Direct Fit to Nature: An Evolutionary Perspective on Biological and Artificial Neural Networks

Paper by Uri Hasson, Samuel A. Nastase, Ariel Goldstein
Paper Overview by Taavi Kivisik

University of Tartu
2020-11-12

What to Expect From a Model?

Models in Neuroscience

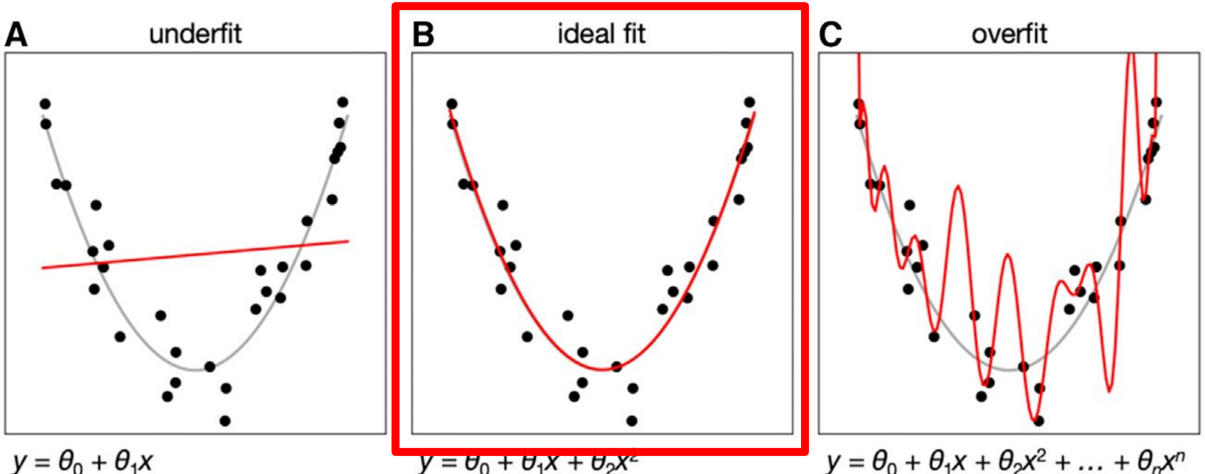
Value (and assume brain does it)

- Interpretability
- Generalization (novel contexts)
- Aim for **ideal fit**

Models in Neuroscience

Value (and assume brain does it)

- Interpretability
- Generalization (novel contexts)
- Aim for **ideal fit**



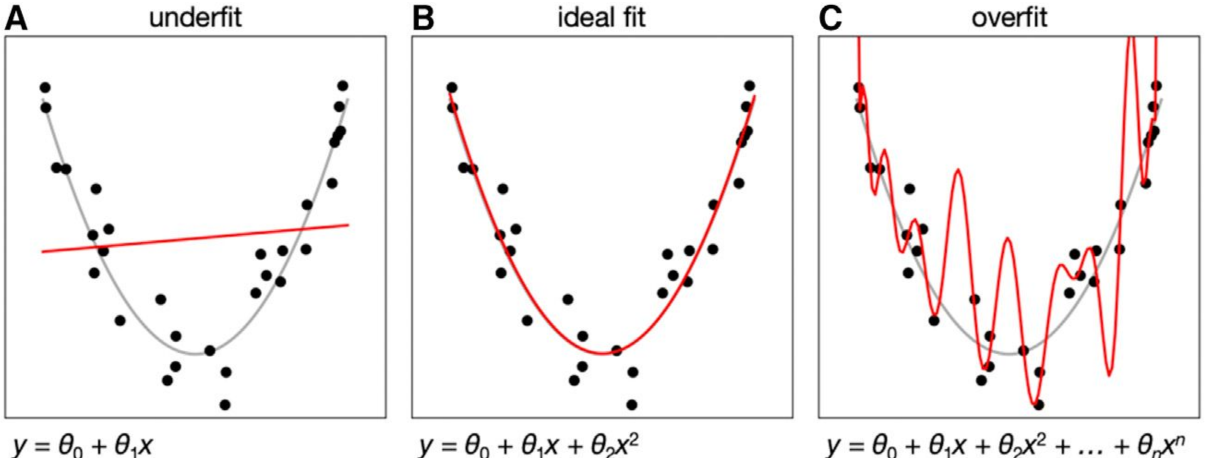
Models in Neuroscience **VS** Machine Learning

Value

- Interpretability
- Generalization (novel contexts)
- Aim for **ideal fit**

Value

- Behavior (task **performance**)
- Embrace complexity
- Aim for **direct fit**



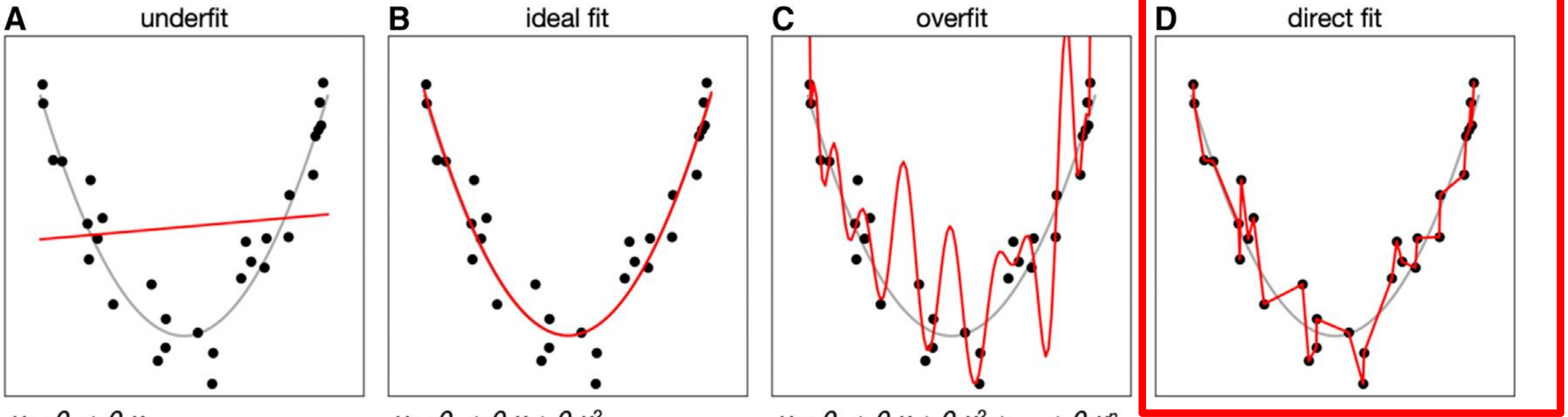
Models in
Neuroscience VS Machine Learning

