BIG DATA AND CO

A glimpse on possibilities in Accounting

WHO AM I?

Assistant Professor of Accounting at Rady

- PhD in Finance - Dauphine University, France
- Master in Finance - Dauphine University, France
- Master in AI - Toulouse University, France
- Master in CS Engineering - INSA, France

RESEARCH

BEFORE

- AI - Reinforcement Learning

RECENTLY

- Banking Regulation
- Impact of data on financing decisions

NOW

- Real effects implications of accounting rules
- Public disclosures and Information Processing

INTERNATIONAL EXPERIENCE

Hong Kong University

Harvard

Imperial College Business School

CONTACT

mmilone@ucsd.edu

858-534-8511

Office 3W102

FINANCIAL ACCOUNTING

FINANCIAL ACCOUNTING

Aggregation of information into financial statements

FINANCIAL ACCOUNTING

Aggregation of information into financial statements

Provides information about past transactions

MANAGERIAL ACCOUNTING

MANAGERIAL ACCOUNTING

Produce information for management decisions

MANAGERIAL ACCOUNTING

Produce information for management decisions

Focuses on the future

MANAGERIAL ACCOUNTING

Produce information for management decisions

Needs a lot of Data

Focuses on the future

MANAGERIAL ACCOUNTING

Produce information for management decisions

Needs a lot of Data

Focuses on the future

Need for prediction

ON THE NEED FOR DATA

ON THE NEED FOR DATA

Cost estimations

ON THE NEED FOR DATA

Cost estimations

Cost allocations

ON THE NEED FOR DATA

Cost estimations

Cost allocations

Breakeven Point

ON THE NEED FOR DATA

Cost estimations

Cost allocations

Breakeven Point

Sensitivity Analysis

ON THE NEED FOR DATA

Cost estimations

Cost allocations

Breakeven Point

Sensitivity Analysis

Transfer Pricing

ON THE NEED FOR PREDICTION

Forecast of Cash Flows

Risk Management

...

Cost Analysis

Fraud detection

...

ON THE NEED FOR PREDICTION

ABOUT THE FUTURE

Forecast of Cash Flows

Risk Management

...

Cost Analysis

Fraud detection

...

ON THE NEED FOR PREDICTION

ABOUT THE FUTURE

Forecast of Cash Flows

Risk Management

...

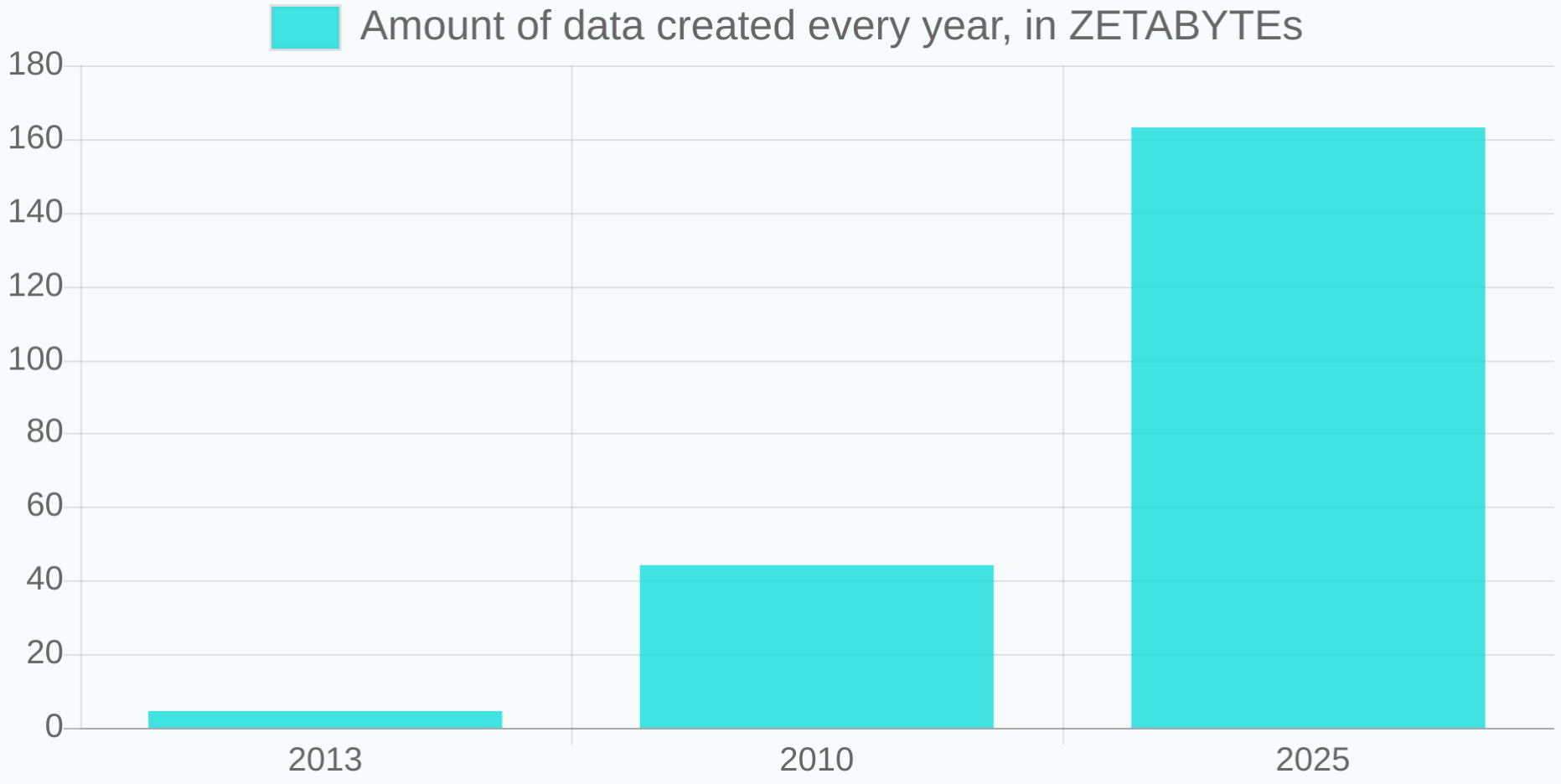
ABOUT THINGS WE DO NOT KNOW

Cost Analysis

Fraud detection

...

