Deep Learning

with No math

Philipp Gabler

2022-06-28

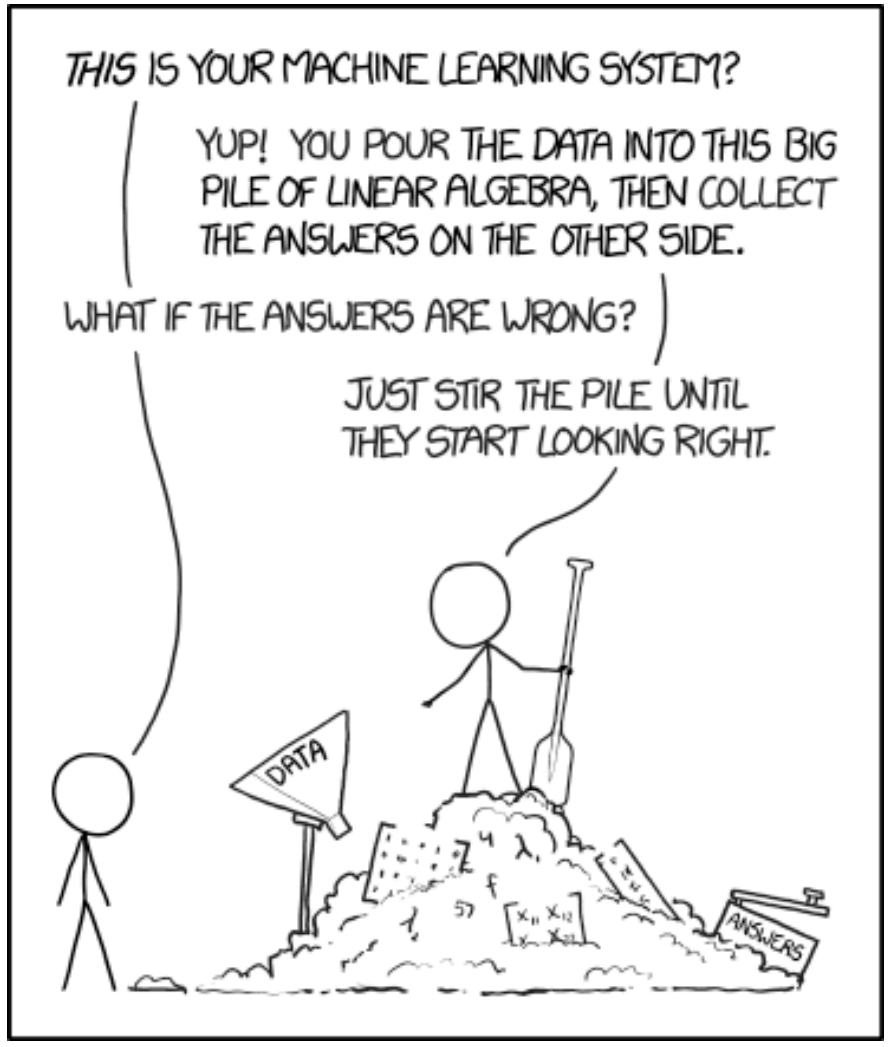
DEEP LEARNING

WITH ~~NO~~ MATH

AS LITTLE AS POSSIBLE

PHILIPP GABLER

2022-06-28



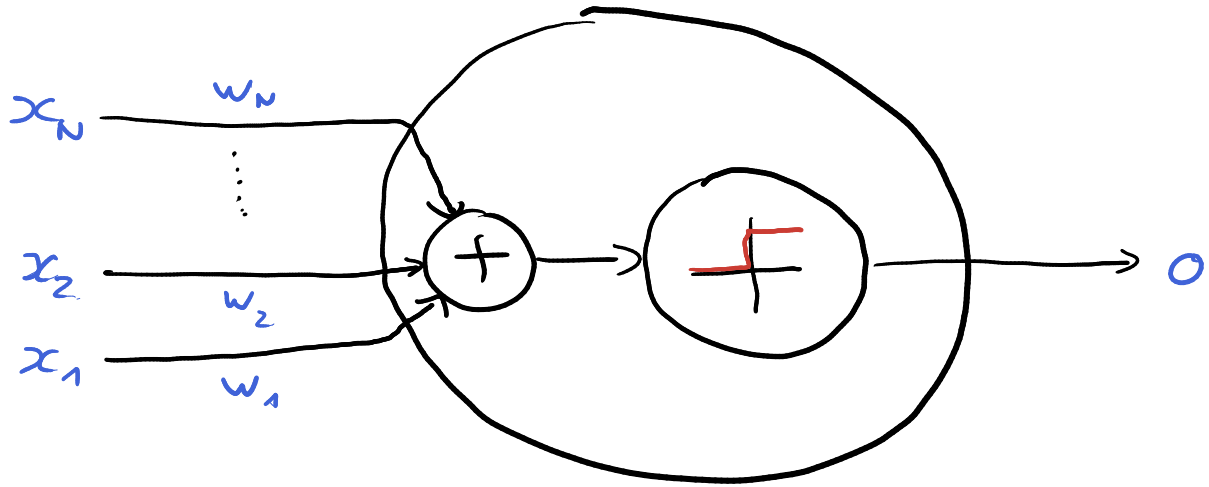
https://xkcd.com/1838/

OUTLINE

1. A BIT OF HISTORY
2. WHAT PROBLEM(S) DOES DL SOLVE? — COMPUTATIONAL
3. HOW DO SOLUTIONS LOOK LIKE? — ALGORITHMIC
4. HOW ARE SOLUTIONS IMPLEMENTED? — IMPLEMENTATIONAL
5. SOME LANGUAGE-RELATED USE CASES

