8th Monday Brown Bag Seminar, 2024

Faculty of Economics, Thammasat University

The Use of Generative Adversarial Networks (GANs) for Economic Research

Theepakorn Jithitikulchai

Faculty of Economics, Thammasat University

March 11, 2024

Abstract

- Generative Adversarial Networks (GANs) are a powerful machine learning technique with exciting applications in economic research.
- This presentation starts with the exciting growing literature from intersection of machine learning (ML) and economics.
- We’ll review a foundation in core ML concepts, followed by a gentle introduction to deep learning, a powerful subfield.
- We'll then delve into Generative Adversarial Networks (GANs), a fascinating deep learning technique.
- Finally, the presentation will discuss the potential applications of GANs that can contribute to new economic insights and advancements.

Presentation Outline

- Recent Applications of Machine Learning in Economics
- Some Basics of Machine Learning
- Gentle Overview of Deep Learning
- Generative Adversarial Networks (GANs)
- Potential Uses of GANs in Economics

Recent Applications of Machine Learning in Economics

Recent Applications of Machine Learning in Economics

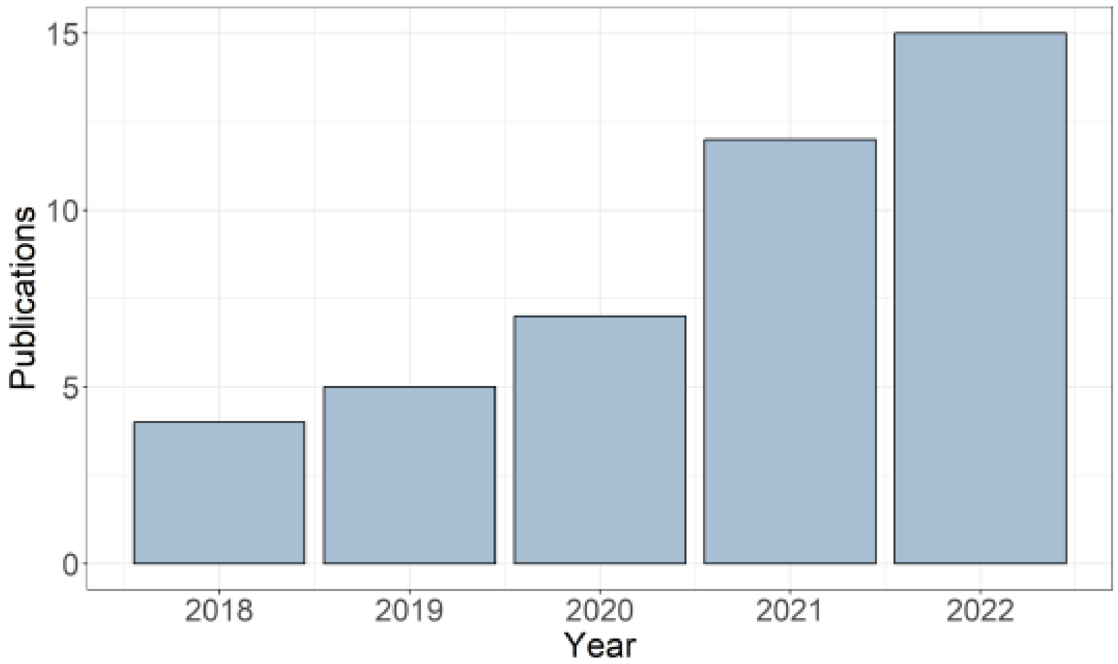
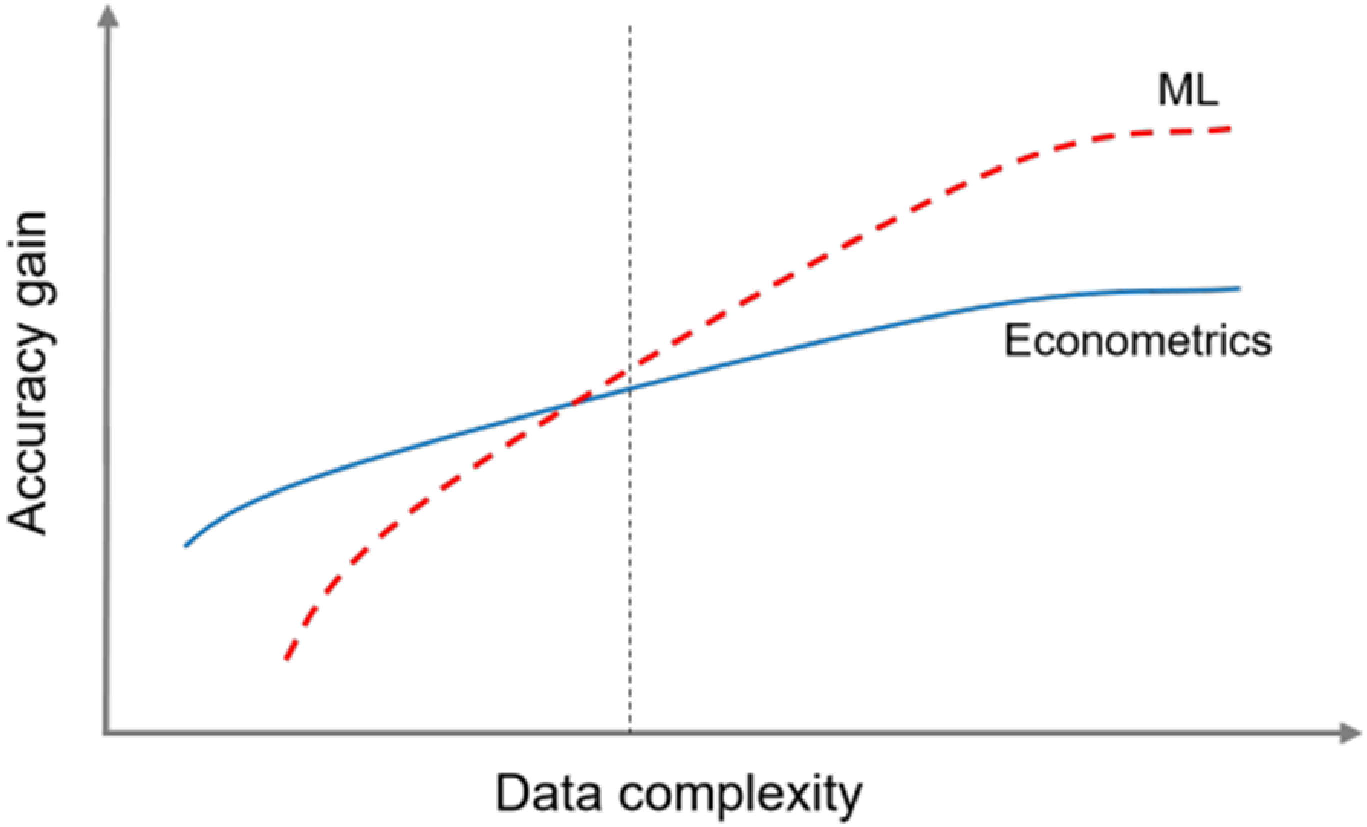


Figure 1: The number of publications over five years (between 2018-2022) in the leading economics journals that use ML. The data includes articles from the following ten journals: American Economic Review (AER), Econometrica, Journal of Economic Perspectives (JEP), Journal of Monetary Economics (JME), Journal of Political Economy (JPE), Journal of Econometrics (JoE), Quarterly Journal of Economics (QJE), Review of Economic Studies (RES), American Economic Journal (AJE): Macroeconomics and Microeconomics. The relevant papers are identified using the following search terms: Machine learning, Ensemble learning, Deep learning, Statistical learning, Reinforcement learning, and Natural language processing.

Desai, A., 2023. Machine Learning for Economics Research: When What and How?. Available at arXiv:2304.00086.
https://arxiv.org/abs/2304.00086

Recent Applications of Machine Learning in Economics

Schematic diagram representing the relative merits of ML and traditional econometric methods



Desai, A., 2023. Machine Learning for Economics Research: When What and How?. Available at arXiv:2304.00086.
https://arxiv.org/abs/2304.00086

Recent Applications of Machine Learning in Economics

ScienceDirect

Journals & Books

Joe TJ JT

Find articles with these terms

"machine learning" econometrics

Advanced search

5,496 results

Refine by:

Years

- 2024 (526)
- 2023 (1,232)
- 2022 (950)
- 2021 (735)
- 2020 (450)
- 2019 (349)
- 2018 (241)

Download selected articles | Export | sorted by *relevance* | date

Research article • *Open access*

1 Modeling and forecasting electricity consumption amid the COVID-19 pandemic: Machine learning vs. nonlinear econometric time series models

Sustainable Cities and Society, 14 August 2023

Lanouar Charfeddine, Esmat Zaidan, ... Ammar Abulibdeh

View PDF | Abstract | Graphical Abstract | Export

Research article

2 Using econometric and machine learning models to forecast crude oil prices: Insights from economic history

Resources Policy, 12 May 2023

Zilin Xu, Muhammad Mohsin, ... Xiaoyu Ma

https://www.sciencedirect.com/search?qs=%22machine%20learning%22%20econometrics

Recent Applications of Machine Learning in Economics

ScienceDirect

Journals & Books ? Joe TJ JT

Find articles with these terms

"deep learning" econometrics

Advanced search

1,544 results

Refine by:

Years

- 2024 (208)
- 2023 (436)
- 2022 (322)
- 2021 (228)
- 2020 (154)
- 2019 (88)
- 2018 (62)

Download selected articles Export

sorted by *relevance* | date

Research article

1 A novel perspective on forecasting non-ferrous metals' volatility: Integrating deep learning techniques with econometric models

Finance Research Letters, 28 September 2023

Qi Shu, Heng Xiong, ... Rogemar Mamon

Abstract Export

Research article

2 Synthesizing the market clearing mechanism based on the national power grid using hybrid of deep learning and econometric models: Evidence from the Japan Electric Power Exchange (JEPX) market

Journal of Cleaner Production, 29 April 2023

Mohamed Saad Suliman, Hooman Farzaneh

Abstract Export

https://www.sciencedirect.com/search?qs=%22deep%20learning%22%20econometrics

Recent Applications of Machine Learning in Economics

Annual Review of Economics

Machine Learning Methods That Economists Should Know About

Susan Athey[1,2,3] and Guido W. Imbens[1,2,3,4]

[1]Graduate School of Business, Stanford University, Stanford, California 94305, USA; email: athey@stanford.edu, imbens@stanford.edu

[2]Stanford Institute for Economic Policy Research, Stanford University, Stanford, California 94305, USA

[3]National Bureau of Economic Research, Cambridge, Massachusetts 02138, USA

[4]Department of Economics, Stanford University, Stanford, California 94305, USA

Athey, S. and Imbens, G.W., 2019. Machine learning methods that economists should know about. *Annual Review of Economics*, 11, pp.685-725.
https://www.annualreviews.org/doi/pdf/10.1146/annurev-economics-080217-053433

Recent Applications of Machine Learning in Economics

ELSEVIER

Journal of Monetary Economics
Volume 122, September 2021, Pages 76-101

Deep learning for solving dynamic economic models.

Lilia Maliar [a], Serguei Maliar [b], Pablo Winant [c]

[a] The Graduate Center, City University of New York, CEPR, and Hoover Institution, Stanford University
[b] Santa Clara University
[c] ESCP Business School and CREST/Ecole Polytechnique

Highlights

- We introduce a deep learning (DL) method that solves dynamic economic models by casting them into nonlinear regression equations.
- We derive such equations for three fundamental objects of economic dynamics - lifetime reward, Bellman equation and Euler equation.
- We propose all-in-one integration technique that facilitates construction of high-dimensional expectation functions.
- We use deep neural network to deal with multicollinearity and to perform model reduction.
- Taken together, these techniques enable us to solve economic models with thousands of state variables, such as Krusell and Smith (1998) model.
- We provide a TensorFlow code that accommodates a variety of applications.

https://www.sciencedirect.com/science/article/abs/pii/S0304393221000799

Recent Applications of Machine Learning in Economics

Journal of
APPLIED ECONOMETRICS

RESEARCH ARTICLE

How is machine learning useful for macroeconomic forecasting?

Philippe Goulet Coulombe, Maxime Leroux, Dalibor Stevanovic, Stéphane Surprenant

First published: 13 May 2022 | https://doi.org/10.1002/jae.2910 | Citations: 11

Summary

We move beyond *Is Machine Learning Useful for Macroeconomic Forecasting?* by adding the *how*. The current forecasting literature has focused on matching specific variables and horizons with a particularly successful algorithm. To the contrary, we study the usefulness of the underlying features driving ML gains over standard macroeconometric methods. We distinguish four so-called features (nonlinearities, regularization, cross-validation, and alternative loss function) and study their behavior in both the data-rich and data-poor environments. To do so, we design experiments that allow to identify the "treatment" effects of interest. We conclude that **(i)** nonlinearity is the true game changer for macroeconomic prediction, **(ii)** the standard factor model remains the best regularization, **(iii)** K-fold cross-validation is the best practice, and **(iv)** the L_2 is preferred to the $\bar{\epsilon}$-insensitive in-sample loss. The forecasting gains of nonlinear techniques are associated with high macroeconomic uncertainty, financial stress and housing bubble bursts. Furthermore, ML nonlinearities are helpful when considering density forecasts.

https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2910

Recent Applications of Machine Learning in Economics

European Review of Agricultural Economics

European Review of Agricultural Economics Vol **47** **(3)** (2020) pp. 849–892
doi:10.1093/erae/jbz033
Advance Access Publication 21 August 2019

Machine learning in agricultural and applied economics

Hugo Storm[†,*], Kathy Baylis[‡] and Thomas Heckelei[†]

[†]*Institute for Food and Resource Economics, University of Bonn, Germany;* [‡]*Agricultural and Consumer Economics, University of Illinois, USA*

Abstract

This review presents machine learning (ML) approaches from an applied economist's perspective. We first introduce the key ML methods drawing connections to econometric practice. We then identify current limitations of the econometric and simulation model toolbox in applied economics and explore potential solutions afforded by ML. We dive into cases such as inflexible functional forms, unstructured data sources and large numbers of explanatory variables in both prediction and causal analysis, and highlight the challenges of complex simulation models. Finally, we argue that economists have a vital role in addressing the shortcomings of ML when used for quantitative economic analysis.

https://academic.oup.com/erae/article/47/3/849/5552525

Recent Applications of Machine Learning in Economics

ELSEVIER

Energy Economics

Volume 81, June 2019, Pages 709-727

Machine learning in energy economics and finance: A review

Hamed Ghoddusi [a], Germán G. Creamer [a], Nima Rafizadeh [b]

[a] School of Business, Stevens Institute of Technology, Hoboken, NJ, USA

[b] Independent Researcher, Tehran, Iran

Highlights

- First comprehensive review of machine learning in energy economics
- Identified more than than 120 papers published in this area
- Support-Vector-Machine and Artificial Neural Networks found to be the most popular methods
- Crude oil and electricity price predictions are the two most frequent domain applications.
- Opportunities to apply ML techniques to energy-related volatility prediction, social network analysis, and text processing

https://www.sciencedirect.com/science/article/abs/pii/S0140988319301513?via%3Dihub

Recent Applications of Machine Learning in Economics

The Econometrics Journal

Econometrics Journal (2018), volume **21**, pp. C1–C68.
doi: 10.1111/ectj.12097

Cited by 2625 (as of 3/10/2024)

Double/debiased machine learning for treatment and structural parameters

VICTOR CHERNOZHUKOV[†], DENIS CHETVERIKOV[‡], MERT DEMIRER[†], ESTHER DUFLO[†], CHRISTIAN HANSEN[§], WHITNEY NEWEY[†] AND JAMES ROBINS[‖]

[†]*Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, USA.*
E-mail: vchern@mit.edu, mdemirer@mit.edu, duflo@mit.edu, wnewey@mit.edu

[‡]*University of California Los Angeles, 315 Portola Plaza, Los Angeles, CA 90095, USA.*
E-mail: chetverikov@econ.ucla.edu

[§]*University of Chicago, 5807 S. Woodlawn Ave., Chicago, IL 60637, USA.*
E-mail: chansen1@chicagobooth.edu

[‖]*Harvard University, 677 Huntington Avenue, Boston, MA 02115, USA.*
E-mail: robins@hsph.harvard.edu

First version received: October 2016; final version accepted: June 2017

Double/Debiased Machine Learning (DML) offers a solution to the challenge of estimating low-dimensional parameters (θ_0) in the presence of high-dimensional nuisance parameters (η_0) that traditional methods struggle with. DML achieves this by employing two key techniques: Neyman-orthogonal moments/scores and cross-fitting. This approach mitigates the bias introduced by regularization and overfitting in standard machine learning methods. As a result, DML delivers accurate and reliable point estimates with theoretical guarantees of approximately unbiased and normally distributed, allowing for valid statistical inference. The framework is flexible and can be applied with various modern machine learning algorithms such as deep neural networks.

https://academic.oup.com/ectj/article/21/1/C1/5056401

Recent Applications of Machine Learning in Economics

Among many others

Google Scholar "Machine Learning" "Economics"

Articles About 2,110,000 results (0.06 sec)

Any time
Since 2024
Since 2023
Since 2020
Custom range...

Sort by relevance
Sort by date

Any type
Review articles

[HTML] **Machine learning** in energy **economics** and finance: A review
H Ghoddusi, GG Creamer, N Rafizadeh - Energy **Economics**, 2019 - Elsevier
… **Machine learning** (ML) is generating new opportunities for innovative research in energy **economics** … We critically review the burgeoning literature dedicated to Energy **Economics**/…
☆ Save 99 Cite Cited by 335 Related articles All 9 versions

The impact of **machine learning** on **economics**
S Athey - The **economics** of artificial intelligence: An agenda, 2018 - degruyter.com
… I believe that **machine learning** (ML) will have a dramatic impact on the field of **economics** … Next, I review some applications of ML in **economics** where ML can be used off the shelf: the …
☆ Save 99 Cite Cited by 890 Related articles All 16 versions

Google Scholar "Machine Learning" "Economics"

Articles About 15,900 results (0.06 sec)

Any time
Since 2024
Since 2023
Since 2020
Custom range...

Sort by relevance
Sort by date

Any type
Review articles

Law & **Economics** at sixty: Mapping the field with bibliometric and **machine learning** tools
E Kantorowicz-Reznichenko… - Journal of Economic …, 2024 - Wiley Online Library
… We demonstrate that Law & **Economics** shifted from more theory driven work to empirical … Law & **Economics** field tends to be dominated by authors affiliated with **economics** departments…
☆ Save 99 Cite Related articles

Computer and Mathematical Modeling: Translational Research and **Economics** in Clinical Diagnostics
J Wang, Y Qin, W Si, L Li - Electronics Science Technology and …, 2024 - ojs.as-pub.com
… One way that **machine learning** can help to overcome these … gene expression, and **machine learning** models to predict drug … phase in which a **machine-learning** approach yields a list of …
☆ Save 99 Cite Related articles

Note: While these numbers provide initial insights, further investigation is necessary for conclusive interpretation from the data's limitations and representation.

Recent Applications of Machine Learning in Economics

AMERICAN ECONOMIC ASSOCIATION

Home › ASSA Annual Meeting › AEA Continuing Education › 2023 Continuing Education Webcasts

2023 Continuing Education Webcasts

January 8-10, 2023 New Orleans, Lousiana

The AEA's 2023 Continuing Education Program was held on January 8-10, 2023, at the Sheraton New Orleans Hotel.

Machine Learning and Big Data—Melissa Dell and Matthew Harding

Webcasts Day One:
Part 1

Webcasts Day Two:
Part 2 | Part 3 | Part 4 | Part 5 | Part 6

Webcasts Day Three:
Part 7 | Part 8 | Part 9

Melissa Dell | Harvard University

Melissa Dell is the Andrew E. Furer Professor of Economics at Harvard University. She is the 2020 recipient of the John Bates Clark Medal, awarded each year to an American economist under the age of forty who is judged to have made the most significant contribution to economic thought and knowledge. In 2018, *The Economist* named her one of the decade's eight best young economists, and in 2014 she was named by the IMF as the youngest of 25 economists under the age of 45 shaping thought about the global economy. Her research focuses on economic growth and political economy. She has examined the factors leading to the persistence of poverty and prosperity in the long run, the effects of trade-induced job loss on crime, the impacts of U.S. foreign intervention, and the effects of weather on economic growth. She has also developed deep learning powered methods for curating social science data at scale, released in the open-source package Layout Parser. This work supports many of her current projects, which rely on digitizing historical sources far too large for manual digitization. Professor Dell is a senior scholar at the Harvard Academy for Area and International Studies and a research associate at the National Bureau of Economic Research. She received an AB in Economics from Harvard in 2005, an MPhil in Economics from Oxford in 2007, and a PhD in Economics from MIT in 2012. Before joining the Harvard Economics department in 2014, she was a Junior Fellow at the Harvard Society of Fellows.

