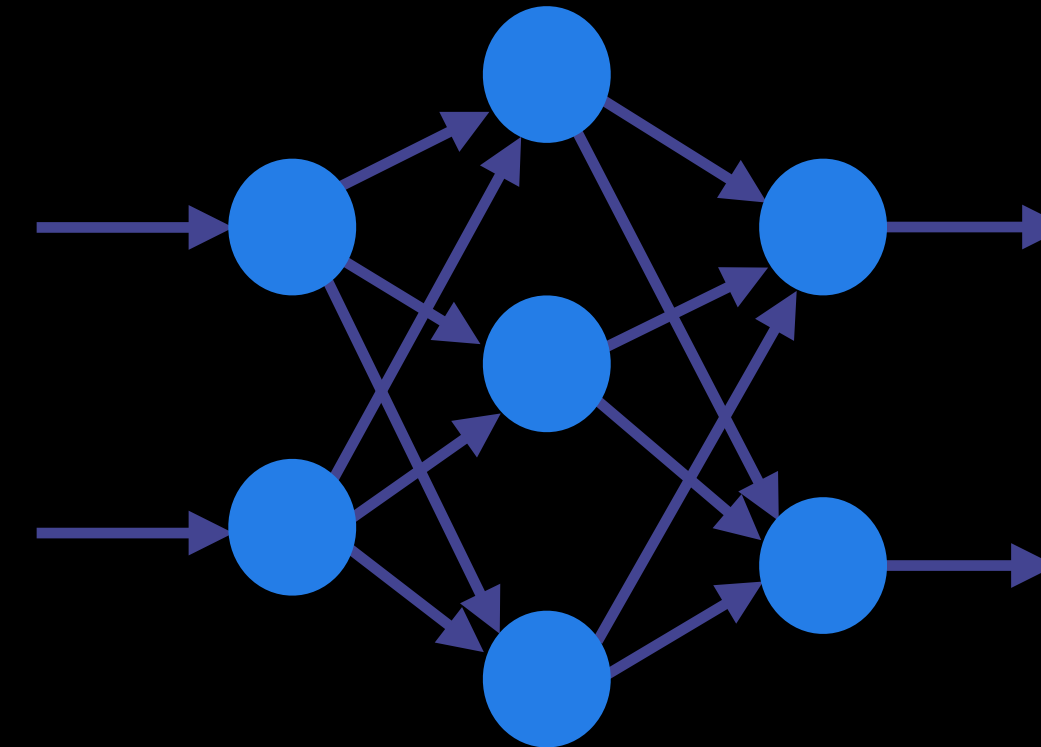
- Three common types:
 - MLP (Multi layer perceptron)
 - CNN (Convolutional neural network)
 - RNN (Recurrent neural network)
- Hybrid networks: use layers or subnets of different types



MODEL CHARACTERISTICS

DEEP LEARNING - MULTI LAYER PERCEPTRON

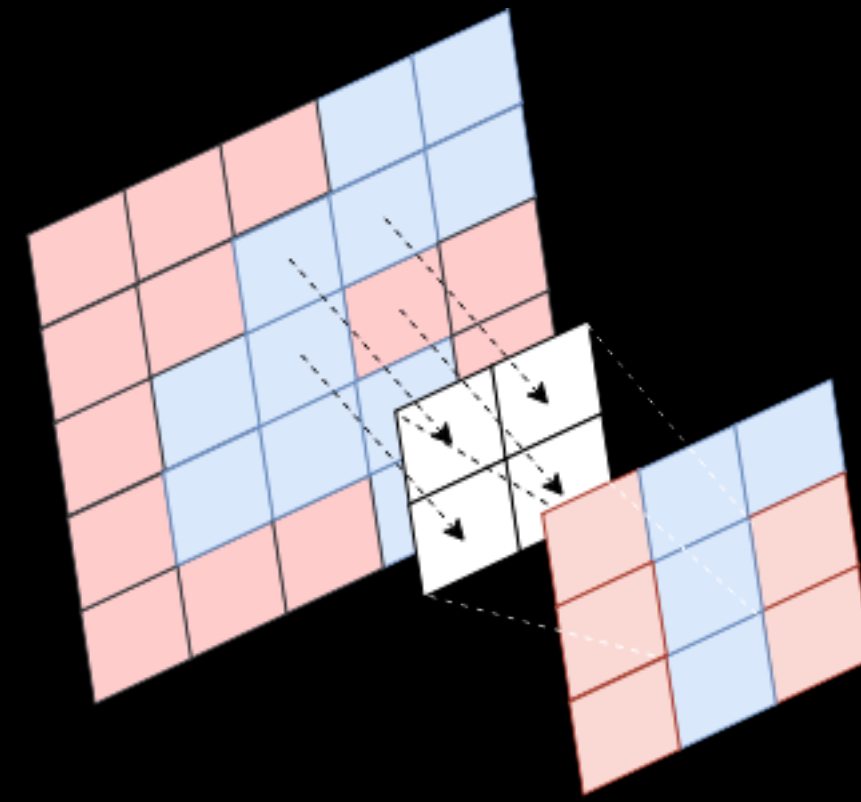
- General purpose architecture
- Particularly useful for tabular data, e.g. csv-files.