Value

- Interpretability
- Generalization (novel contexts)
- Aim for **ideal fit**

Value

- Behavior (task **performance**)
- Embrace complexity
- Aim for **direct fit**



$y = \theta_0 + \theta_1 x$

$y = \theta_0 + \theta_1 x + \theta_2 x^2$

$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \ldots + \theta_n x^n$

Models in
Neuroscience VS Machine L

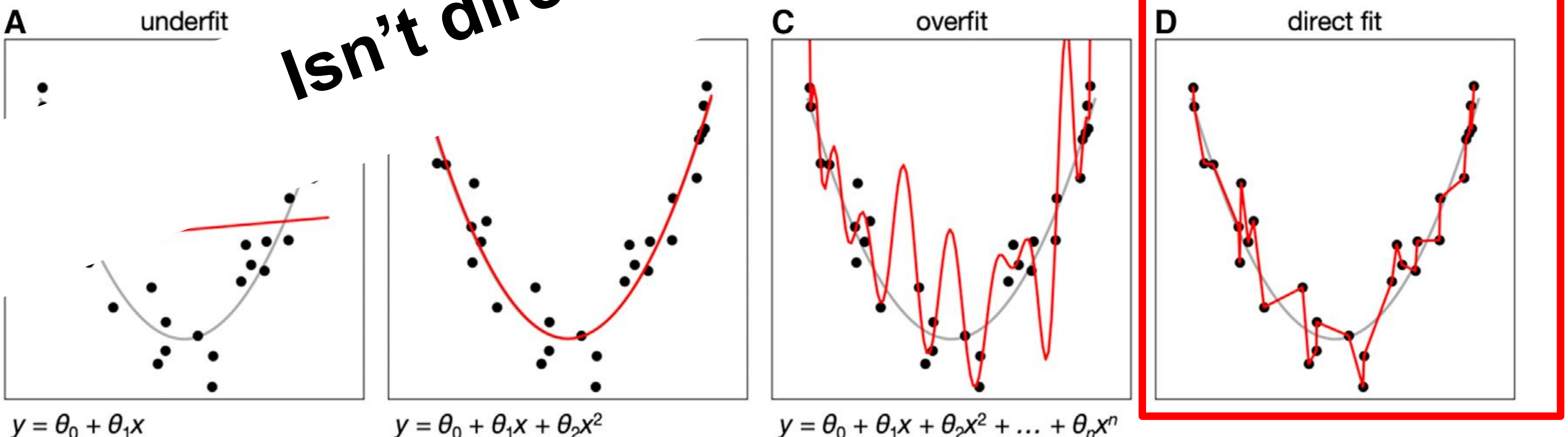
Value

- Interpretability
- Generalization (novel context
- Aim for **ideal fit**

Value

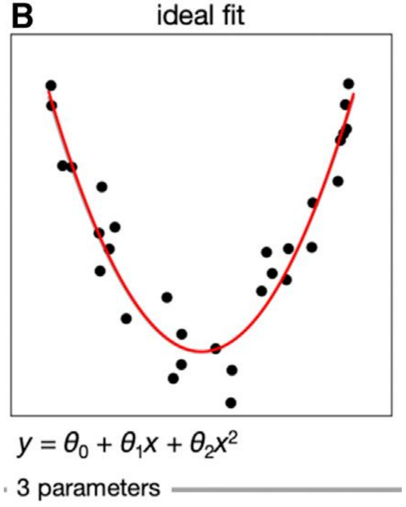
- **performance**)
- ce complexity
- Aim for **direct fit**

Isn't direct fit just overfitting?



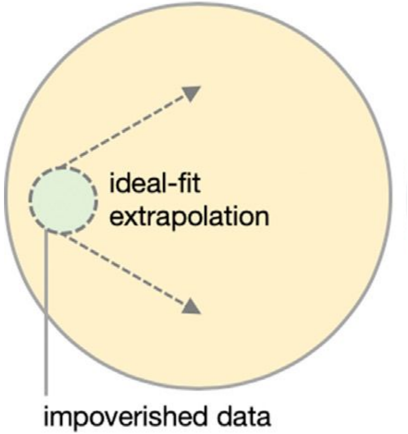
Two Forms of Generalization

B



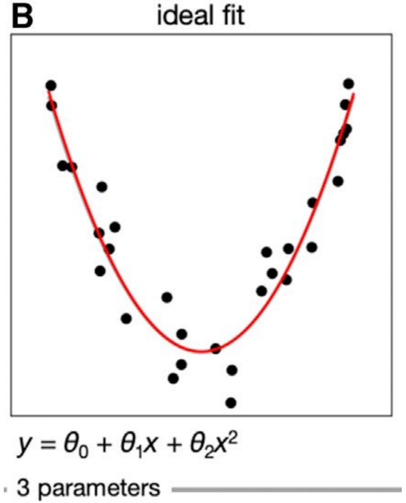
ideal fit
$y = \theta_0 + \theta_1 x + \theta_2 x^2$
3 parameters