Matthew Harding | University of California, Irvine

Matthew Harding is a Professor of Economics and Statistics at the University of California, Irvine. He is an Econometrician and Data Scientist who develops techniques at the intersection of Machine Learning and Econometrics to answer Big Data questions related to individual and firm behavior in areas such as energy, consumer finance, and health. He directs the Deep Data Lab which conducts research into cutting edge methods, inspired by recent advances in machine learning, for the analysis of "deep data", large and information-rich data sets derived from many seemingly unrelated sources to provide novel economic insights to economics and business. Professor Harding focuses on developing triple-win strategies that balance individual and corporate welfare with broader policy and societal goals while maintaining a core focus on evaluating causal determinants. He has extensive experience working with industry leaders and regulators and advised major companies, agencies, and organizations. He is a Co-Founder of FASTLab.Global, a non-profit thought academy aimed at providing research and mentoring opportunities in areas of fair access and sustainable technologies for underprivileged minorities. Professor Harding is also a co-author of the best-selling textbook *Modern Business Analytics: Practical Data Science for Decision-Making* published by McGraw-Hill. He received a BA in Philosophy and Economics from University College London, an MPhil in Economics from the University of Oxford, and a PhD in Economics from MIT. Before joining UC Irvine, he taught at Stanford University and Duke University.

https://www.aeaweb.org/conference/cont-ed/2023-webcasts

Recent Applications of Machine Learning in Economics

ES Society Econometric The Econometric Society

Log In Become A Member Make A Donation

Home Society Publications 2025 World Congress Regional Activities Job Postings Membership Contact Sponsors And Donors

Home > Society > News > Announcing the First Conference in a New Interdisciplinary Series: The 2024 ESIF Economics and AI+ML Meeting

Society

About the Society Video
About
News
Organization And Governance
Special Lectures
Awards

Follow Us On...

@econometricsoc
@ecmaEditors
@EconTheory
@qe_editors

@TheEconometricSociety

@econometricsociety

@the-econometric-society

Announcing the First Conference in a New Interdisciplinary Series: The 2024 ESIF Economics and AI+ML Meeting

Share:

The 2024 ESIF Economics and AI+ML Meeting
August 13 - 14, 2024,
Cornell University, Ithaca, United States

The Econometric Society is pleased to announce an interdisciplinary conference on Economics and AI+ML, the first in a series of Econometric Society Interdisciplinary Frontiers (ESIF) conferences.

The purpose of the Economics and AI+ML Meeting, which will be held at Cornell University, in Ithaca NY (USA) on August 13-14, 2024, is to foster interaction of ideas and methodologies from the areas of Computer Science and Economics (broadly defined, but with emphasis on AI and ML). The conference will feature keynote lectures and parallel sessions, bringing together scholars from both fields.

The ESIF series more broadly will promote interdisciplinary approaches to important economic issues and global challenges, with each conference hosted or co-hosted by various regions of the Econometric Society. The 2024 ESIF Economics and AI+ML Meeting is hosted by the North America region.

Important Dates for the 2024 ESIF Economics and AI+ML Meeting

Submissions open: December 20, 2023
Paper Submission Period: December 20, 2023 – February 25, 2024
Decision Notification Deadline: April 21, 2024
Registration Period (for presenters): April 21, 2024 – May 4, 2024
Preliminary Program Announcement: May 25, 2024

Keynote Speakers

Susan Athey (Stanford University)
David Blei (Columbia University)
Avrim Blum (Toyota Technological Institute at Chicago)
Jesus Fernandez-Villaverde (University of Pennsylvania)
Michael I. Jordan (University of California, Berkeley)
Whitney Newey (MIT)

Recent Applications of Machine Learning in Economics

Lyn Hogan, Executive Director Sat, Mar 9, 2024 at 3:19 AM
<announcement@econometricsociety.org>
Reply-To: announcement@econometricsociety.org
To: TJ <theepakorn@econ.tu.ac.th>

Having trouble reading this email? View it on your browser.
Forward this email to a friend. • Share this newsletter on Facebook

The Econometric Society
An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

March 8, 2024 www.econometricsociety.org

AI Replication Game @ Cornell University 2024
August 12, 2024

We are pleased to announce that the Institute for Replication (I4R) and the Labor Dynamics Institute @ Cornell will be organizing the AI Replication Game @ Cornell 2024 on **August 12, 2024, in Ithaca, NY.** The event will take place on the day before the start of the 2024 ESIF Economics and AI+ML Meeting (ESIF-AIML2024). For more information and to register for the AI Replication Game, please see the information below. To separately register for the 2024 ESIF Economics and AI+ML Meeting (ESIF-AIML2024), please see the event page.

The AI RG is organized by I4R and the Labor Dynamics Institute @ Cornell. The AI Replication Game @ Cornell University is a one-day event on **August 12, 2024** that brings researchers together to collaborate on reproducing quantitative papers published in high-ranking social science journals. Replication is a crucial aspect of scientific research, ensuring that results are reliable and reproducible. By participating in the Replication Games, you will not only contribute to the integrity of research in your field, but also have the opportunity to network with fellow researchers and develop your coding skills. Replication Games are open to faculty-level researchers, post-docs, and graduate students.

Participation: Researchers participating in the AI Replication Game @ Cornell University will be randomly assigned to one of three teams: Skynet, Cyborg or John

https://www.econometricsociety.org/regional-activities/schedule/2024/08/13/2024-ESIFEconomics-and-AIML-Meeting#ai-replication-game

Connor. Skynet and Cyborg teams will have access to (commercially available) LLM models to conduct their work; John Connor teams of course rely only on unaugmented human skills. Each team consists of 3 members with similar research interests and varying skill levels. Teams will be asked to check for coding errors and conduct a robustness reproduction, which is the ability to duplicate the results of a prior study using the same data but different procedures as were used by the original investigator.

During the event, participants are expected to read the paper and familiarize themselves with the replication package. Teams then will work together to check for coding errors and conduct sensitivity analysis. No work is conducted before/after the event other than answering a short survey.

Schedule: The AI Replication Game @ Cornell University will start at 8:45 AM local time, and end at 4PM local time.

Please register here: before July 15th, 2024.

The Econometric Society
www.econometricsociety.org

Some Basics of Machine Learning

What is Machine Learning ?

Machine Learning

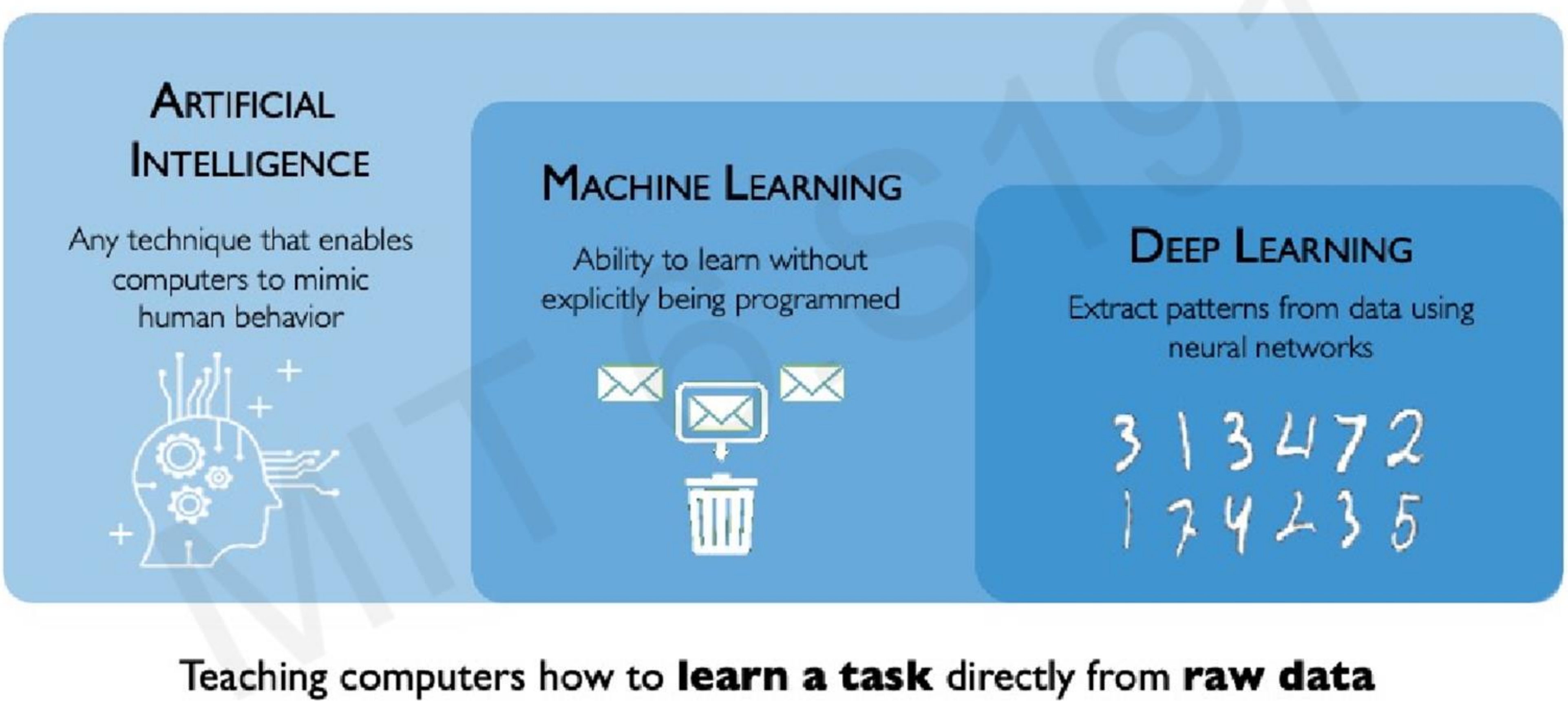
A relatively new approach to **data analytics**, which places itself in the intersection between **statistics**, **computer science**, and **artificial intelligence**

ML objective

Turning **information** into **knowledge** and **value** by "**letting the data speak**"

Giovanni Cerulli. *A review of machine learning commands in Stata: Performance and usability evaluation*. UK Stata Conference 2023.
https://www.stata.com/meeting/uk23/slides/UK23_Cerulli.pdf

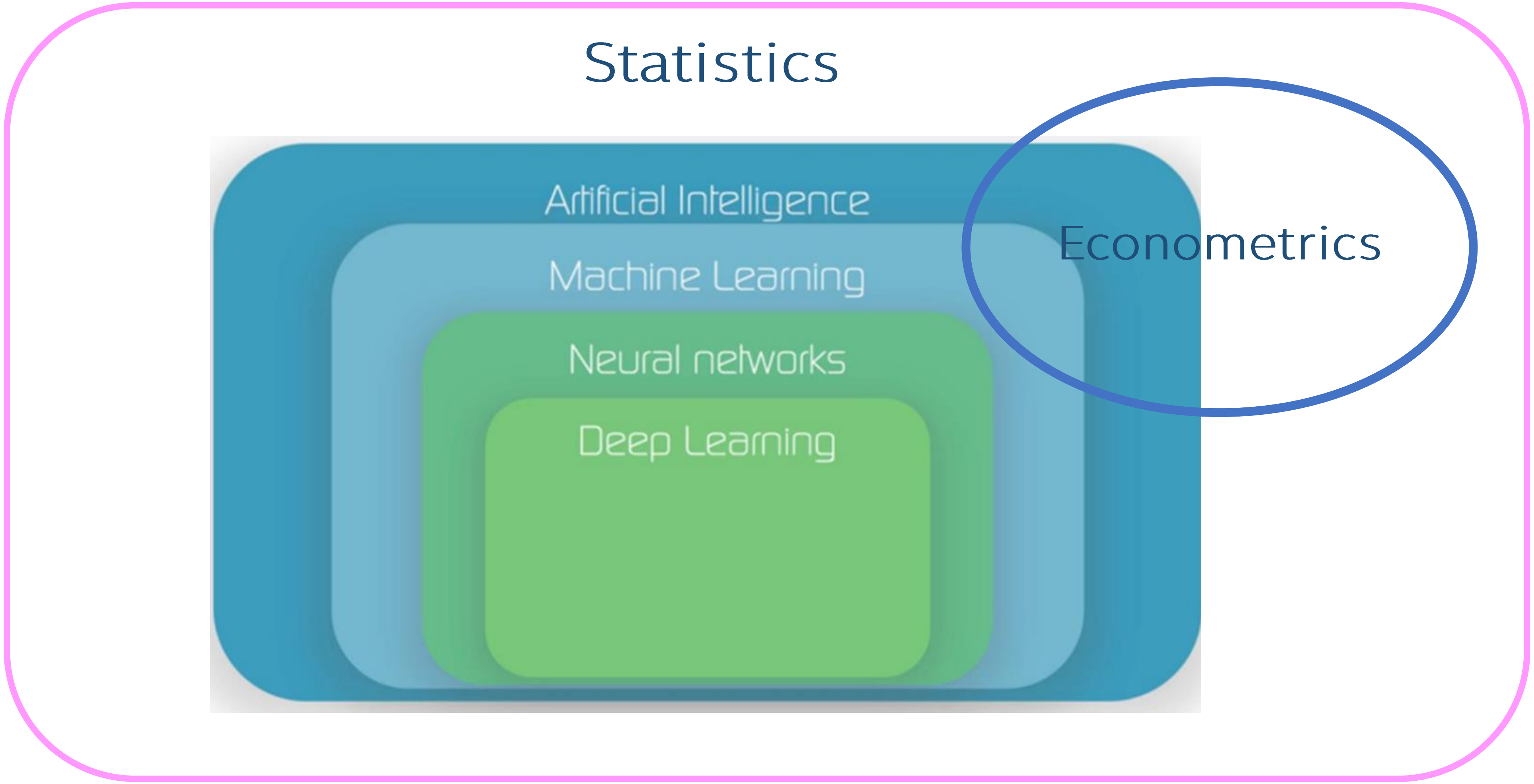
Some Basics of Machine Learning



Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Some Basics of Machine Learning

My personal view about the universe



Adapted from
https://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf

Some Basics of Machine Learning

ML purposes

Limiting **prior assumptions**

Model-free philosophy

Based on **algorithm computation, graphics**

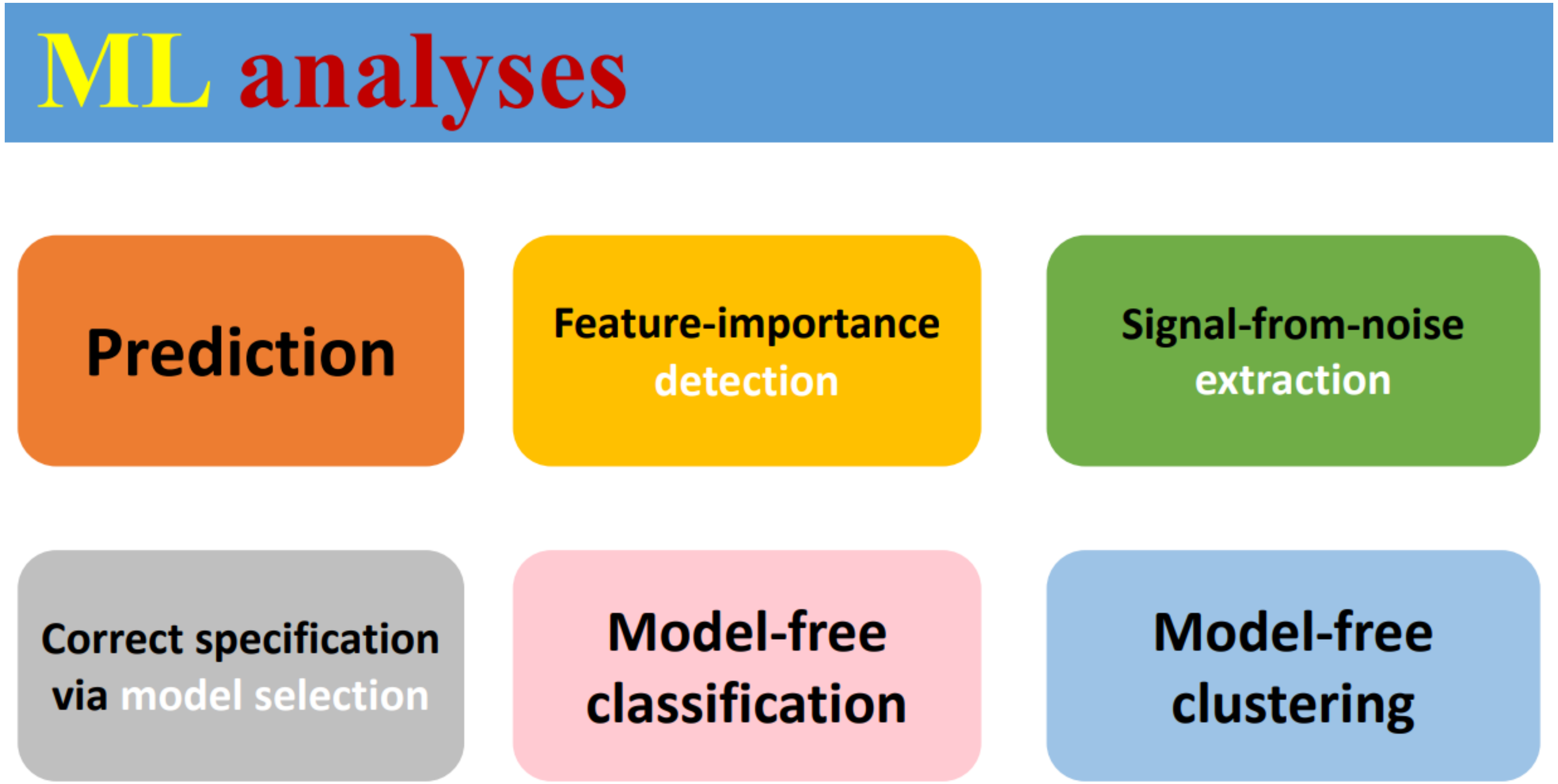
Mostly focused on **prediction** than **inference**

Targeted to **Big Data**

Targeted to **complexity reduction**

Giovanni Cerulli. *A review of machine learning commands in Stata: Performance and usability evaluation.* UK Stata Conference 2023.
https://www.stata.com/meeting/uk23/slides/UK23_Cerulli.pdf

Some Basics of Machine Learning



Giovanni Cerulli. *A review of machine learning commands in Stata: Performance and usability evaluation.* UK Stata Conference 2023.
https://www.stata.com/meeting/uk23/slides/UK23_Cerulli.pdf

Some Basics of Machine Learning

Supervised vs Unsupervised Learning

Supervised Learning

Data: (x, y)
x is data, y is label

Goal: Learn a *function* to map x -> y

Examples: Classification, regression, object detection, semantic segmentation, image captioning, etc.

Unsupervised Learning

Data: x
Just data, no labels!

Goal: Learn some underlying hidden *structure* of the data

Examples: Clustering, dimensionality reduction, feature learning, density estimation, etc.