DATA IS COMING ...



1 ZB = 1 Billion Terabytes

BIG DATA

ARE COMPANIES THERE YET?

WHAT IS BIG DATA?

WHAT IS BIG DATA?

Just a Big Buzzword?

WHAT IS BIG DATA?

Just a Big Buzzword?

Just a lot more data?

WHAT IS BIG DATA?

Just a Big Buzzword?

Just a lot more data?

The combination of previously separate datasets?

WHAT IS BIG DATA?

Just a Big Buzzword?

Just a lot more data?

The combination of previously separate datasets?

Structure vs unstructure data?

WHAT ARE COMPANIES DOING TODAY?

Microsoft BI and co

Slowly Starting

WHAT ARE COMPANIES DOING TODAY?

BUSINESS ANALYTICS

Microsoft BI and co

Slowly Starting

WHAT ARE COMPANIES DOING TODAY?

BUSINESS ANALYTICS

Microsoft BI and co

PREDICTIVE ANALYTICS

Slowly Starting

RPA

= ROBOT PROCESSING AUTOMATION

RPA

= ROBOT PROCESSING AUTOMATION

Automation of routine rules-based business processes

RPA

= ROBOT PROCESSING AUTOMATION

Automation of routine rules-based business processes

Embedded into Enterprise Resource Planning (ERP) softwares

RPA

= ROBOT PROCESSING AUTOMATION

Automation of routine rules-based business processes

Embedded into Enterprise Resource Planning (ERP) softwares

UiPath, Blue Prism, Automation Anywhere

RPA

EXAMPLES

Invoice processing

Data validation

Customer Relationship

HR processes (payroll, ...)

RPA

EXAMPLES

Invoice processing

Data validation

Customer Relationship

HR processes (payroll, ...)

RPA

WHY DOES IT MATTER?

RPA

WHY DOES IT MATTER?

Digitization of processes

RPA

WHY DOES IT MATTER?

Digitization of processes

Enhances the creation and gathering of data

RPA

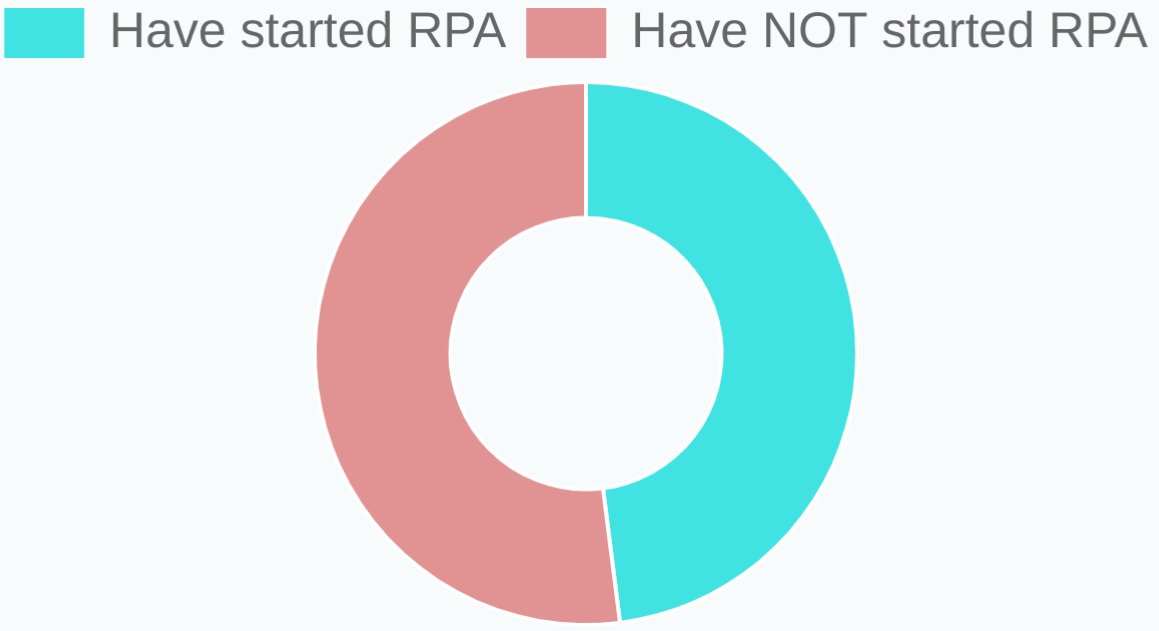
WHY DOES IT MATTER?