E

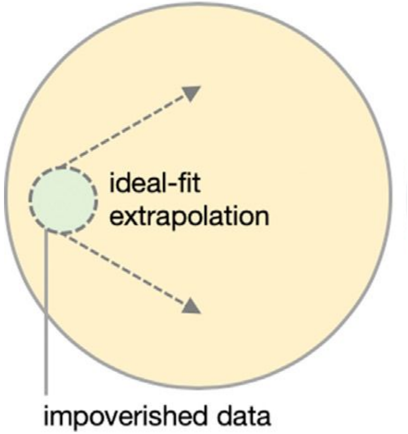


ideal-fit
extrapolation
impoverished data

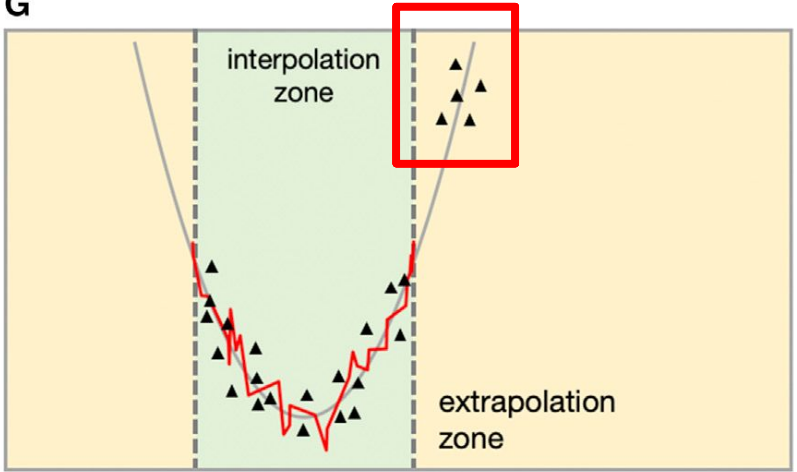
B

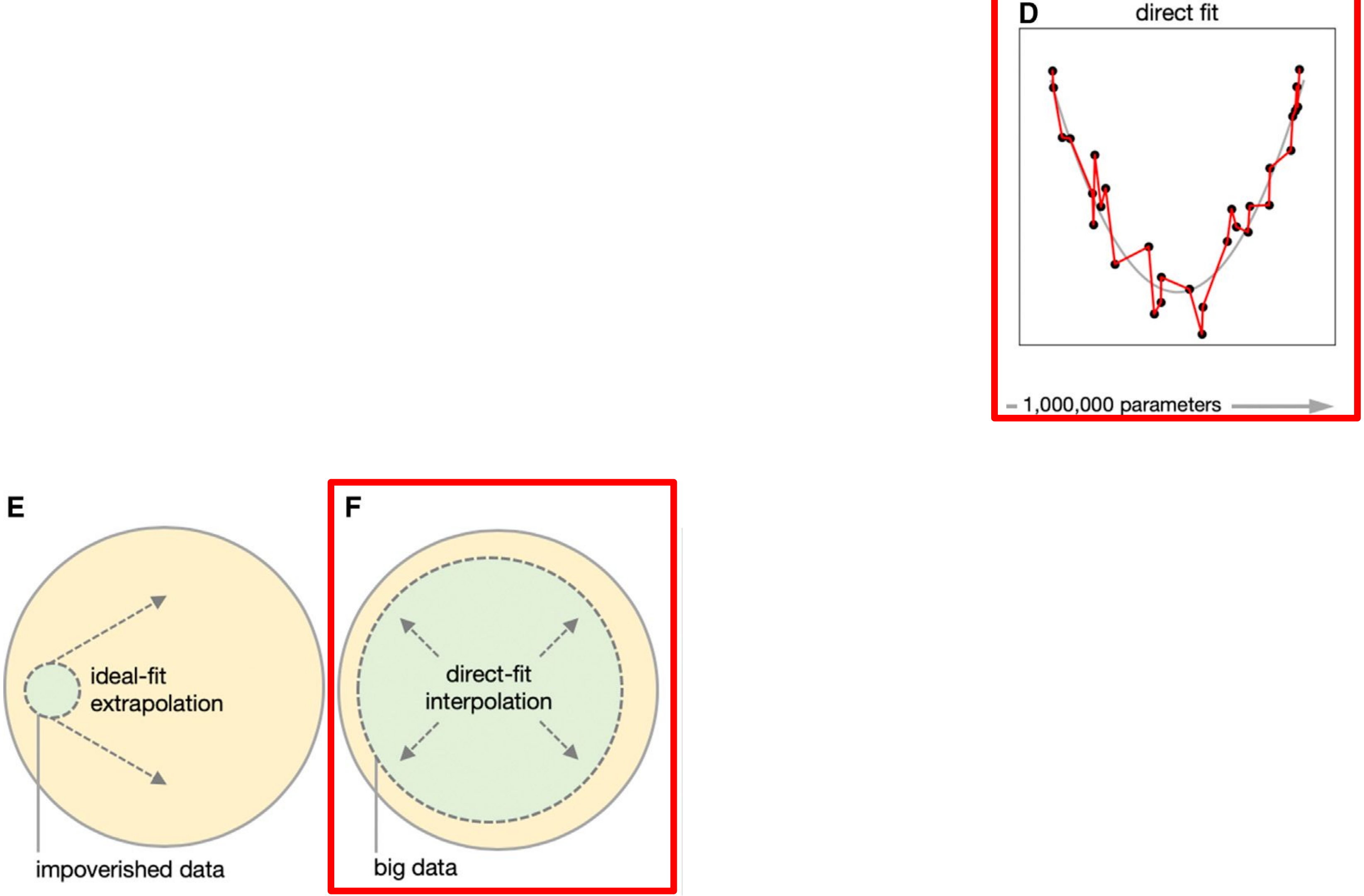

ideal fit
$y = \theta_0 + \theta_1 x + \theta_2 x^2$
3 parameters

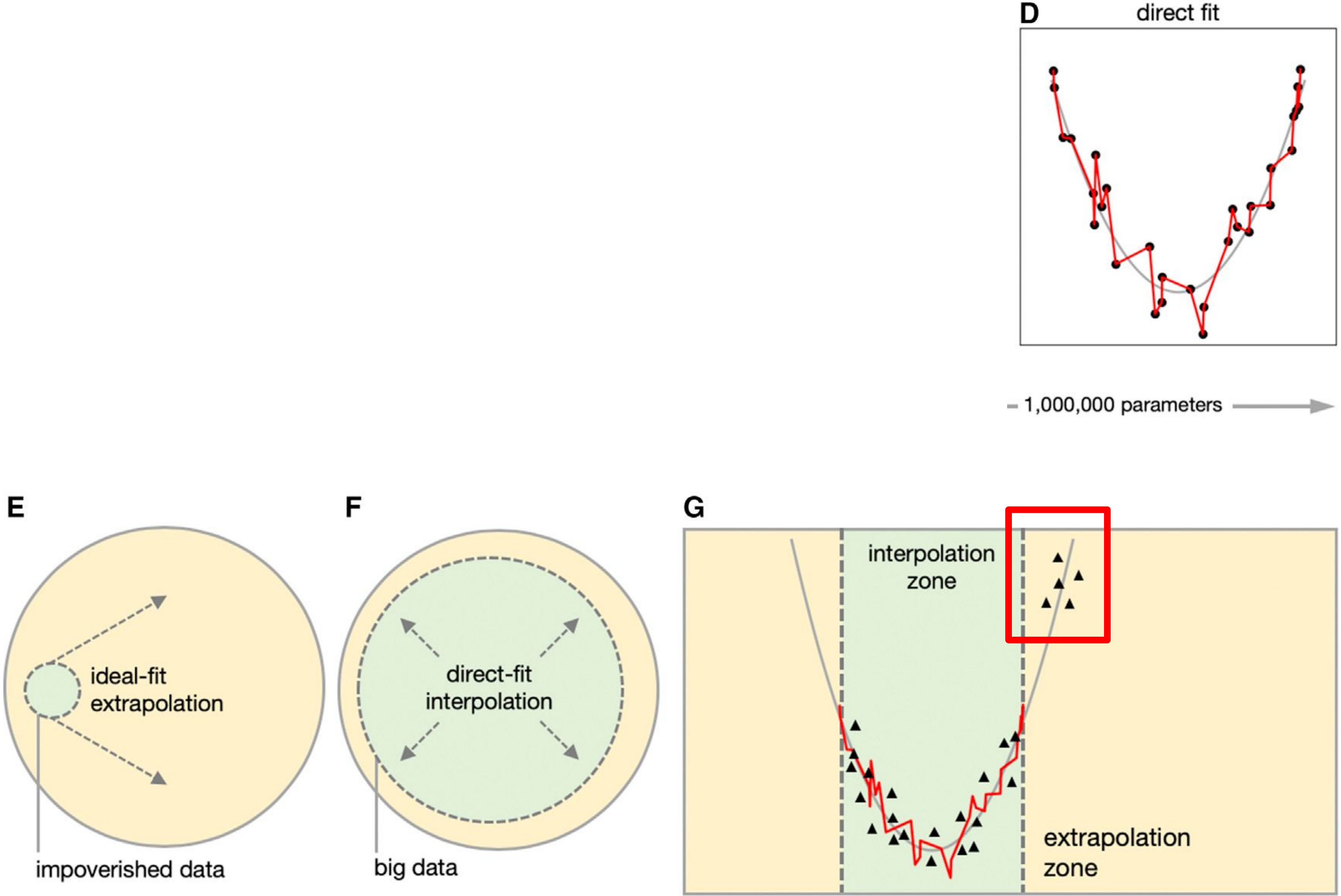
E


ideal-fit
extrapolation
impoverished data

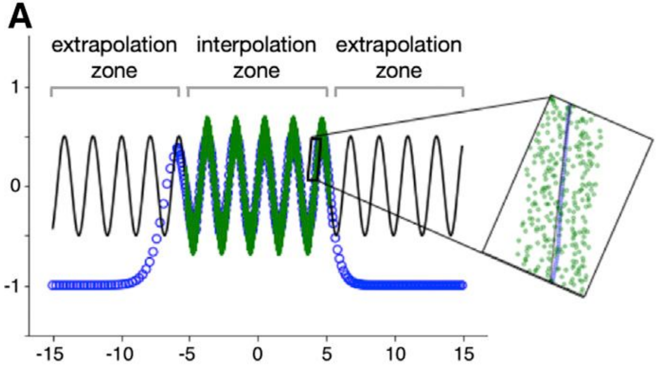
G


interpolation
zone
extrapolation
zone


D
direct fit
1,000,000 parameters
E
ideal-fit
extrapolation
impoverished data
F
direct-fit
interpolation
big data

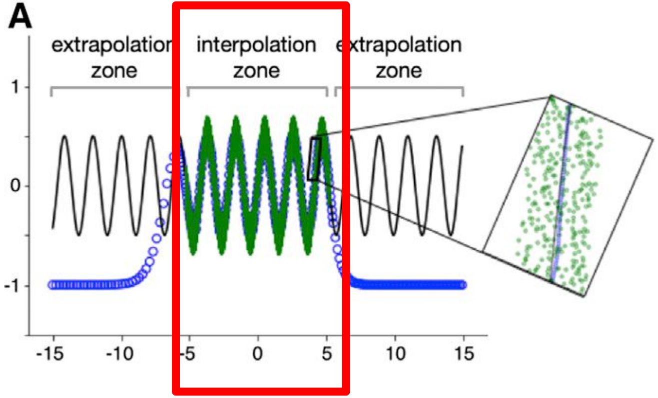

D
direct fit
1,000,000 parameters
E
ideal-fit
extrapolation
impoverished data
F
direct-fit
interpolation
big data
G
interpolation
zone
extrapolation
zone