Ἐν ἀρχῇ ἦν τὸ PERCEPTRON

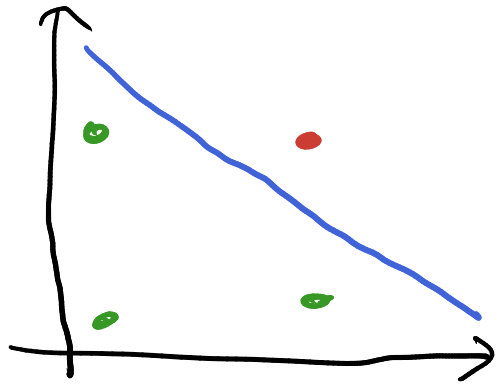
- McCULLOCH & PITTS, 1948: NEURON MODEL

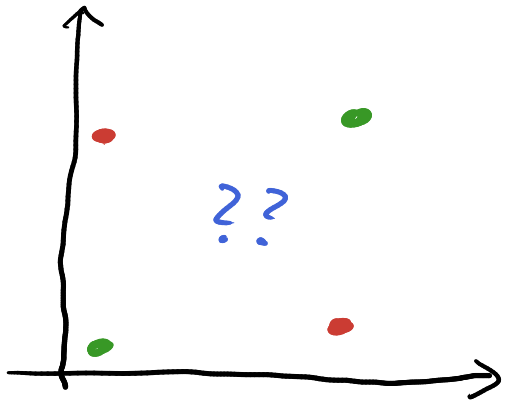


- ROSENBLATT, 1958: ENTHUSIASM

- MINSKY & PAPERT, 1969: SHOCK OF "PERCEPTRONS"

❄ AI WINTER ❄

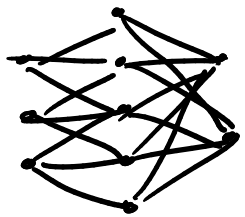




PROBLEM OF LINEAR SEPARABILITY

NEW SPRING

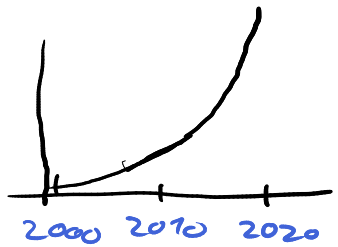
- RUMELHART & MCCLELLAND (1986): PDP
  - ↳ CONNECTIONISM
  - ↳ BACKPROPAGATION
  - ↳ SINGLE-LAYER → MULTI-LAYER

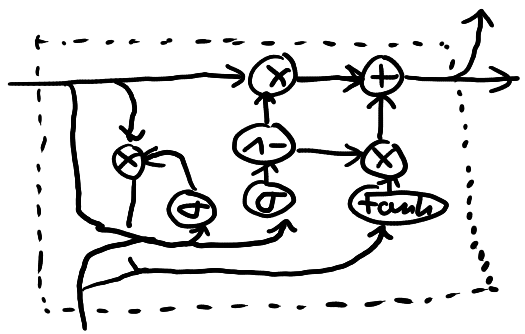


- STATISTICAL LEARNING / MACHINE LEARNING OVER (SYMBOLIC) AI

TOWARDS DEEP LEARNING

- MORE LAYERS
- MORE DATA
- MORE COMPUTE (MOORE'S LAW)
- BETTER TUNING & TRAINING





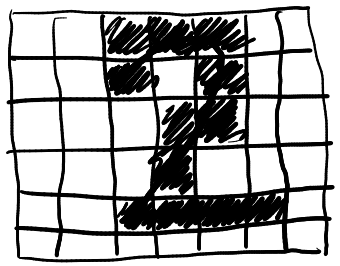
TAKE-AWAYS 1

- ARTIFICIAL – NO NEURONS ANYMORE!
- AI AND ML HAVE SPLIT – DL ≠ INTELLIGENCE
- GROWTH OF TECHNOLOGY IMPORTANT
- DL IS SUB SYMBOLIC (≃ CONNECTIONISM)

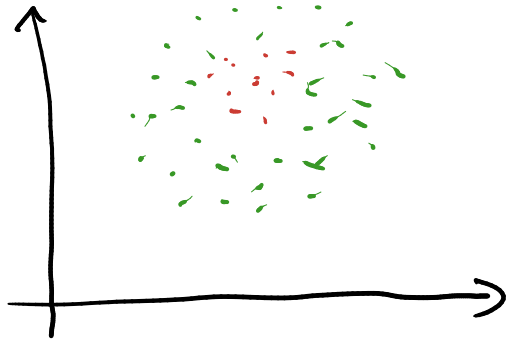
WHAT PROBLEMS DOES DEEP LEARNING SOLVE?