Fei-Fei Li: Lecture 13 Generative Models, Convolutional Neural Networks for Visual Recognition (Spring 2017)
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Some Basics of Machine Learning

Prediction in a Stable Environment

Goal: estimate $\mu(x) = E[Y|X = x]$ and minimize MSE in a new dataset where only X is observed

- MSE: $\frac{1}{I}\sum_i \left(Y_i - \hat{\mu}(X_i)\right)^2$
- No matter how complex the model, the output, the prediction, is a single number
- Can hold out a test set and evaluate the performance of a model
- Ground truth is observed in a test set
- Only assumptions required: independent observations, and joint distribution of (Y,X) same in test set as in training set

Note: minimizing MSE entails bias-variance tradeoff, and always accept some bias

- Idea: if estimator too sensitive to current dataset, then procedure will be variable across datasets
- Models are very rich, and overfitting is a real concern, so approaches to control overfit necessary

Idea of ML algorithms

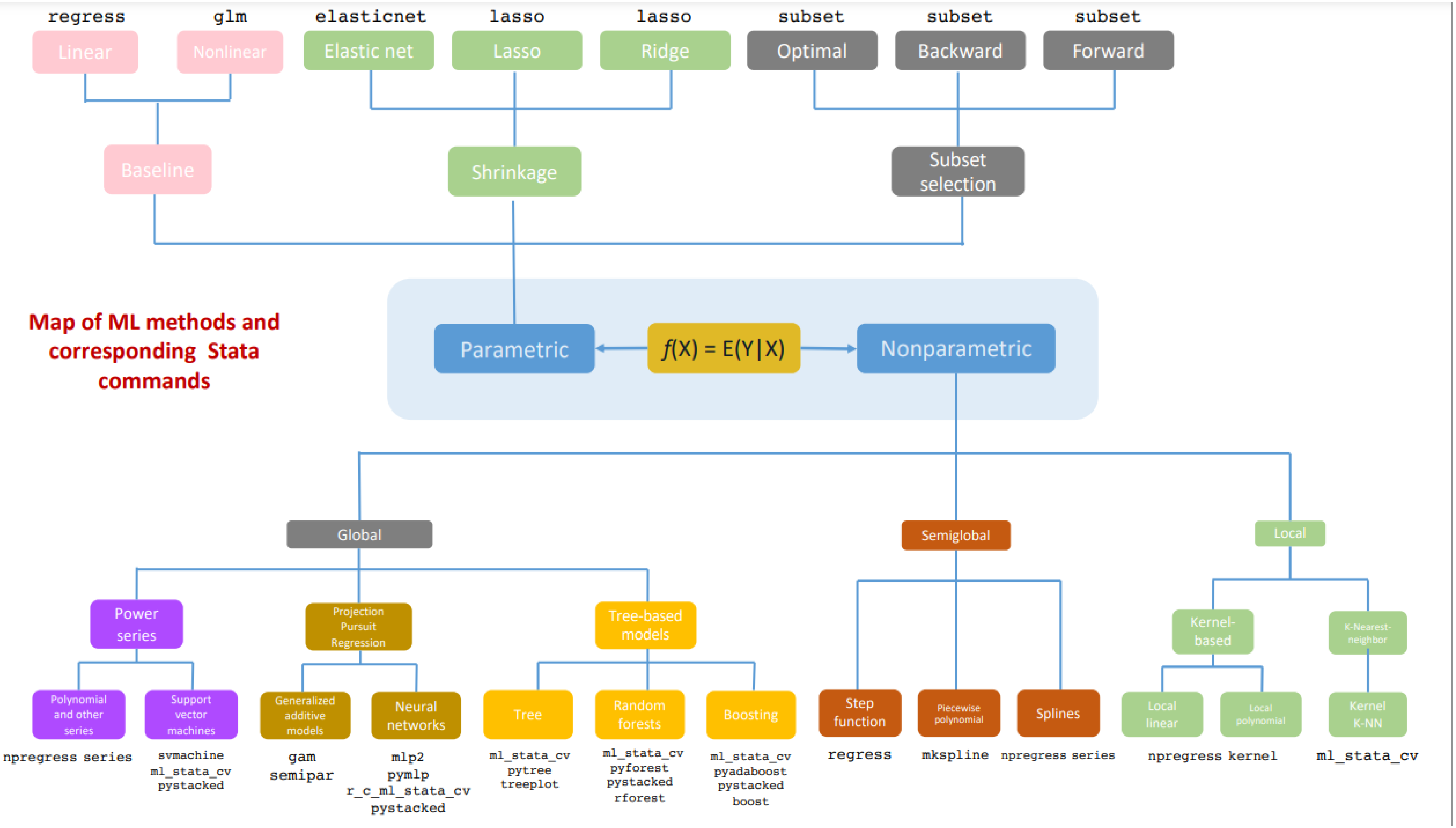
- Consider a family of models
- Use the data to select among the models or choose tuning parameters
- Common approach: cross-validation
  - Break data into 10 folds
  - Estimate on 9/10 of data, estimate MSE on last tenth, for each of a grid of tuning parameters
  - Choose the parameters that minimize MSE

ML works well because you can accurately evaluate performance without add'l assumptions

- Your robotic research assistant then tests many models to see what performs best

Susan Athey: *Machine Learning and Economics: An Introduction* (Machine Learning & Causal Inference: A Short Course, 2021)
https://www.youtube.com/watch?v=Z0ZcsxI-HTs

Supervised Machine learning in Stata



Giovanni Cerulli. *A review of machine learning commands in Stata: Performance and usability evaluation*. UK Stata Conference 2023. https://www.stata.com/meeting/uk23/slides/UK23_Cerulli.pdf

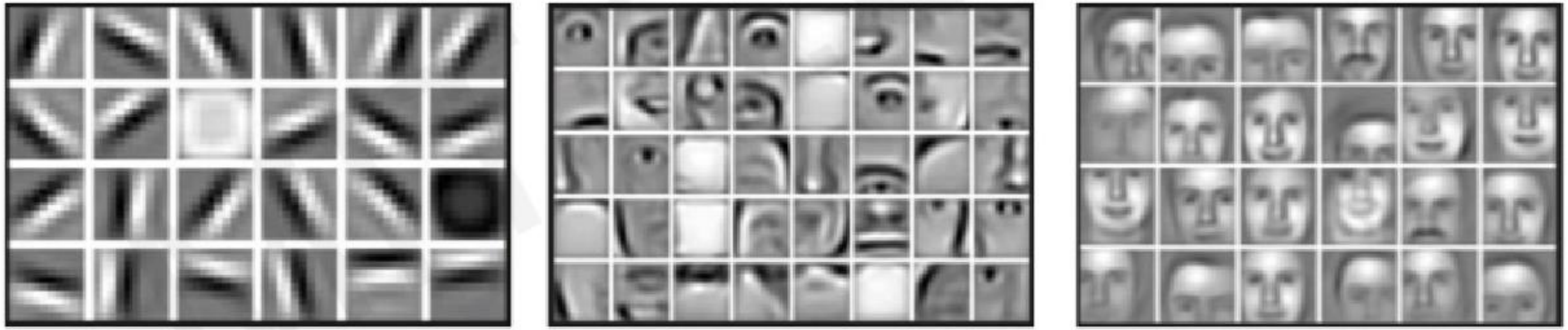
Gentle Overview of Deep Learning

Why deep learning?

Hand engineered features are time consuming, brittle, and not scalable in practice

Can we learn the **underlying features** directly from data?

Low Level Features | **Mid Level Features** | **High Level Features**



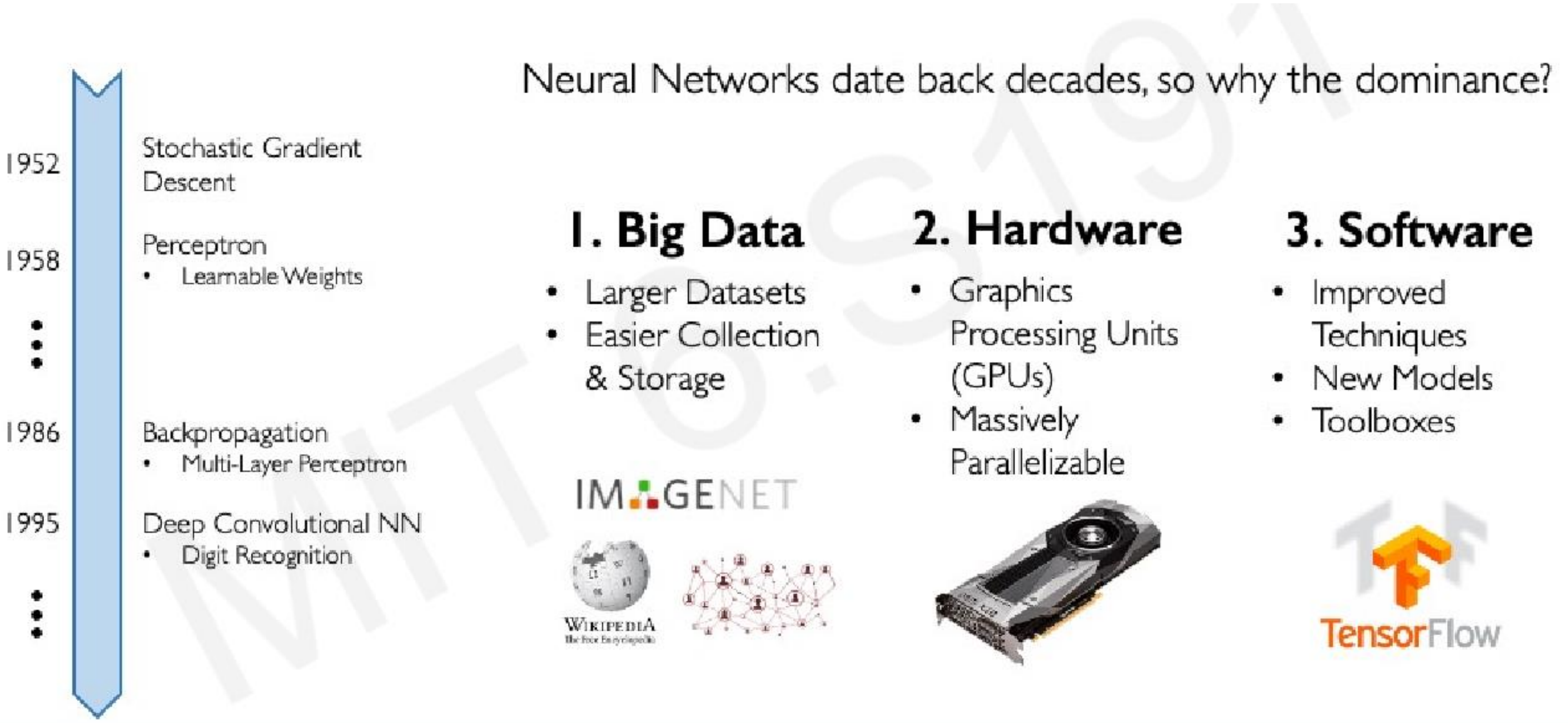
Lines & Edges | Eyes & Nose & Ears | Facial Structure

Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

My undereducated view on feature levels in time series data

- Low-level features
  - Mean, standard deviation, minimum, maximum, skewness, kurtosis
  - Trends like slopes or moving averages, and cyclical/seasonal patterns
- Mid-level features
  - Autocorrelation or partial autocorrelation
  - Empirical density functions
  - Event detection flags e.g., peaks, anomalies
  - Brownian motion of noise behaviors
- High-level features
  - Hidden states or latent factors representing the core dynamics of the series
  - Fermi-Dirac superdistribution, Wilson loop perceptron

Why now?

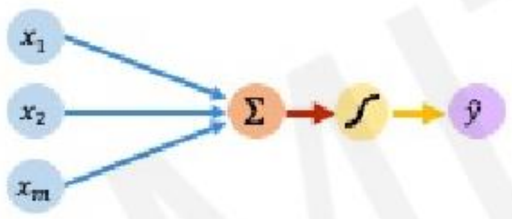


Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Core Foundation

The Perceptron

- Structural building blocks
- Nonlinear activation functions

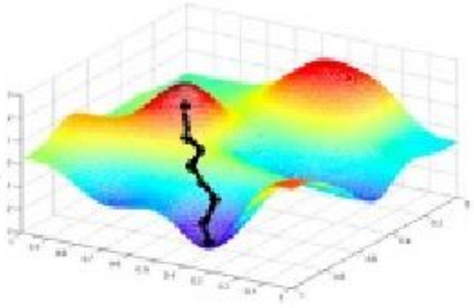


Neural Networks

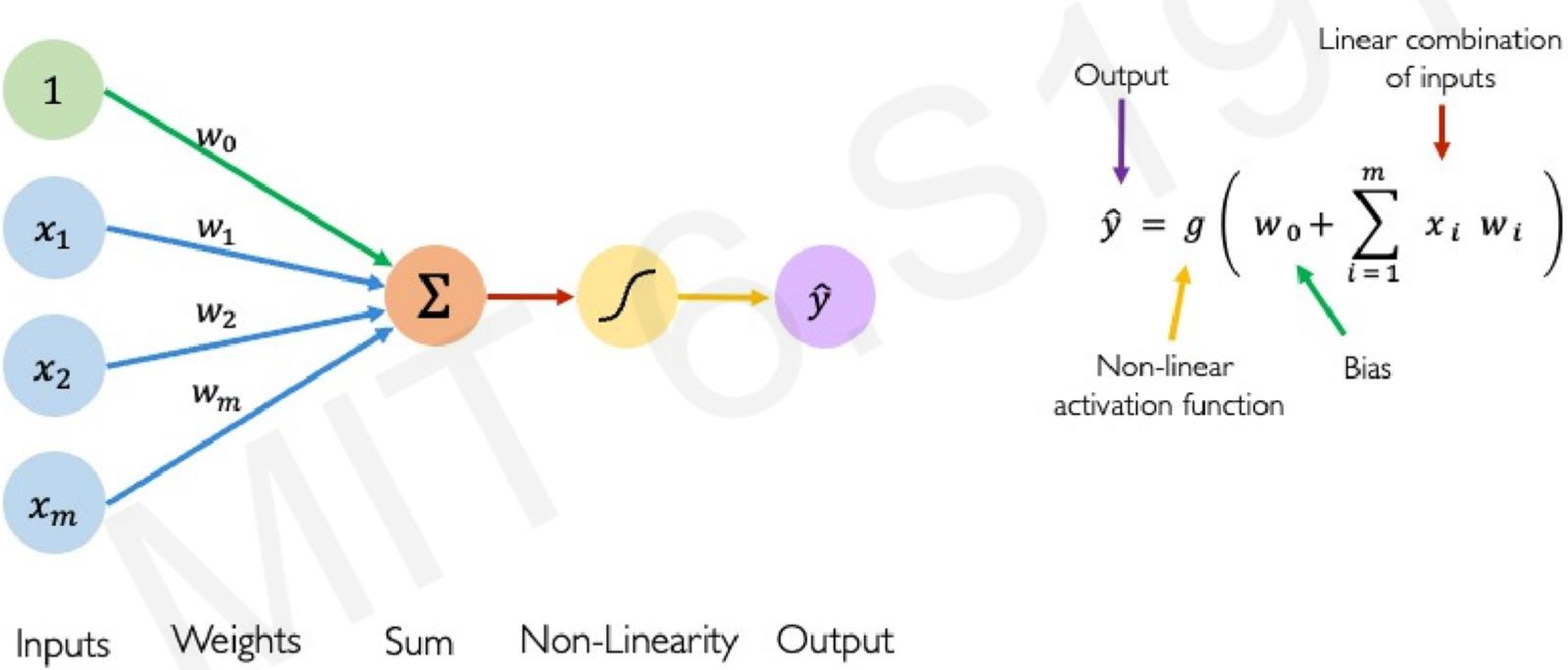
- Stacking Perceptrons to form neural networks
- Optimization through backpropagation

Training in Practice

- Adaptive learning
- Batching
- Regularization

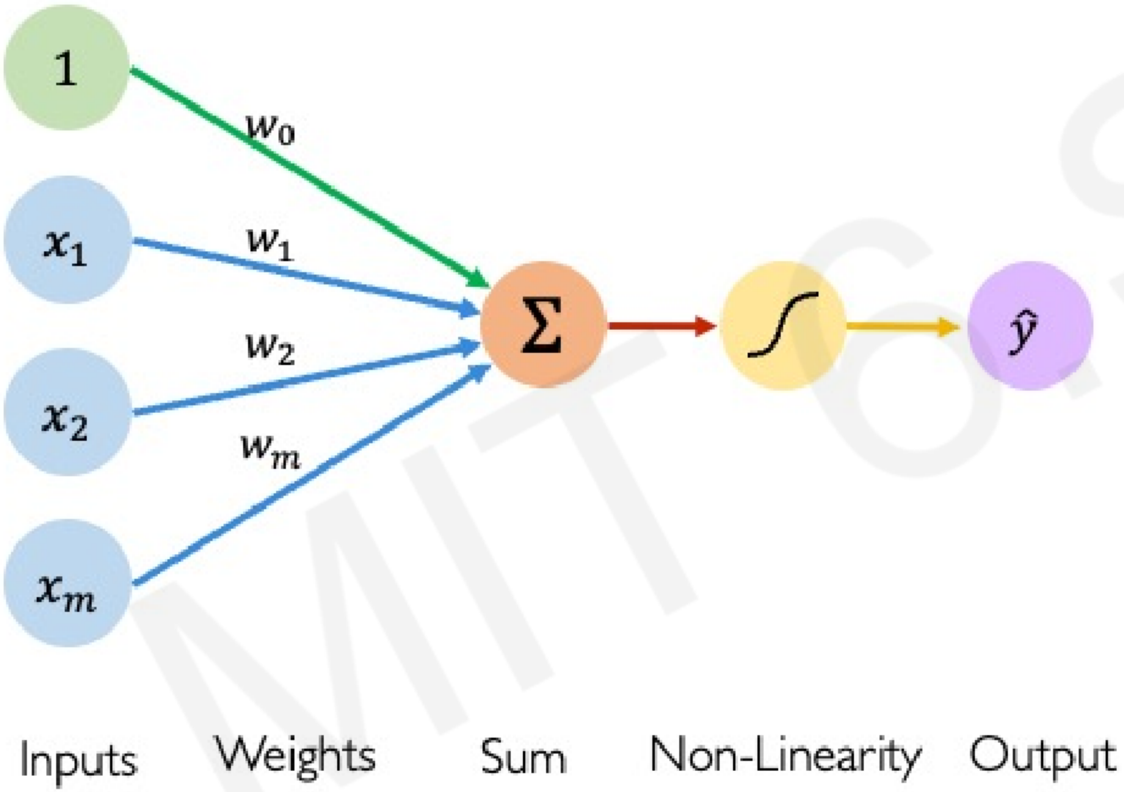


The Perceptron: Forward Propagation



The Perceptron: The structural building block of deep learning

The Perceptron: Forward Propagation

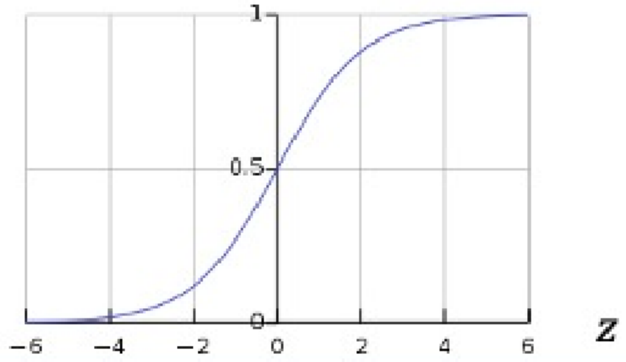


Activation Functions

$$\hat{y} = g\left(w_0 + \boldsymbol{X}^T \boldsymbol{W} \right)$$

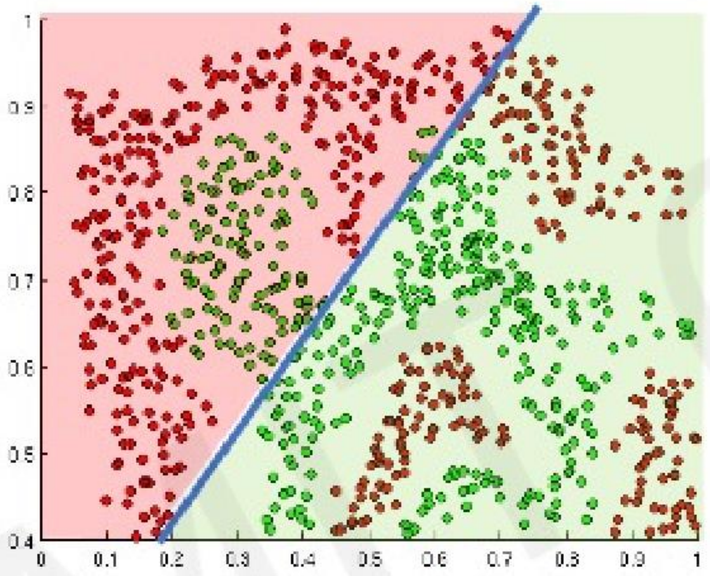
- Example: sigmoid function

$$g(z) = \sigma(z) = \frac{1}{1 + e^{-z}}$$

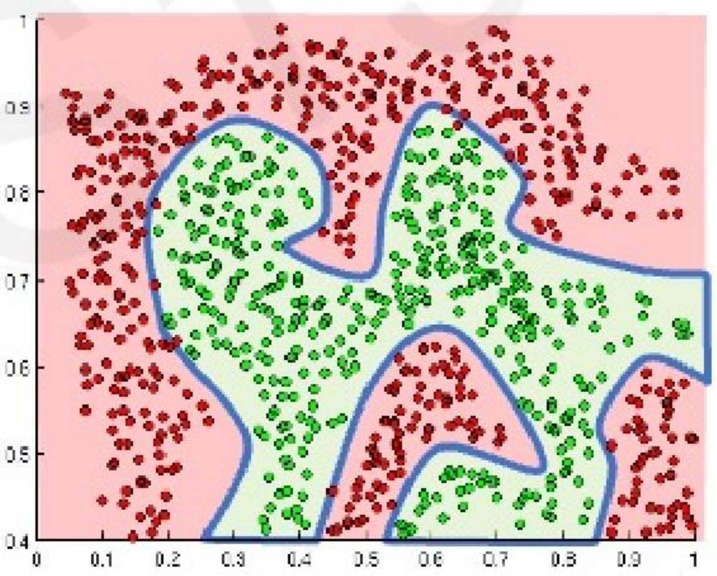


Importance of Activation Functions

The purpose of activation functions is to ***introduce non-linearities*** *into the network*



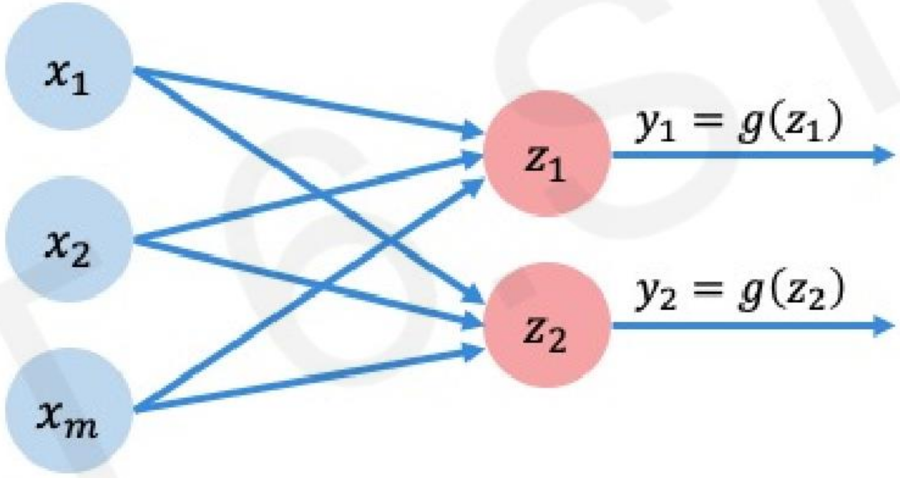
Linear activation functions produce linear decisions no matter the network size