Examples of Direct Fits



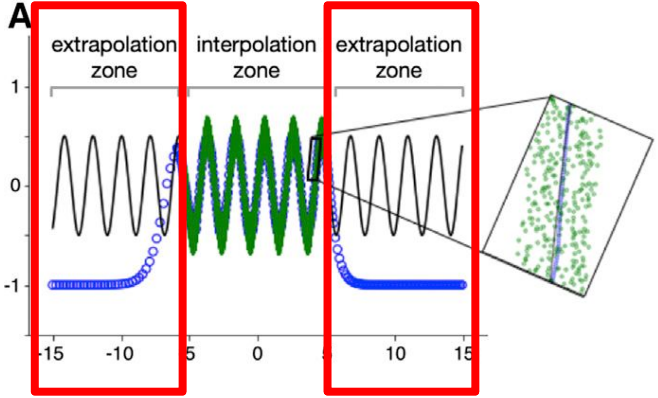
Artificial Neural Network (ANN)

- **Sine wave** - Training sample of size 10000 (-5 < x < 5)
- 1 input neuron
- 3 hidden layers (300 neurons each)
- 1 output neuron
- → 902 neurons
- → 180600 parameters (over-parameterized)



Artificial Neural Network (ANN)

- **Sine wave** - Training sample of size 10000 (-5 < x < 5)
- 1 input neuron
- 3 hidden layers (300 neurons each)
- 1 output neuron
- → 902 neurons
- → 180600 parameters (over-parameterized)

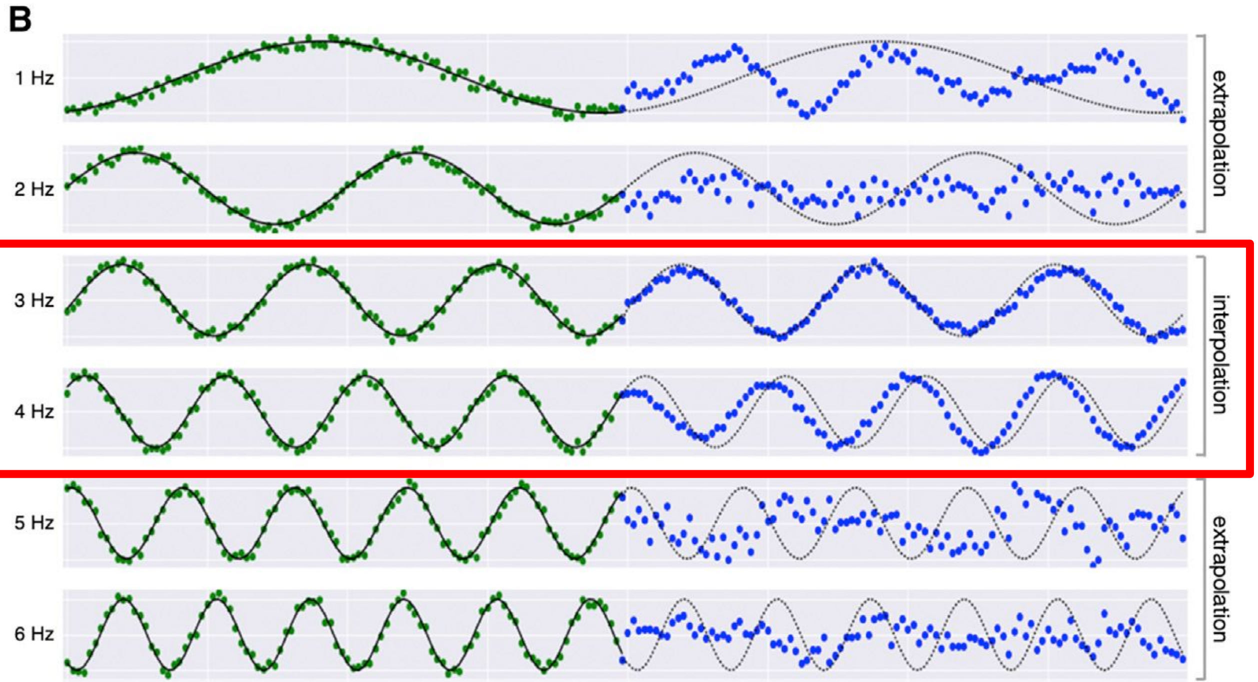


Artificial Neural Network (ANN)

- **Sine wave** - Training sample of size 10000 (-5 < x < 5)
- 1 input neuron
- 3 hidden layers (300 neurons each)
- 1 output neuron
- → 902 neurons
- → 180600 parameters (over-parameterized)

Recurrent Long-Short Term Memory (LSTM)

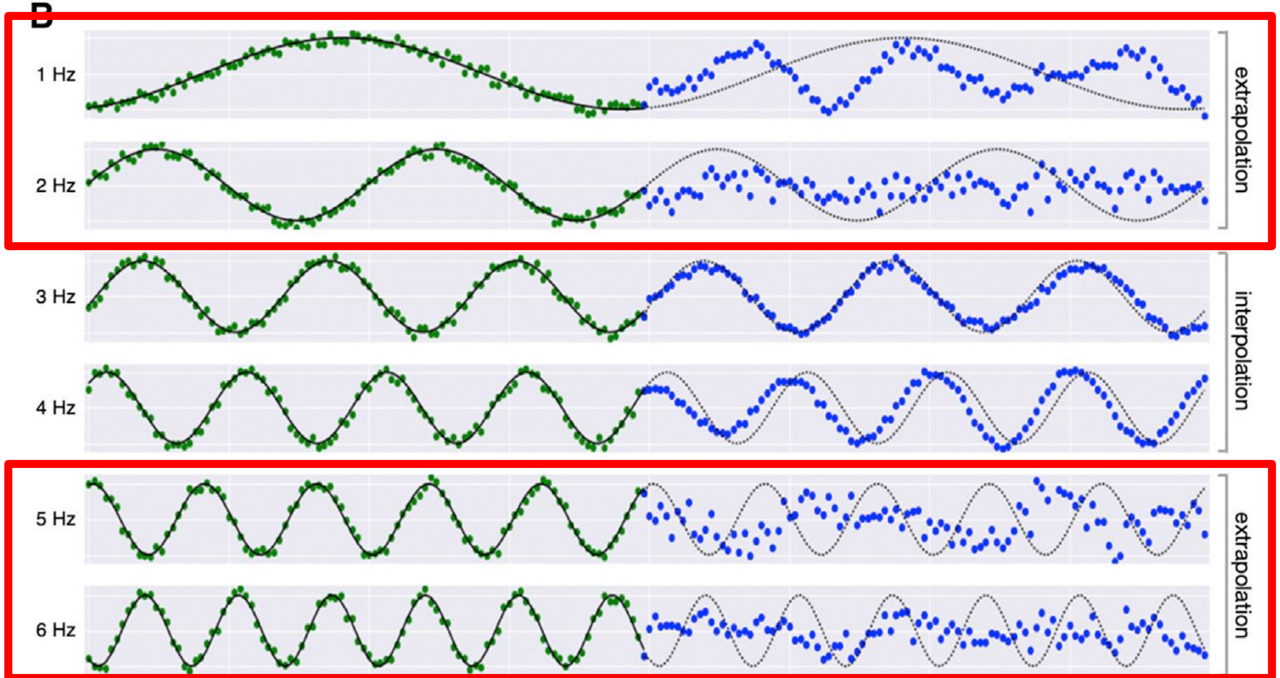
- **Sine wave** - Training sample of size 10000 (2.5Hz < x < 4.5Hz)
- Task - predict the next 100 values