CHARACTERISTICS OF DEEP LEARNING

HIGH-DIMENSIONAL

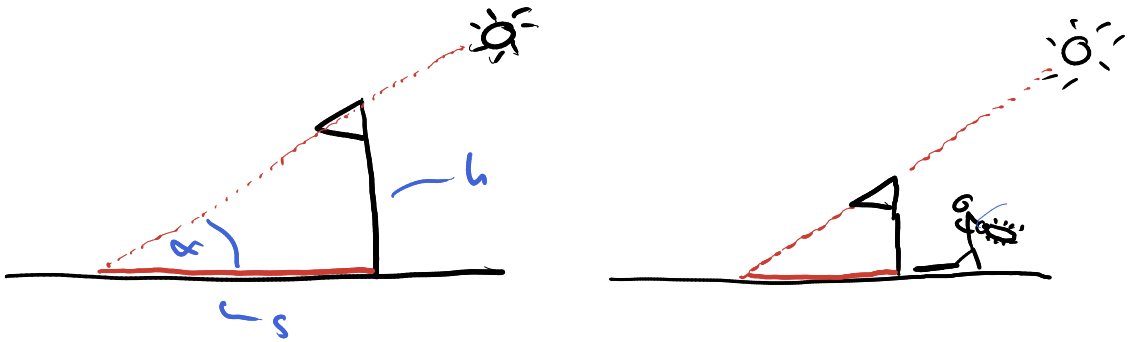


NON-LINEAR

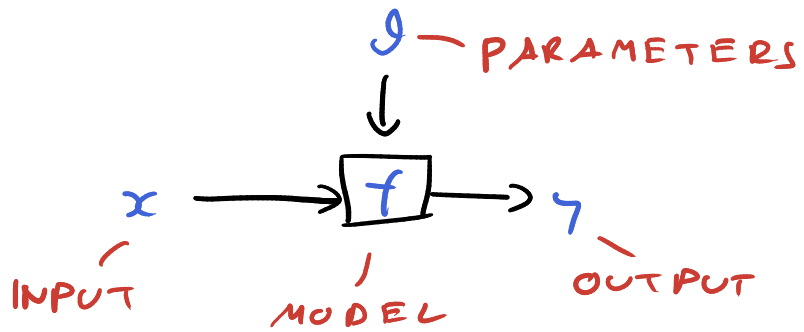


PREDICTION

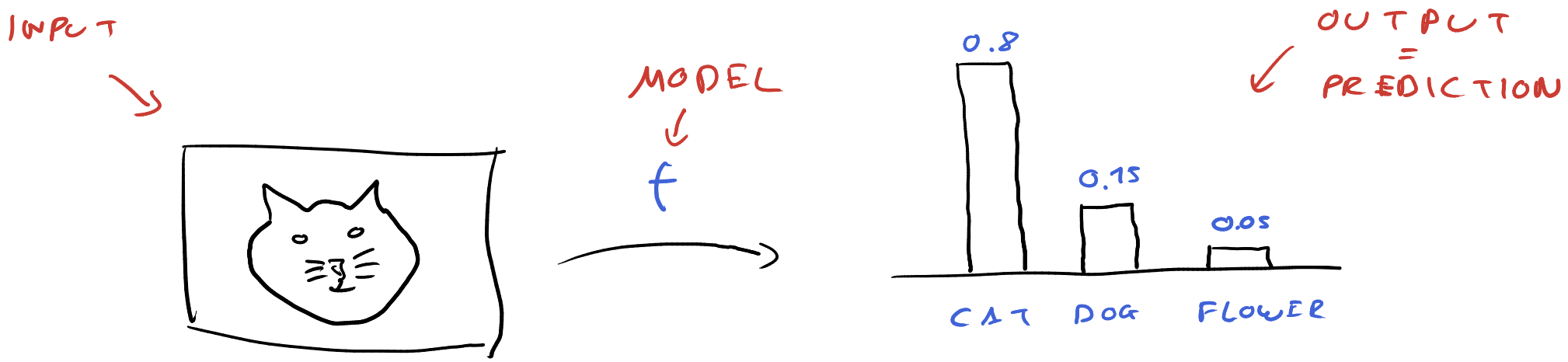
ASSOCIATIVE



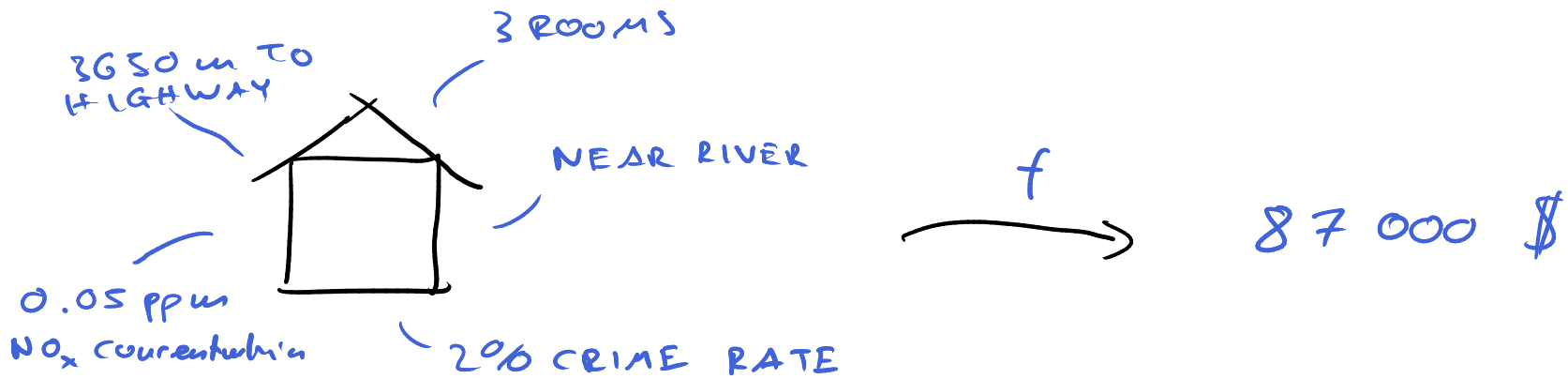