Digitization of processes

Enhances the creation and gathering of data

First step towards *Intelligent Automation*

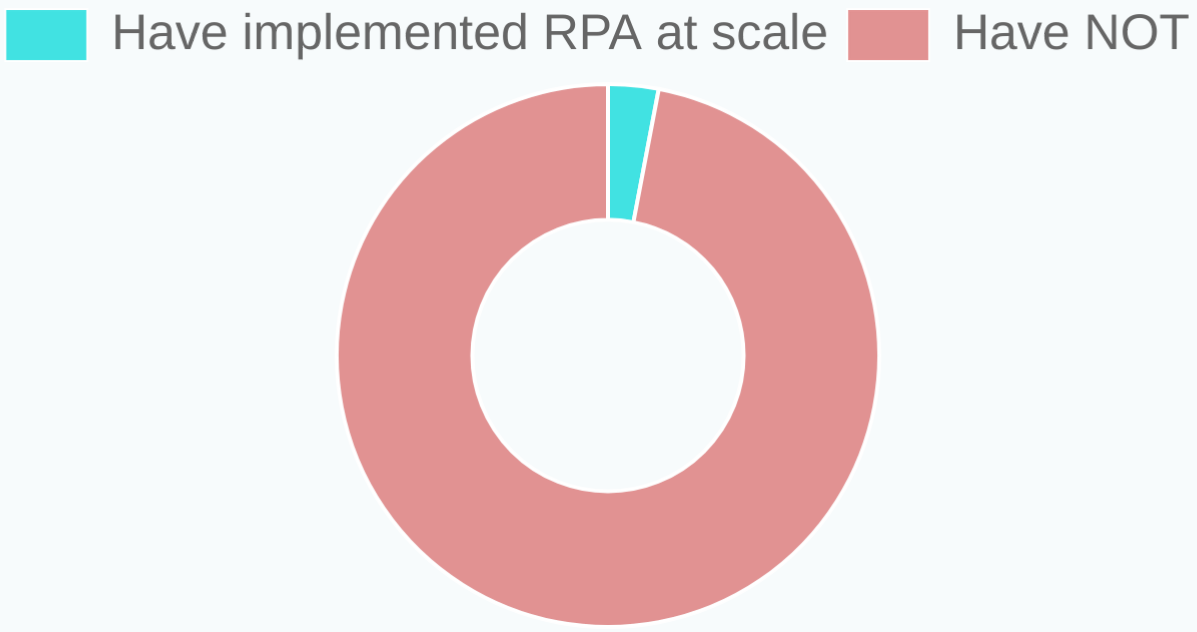
RPA



< 50% of companies have started RPA

Data from EY, 2018

RPA



3% of enterprises have achieved scale

Data from EY, 2018

MACHINE LEARNING

AND HOW IT AFFECTS ACCOUNTING

HOW DO WE SOLVE PROBLEMS?

HOW DO WE SOLVE PROBLEMS?

Need to find an *Algorithm*

HOW DO WE SOLVE PROBLEMS?

Need to find an *Algorithm*

= Steps to solve a problem

HOW ABOUT ML?

Supervised / Unsupervised Learning

Reinforcement Learning

HOW ABOUT ML?

We want the machine to find the algorithm by itself

Supervised / Unsupervised Learning

Reinforcement Learning

HOW ABOUT ML?

We want the machine to find the algorithm by itself

HOW DO WE DO THAT?

Supervised / Unsupervised Learning

Reinforcement Learning

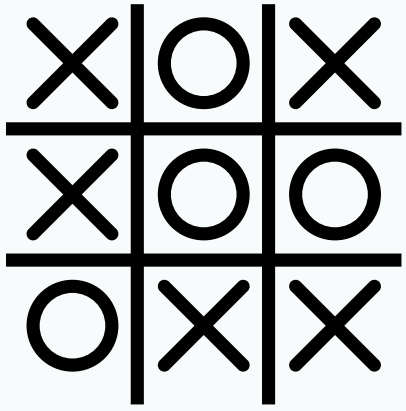
HOW ABOUT ML?

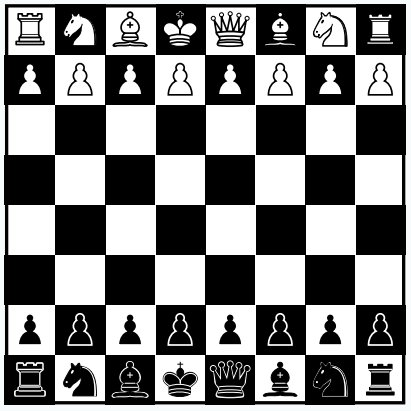
We want the machine to find the algorithm by itself

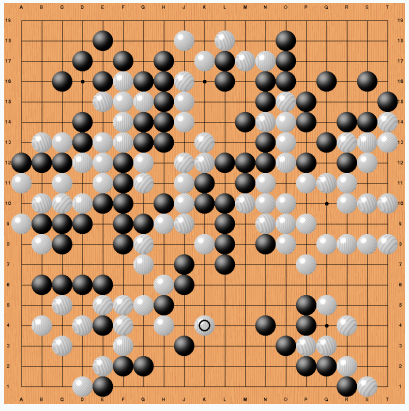
HOW DO WE DO THAT?

Supervised / Unsupervised Learning

Reinforcement Learning







HOW DOES THE MACHINE LEARN?

Input: board, image

Output: move, number

Highly non-linear math function

HOW DOES THE MACHINE LEARN?

NEED TO MAP AN INPUT TO AN OUTPUT

Input: board, image

Output: move, number

Highly non-linear math function

HOW DOES THE MACHINE LEARN?

NEED TO MAP AN INPUT TO AN OUTPUT

Input: board, image

Output: move, number

HOW DO WE MAP THESE?!

Highly non-linear math function

HIGHLY WHAT?!

HIGHLY WHAT?!

POSSIBLE METHODS

HIGHLY WHAT?!

POSSIBLE METHODS

Regression / Classification Trees

HIGHLY WHAT?!

POSSIBLE METHODS

Regression / Classification Trees

Random Forests

HIGHLY WHAT?!

POSSIBLE METHODS

Regression / Classification Trees

Random Forests

K-nearest neighbor

HIGHLY WHAT?!