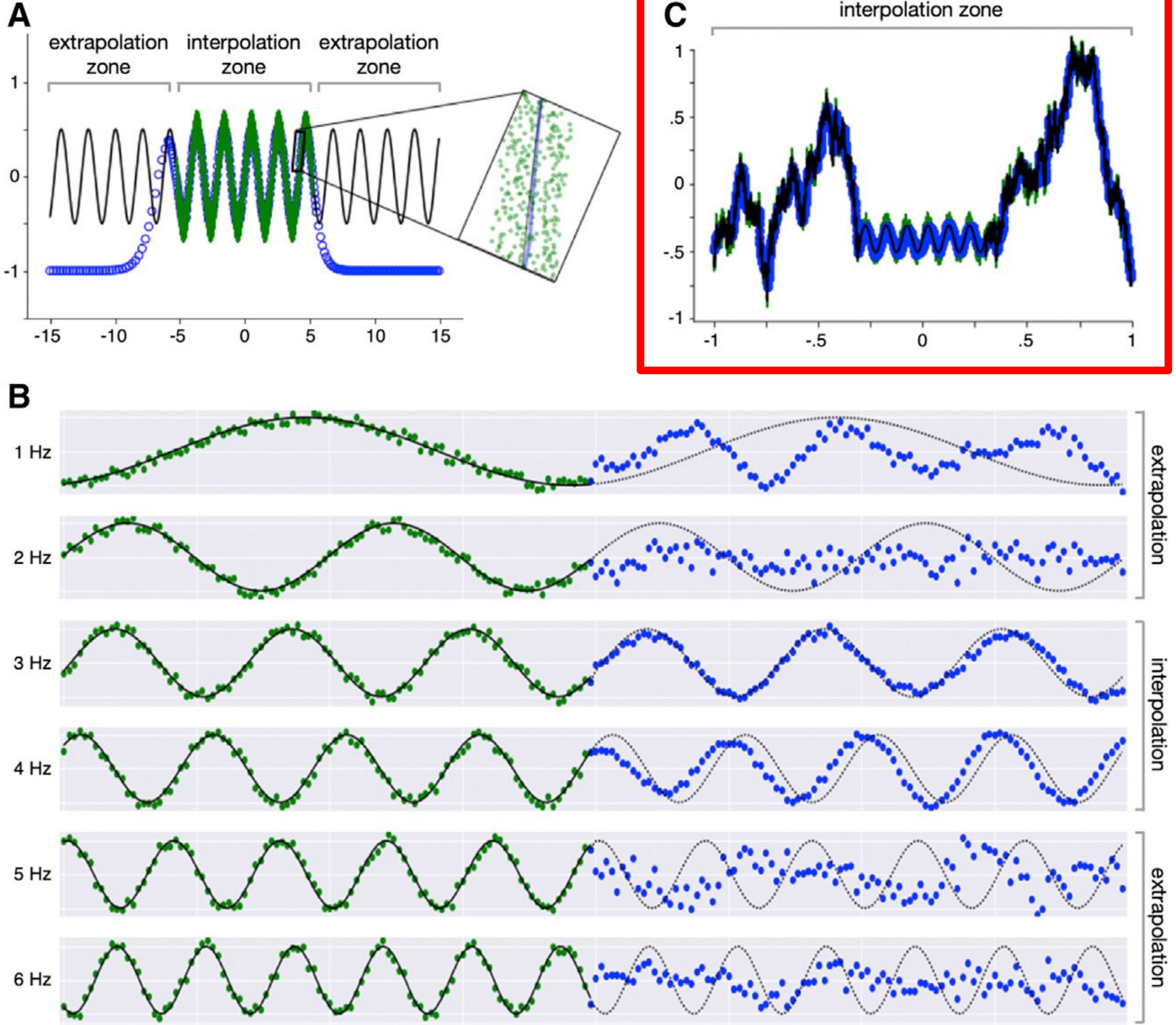
Recurrent Long-Short Term Memory (LSTM)

- **Sine wave** - Training sample of size 10000 (2.5Hz < x < 4.5Hz)
- Task - predict the next 100 values

LSTM



FCNN



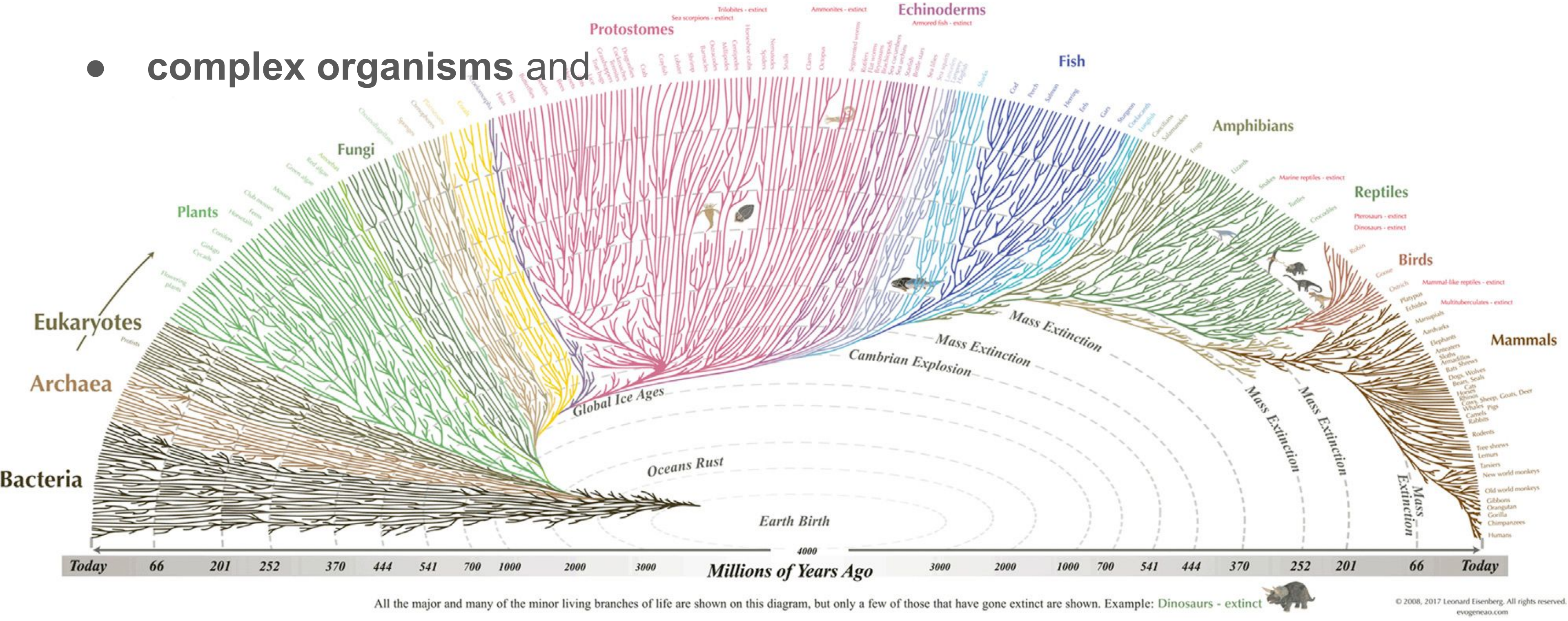
FCNN

LSTM

Evolution and ANNs

Evolutionary theory aims to explain how

- **complex organisms** and



Evolutionary theory aims to explain how

- **complex organisms** and
- **complex biological mechanisms** (e.g. photosynthesis, wings, and retinas)

Evolutionary theory aims to explain how

- **complex organisms** and
- **complex biological mechanisms** (e.g. photosynthesis, wings, and retinas)
- evolved to fit their **local ecological niches**

Evolutionary theory aims to explain how

- **complex organisms** and
- **complex biological mechanisms** (e.g. photosynthesis, wings, and retinas)
- evolved to fit their **local ecological niches**
- **without any explicit comprehension of the problems** at hand and

Evolutionary theory aims to explain how

- **complex organisms** and
- **complex biological mechanisms** (e.g. photosynthesis, wings, and retinas)
- evolved to fit their **local ecological niches**
- **without any explicit comprehension of the problems** at hand and
- **without any understanding of the solutions** to overcome them

(Darwin, 1859, via Hasson et al, 2020)

~~Evolutionary theory~~ Direct fit ~~aims to explain how~~

- **Is possible via** complex **models** ~~organisms and~~
- **Leads to** complex **behavior** ~~biological mechanisms~~
- **Trains** ~~evolved~~ to fit their **interpolation zone** ~~local ecological niches~~
- **Learns without any explicit comprehension of the problems** at hand and
- **Learns without any understanding of the solutions** to overcome them

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

- Simple and parsimonious
- Yet inefficient and costly in implementation
- Allows no extrapolation
  - to far future
  - another planet

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natura'
- Time

- Simpl-
- extrapolation
    - to far future
    - another planet

Lack of extrapolation undermines the theory of evolution?