END-TO-END TRAINING, PARAMETRIC



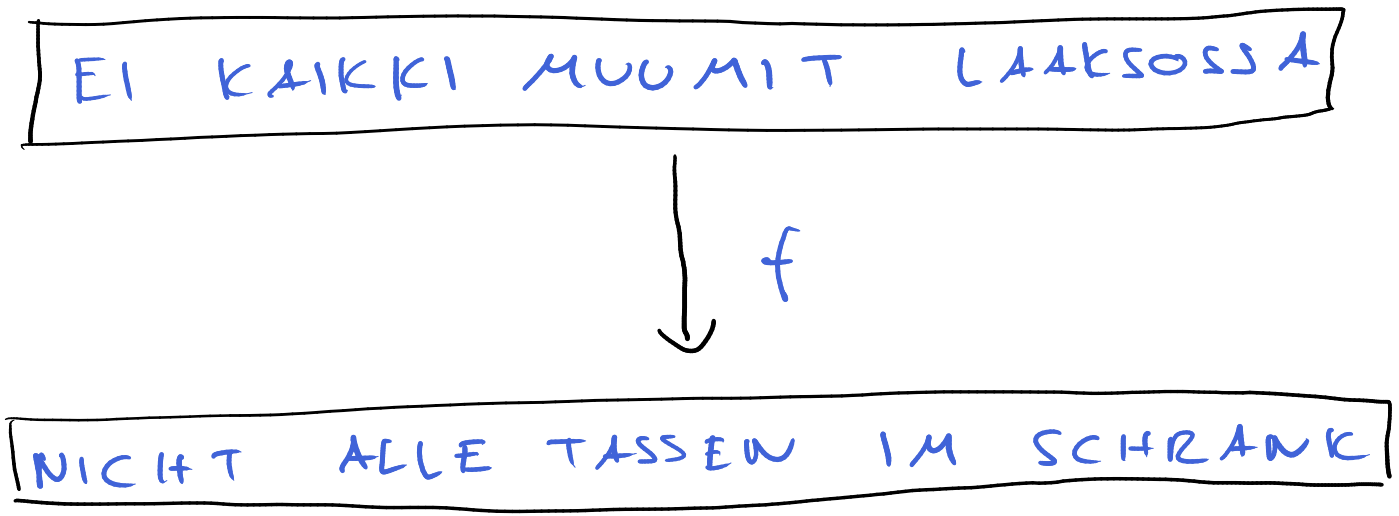
CLASSIFICATION



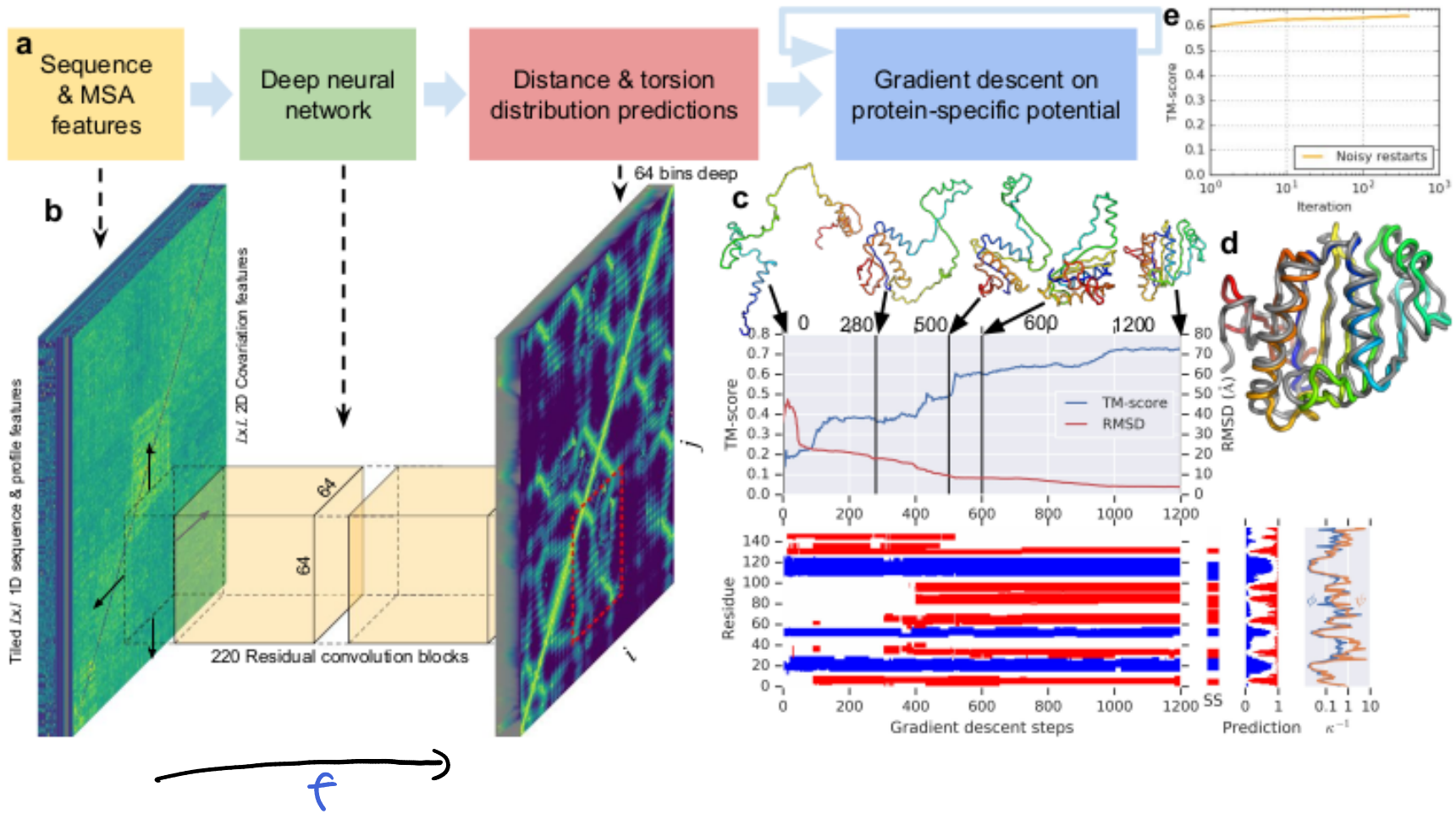
REGRESSION


3650 m TO HIGHWAY
3 ROOMS
NEAR RIVER
0.05 ppm
NO_x concentration
20% CRIME RATE
f
87 000 $

SEQUENCE PREDICTION