DATABASE/TABULAR DATA

DEEP LEARNING - CONVOLUTIONAL NEURAL NETWORK

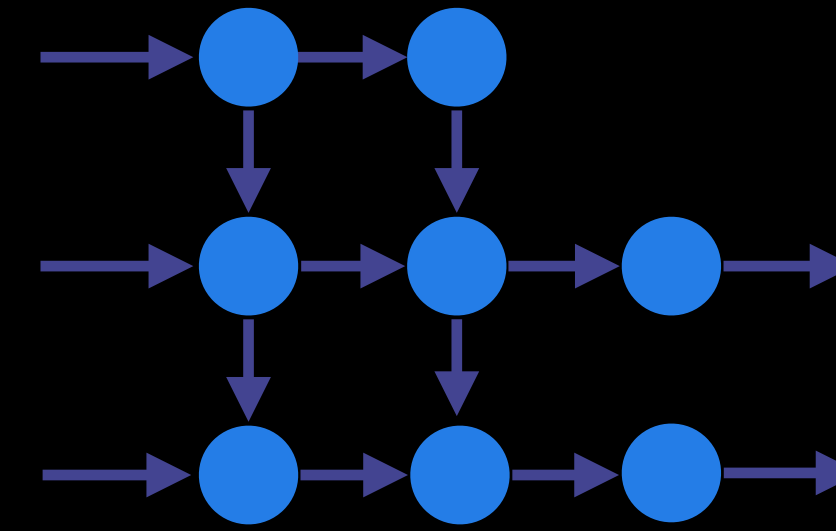
- Useful to make generalisations of the input (has/has not)
- Particularly useful for identifying patterns in images
- Not good for anomaly detection



IMAGES

DEEP LEARNING - RECURRENT NEURAL NETWORK

- Complex, often difficult to train
- LSTM (Long-short term memory) successful exception
- Useful for time-series such as sound or text