Non-linearities allow us to approximate arbitrarily complex functions

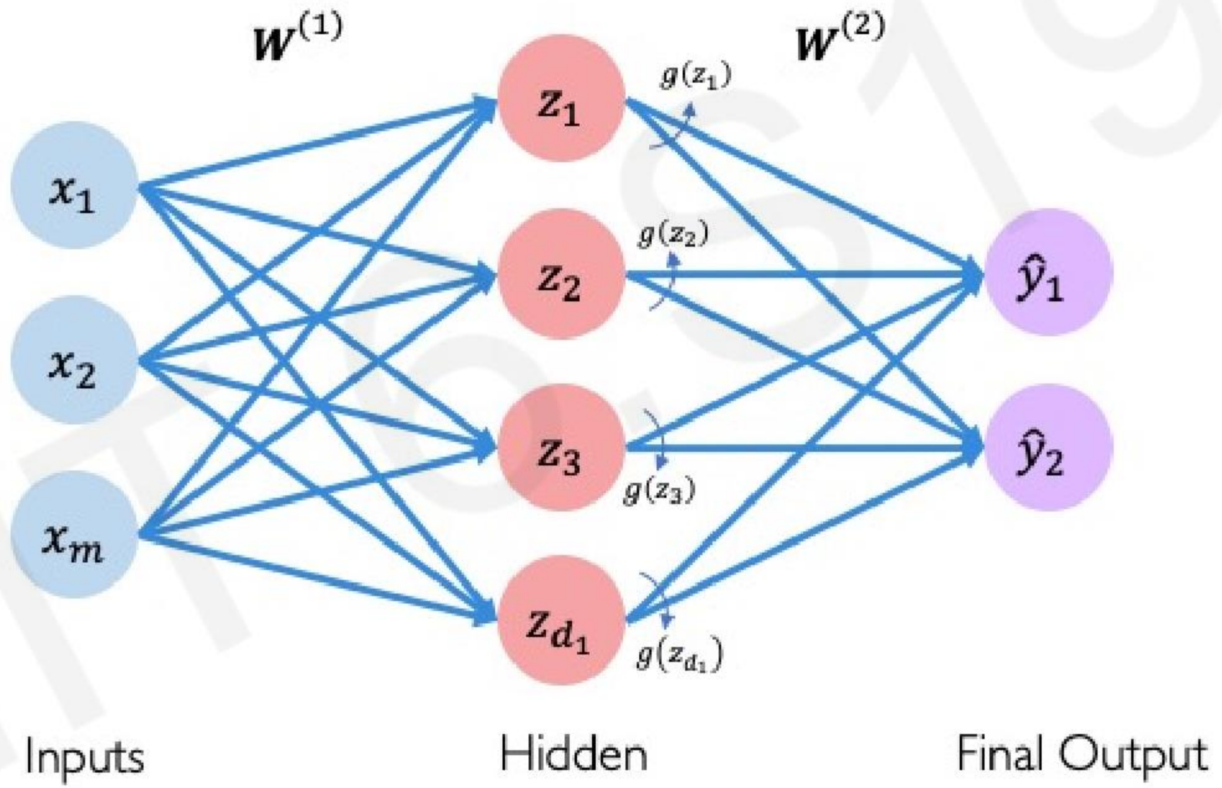
Multi Output Perceptron

Because all inputs are densely connected to all outputs, these layers are called **Dense** layers



$$z_i = w_{0,i} + \sum_{j=1}^{m} x_j \, w_{j,i}$$

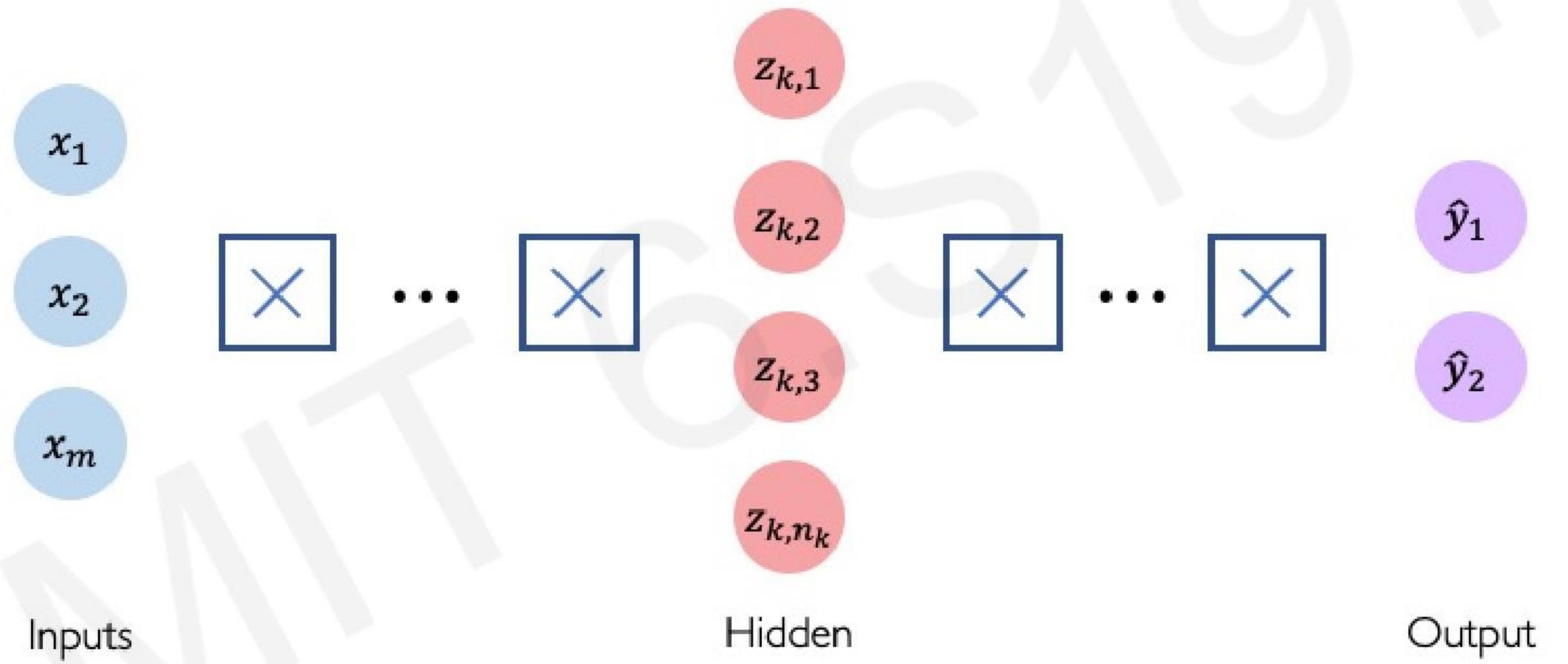
Single Layer Neural Network



$$z_i = w_{0,i}^{(1)} + \sum_{j=1}^{m} x_j \, w_{j,i}^{(1)} \qquad \hat{y}_i = g\left(w_{0,i}^{(2)} + \sum_{j=1}^{d_1} g(z_j) \, w_{j,i}^{(2)}\right)$$

Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

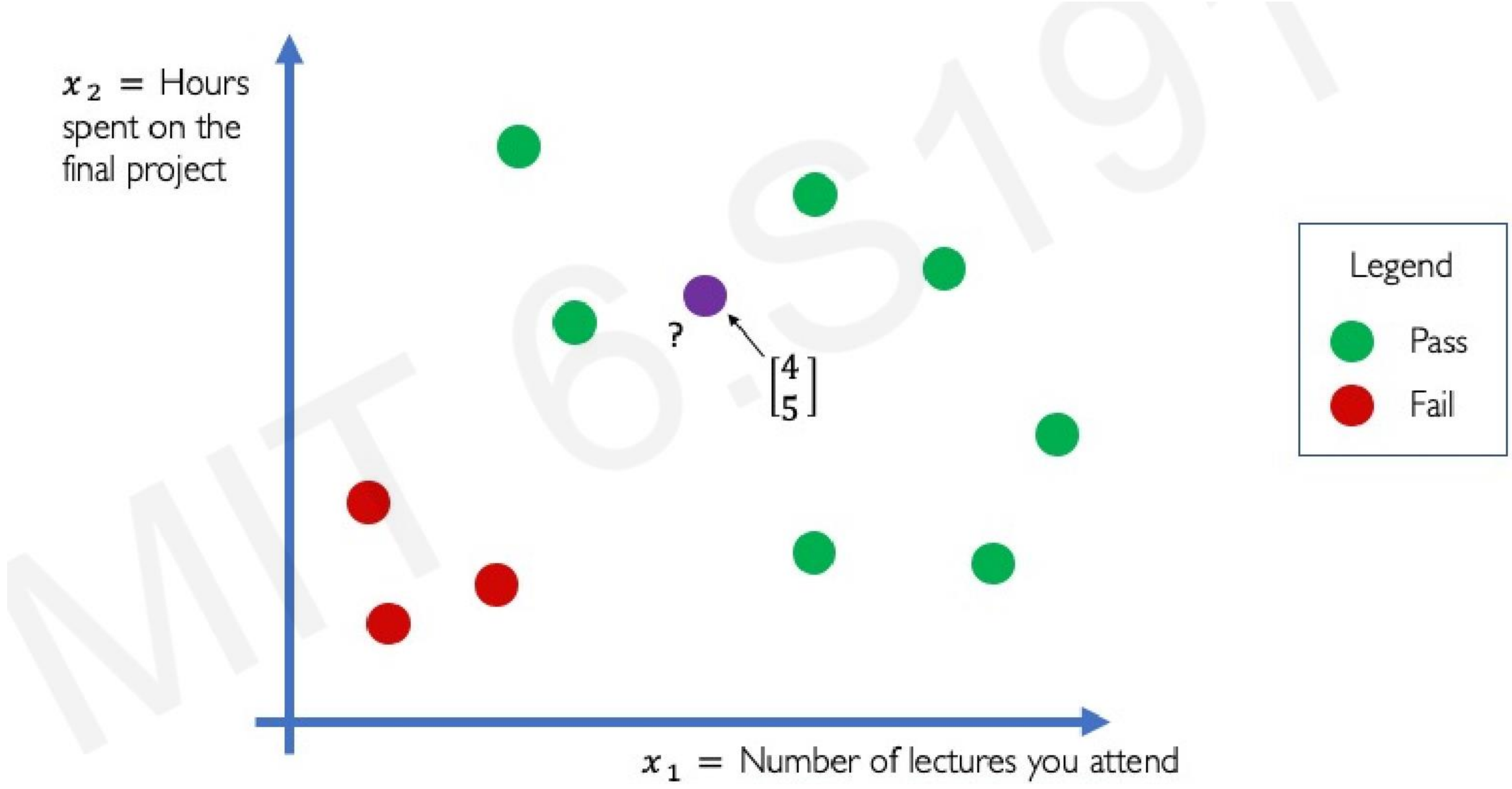
Deep Neural Network



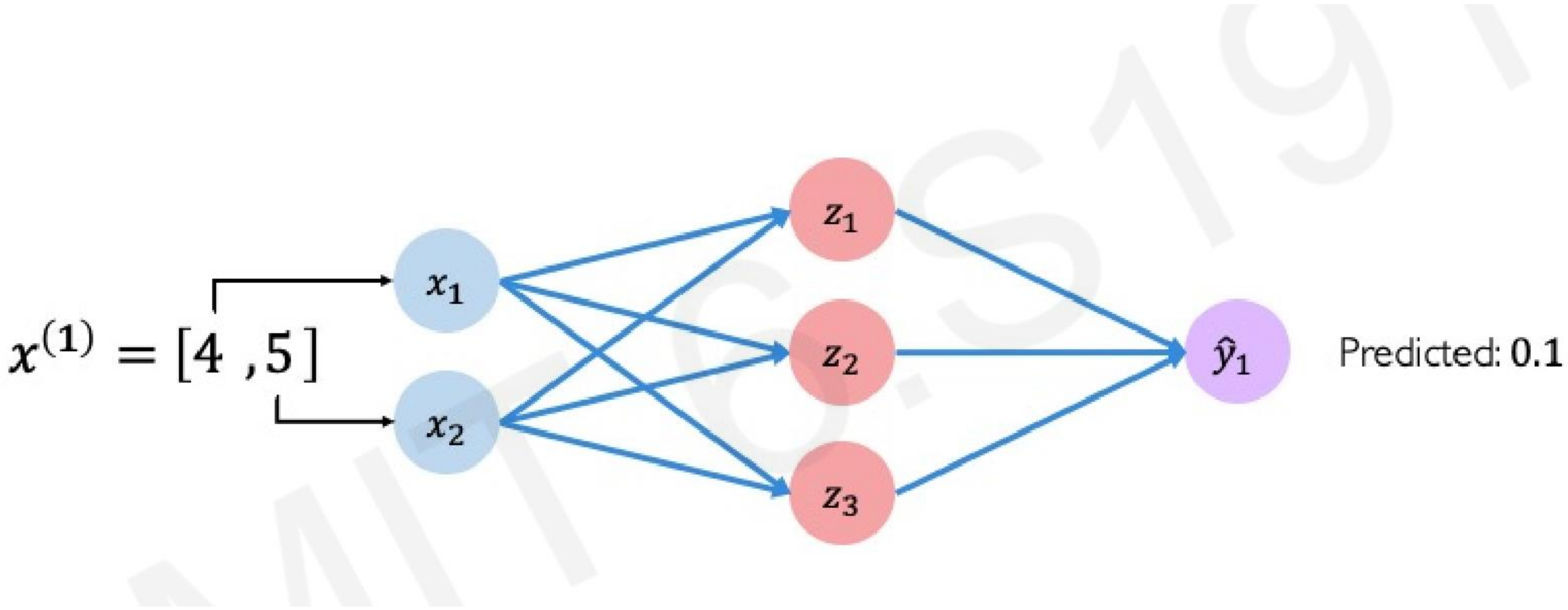
$$z_{k,i} = w_{0,i}^{(k)} + \sum_{j=1}^{n_{k-1}} g(z_{k-1,j})\, w_{j,i}^{(k)}$$

Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Example Problem: Will I pass this class?



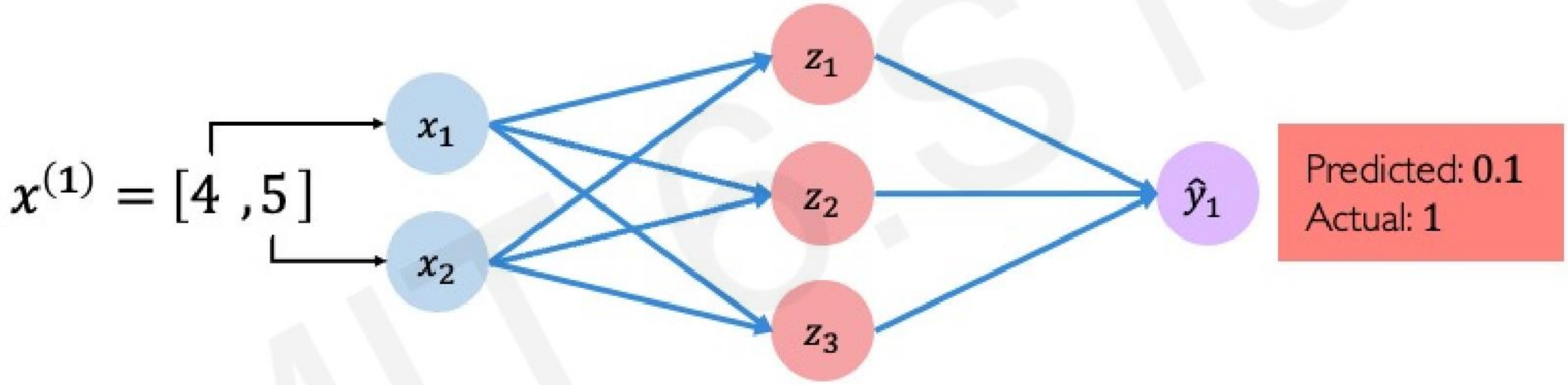
Example Problem: Will I pass this class?



Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Quantifying Loss

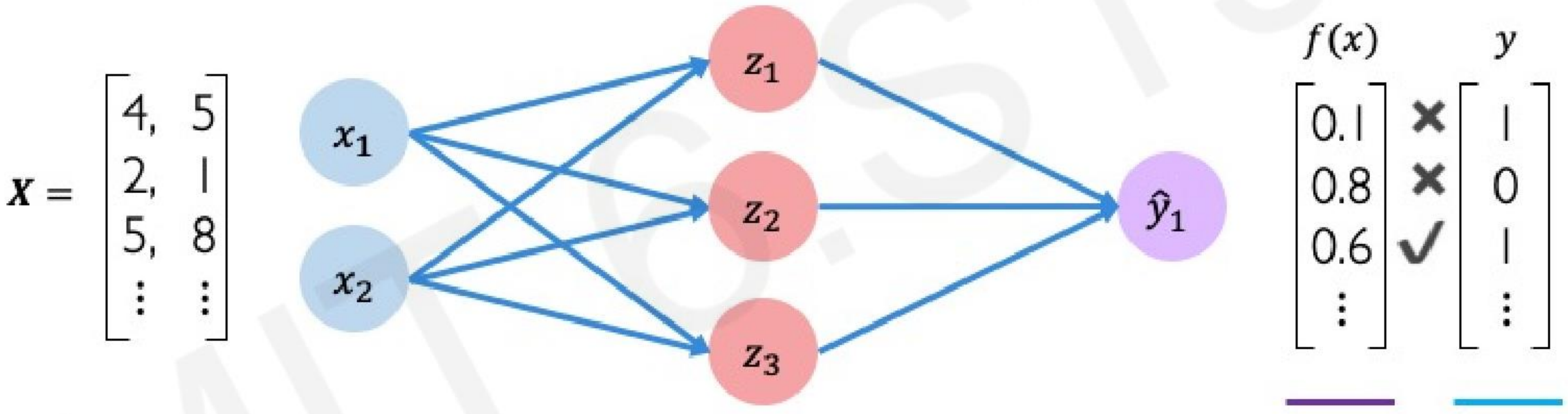
*The **loss** of our network measures the cost incurred from incorrect predictions*



$$\underbrace{\mathcal{L}(f(x^{(i)}; \boldsymbol{W})}_{\text{Predicted}}, \underbrace{y^{(i)}}_{\text{Actual}})$$

Empirical Loss

The ***empirical loss*** *measures the total loss over our entire dataset*



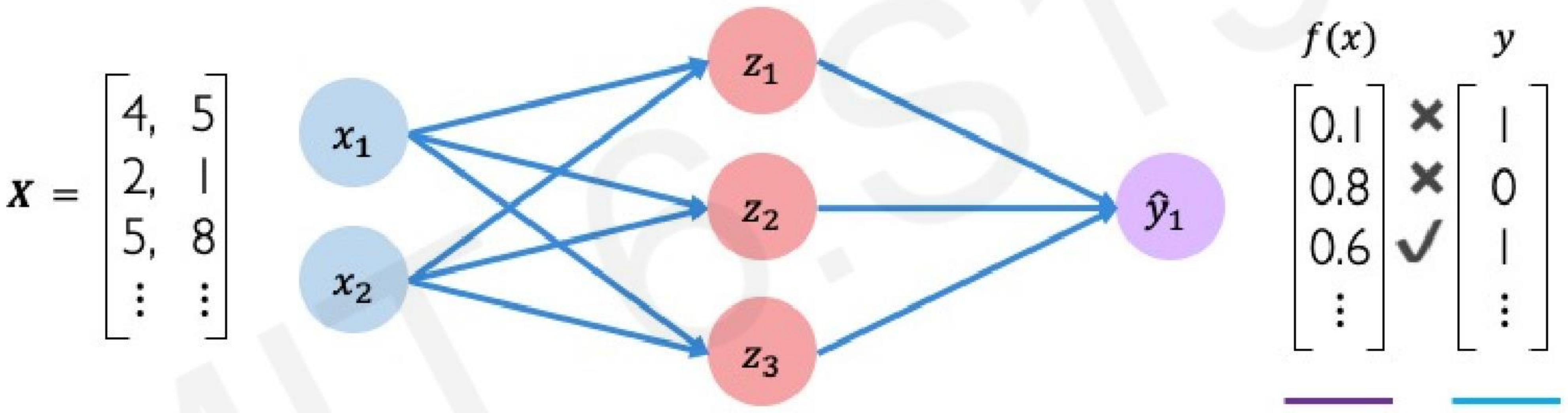
$$J(\boldsymbol{W}) = \frac{1}{n}\sum_{i=1}^{n} \mathcal{L}\big(\underbrace{f(x^{(i)}; \boldsymbol{W})}_{\text{Predicted}}, \underbrace{y^{(i)}}_{\text{Actual}}\big)$$

Also known as:

- Objective function
- Cost function
- Empirical Risk

Binary Cross Entropy Loss

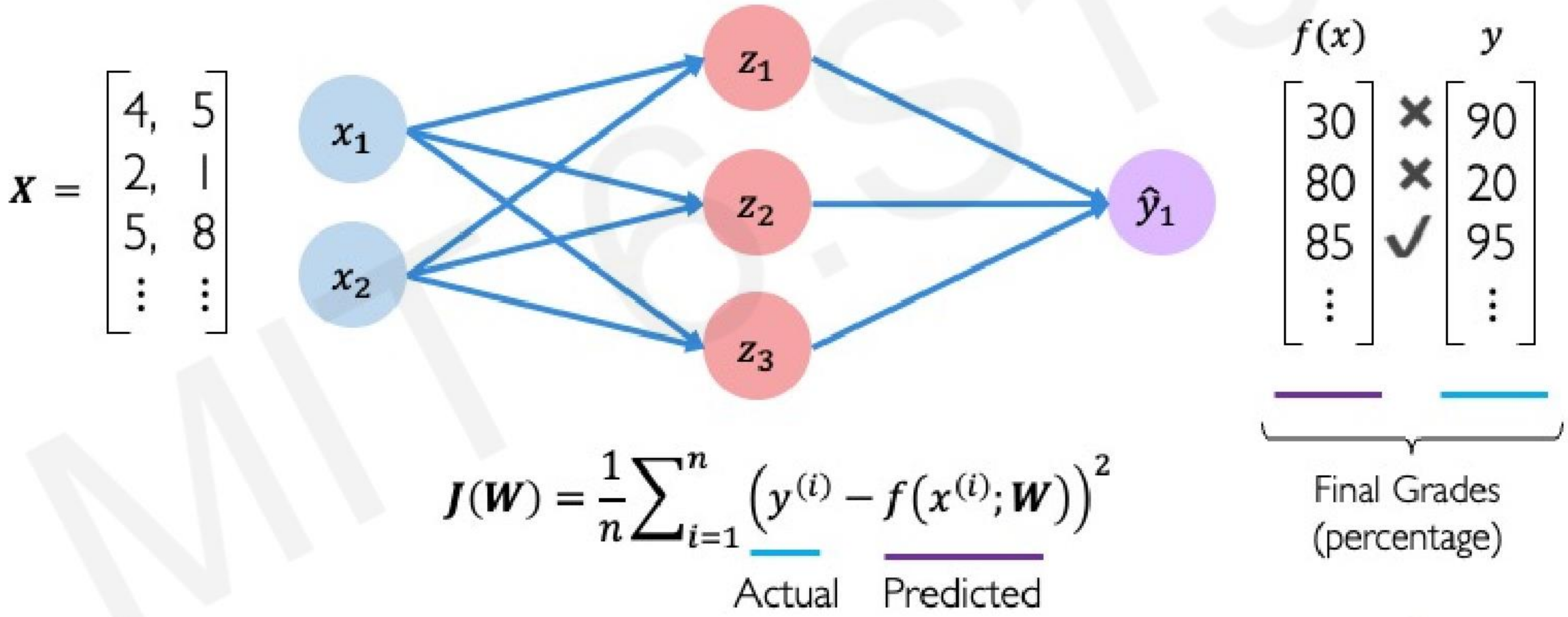
Cross entropy loss can be used with models that output a probability between 0 and 1



$$J(\boldsymbol{W}) = -\frac{1}{n}\sum_{i=1}^{n} \underbrace{y^{(i)}}_{\text{Actual}} \log\Big(\underbrace{f\big(x^{(i)};\boldsymbol{W}\big)}_{\text{Predicted}}\Big) + (1 - \underbrace{y^{(i)}}_{\text{Actual}}) \log\Big(1 - \underbrace{f\big(x^{(i)};\boldsymbol{W}\big)}_{\text{Predicted}}\Big)$$

Mean Squared Error Loss

Mean squared error loss *can be used with regression models that output continuous real numbers*



$$J(\boldsymbol{W}) = \frac{1}{n}\sum_{i=1}^{n}\left(\underbrace{y^{(i)}}_{\text{Actual}} - \underbrace{f\left(x^{(i)};\boldsymbol{W}\right)}_{\text{Predicted}}\right)^2$$

Loss Optimization

We want to find the network weights that **achieve the lowest loss**

$$W^* = \underset{W}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}\left(f\left(x^{(i)}; W\right), y^{(i)}\right)$$

$$W^* = \underset{W}{\operatorname{argmin}}\, J(W)$$

Remember:

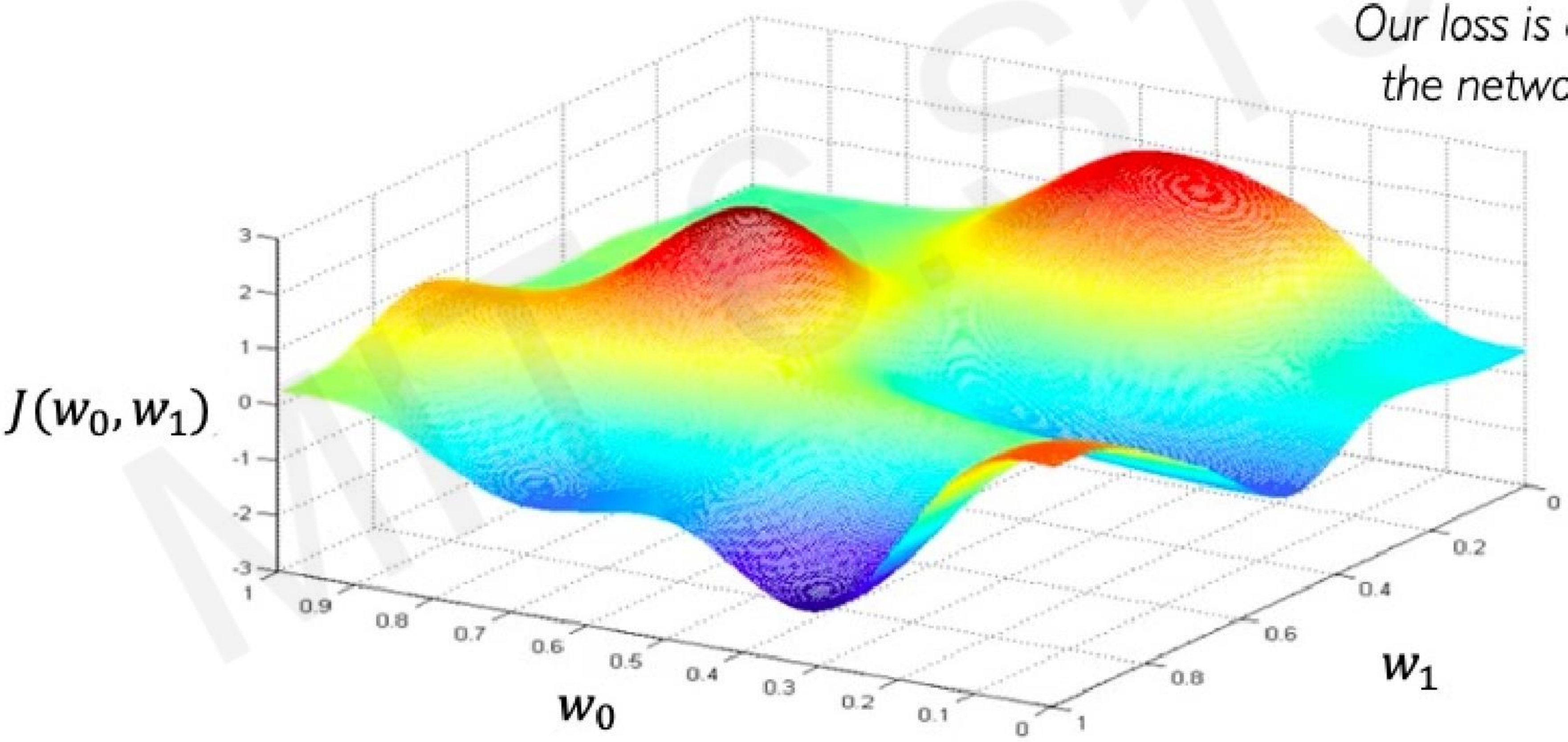
$$W = \left\{W^{(0)}, W^{(1)}, \cdots\right\}$$

Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Loss Optimization

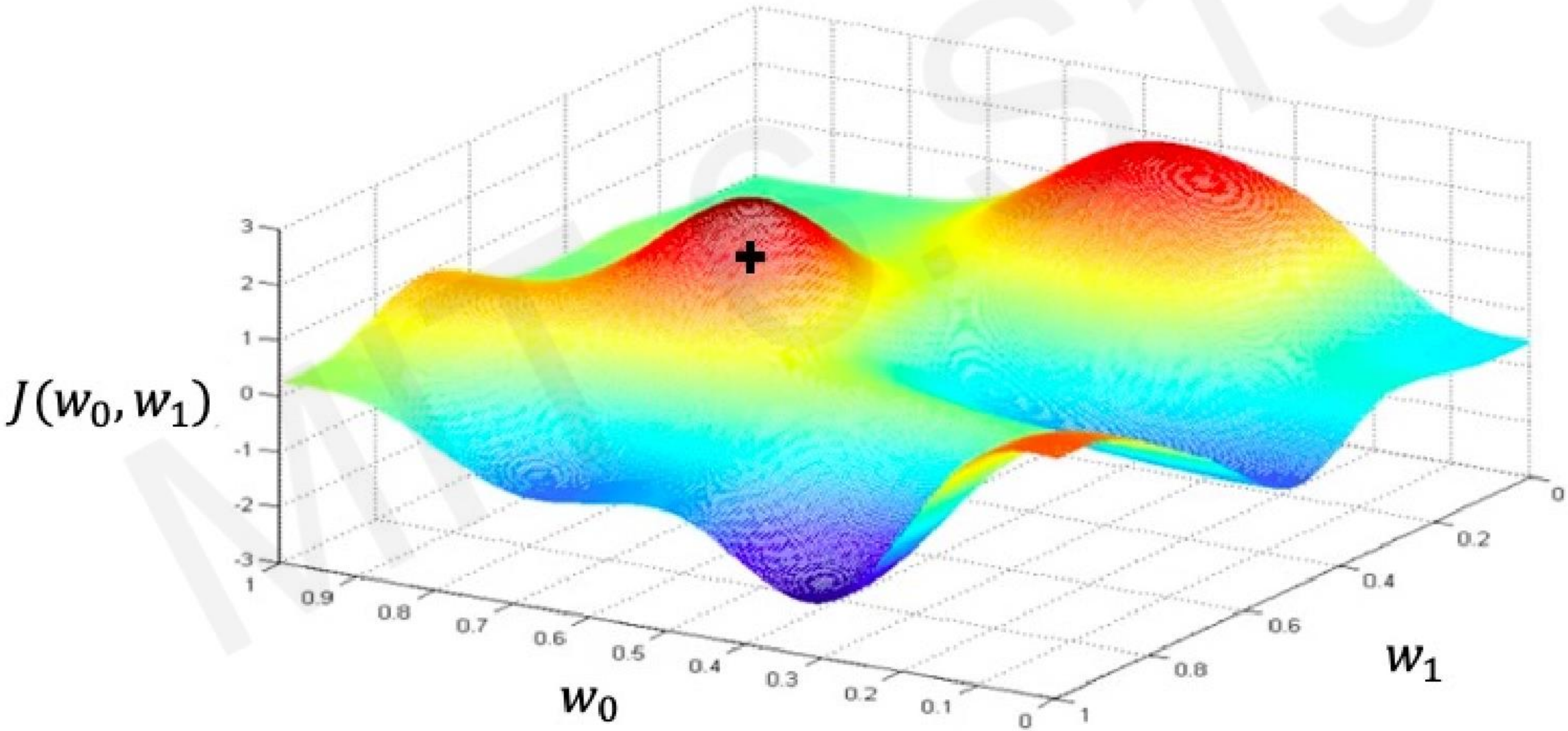
$$W^* = \underset{W}{\operatorname{argmin}} J(W)$$

Remember:
Our loss is a function of the network weights!



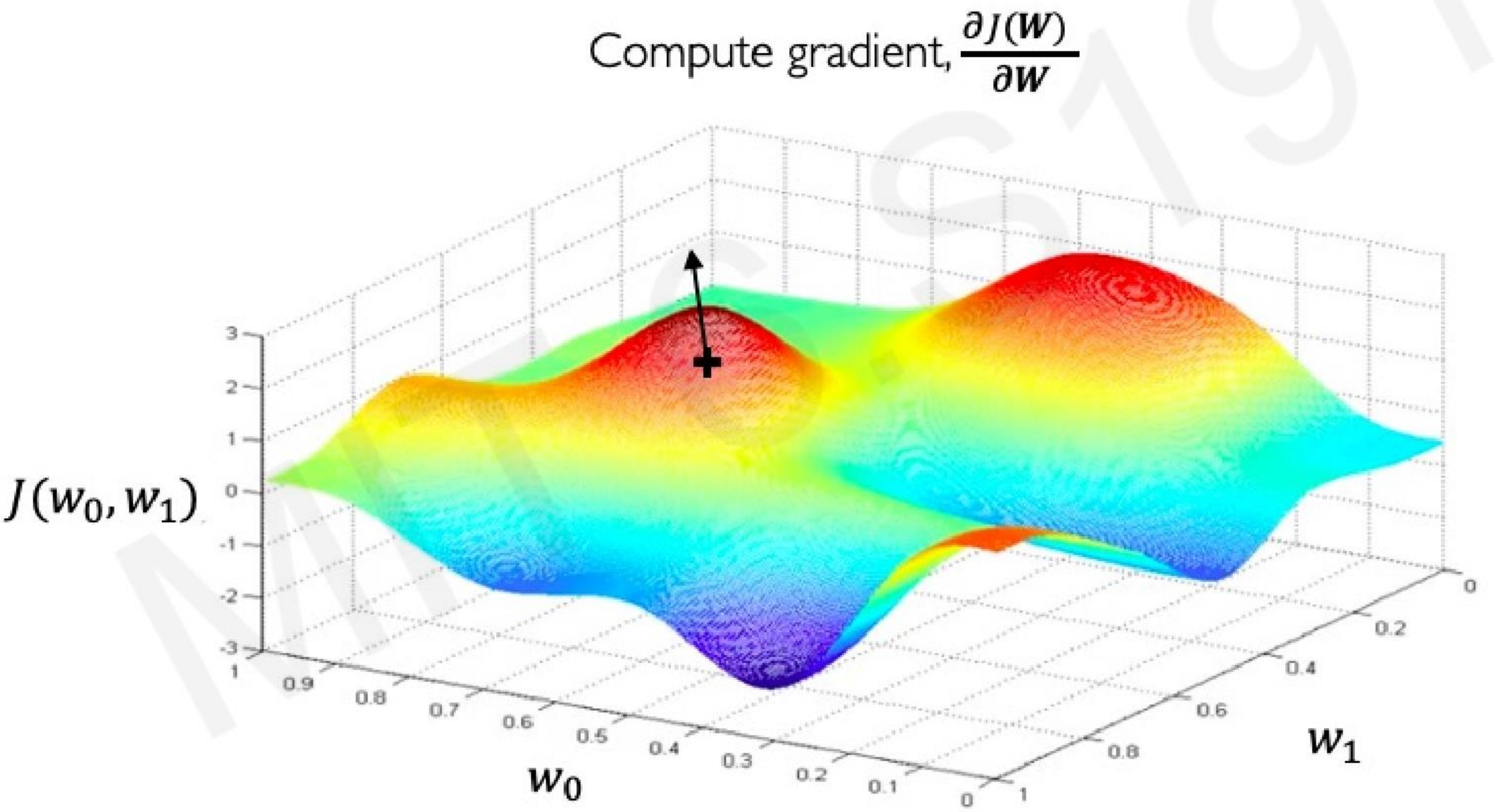
Loss Optimization

Randomly pick an initial (w_0, w_1)



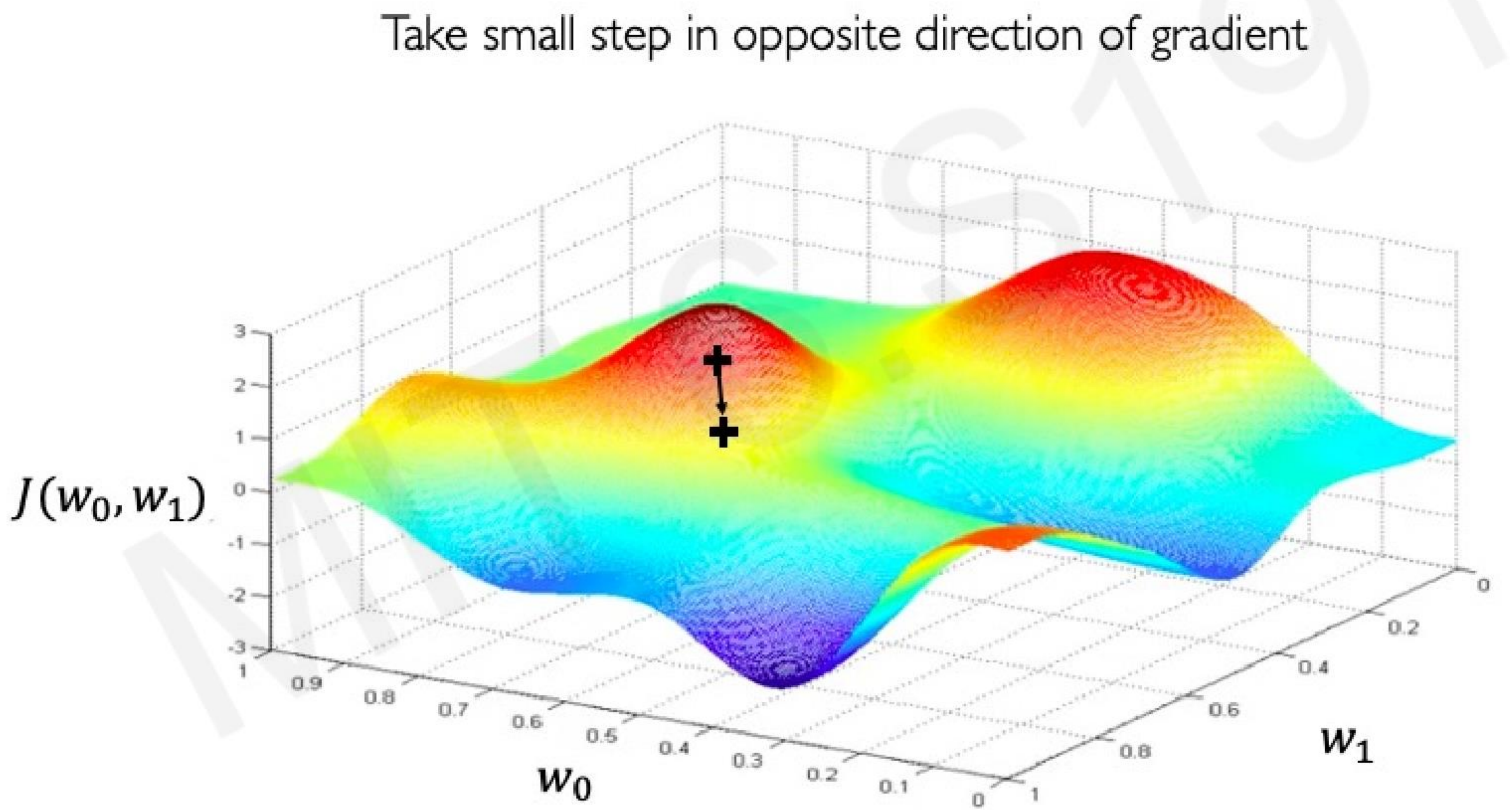
Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Loss Optimization