HIGH-DIMENSIONAL MODELLING



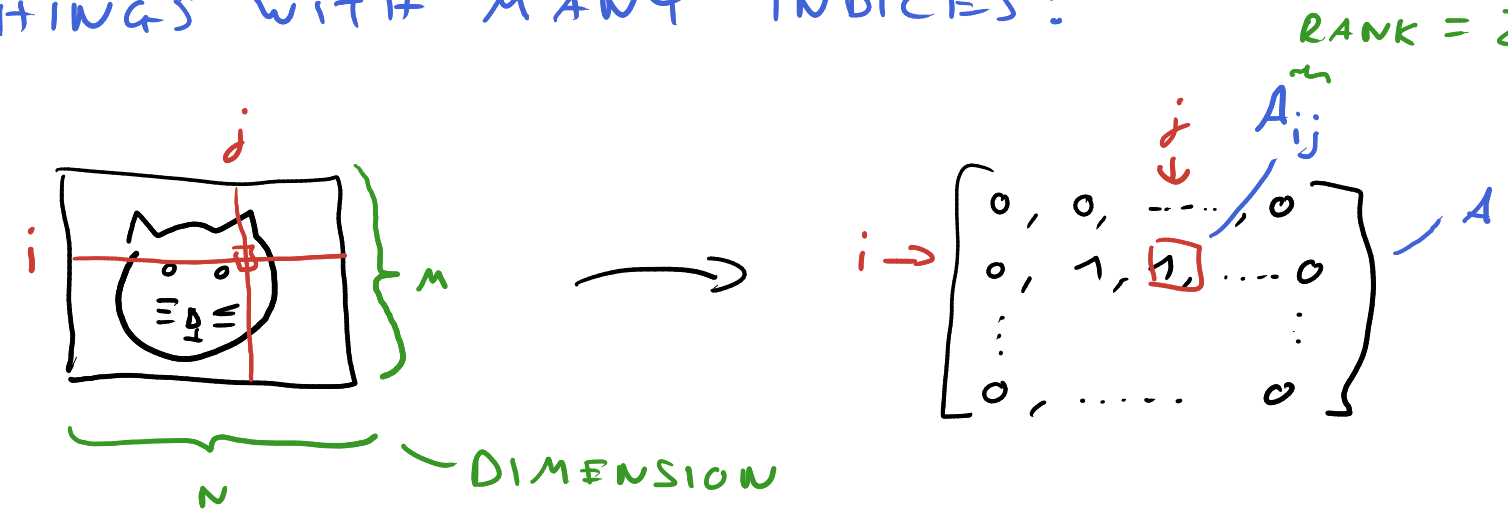
f

https://doi.org/10.1038/s41586-019-1923-7

TENSORS

TensorFlow

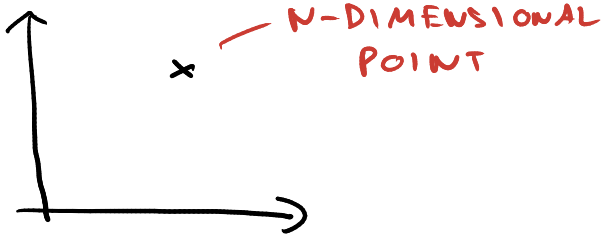
JUST THINGS WITH MANY INDICES:



EVERYTHING IS LINEAR ALGEBRA!