DOCUMENTS/TEXT

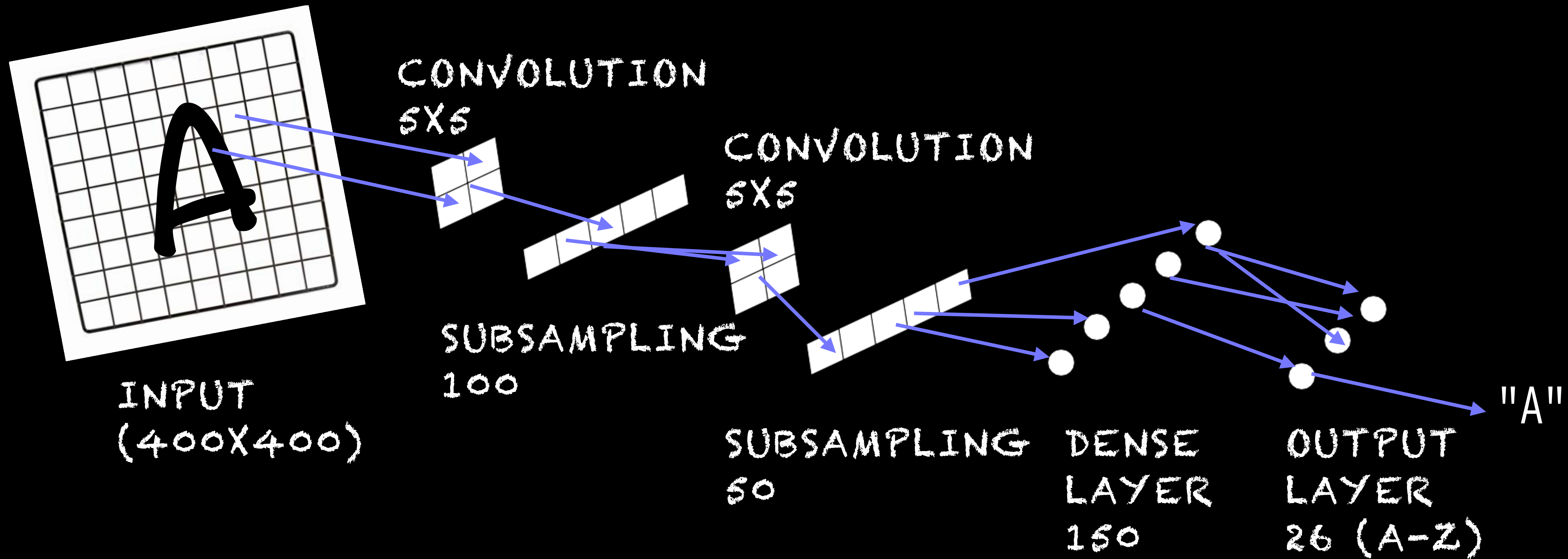


SOUND

HOW TO BUILD?

HOW BUILD - CONVOLUTIONAL NEURAL NETWORK

- We will build a convolutional neural network using Deeplearning4J



HOW BUILD - CONVOLUTIONAL NEURAL NETWORK

- We will build a convolutional neural network using Deeplearning4J
- Use Deeplearning4J MultilayerConfiguration to build a MultiLayerNetwork

HOW TO BUILD - MULTILAYERCONFIGURATION

```
// ...
MultiLayerConfiguration conf = new NeuralNetConfiguration.Builder()
    .seed(132);
    .optimizationAlgo(STOCHASTIC_GRADIENT_DESCENT)
    .weightInit(XAVIER)
    .updater(new Nesterov(learningRate, momentum))
    .list()
    .layer(0, new ConvolutionLayer.Builder(5, 5)
        .nIn(1).nOut(100).activation(IDENTITY).build())
// ... More layers here
    .layer(5, new OutputLayer.Builder()
        .nOut(26).activation(SOFTMAX).build())
    .backprop(true)
    .build();
// ...
```

HOW TO BUILD - CONVOLUTIONAL NEURAL NETWORK

- We will build a convolutional neural network using Deeplearning4J
- Use Deeplearning4J MultilayerConfiguration to build a MultiLayerNetwork
- Use Deeplearning4J EarlyStoppingTrainer to train and save the network

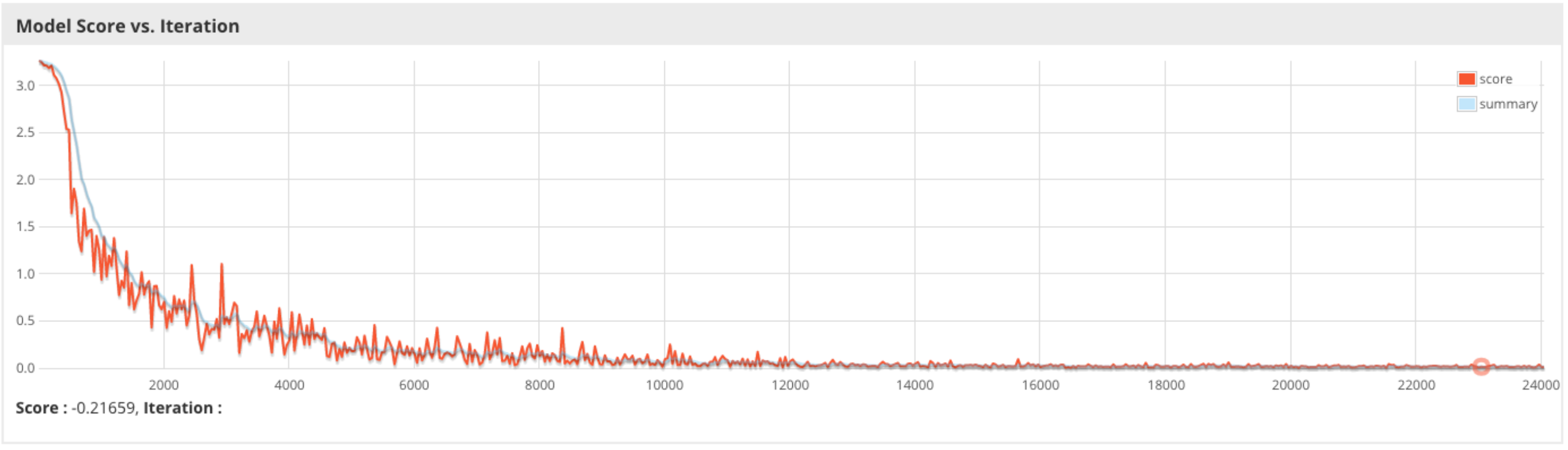
HOW TO BUILD - EARLYSTOPPINGTRAINER

```
private static EarlyStoppingTrainer trainer(DataSetIterator iter,
MultiLayerNetwork model) {
    EarlyStoppingConfiguration conf = new EarlyStoppingConfiguration.Builder()
        .epochTerminationConditions(new MaxEpochsTerminationCondition(30));
        .iterationTerminationConditions(new MaxTimeIterationTerminationCondition
(2, HOURS))
        .scoreCalculator(new ClassificationScoreCalculator(ACCURACY, iter))
        .evaluateEveryNEpochs(1)
        .modelSaver(new LocalFileModelSaver(modelSaveDirectory))
        .build();
    return new EarlyStoppingTrainer(conf, model, iter);
}
```