POSSIBLE METHODS

Regression / Classification Trees

Random Forests

K-nearest neighbor

Neural Networks

HIGHLY WHAT?!

POSSIBLE METHODS

Regression / Classification Trees

Random Forests

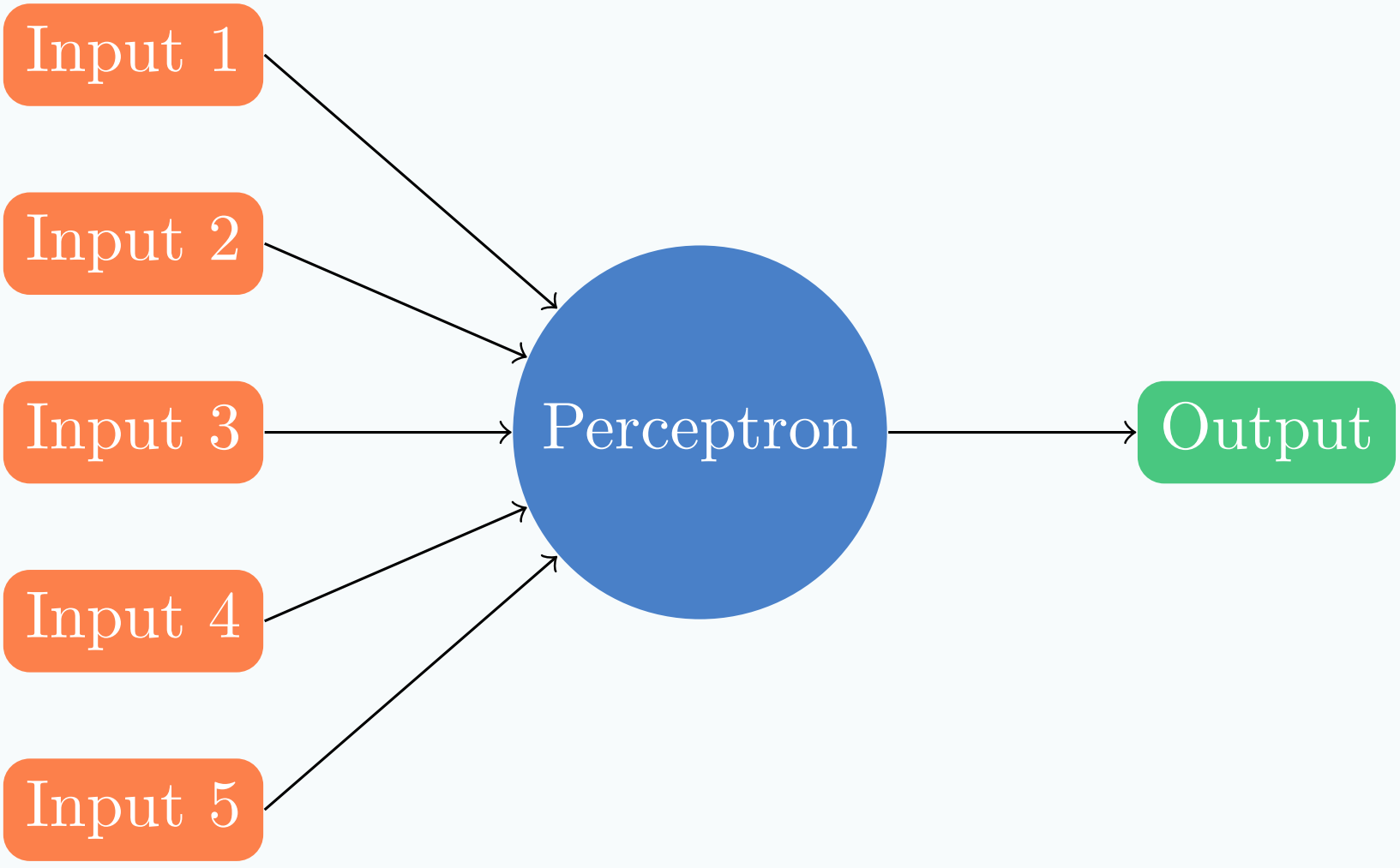
K-nearest neighbor

Neural Networks

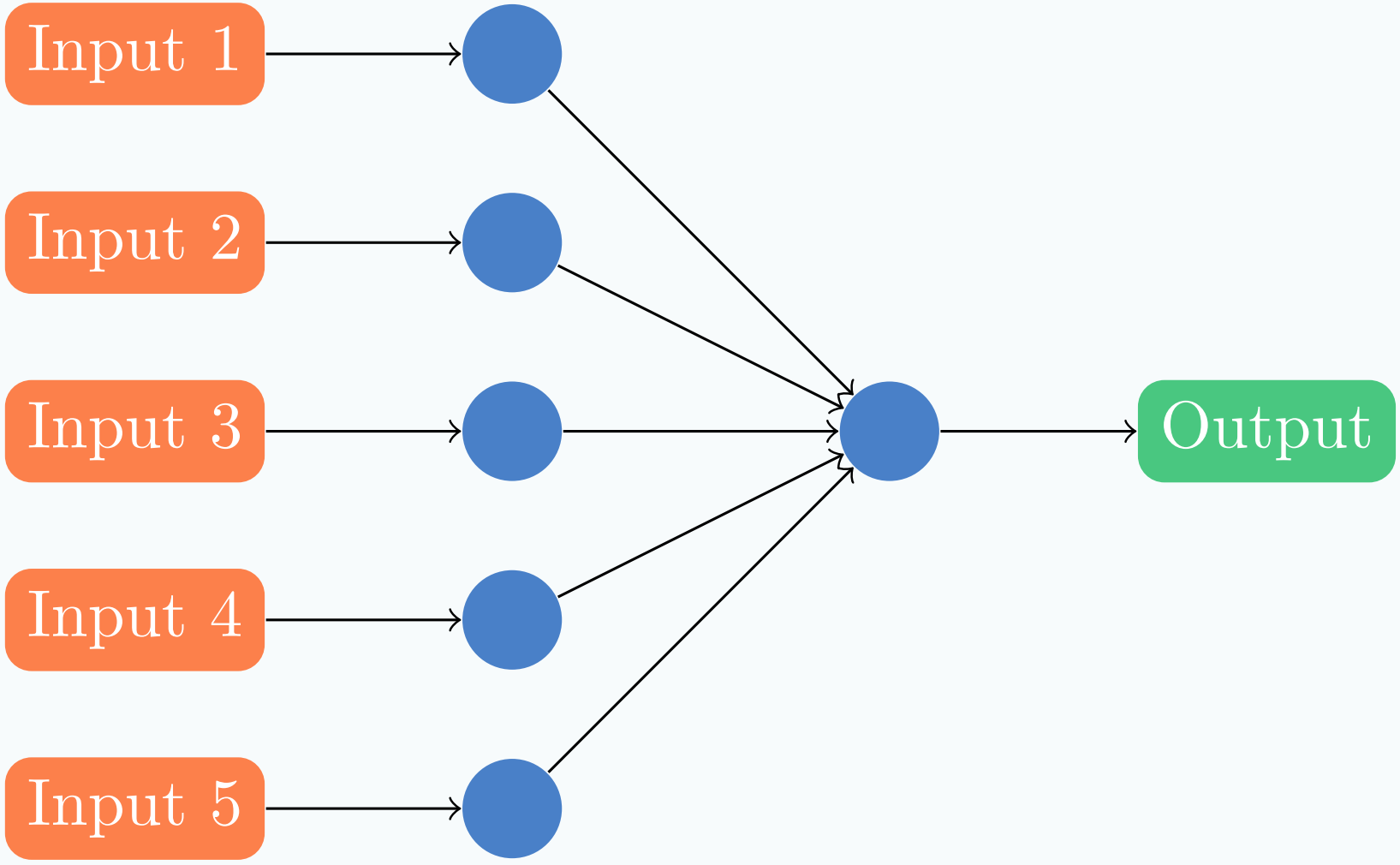
...