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

- Simple and parsimonious
- Yet inefficient and costly in implementation
- Allows no extrapolation
    - to far future
    - another planet
    - **TK: explains past and present**

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

Direct fit learning

- Over-sampling with variation

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

Direct fit learning

- Over-sampling with variation
- plasticity

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

Direct fit learning

- Over-sampling with variation
- plasticity
- Combinatorial neural code

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

Direct fit learning

- Over-sampling with variation
- plasticity
- Combinatorial neural code

- Objective functions

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

Direct fit learning

- Over-sampling with variation
- plasticity
- Combinatorial neural code

- Objective functions
- Iteration over samples

Evolution does it using...

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

No need for “intelligent” force to guide the change.

Direct fit learning

- Over-sampling with variation
- plasticity
- Combinatorial neural code
- Objective functions
- Iteration over samples

No need for intentional or interpretable rules to guide learning.

Are Direct Fit Models Biological (BNN) and Artificial Neural Networks (ANN) Black Boxes?

Black Box

- We build ANNs
  - **explicit architectural specifications**
  - **explicit learning rules** and **finite training samples**
  - well-specified **objective functions**
  - **direct access to each weight**

Glass Box instead?

- We build ANNs
  - **explicit architectural specifications**
  - **explicit learning rules** and **finite training samples**
  - well-specified **objective functions**
  - **direct access to each weight**
- Understand via
  - Network architectures
  - Learning rules
  - Objective functions

How to Achieve Successful Direct Fit

Successful Direct Fit

1. Must be fit to a **structured world** - faces around you

Successful Direct Fit

1. Must be fit to a structured world
2. World must be **sampled densely and widely** - model of western faces

Successful Direct Fit

1. Must be fit to a structured world
2. World must be sampled densely and widely
3. Model must support **high-dimensional encoding space**

Successful Direct Fit

1. Must be fit to a structured world
2. World must be sampled densely and widely
3. Model must support high-dimensional encoding space
4. Model must have a **correct objective function**(s) - adaptive advantage

Successful Direct Fit

1. Must be fit to a structured world
2. World must be sampled densely and widely
3. Model must support high-dimensional encoding space
4. Model must have a correct objective function(s)
5. Model must implement **effective regularization** during optimization
    a. (to avoid explosive overfit)
    b. Like evolution - genetic priors on learning.

What Does it Mean for Cognitive Psychology?

Computational Resources are not Scarce

Brain as a Direct Fit model

1. Mm^3 100K+ neurons
2. M+ adjustable synaptic weights
3. Great interpolation

Relative to the Brain, ANNs are simplistic and minuscule

Input is not Impoverished

We may be exposed to

- Thousands of exemplars of daily categories / year
- Thousands of views in each encounter
- → rich training set

Input is not Impoverished

We may be exposed to

- Thousands of exemplars of daily categories / year
- Thousands of views in each encounter
- → rich training set

Children exposed to

- Several million words per year (Roy et al., 2015)

Beware of (impoverished) experiments

Shallow Self-Supervision and External-Supervision are Sufficient for Learning

External supervision may be guided by

- Other social agents
- Human annotators (BNNs)

In the absence of external supervision, BNNs and ANNs can rely on

- Self-supervised objective functions (space, time, relative to self-motion or action)
- ⇒ “predictive coding”

At Which Level Does Psychology Emerge?

At Which Level Does Psychology Emerge?

Instead of imposing

- Efficiency, simplicity, and interpretability
- wholesale across neural systems

Ask

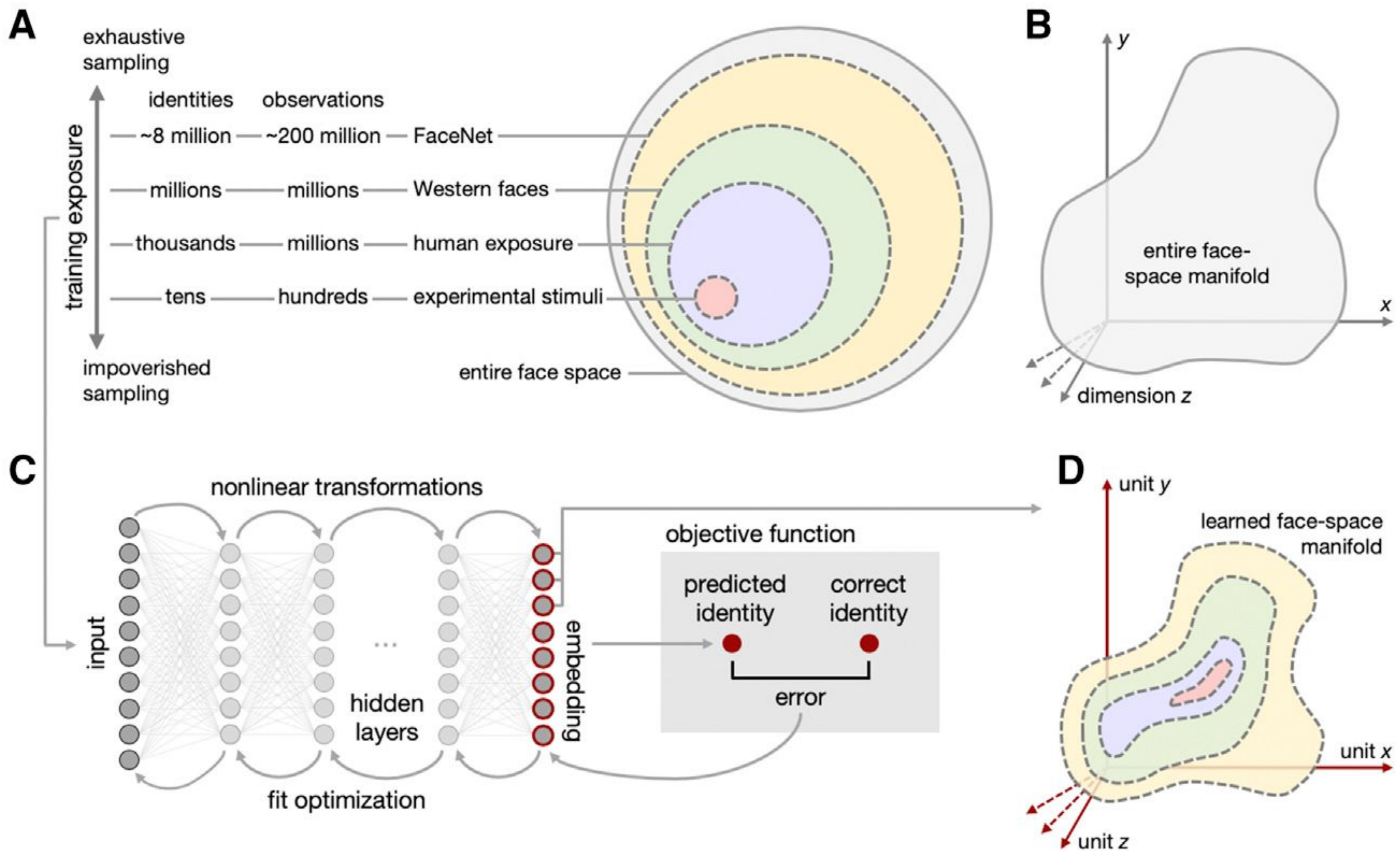
- How uniquely human capabilities
- Can extract explicit and compact knowledge about the outside world
- From the billions of direct fit model weights?