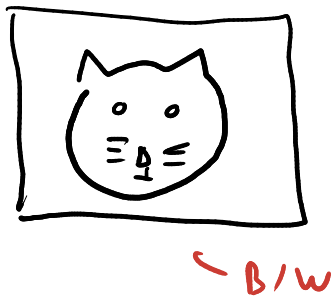
TENSOR SHAPES

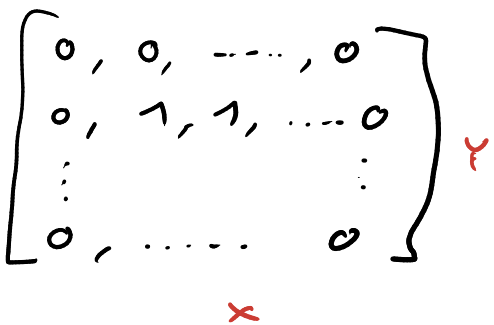
RANK 1



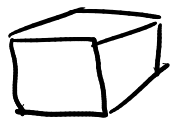
[2.0, -0.3]

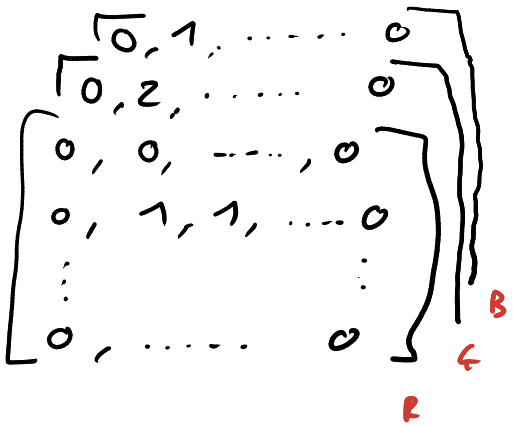
RANK 2



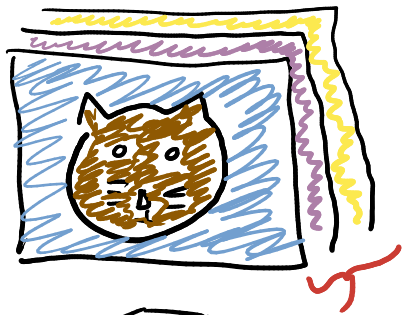


RANK 3



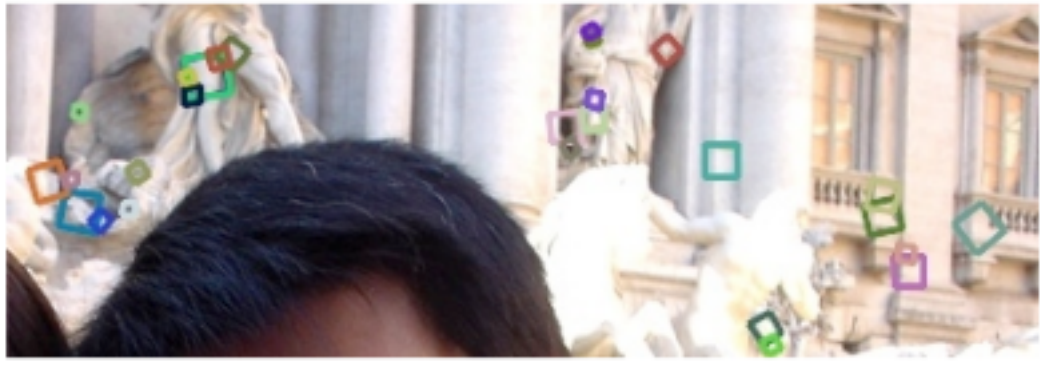


RANK 4

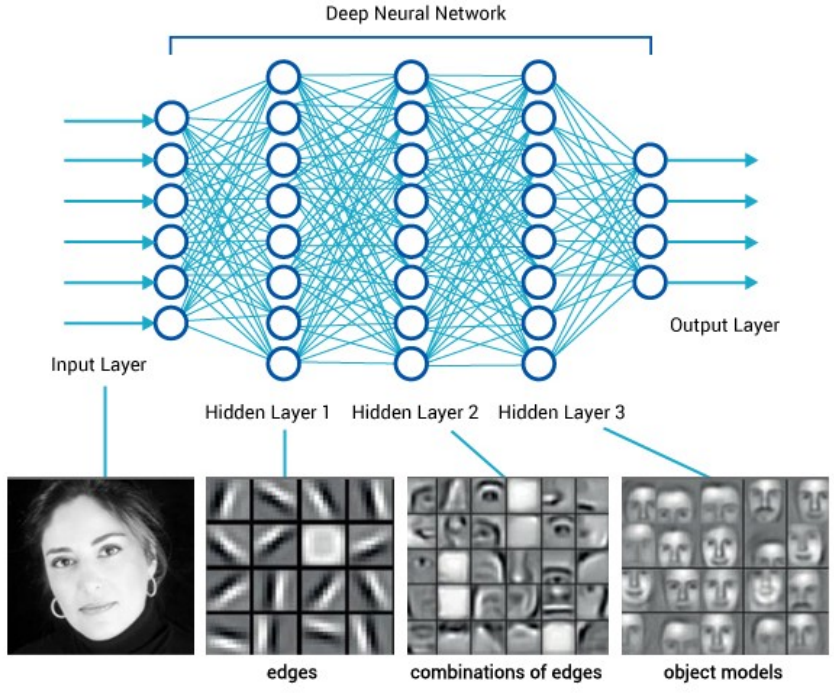


From Features to Layers

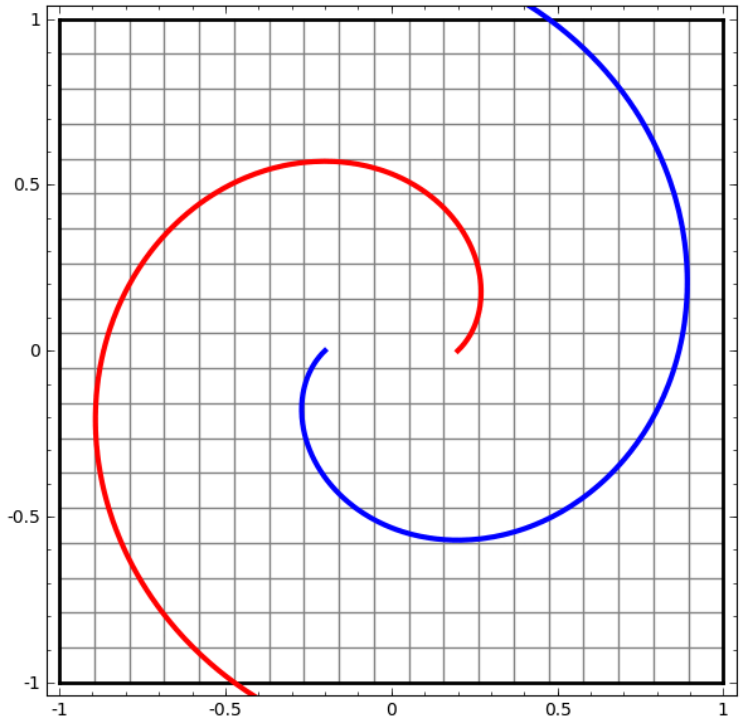
Traditional



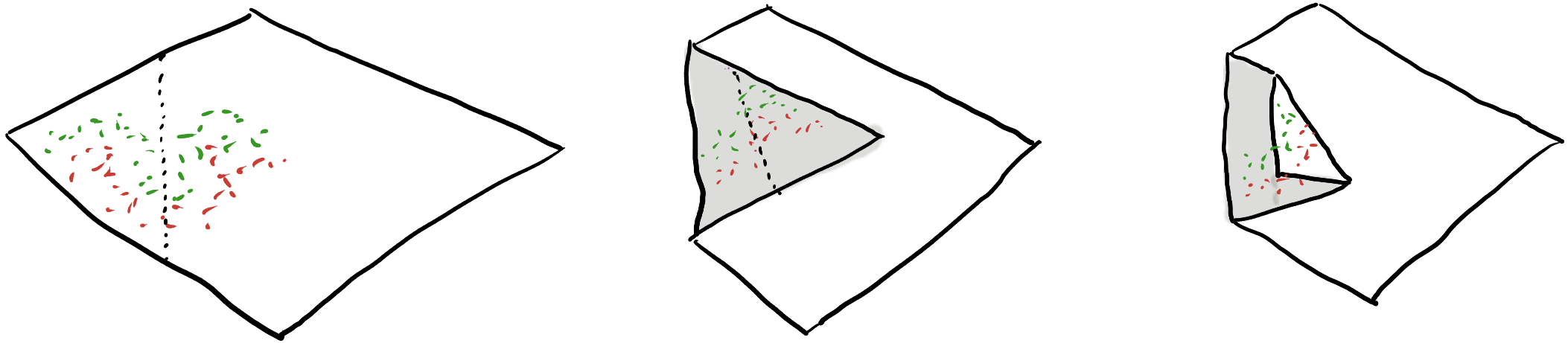
Deep Learning



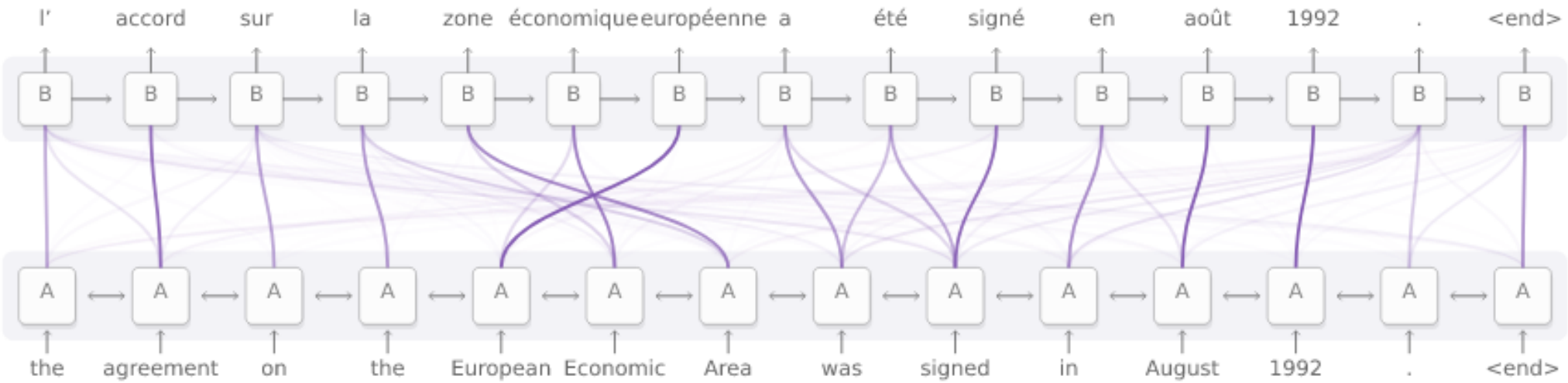
https://medium.com/diaryofawannapreneur/deep-learning-for-computer-vision-for-the-average-person-861661d8aa61