NEURAL NETWORKS

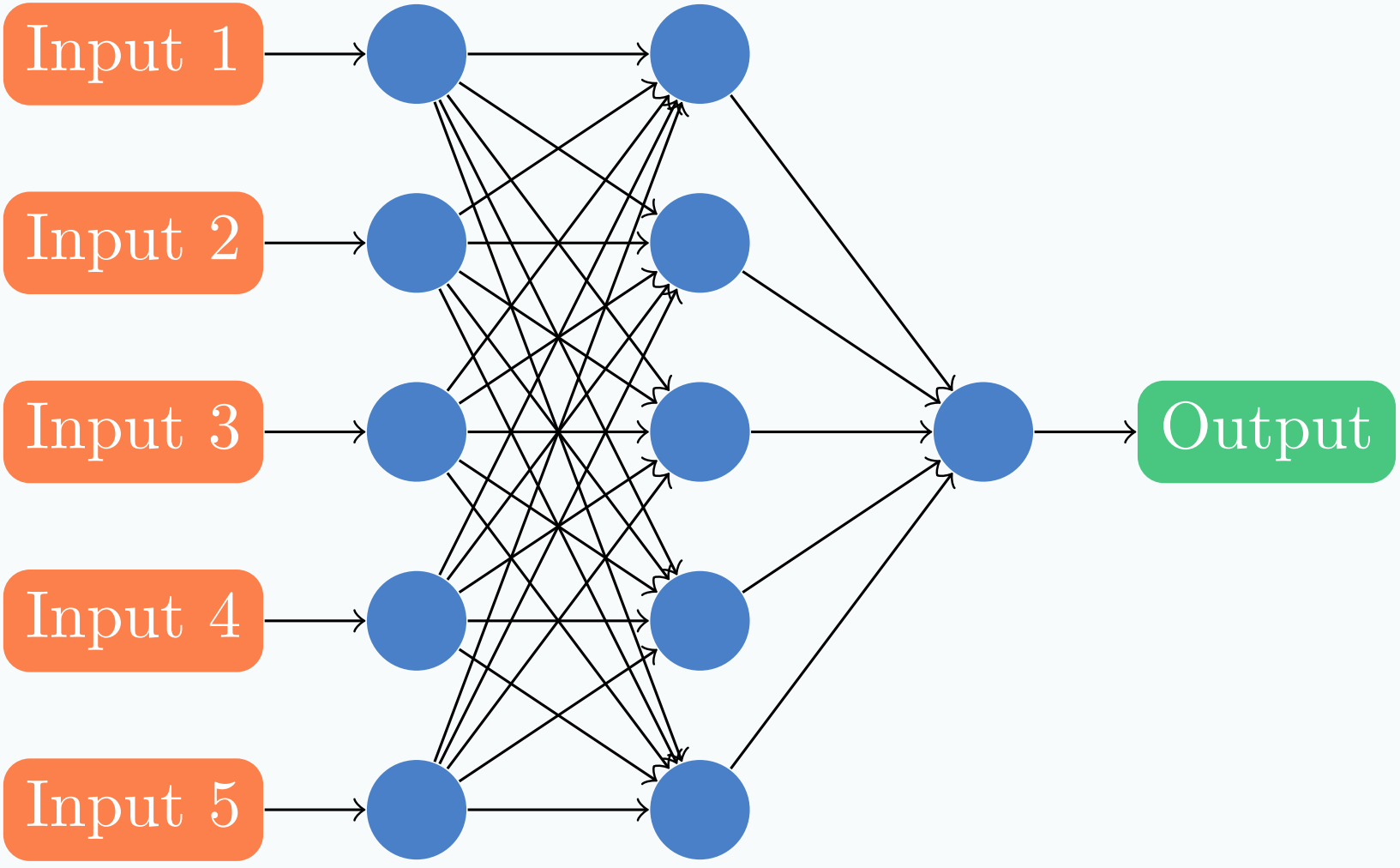
NEURAL NETWORKS



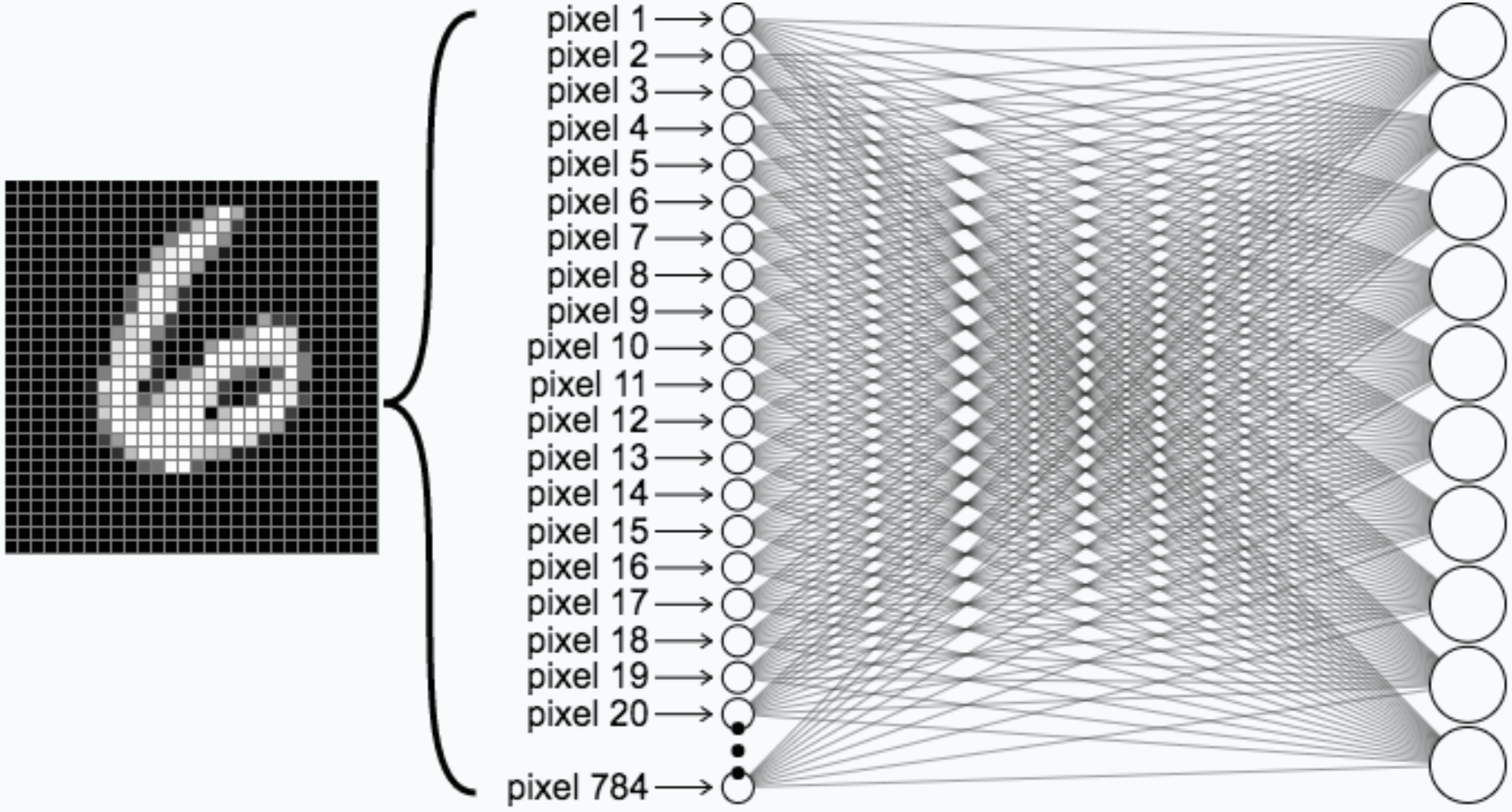
NEURAL NETWORKS



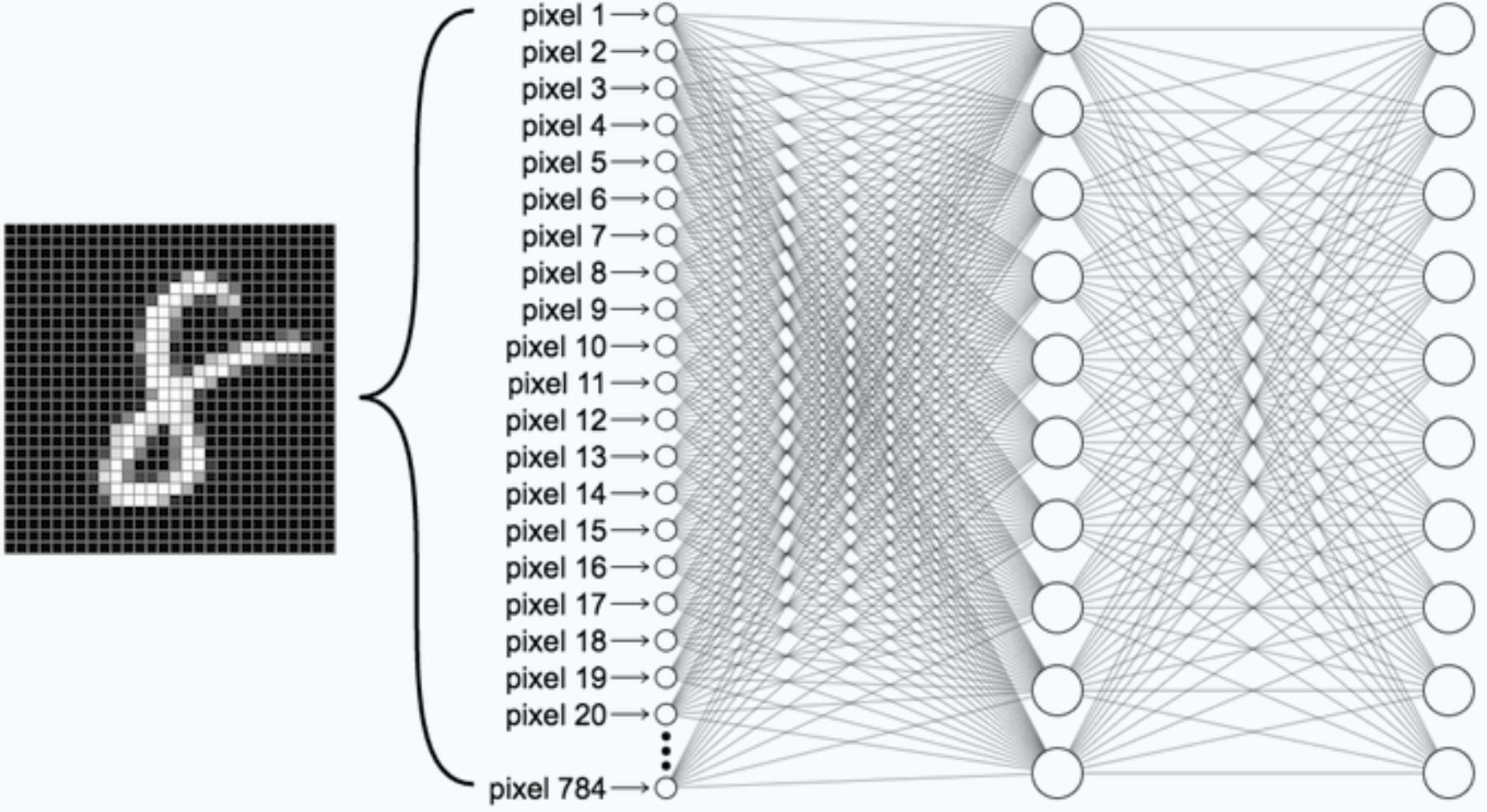
NEURAL NETWORKS



NEURAL NETWORKS

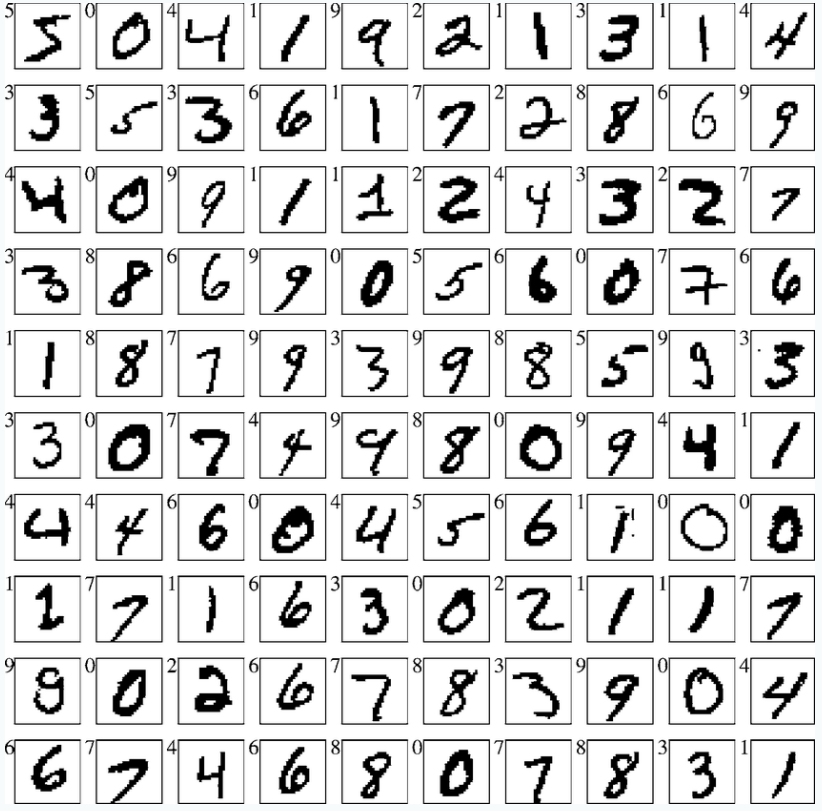


NEURAL NETWORKS

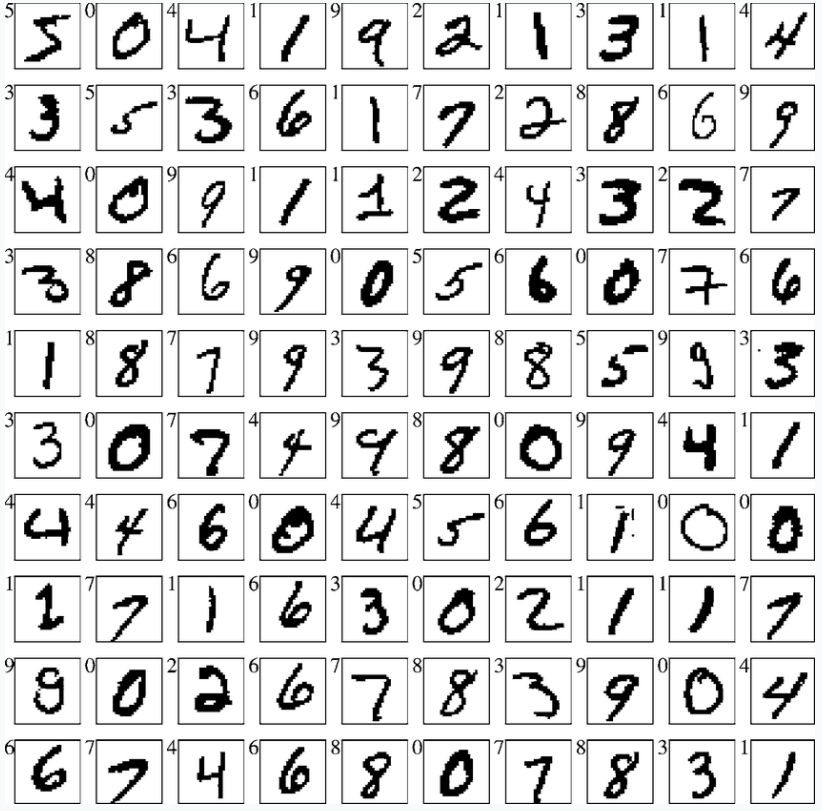


SURE, BUT HOW DOES IT LEARN?!

SURE, BUT HOW DOES IT LEARN?!



SURE, BUT HOW DOES IT LEARN?!



Supervised Learning

WHY DO WE CARE?

Real life problems are highly non-linear

Difficult to find rule based algorithms

EXAMPLES

Image recognition

Translations

Self-Driving

Natural Language Processing

...

MACHINE LEARNING IN ACCOUNTING

Error detections

High risk transactions detection

Transaction allocations

Fraud detections

MACHINE LEARNING IN ACCOUNTING

Error detections

High risk transactions detection

Transaction allocations

Fraud detections

MACHINE LEARNING IN PUBLIC ACCOUNTING

MACHINE LEARNING IN PUBLIC ACCOUNTING

ML requires a lot of data to learn