Direct Fit to Nature

Thank You for Listening

Paper Overview by Taavi Kivisik

University of Tartu
2020-11-12


A
exhaustive sampling
identities
observations
~8 million
~200 million
FaceNet
millions
millions
Western faces
thousands
millions
human exposure
tens
hundreds
experimental stimuli
training exposure
impoverished sampling
entire face space
B
y
x
entire face-space manifold
dimension z
C
nonlinear transformations
input
...
hidden layers
embedding
objective function
predicted identity
correct identity
error
fit optimization
D
unit y
learned face-space manifold
unit x
unit z

BACKLOG

Summary

1. Given enough relevant data, interpolation might be great

No Free Lunch theorem

Black Box

- Criticism - over-parameterized models
  - Given correct input
  - Generate the correct output
  - Without any explanation of their internal workings
- We build ANNs
  - According to explicit architectural specifications
  - Train networks using explicit learning rules and finite training samples
  - With well+specified objective functions
  - We have direct access to each weight in the network
- Given their unprecedented level of transparency, why do we deem ANNs black-box models?
- We should exercise caution in cases in which these models seem to “learn” simple, psychologically interpretable variables.

Evolution does it by

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

(Lewontin, 1970; Gould, 1982; via Hasson et al, 2020)

ANN aiming to direct fit

- is an over-parameterized model
- can learn arbitrarily complicated functions
- given big relevant data, can interpolate well!
- cannot extrapolate outside learning range
- is “mindless” optimization

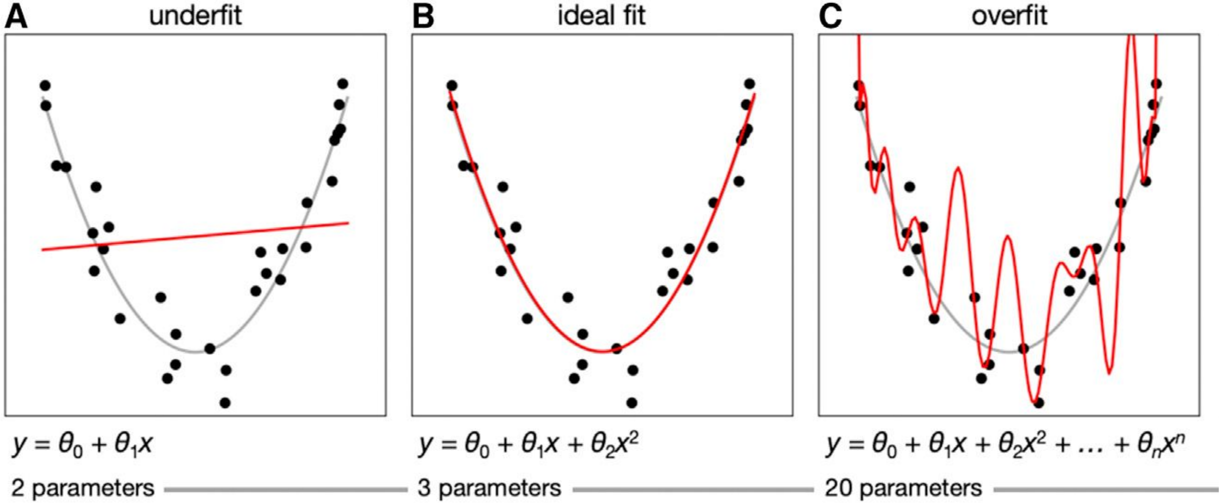
Two Types of Generalization

Interpolation VS Extrapolation

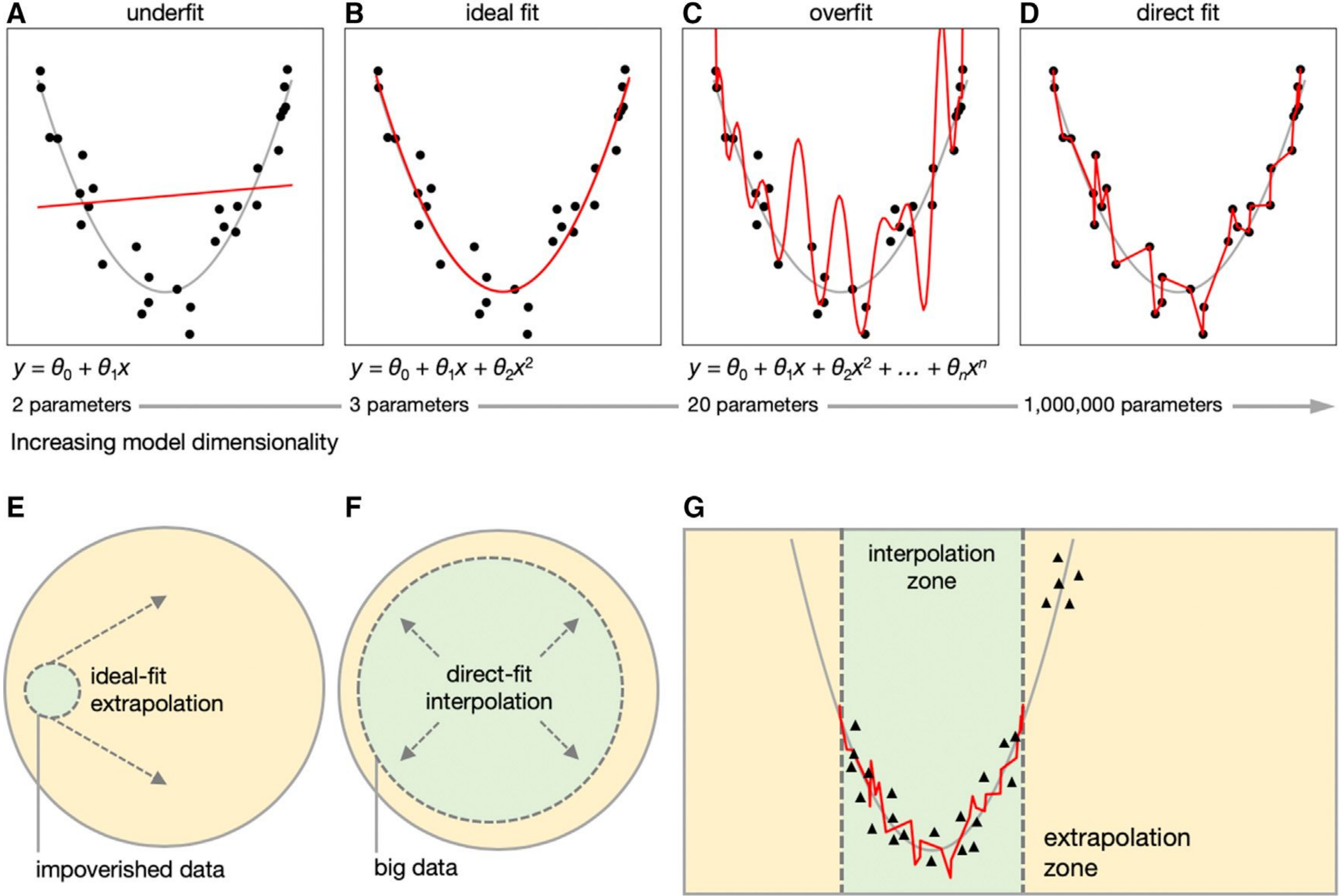
- Within the training data range
-
- Given
  - direct fit models
  - big real world data
  - ⇒ mindless yet powerful form of generalization