1
0.5
0
-0.5
-1
-1
-0.5
0
0.5
1

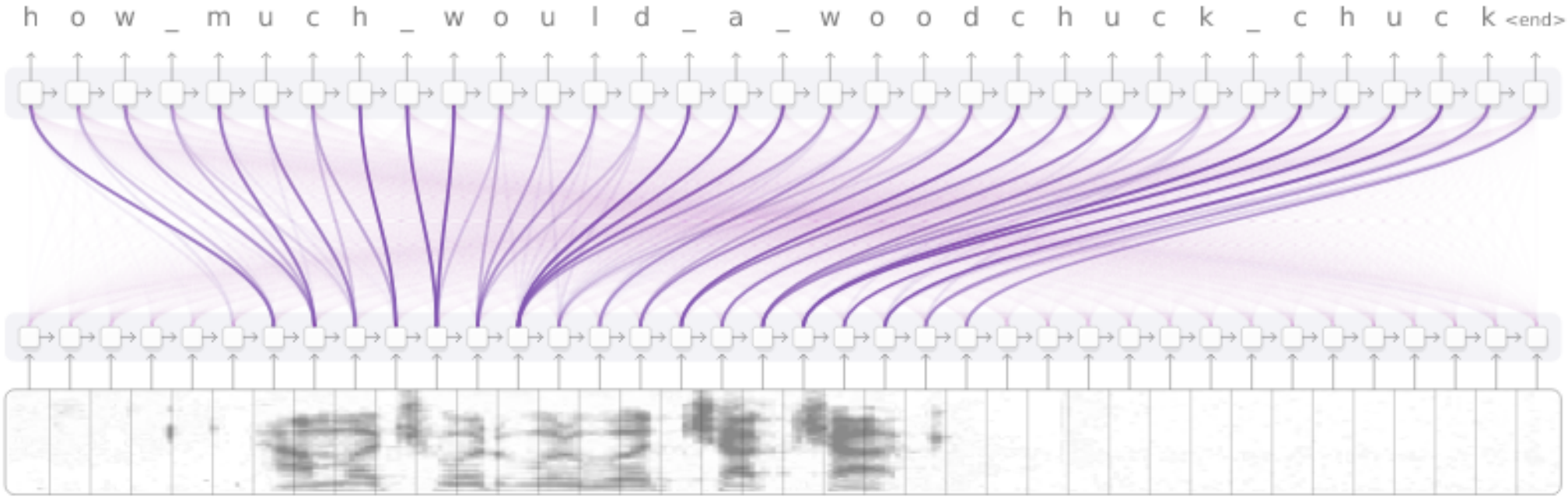
Space Folding



https://doi.org/10.48550/arXiv.1402.1869



distill.pub/2016/augmented-rnns/



distill.pub/2016/augmented-rnns/

References

Introductory:

- Buckner, C. (2019). Deep learning: A philosophical introduction. Philosophy Compass, 14(10), e12625. https://doi.org/10.1111/phc3.12625
- Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. Nature Methods, 15(4), 233.
- Liquet, B., Moka, S., Nazarathy, Y. (2022). The Mathematical Engineering of Deep Learning (draft): Introduction. https://deeplearningmath.org/drafts/chap1.pdf

General:

- Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.
- Murphy, K. P. (2012). Machine learning: A probabilistic perspective. MIT Press.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. http://www.deeplearningbook.org

REFERENCES

OPTIMIZATION & LEARNING:

- Kochenderfer, M. J., & Wheeler, T. A. (2019). Algorithms for Optimization. MIT Press.
- LeCun, Y., Chopra, S., Hadsell, R., Ranzato, M., & Huang, F. J. (2006). A Tutorial on Energy-Based Learning. http://yann.lecun.org/exdb/publis/pdf/lecun-06.pdf