MovieTime

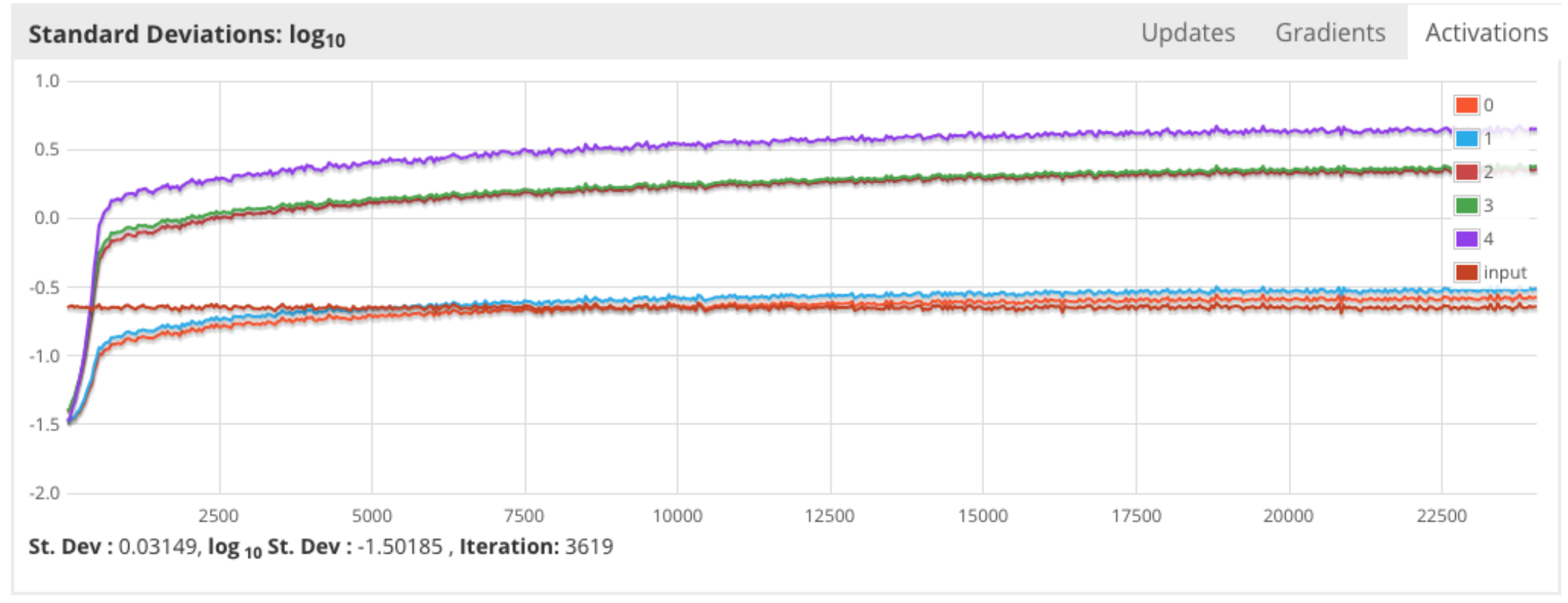
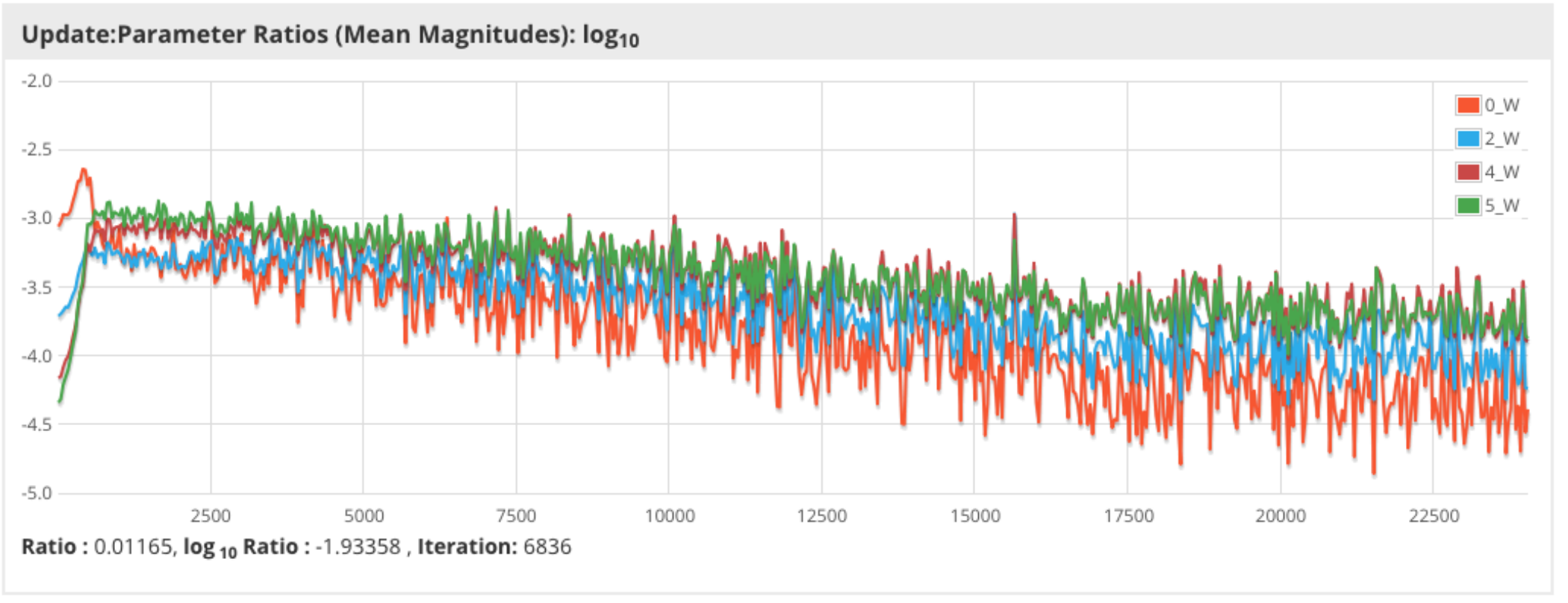