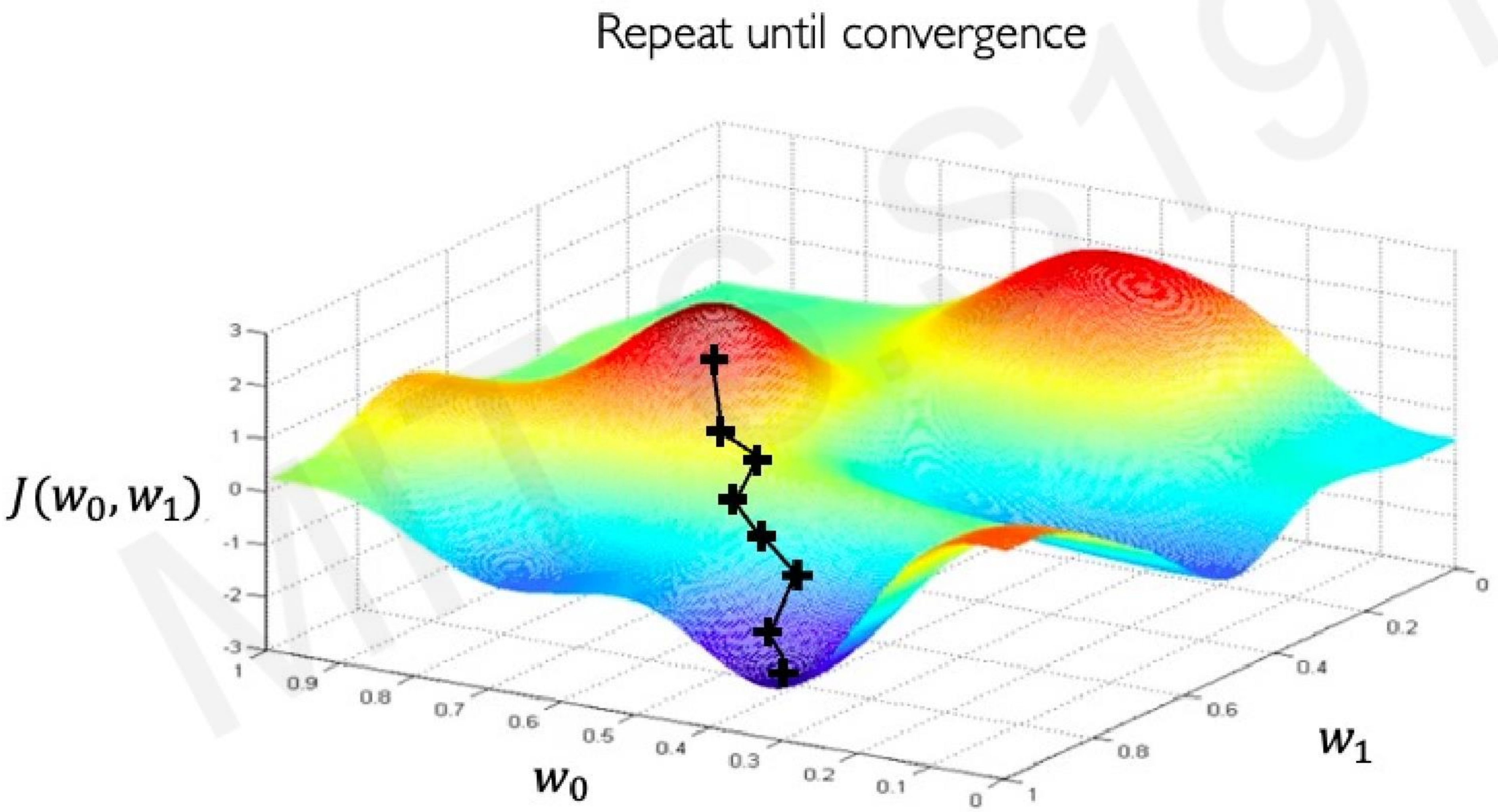


Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Loss Optimization



Loss Optimization

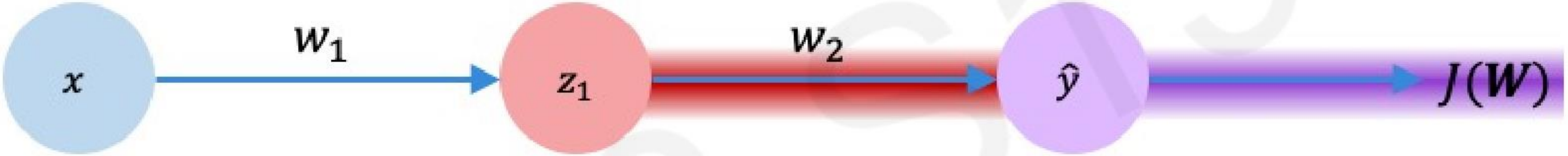


Gradient Descent

Algorithm

1. Initialize weights randomly $\sim\mathcal{N}(0, \sigma^2)$
2. Loop until convergence:
3. Compute gradient, $\boxed{\frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}}}$
4. Update weights, $\boldsymbol{W} \leftarrow \boldsymbol{W} - \eta \frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}}$
5. Return weights

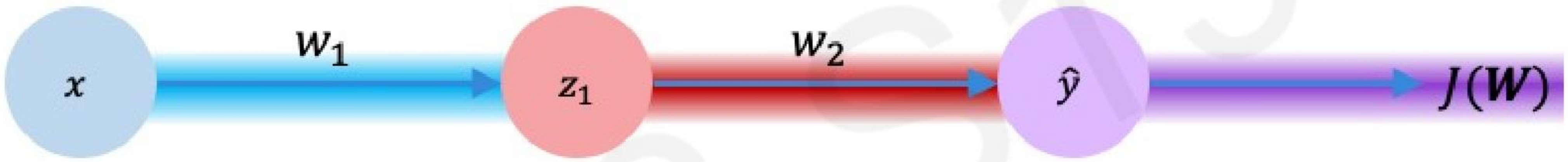
Computing Gradients: Backpropagation



$$\frac{\partial J(\boldsymbol{W})}{\partial w_2} = \frac{\partial J(\boldsymbol{W})}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial w_2}$$

How does a small change in one weight (e.g. $\boldsymbol{W_2}$) affect the final loss $\boldsymbol{J(W)}$

Computing Gradients: Backpropagation



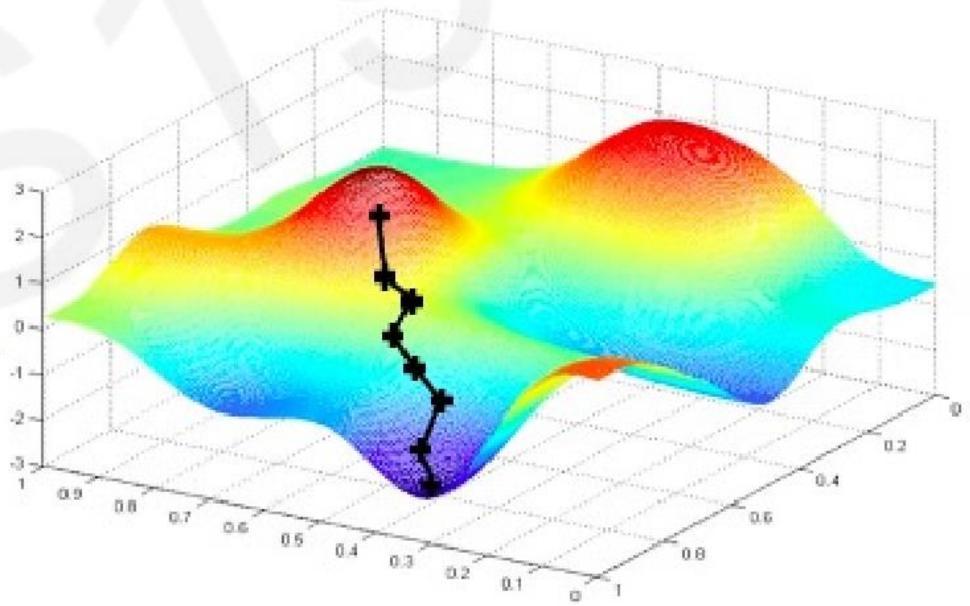
$$\frac{\partial J(\boldsymbol{W})}{\partial w_1} = \frac{\partial J(\boldsymbol{W})}{\partial \hat{y}} * \frac{\partial \hat{y}}{\partial z_1} * \frac{\partial z_1}{\partial w_1}$$

Repeat this for **every weight in the network** using gradients from later layers

Gradient Descent

Algorithm

1. Initialize weights randomly $\sim\mathcal{N}(0,\sigma^2)$
2. Loop until convergence:
3. Compute gradient, $\frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}}$
4. Update weights, $\boldsymbol{W} \leftarrow \boldsymbol{W} - \eta \frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}}$
5. Return weights

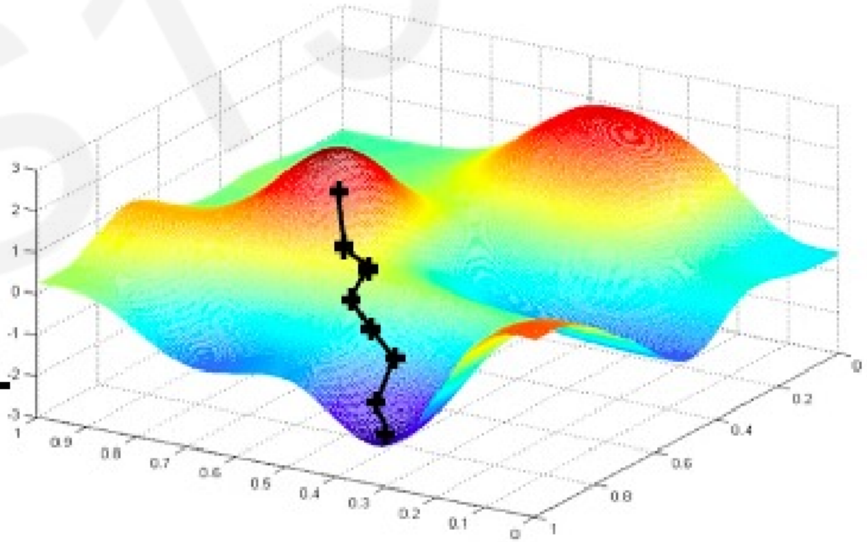


Potentially very computationally intensive to compute!

Stochastic Gradient Descent

Algorithm

1. Initialize weights randomly $\sim\mathcal{N}(0,\sigma^2)$
2. Loop until convergence:
3. Pick batch of B data points
4. Compute gradient, $\frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}} = \frac{1}{B}\sum_{k=1}^{B}\frac{\partial J_k(\boldsymbol{W})}{\partial \boldsymbol{W}}$
5. Update weights, $\boldsymbol{W} \leftarrow \boldsymbol{W} - \eta\frac{\partial J(\boldsymbol{W})}{\partial \boldsymbol{W}}$
6. Return weights



Fast to compute and a much better estimate of the true gradient!

Mini-batches while training

More accurate estimation of gradient

Smoother convergence

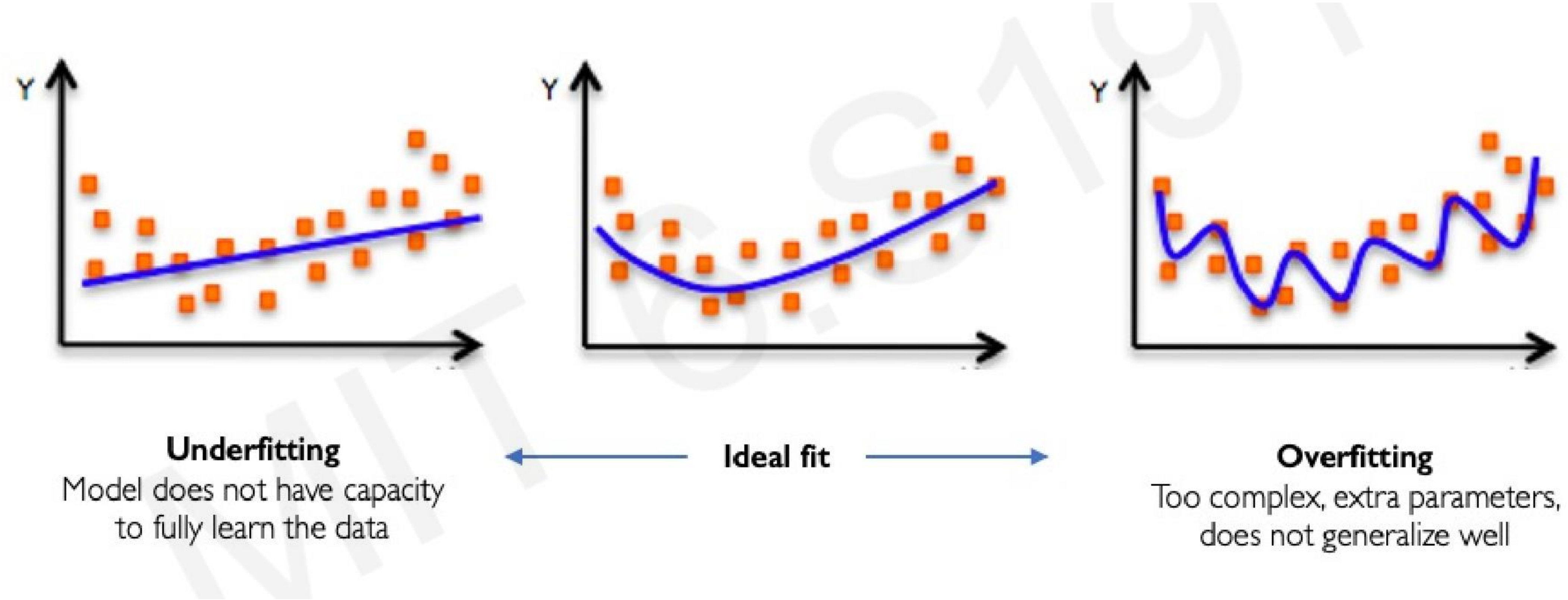
Allows for larger learning rates

Mini-batches lead to fast training!

Can parallelize computation + achieve significant speed increases on GPU's

Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

The Problem of Overfitting



Regularization

What is it?

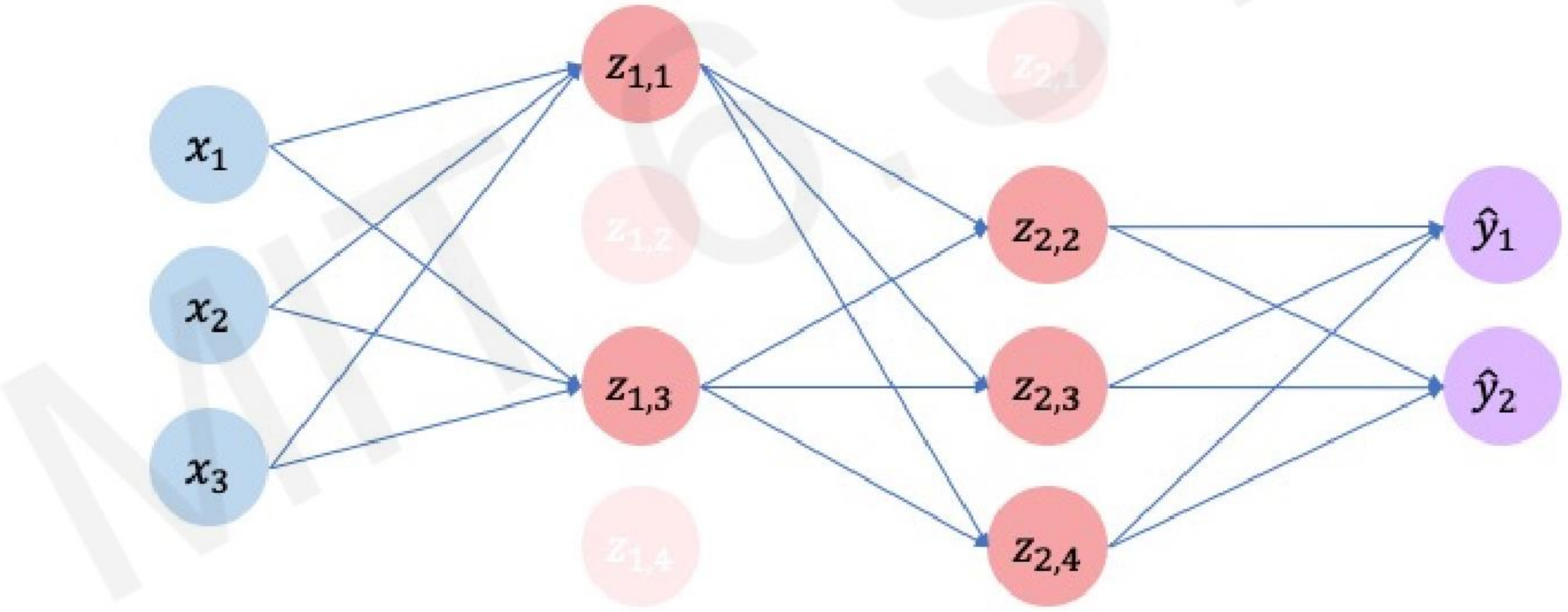
Technique that constrains our optimization problem to discourage complex models

Why do we need it?

Improve generalization of our model on unseen data

Regularization I: Dropout

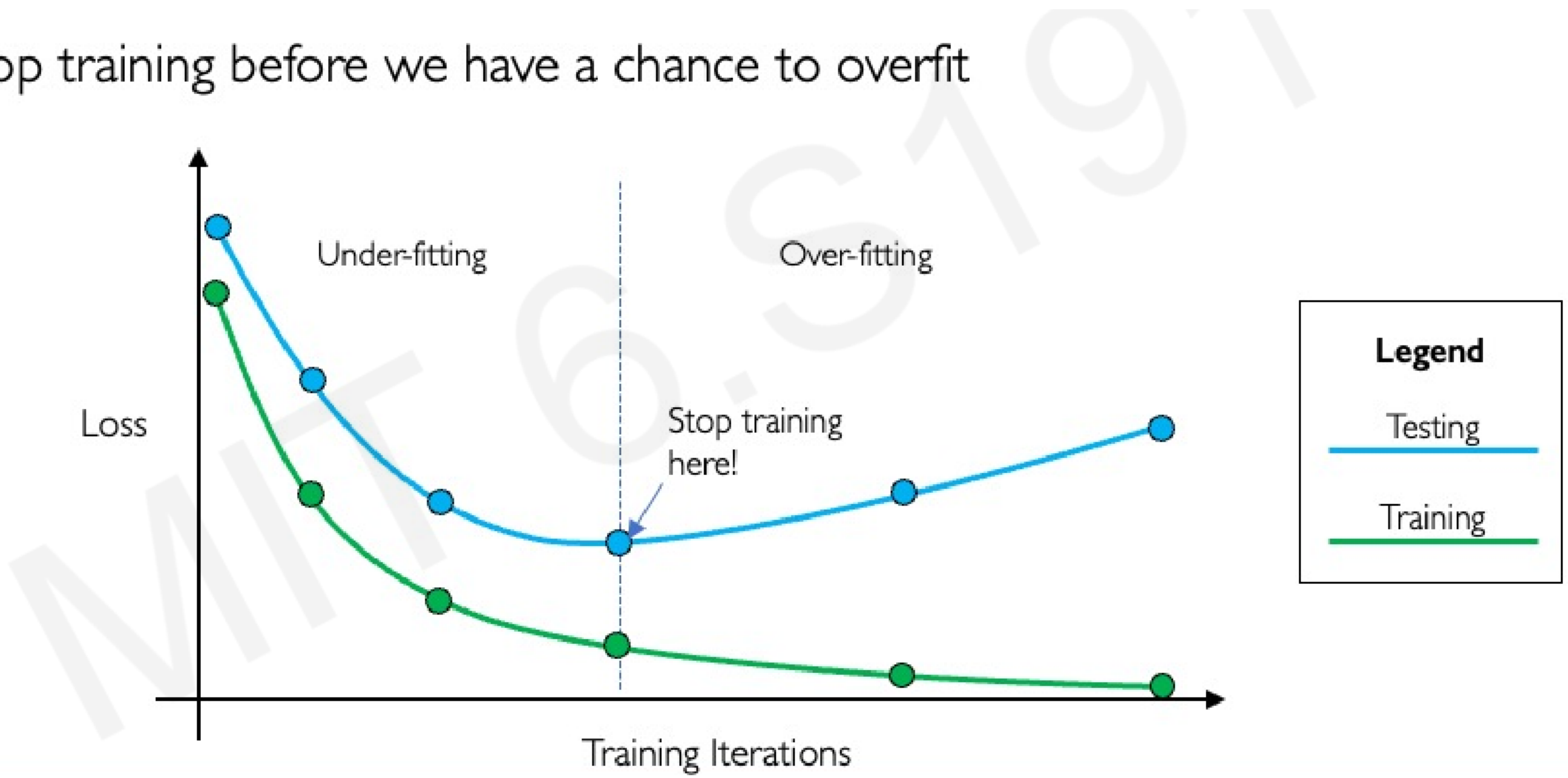
- During training, randomly set some activations to 0
  - Typically 'drop' 50% of activations in layer
  - Forces network to not rely on any 1 node



Alexander Amini: MIT Introduction to Deep Learning 6.S191: Lecture 1
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Regularization I: Early Stopping

- Stop training before we have a chance to overfit



Gentle Overview of Deep Learning

Generative Models

Given training data, generate new samples from same distribution

Training data ~ $p_{data}(x)$

Generated samples ~ $p_{model}(x)$

Want to learn $p_{model}(x)$ similar to $p_{data}(x)$

Addresses density estimation, a core problem in unsupervised learning

Several flavors:

- Explicit density estimation: explicitly define and solve for $p_{model}(x)$
- Implicit density estimation: learn model that can sample from $p_{model}(x)$ w/o explicitly defining it

Fei-Fei Li: Lecture 13 Generative Models, Convolutional Neural Networks for Visual Recognition (Spring 2017)
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Gentle Overview of Deep Learning

Why Generative Models?

- Realistic samples for artwork, super-resolution, colorization, etc.

- Generative models of time-series data can be used for simulation and planning (reinforcement learning applications!)
- Training generative models can also enable inference of latent representations that can be useful as general features

Fei-Fei Li: Lecture 13 Generative Models, Convolutional Neural Networks for Visual Recognition (Spring 2017)
https://www.youtube.com/watch?v=5WoItGTWV54&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=13

Gentle Overview of Deep Learning

Generative AI use cases

- Writing or improving content by producing a draft text in a specific style or length
- Adding subtitles or dubbing educational content, films, and other content in different languages
- Outlining briefs, resumes, term papers, and more
- Receiving a generic code to edit or improve upon
- Summarizing articles, emails, and reports
- Improving demonstration or explanation videos
- Creating music in a specific tone or style

What Is Generative AI? Definition, Applications, and Impact
https://www.coursera.org/articles/what-is-generative-ai

Concerns about generative AI

Generative AI's popularity is accompanied by concerns of ethics, misuse, and quality control. Because it is trained on existing sources, including those that are unverified on the internet, generative AI can provide misleading, inaccurate, and fake information. Even when a source is provided, that source might have incorrect information or may be falsely linked.