Value

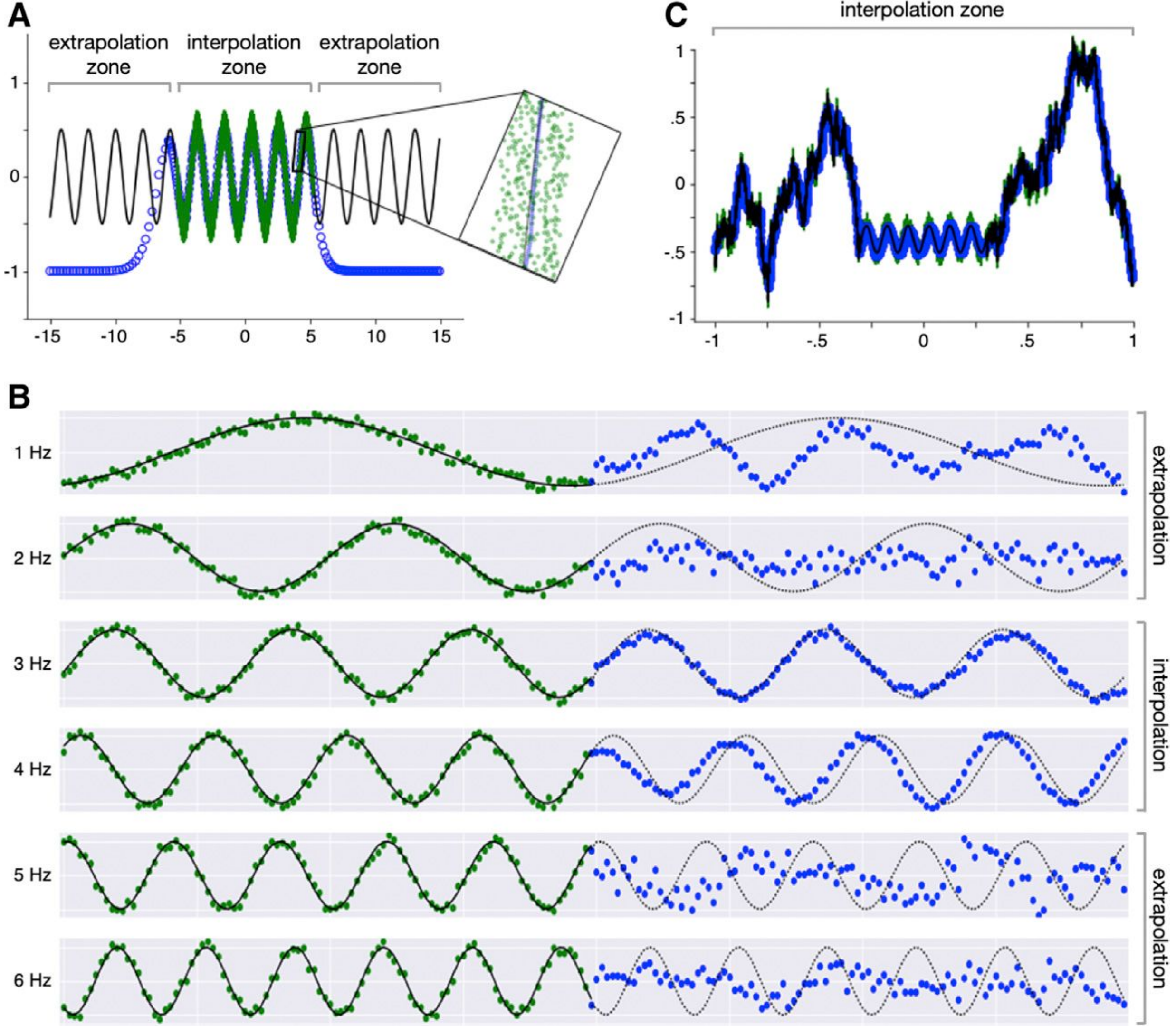
- Behavior (task **performance**)
- Embrace complexity
- Aim for **direct fit**



- Neuroscientific models
  - **interpretability** and **generalization** valued



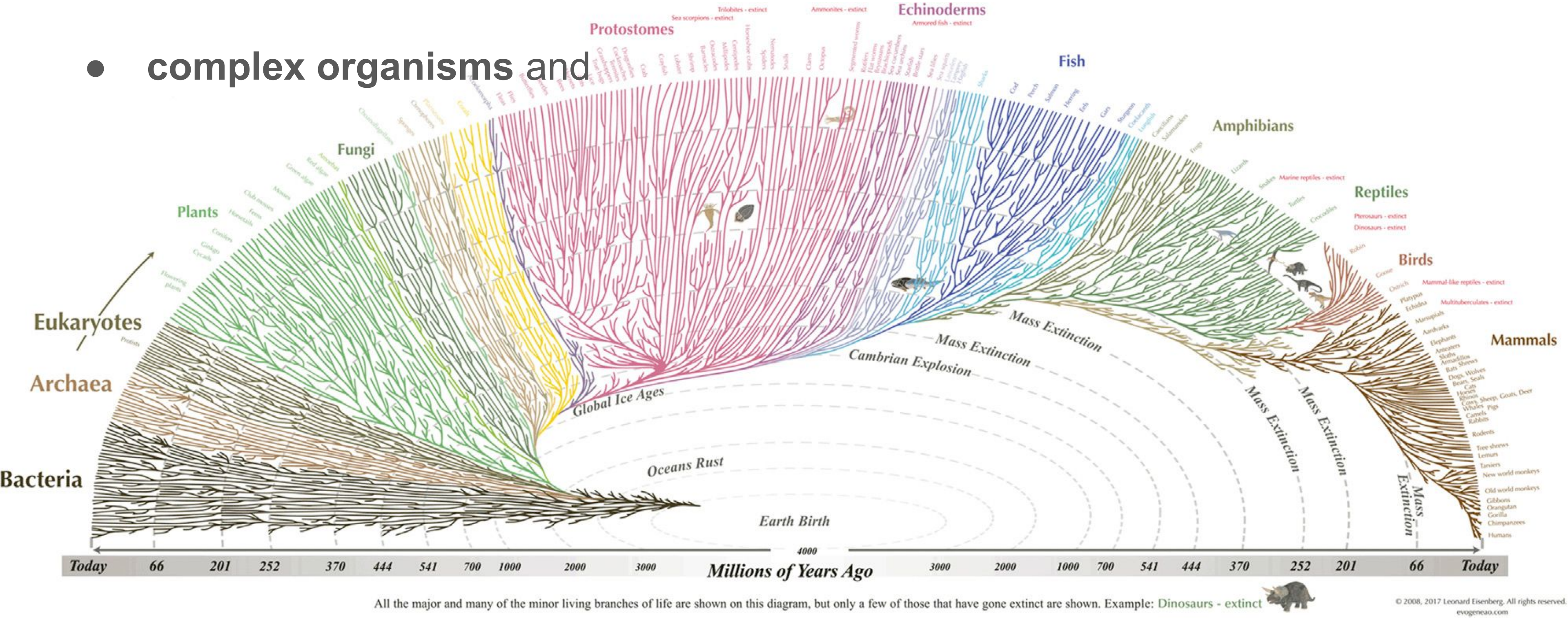
A
underfit
B
ideal fit
C
overfit
D
direct fit
$y = \theta_0 + \theta_1 x$
$y = \theta_0 + \theta_1 x + \theta_2 x^2$
$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \ldots + \theta_n x^n$
2 parameters
3 parameters
20 parameters
1,000,000 parameters
Increasing model dimensionality
E
ideal-fit
extrapolation
impoverished data
F
direct-fit
interpolation
big data
G
interpolation
zone
extrapolation
zone



A
extrapolation zone
interpolation zone
extrapolation zone
1
0
-1
-15
-10
-5
0
5
10
15
C
interpolation zone
1
.5
0
-.5
-1
-1
-.5
0
.5
1
B
1 Hz
2 Hz
3 Hz
4 Hz
5 Hz
6 Hz
extrapolation
interpolation
extrapolation

Evolutionary theory aims to explain how

- **complex organisms** and



Evolutionary theory aims to explain how

- **complex organisms** and
- **complex biological mechanisms** (e.g. photosynthesis, wings, and retinas)
- evolved to fit their **local ecological niches**
- **without any explicit comprehension of the problems** at hand and
- **without any understanding of the solutions** to overcome them

(Darwin, 1859, via Hasson et al, 2020)

Evolution does it by

- Over-production with variance
- Inheritance
- Combinatorial power (of the genetic code)
- Selection (natural and artificial)
- Time

(Lewontin, 1970; Gould, 1982; via Hasson et al, 2020)