- Overview
- Model
- System
- Language



Model and Training Information

Model Type	MultiLayerNetwork
Layers	6
Total Parameters	499226
Start Time	
Total Runtime	
Last Update	2019-01-06 14:20:48
Total Parameter Updates	24051
Updates/sec	5,70
Examples/sec	182,54



HOW TO USE?

DEMO

ALL THE ANSWERS

- What is it?
 - *Self-adapting* algorithms to identify *patterns*
- Why machine learning?
 - *Patterns are everywhere*, you probably have patterns in your data, how can you use that?
- Why use Java for machine learning?
 - Java has a *great ecosystem* of tools and frameworks
 - Sheer force of numbers (systems and developers)
- How can we use Java for machine learning?
 - I think DL4J solves many of the problems related to working with ML...

| SOME LINKS

- <https://deeplearning4j.org/>
- <https://skymind.ai/wiki/>
- <https://archive.ics.uci.edu/ml/datasets.html>
- <https://www.analyticsvidhya.com/blog/>
- <https://machinelearningmastery.com/blog/>

THANK YOU!

ALL CODE USED IN DEMOS: [HTTPS://GITHUB.COM/CALLISTAENTERPRISE/CADEC-2019-DEEPLARNING4J.GIT](https://github.com/callistaenterprise/caDEC-2019-deeplearning4j.git)