Since generators such as ChatGPT allow humans to input prompts with everyday language, it has become easier to use--so much so, that university students might use it to plagiarize or generate essays, and content creators may be accused of stealing from original artists. Falsified information can make it easier to impersonate people for cyber attacks.

What Is Generative AI? Definition, Applications, and Impact
https://www.coursera.org/articles/what-is-generative-ai

Generative Adversarial Networks (GANs)

Generative Adversarial Networks

Ian Goodfellow et al., "Generative Adversarial Nets", NIPS 2014

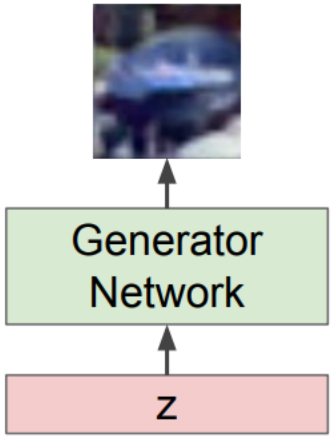
Problem: Want to sample from complex, high-dimensional training distribution. No direct way to do this!

Solution: Sample from a simple distribution, e.g. random noise. Learn transformation to training distribution.

Q: What can we use to represent this complex transformation?

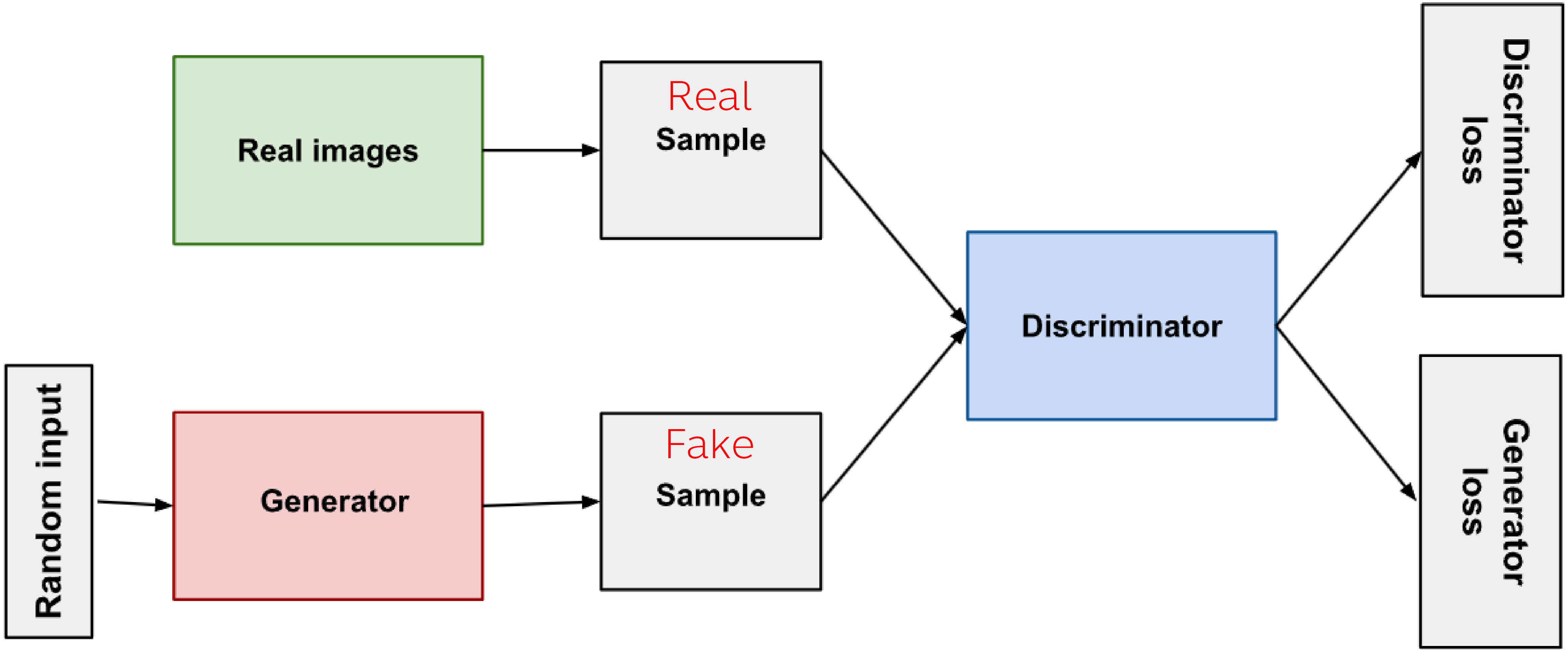
A: A neural network!

Output: Sample from training distribution



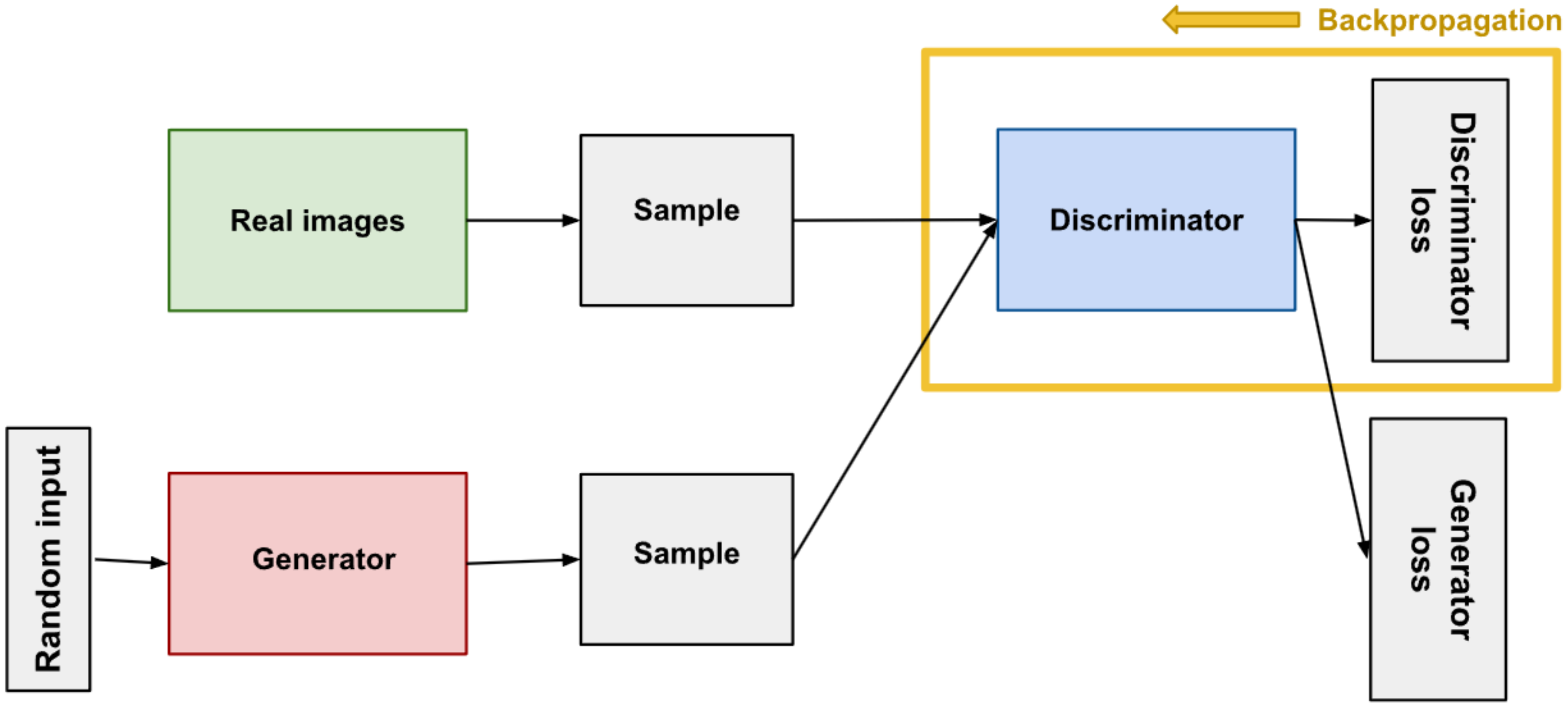
Input: Random noise

Generative Adversarial Networks (GANs)



Overview of GAN Structure
https://developers.google.com/machine-learning/gan/gan_structure

Generative Adversarial Networks (GANs)

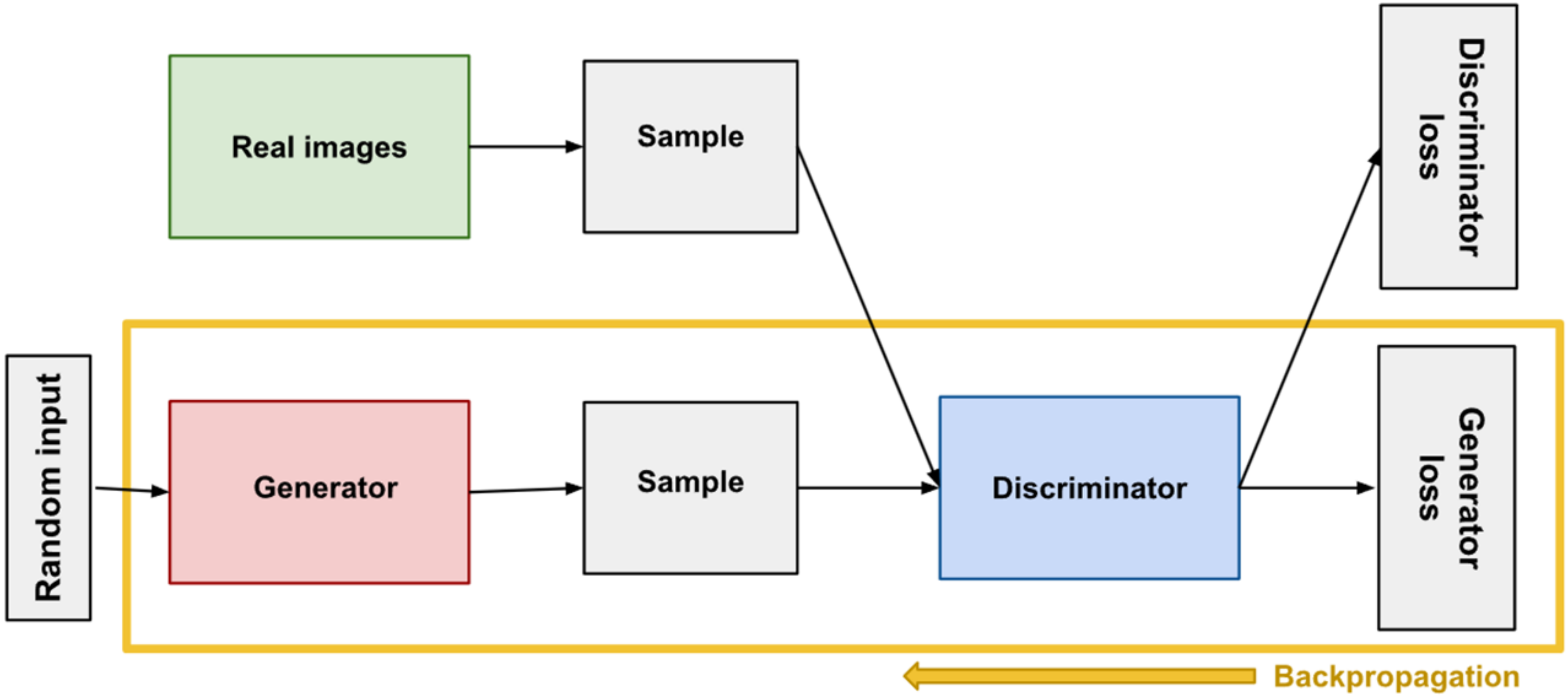


During discriminator training:

1. The discriminator classifies both real data and fake data from the generator.
2. The discriminator loss penalizes the discriminator for misclassifying a real instance as fake or a fake instance as real.
3. The discriminator updates its weights through backpropagation from the discriminator loss through the discriminator network.

The Discriminator
https://developers.google.com/machine-learning/gan/discriminator

Generative Adversarial Networks (GANs)



we train the generator with the following procedure:

1. Sample random noise.
2. Produce generator output from sampled random noise.
3. Get discriminator "Real" or "Fake" classification for generator output.
4. Calculate loss from discriminator classification.
5. Backpropagate through both the discriminator and generator to obtain gradients.
6. Use gradients to change only the generator weights.

The Generator
https://developers.google.com/machine-learning/gan/generator

Generative Adversarial Networks (GANs)

Training GANs: Two-player game

Ian Goodfellow et al., "Generative Adversarial Nets", NIPS 2014

Generator network: try to fool the discriminator by generating real-looking images
Discriminator network: try to distinguish between real and fake images

Train jointly in **minimax game**

Minimax objective function:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z))) \right]$$

Generative Adversarial Networks (GANs)

Training GANs: Two-player game

Ian Goodfellow et al., "Generative Adversarial Nets", NIPS 2014

Generator network: try to fool the discriminator by generating real-looking images
Discriminator network: try to distinguish between real and fake images

Train jointly in **minimax game**

Minimax objective function:

Discriminator outputs likelihood in (0,1) of real image

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log \underbrace{D_{\theta_d}(x)}_{\text{Discriminator output for real data x}} + \mathbb{E}_{z \sim p(z)} \log(1 - \underbrace{D_{\theta_d}(G_{\theta_g}(z))}_{\text{Discriminator output for generated fake data G(z)}}) \right]$$

Generative Adversarial Networks (GANs)

Training GANs: Two-player game

Ian Goodfellow et al., "Generative Adversarial Nets", NIPS 2014

Generator network: try to fool the discriminator by generating real-looking images
Discriminator network: try to distinguish between real and fake images

Train jointly in **minimax game**

Discriminator outputs likelihood in (0,1) of real image

Minimax objective function:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log \underbrace{D_{\theta_d}(x)}_{\text{Discriminator output for real data x}} + \mathbb{E}_{z \sim p(z)} \log(1 - \underbrace{D_{\theta_d}(G_{\theta_g}(z))}_{\text{Discriminator output for generated fake data G(z)}}) \right]$$

- Discriminator (θ_d) wants to **maximize objective** such that D(x) is close to 1 (real) and D(G(z)) is close to 0 (fake)
- Generator (θ_g) wants to **minimize objective** such that D(G(z)) is close to 1 (discriminator is fooled into thinking generated G(z) is real)

Generative Adversarial Networks (GANs)

Training GANs: Two-player game

Ian Goodfellow et al., "Generative Adversarial Nets", NIPS 2014

Minimax objective function:

$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z))) \right]$$

Alternate between:

1. **Gradient ascent** on discriminator

$$\max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z))) \right]$$

2. **Gradient descent** on generator

$$\min_{\theta_g} \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z)))$$

Generative Adversarial Networks (GANs)

Training GANs: Two-player game

Ian Goodfellow et al., "Generative Adversarial Nets", NIPS 2014

Putting it together: GAN training algorithm

for number of training iterations **do**
  for k steps **do**
  - Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(m)}\}$ from noise prior $p_g(\boldsymbol{z})$.
  - Sample minibatch of m examples $\{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(m)}\}$ from data generating distribution $p_{\text{data}}(\boldsymbol{x})$.
  - Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D_{\theta_d}(x^{(i)}) + \log(1 - D_{\theta_d}(G_{\theta_g}(z^{(i)}))) \right]$$

  end for
  - Sample minibatch of m noise samples $\{\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(m)}\}$ from noise prior $p_g(\boldsymbol{z})$.
  - Update the generator by ascending its stochastic gradient (improved objective):

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log(D_{\theta_d}(G_{\theta_g}(z^{(i)})))$$

end for

Generative Adversarial Networks (GANs)

Pros and cons of GANs

Advantages of GANs	Disadvantages of GANs
GANs are considered unsupervised learning models, continuing to train themselves after the initial input and capable of learning from unlabeled data.	They can be difficult to train due to the need for large, varied, and advanced data sets.
GANs are capable of identifying anomalies based on measurements that indicate how well the generator and discriminator are able to model the data.	It can be challenging to evaluate results depending on the complexity of a given task.
Ability to create realistic data samples	GANs suffer from mode collapse, or learning to produce only one output due to its high plausibility and ability to trick the discriminator.

What is GAN? Generative Adversarial Networks Explained
https://www.coursera.org/articles/what-is-gan

Generative Adversarial Networks (GANs)

See also: https://github.com/soumith/ganhacks for tips and tricks for trainings GANs

"The GAN Zoo"

- GAN - Generative Adversarial Networks
- 3D-GAN - Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling
- acGAN - Face Aging With Conditional Generative Adversarial Networks
- AC-GAN - Conditional Image Synthesis With Auxiliary Classifier GANs
- AdaGAN - AdaGAN: Boosting Generative Models
- AEGAN - Learning Inverse Mapping by Autoencoder based Generative Adversarial Nets
- AffGAN - Amortised MAP Inference for Image Super-resolution
- AL-CGAN - Learning to Generate Images of Outdoor Scenes from Attributes and Semantic Layouts
- ALI - Adversarially Learned Inference
- AM-GAN - Generative Adversarial Nets with Labeled Data by Activation Maximization
- AnoGAN - Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery
- ArtGAN - ArtGAN: Artwork Synthesis with Conditional Categorial GANs
- b-GAN - b-GAN: Unified Framework of Generative Adversarial Networks
- Bayesian GAN - Deep and Hierarchical Implicit Models
- BEGAN - BEGAN: Boundary Equilibrium Generative Adversarial Networks
- BiGAN - Adversarial Feature Learning
- BS-GAN - Boundary-Seeking Generative Adversarial Networks
- CGAN - Conditional Generative Adversarial Nets
- CaloGAN - CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks
- CCGAN - Semi-Supervised Learning with Context-Conditional Generative Adversarial Networks
- CatGAN - Unsupervised and Semi-supervised Learning with Categorical Generative Adversarial Networks
- CoGAN - Coupled Generative Adversarial Networks
- Context-RNN-GAN - Contextual RNN-GANs for Abstract Reasoning Diagram Generation
- C-RNN-GAN - C-RNN-GAN: Continuous recurrent neural networks with adversarial training
- CS-GAN - Improving Neural Machine Translation with Conditional Sequence Generative Adversarial Nets
- CVAE-GAN - CVAE-GAN: Fine-Grained Image Generation through Asymmetric Training
- CycleGAN - Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks
- DTN - Unsupervised Cross-Domain Image Generation
- DCGAN - Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks
- DiscoGAN - Learning to Discover Cross-Domain Relations with Generative Adversarial Networks
- DR-GAN - Disentangled Representation Learning GAN for Pose-Invariant Face Recognition
- DualGAN - DualGAN: Unsupervised Dual Learning for Image-to-Image Translation
- EBGAN - Energy-based Generative Adversarial Network
- f-GAN - f-GAN: Training Generative Neural Samplers using Variational Divergence Minimization
- FF-GAN - Towards Large-Pose Face Frontalization in the Wild
- GAWWN - Learning What and Where to Draw
- GeneGAN - GeneGAN: Learning Object Transfiguration and Attribute Subspace from Unpaired Data
- Geometric GAN - Geometric GAN
- GoGAN - Gang of GANs: Generative Adversarial Networks with Maximum Margin Ranking
- GP-GAN - GP-GAN: Towards Realistic High-Resolution Image Blending
- IAN - Neural Photo Editing with Introspective Adversarial Networks
- iGAN - Generative Visual Manipulation on the Natural Image Manifold
- IcGAN - Invertible Conditional GANs for image editing
- ID-CGAN - Image De-raining Using a Conditional Generative Adversarial Network
- Improved GAN - Improved Techniques for Training GANs
- InfoGAN - InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets
- LAGAN - Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis
- LAPGAN - Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks

https://github.com/hindupuravinash/the-gan-zoo

Potential Uses of GANs in Economics

Potential Uses of GANs in Economics

ELSEVIER

Journal of Econometrics

Available online 20 March 2021, 105076

In Press, Corrected Proof What's this?

Using Wasserstein Generative Adversarial Networks for the design of Monte Carlo simulations ☆

Susan Athey [a b], Guido W. Imbens [a c b], Jonas Metzger [c], Evan Munro [a]

[a] Graduate School of Business, Stanford University, United States of America

[b] NBER, United States of America

[c] Department of Economics, Stanford University, United States of America

A B S T R A C T

When researchers develop new econometric methods it is common practice to compare the performance of the new methods to those of existing methods in Monte Carlo studies. The credibility of such Monte Carlo studies is often limited because of the discretion the researcher has in choosing the Monte Carlo designs reported. To improve the credibility we propose using a class of generative models that has recently been developed in the machine learning literature, termed Generative Adversarial Networks (GANs) which can be used to systematically generate artificial data that closely mimics existing datasets. Thus, in combination with existing real data sets, GANs can be used to limit the degrees of freedom in Monte Carlo study designs for the researcher, making any comparisons more convincing. In addition, if an applied researcher is concerned with the performance of a particular statistical method on a specific data set (beyond its theoretical properties in large samples), she can use such GANs to assess the performance of the proposed method, *e.g.* the coverage rate of confidence intervals or the bias of the estimator, using simulated data which closely resembles the exact setting of interest. To illustrate these methods we apply Wasserstein GANs (WGANs) to the estimation of average treatment effects. In this example, we find that (*i*) there is not a single estimator that outperforms the others in all three settings, so researchers should tailor their analytic approach to a given setting, (*ii*) systematic simulation studies can be helpful for selecting among competing methods in this situation, and (*iii*) the generated data closely resemble the actual data.