MACHINE LEARNING IN PUBLIC ACCOUNTING

ML requires a lot of data to learn

Audit firms observe many firms over time

MACHINE LEARNING IN PUBLIC ACCOUNTING

ML requires a lot of data to learn

Audit firms observe many firms over time

They observe errors in many different situations

MACHINE LEARNING IN PUBLIC ACCOUNTING

ML requires a lot of data to learn

Audit firms observe many firms over time

They observe errors in many different situations

Increased ability to use ML methods

RPA + ML = INTELLIGENT AUTOMATION

THIS IS AWESOME!

THIS IS AWESOME!

NOT REALLY...

Overfitting problems

Regularization methods

Techniques allow to tackle complex problems

But we do not yet have enough ressources

COMPUTERS ARE "TOO" SMART ...

Overfitting problems

Regularization methods

Techniques allow to tackle complex problems

But we do not yet have enough ressources

COMPUTERS ARE "TOO" SMART ...

Overfitting problems

Regularization methods

OR NOT GOOD ENOUGH...

Techniques allow to tackle complex problems

But we do not yet have enough ressources

COMPUTERS ARE "TOO" SMART ...

Overfitting problems

Regularization methods

OR NOT GOOD ENOUGH...

Techniques allow to tackle complex problems

But we do not yet have enough ressources

BASICALLY, IT IS HARD TO DESIGN ...

MACHINE LEARNING PITFALLS (1)

MACHINE LEARNING PITFALLS (1)

ML (supervised) requires a lot of data to learn

MACHINE LEARNING PITFALLS (2)

MACHINE LEARNING PITFALLS (2)

Can only learn information from the given data

MACHINE LEARNING PITFALLS (2)

Can only learn information from the given data

Garbage In = Garbage Out

MACHINE LEARNING PITFALLS (2)

Can only learn information from the given data

Garbage In = Garbage Out

Cannot adapt to unknown situations

MACHINE LEARNING PITFALLS (3)

MACHINE LEARNING PITFALLS (3)

Need to have a way to improve ...

MACHINE LEARNING PITFALLS (3)

Need to have a way to improve ...

... still need humans

IS THAT ALL?

Accounting is about recording transactions

BLOCKCHAIN

BLOCKCHAIN

Created to solve the double spending problem without trusted third party

TRADITIONALLY

Third party records all transactions

When asked for, can say if a new transaction is possible (i.e. pay)

All parties involved trust the third party

BLOCKCHAIN IN ACCOUNTING

Like Accounting, Blockchain is about registering transactions

It seems likely to disrupt accounting

BLOCKCHAIN IN ACCOUNTING

Audit is about verifying transactions

Blockchain provides a way to trust the information

Implementations started in China

SHOULD WE BE WORRIED?

NOT REALLY ...

Accounting involves a lot of reasoning

REMEMBER

REMEMBER

1. Identifying Transactions

REMEMBER

1. Identifying Transactions

2. Recording / Measuring -- Needs concepts (GAPP)

REMEMBER

1. Identifying Transactions

2. Recording / Measuring -- Needs concepts (GAPP)

3. Communicating (BS, OE, IS, SCF)

MAYBE ONE DAY

MAYBE ONE DAY

BLOCKCHAIN + ML

MAYBE ONE DAY

BLOCKCHAIN + ML

But we have time

THANK YOU