Athey, S., Imbens, G.W., Metzger, J. and Munro, E., 2021. Using Wasserstein Generative Adversarial Networks for the design of Monte Carlo simulations. *Journal of Econometrics*, p.105076.

https://www.sciencedirect.com/science/article/pii/S0304407621000440?casa_token=M2cGsbEz9wgAAAAA:czIDA78fGTQ_TC98uTtGKLYOnY-4iz3EX9ubvXjLmDdQQaCh3v0EIC6WdZ-eL7m0-5D3c7-Q

Potential Uses of GANs in Economics

ELSEVIER

Journal of Econometrics

Available online 20 March 2021, 105076

In Press, Corrected Proof ⓘ What's this?

Using Wasserstein Generative Adversarial Networks for the design of Monte Carlo simulations ☆

Susan Athey [a b], Guido W. Imbens [a c b], Jonas Metzger [c], Evan Munro [a]

[a] Graduate School of Business, Stanford University, United States of America

[b] NBER, United States of America

[c] Department of Economics, Stanford University, United States of America

Repository of R and Python codes

https://github.com/evanmunro/dswgan-paper

Athey, S., Imbens, G.W., Metzger, J. and Munro, E., 2021. Using Wasserstein Generative Adversarial Networks for the design of Monte Carlo simulations. *Journal of Econometrics*, p.105076.
https://www.sciencedirect.com/science/article/pii/S0304407621000440?casa_token=M2cGsbEz9wgAAAAA:czIDA78fGTQ_TC98uTtGKLYOnY-4iz3EX9ubvXjLmDdQQaCh3v0EIC6WdZ-eL7m0-5D3c7-Q

Potential Uses of GANs in Economics

Time-series Generative Adversarial Networks

Jinsung Yoon*
University of California, Los Angeles, USA
jsyoon0823@g.ucla.edu

Daniel Jarrett*
University of Cambridge, UK
daniel.jarrett@maths.cam.ac.

Mihaela van der Schaar
University of Cambridge, UK
University of California, Los Angeles, USA
Alan Turing Institute, UK
mv472@cam.ac.uk, mihaela@ee.ucla.edu

Abstract

A good generative model for time-series data should preserve *temporal dynamics*, in the sense that new sequences respect the original relationships between variables across time. Existing methods that bring generative adversarial networks (GANs) into the sequential setting do not adequately attend to the temporal correlations unique to time-series data. At the same time, supervised models for sequence prediction—which allow finer control over network dynamics—are inherently deterministic. We propose a novel framework for generating realistic time-series data that combines the flexibility of the unsupervised paradigm with the control afforded by supervised training. Through a learned embedding space jointly optimized with both supervised and adversarial objectives, we encourage the network to adhere to the dynamics of the training data during sampling. Empirically, we evaluate the ability of our method to generate realistic samples using a variety of real and synthetic time-series datasets. Qualitatively and quantitatively, we find that the proposed framework consistently and significantly outperforms state-of-the-art benchmarks with respect to measures of similarity and predictive ability.

Table 1: Results on Autoregressive Multivariate Gaussian Data (Bold indicates best performance).

	Temporal Correlations (fixing $\sigma = 0.8$)			Feature Correlations (fixing $\phi = 0.8$)		
Settings	$\phi = 0.2$	$\phi = 0.5$	$\phi = 0.8$	$\sigma = 0.2$	$\sigma = 0.5$	$\sigma = 0.8$
Discriminative Score (Lower the better)						
TimeGAN	**.175±.006**	**.174±.012**	**.105±.005**	**.181±.006**	**.152±.011**	**.105±.005**
RCGAN	.177±.012	.190±.011	.133±.019	.186±.012	.190±.012	.133±.019
C-RNN-GAN	.391±.006	.227±.017	.220±.016	.198±.011	.202±.010	.220±.016
T-Forcing	.500±.000	.500±.000	.499±.001	.499±.001	.499±.001	.499±.001
P-Forcing	.498±.002	.472±.008	.396±.018	.460±.003	.408±.016	.396±.018
WaveNet	.337±.005	.235±.009	.229±.013	.217±.010	.226±.011	.229±.013
WaveGAN	.336±.011	.213±.013	.230±.023	.192±.012	.205±.015	.230±.023
Predictive Score (Lower the better)						
TimeGAN	**.640±.003**	**.412±.002**	**.251±.002**	**.282±.005**	**.261±0.002**	**.251±.002**
RCGAN	.652±.003	.435±.002	.263±.003	.292±.003	.279±.002	.263±.003
C-RNN-GAN	.696±.002	.490±.005	.299±.002	.293±.005	.280±.006	.299±.002
T-Forcing	.737±.022	.732±.012	.503±.037	.515±.034	.543±.023	.503±.037
P-Forcing	.665±.004	.571±.005	.289±.003	.406±.005	.317±.001	.289±.003
WaveNet	.718±.002	.508±.003	.321±.005	.331±.004	.297±.003	.321±.005
WaveGAN	.712±.003	.489±.001	.290±.002	.325±.003	.353±.001	.290±.002

Yoon et al. 2019. Time-series Generative Adversarial Networks. 33rd Conference on Neural Information Processing Systems (NeurIPS 2019)
https://proceedings.neurips.cc/paper_files/paper/2019/file/c9efe5f26cd17ba6216bbe2a7d26d490-Paper.pdf

Potential Uses of GANs in Economics

Time-series Generative Adversarial Networks

Jinsung Yoon*
University of California, Los Angeles, USA
jsyoon0823@g.ucla.edu

Daniel Jarrett*
University of Cambridge, UK
daniel.jarrett@maths.cam.ac.uk

Mihaela van der Schaar
University of Cambridge, UK
University of California, Los Angeles, USA
Alan Turing Institute, UK
mv472@cam.ac.uk, mihaela@ee.ucla.edu

Abstract

A good generative model for time-series data should preserve *temporal dynamics*, in the sense that new sequences respect the original relationships between variables across time. Existing methods that bring generative adversarial networks (GANs) into the sequential setting do not adequately attend to the temporal correlations unique to time-series data. At the same time, supervised models for sequence prediction—which allow finer control over network dynamics—are inherently deterministic. We propose a novel framework for generating realistic time-series data that combines the flexibility of the unsupervised paradigm with the control afforded by supervised training. Through a learned embedding space jointly optimized with both supervised and adversarial objectives, we encourage the network to adhere to the dynamics of the training data during sampling. Empirically, we evaluate the ability of our method to generate realistic samples using a variety of real and synthetic time-series datasets. Qualitatively and quantitatively, we find that the proposed framework consistently and significantly outperforms state-of-the-art benchmarks with respect to measures of similarity and predictive ability.

Table 2: Results on Multiple Time-Series Datasets (Bold indicates best performance).

Metric	Method	Sines	Stocks	Energy	Events
Discriminative Score (Lower the Better)	TimeGAN	**.011±.008**	**.102±.021**	**.236±.012**	**.161±.018**
	RCGAN	.022±.008	.196±.027	.336±.017	.380±.021
	C-RNN-GAN	.229±.040	.399±.028	.499±.001	.462±.011
	T-Forcing	.495±.001	.226±.035	.483±.004	.387±.012
	P-Forcing	.430±.027	.257±.026	.412±.006	.489±.001
	WaveNet	.158±.011	.232±.028	.397±.010	.385±.025
	WaveGAN	.277±.013	.217±.022	.363±.012	.357±.017
Predictive Score (Lower the Better)	TimeGAN	**.093±.019**	**.038±.001**	**.273±.004**	**.303±.006**
	RCGAN	.097±.001	.040±.001	.292±.005	.345±.010
	C-RNN-GAN	.127±.004	**.038±.000**	.483±.005	.360±.010
	T-Forcing	.150±.022	**.038±.001**	.315±.005	.310±.003
	P-Forcing	.116±.004	.043±.001	.303±.006	.320±.008
	WaveNet	.117±.008	.042±.001	.311±.005	.333±.004
	WaveGAN	.134±.013	.041±.001	.307±.007	.324±.006
	Original	.094±.001	.036±.001	.250±.003	.293±.000

Yoon et al. 2019. Time-series Generative Adversarial Networks. 33rd Conference on Neural Information Processing Systems (NeurIPS 2019)
https://proceedings.neurips.cc/paper_files/paper/2019/file/c9efe5f26cd17ba6216bbe2a7d26d490-Paper.pdf

Potential Uses of GANs in Economics

GENERATIVE ADVERSARIAL NETWORKS IN FINANCE: AN OVERVIEW

A PREPRINT

Florian Eckerli *
School of Engineering
Zurich University of Applied Sciences
Winterthur, Switzerland
e.florian@hotmail.com

Joerg Osterrieder*
School of Engineering
Zurich University of Applied Sciences
Winterthur, Switzerland
joerg.osterrieder@zhaw.ch

The Hightech Business and Entrepreneurship Group
Faculty of Behavioural, Management and Social Sciences
University of Twente
Enschede, Netherlands
joerg.osterrieder@utwente.nl

June 11, 2021

ABSTRACT

Modelling in finance is a challenging task: the data often has complex statistical properties and its inner workings are largely unknown. Deep learning algorithms are making progress in the field of data-driven modelling, but the lack of sufficient data to train these models is currently holding back several new applications. Generative Adversarial Networks (GANs) are a neural network architecture family that has achieved good results in image generation and is being successfully applied to generate time series and other types of financial data. The purpose of this study is to present an overview of how these GANs work, their capabilities and limitations in the current state of research with financial data and present some practical applications in the industry. As a proof of concept, three known GAN architectures were tested on financial time series, and the generated data was evaluated on its statistical properties, yielding solid results. Finally, it was shown that GANs have made considerable progress in their finance applications and can be a solid additional tool for data scientists in this field.

Keywords Generative Adversarial Networks, GANs, Time Series, Synthetic Data

Table 1: GANs in finance research

Field	Application	Method
Time Series Forecasting	Market Prediction	GAN-FD [9], ST-GAN [19], MTSGAN [20]
	Fine-Tuning of trading models	C-GAN [10], MAS-GAN[21]
Portfolio Management	Porfolio Optimization	PAGAN[11], GAN-MP[22], DAT-CGAN[23], CorrGAN[12]
Time Series Generation	Synthetic time series generation and Finance Data Augmentation	TimeGAN[24], WGAN-GP[25], FIN-GAN[3], Quant GAN[14], RA-GAN[26], CDRAGAN[27], SigCWGAN[28], ST-GAN[19]
Fraud Detection	Detection of market manipulation	LSTM-GAN[13]
	Detection of Credit Card Fraud	RWGAN[29], LSTM-GAN-2[30]

Eckerli, F. and Osterrieder, J., 2021. Generative adversarial networks in finance: an overview. Available at arXiv: 2106.06364.
https://arxiv.org/abs/2106.06364

Potential Uses of GANs in Economics

GENERATIVE ADVERSARIAL NETWORKS IN TIME SERIES: A SURVEY AND TAXONOMY

A PREPRINT

Eoin Brophy
Infant Research Centre & School of Computing
Dublin City University
Ireland
eoin.brophy7@mail.dcu.ie

Zhengwei Wang
ByteDance AI Lab
China

Qi She
ByteDance AI Lab
China

Tomás Ward
Insight SFI Research Centre for Data Analytics
Dublin City University
Ireland

July 26, 2021

ABSTRACT

Generative adversarial networks (GANs) studies have grown exponentially in the past few years. Their impact has been seen mainly in the computer vision field with realistic image and video manipulation, especially generation, making significant advancements. While these computer vision advances have garnered much attention, GAN applications have diversified across disciplines such as time series and sequence generation. As a relatively new niche for GANs, fieldwork is ongoing to develop high quality, diverse and private time series data. In this paper, we review GAN variants designed for time series related applications. We propose a taxonomy of discrete-variant GANs and continuous-variant GANs, in which GANs deal with discrete time series and continuous time series data. Here we showcase the latest and most popular literature in this field; their architectures, results, and applications. We also provide a list of the most popular evaluation metrics and their suitability across applications. Also presented is a discussion of privacy measures for these GANs and further protections and directions for dealing with sensitive data. We aim to frame clearly and concisely the latest and state-of-the-art research in this area and their applications to real-world technologies.

Keywords Generative Adversarial Networks · Time Series · Discrete-variant GANs · Continuous-variant GANs

Application	GAN Architecture(s)	Dataset(s)	Evaluation Metrics
Medical/Physiological Generation	LSTM-LSTM, [23] [62] [63], [64], [85], [50] LSTM-CNN, [71] [68] BiLSTM-CNN, [69] BiGridLSTM-CNN, [57] CNN-CNN, [86], [87] AE-CNN, [88] FCNN [89]	EEG, ECG, EHRs, PPG, EMG, Speech, NAF, MNIST, Synthetic sets	TSTR, MMD, Reconstruction error, DTW, PCC, IS, FID, ED, S-WD, RMSE, MAE, FD, PRD, Averaging Samples, WA, UAR, MV-DTW
Financial time series generation/prediction	TimeGAN [21] SigCWGAN [53] DAT-GAN [56] QuantGAN [46]	S&P500 index (SPX), Dow Jones Index (DJI), ETFs	Marginal Distributions, Dependencies, TSTR, Wasserstein Distance, EM distance, DY Metric, ACF score, leverage effect score, discriminative score, predictive score
Time series Estimation/Prediction	LSTM-NN [73] LSTM-CNN [74] LSTM-MLP [74]	Meteorological data, Truven MarketScan dataset	RMSE, MAE, NS, WI, LMI
Audio Generation	C-RNN-GAN [48] TGAN (variant) [81] RNN-FCN [90] DCGAN (variant) [80] CNN-CNN [72]	Nottingham dataset, Midi music files, MIR-1K, TheSession, Speech	Human perception, Polyphony, Scale Consistency, Tone Span, Repetitions, NSDR, SIR, SAR, FD, t-SNE, Distribution of notes
Time series Imputation/Repairing	MTS-GAN [75] CNN-CNN [91] DCGAN variant [92] AE-GRUI [93] RGAN [94] FCN-FCN [95] GRUI-GRUI [96]	TEP, Point Machine, Wind Turbin data, PeMS, PhysioNet Challenge 2012, KDD CUP 2018, Parking lot data,	Visually, MMD, MAE, MSE, RMSE, MRE, Spatial Similarity, AUC score
Anomaly Detection	LSTM-LSTM [78] LSTM-(LSTM&CNN) [97] LSTM-LSTM (MAD-GAN) [77]	SET50, NYC Taxi data, ECG, SWaT, WADI	Manipulated data used as a test set, ROC Curve, Precision, Recall, F1, Accuracy

Brophy et al. 2021. Generative adversarial networks in time series: A survey and taxonomy. Available at arXiv: 2107.11098.
https://arxiv.org/abs/2107.11098

Potential Uses of GANs in Economics

Fin-GAN: forecasting and classifying financial time series via generative adversarial networks

MILENA VULETIĆ*†‡, FELIX PRENZEL† and MIHAI CUCURINGU†‡§¶

†Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Rd, Oxford, OX2 6GC
‡Oxford-Man Institute of Quantitative Finance, University of Oxford, Eagle House, Walton Well Rd, Oxfor
§Department of Statistics, University of Oxford, 24-29 St Giles', Oxford, OX1 3LB, UK
¶The Alan Turing Institute, 96 Euston Rd, London, NW1 2DB, UK

(Received 18 January 2023; accepted 16 December 2023; published online 31 January 2024)

We investigate the use of Generative Adversarial Networks (GANs) for probabilistic forecasting of financial time series. To this end, we introduce a novel economics-driven loss function for the generator. This newly designed loss function renders GANs more suitable for a classification task, and places them into a supervised learning setting, whilst producing full conditional probability distributions of price returns given previous historical values. Our approach moves beyond the point estimates traditionally employed in the forecasting literature, and allows for uncertainty estimates. Numerical experiments on equity data showcase the effectiveness of our proposed methodology, which achieves higher Sharpe Ratios compared to classical supervised learning models, such as LSTMs and ARIMA.

Keywords: GANs; Financial returns; Time series forecasting; Classification

JEL Classifications: G17, C15, C22, C32, C45, C53

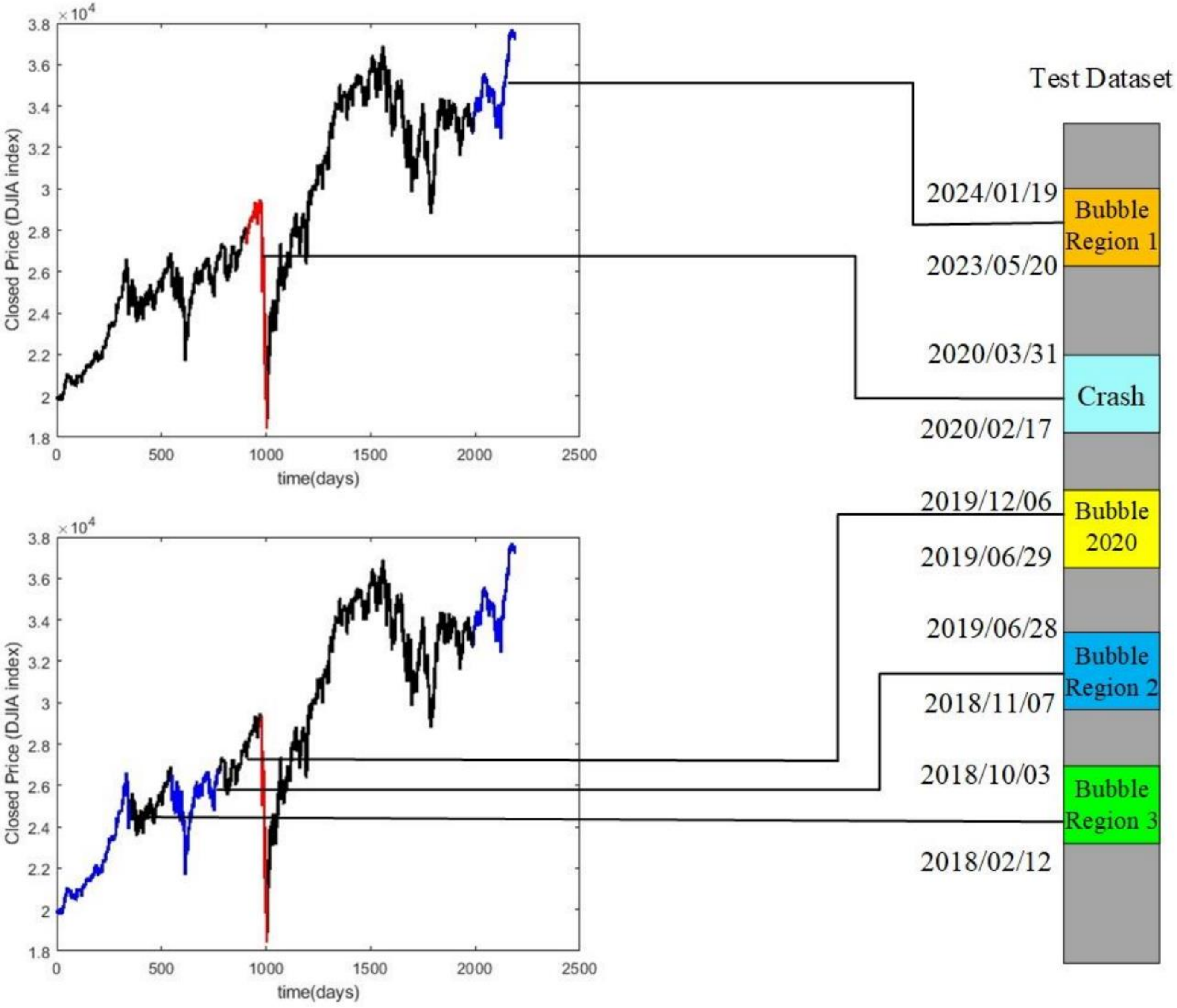
Table 2. Summary of performance metrics over the models across the stocks and ETFs.

	Fin-GAN	ForGAN	LSTM	LSTM-Fin	ARIMA	Long-only
Mean SR	**0.540**	0.033	0.467	0.341	0.206	0.182
Median SR	**0.413**	-0.092	0.214	0.170	0.204	0.194
Portfolio SR	**2.107**	0.172	2.087	0.942	0.612	0.618
Mean PnL	2.978	0.25	**4.123**	2.361	2.059	2.350
Median PnL	1.890	-0.673	1.959	1.735	**2.245**	1.975
Mean MAE	0.044	0.052	0.007	0.007	**0.007**	
Median MAE	0.008	0.009	**0.007**	0.007	0.007	
Mean RMSE	0.049	0.056	0.012	0.012	**0.012**	
Median RMSE	0.012	0.014	0.011	0.011	**0.011**	

Notes: SR refers to the annualized Sharpe Ratio, and PnL refers to the mean daily PnL. MAE and RMSE represent the mean absolute error and the mean root squared error, respectively. Highlighted are the best-performing results according to each metric.

Vuletić et al., 2024. Fin-GAN: Forecasting and classifying financial time series via generative adversarial networks. Quantitative Finance, pp.1-25.
https://www.tandfonline.com/doi/pdf/10.1080/14697688.2023.2299466

Potential Uses of GANs in Economics



Jithitikulchai and Kanjamapornkul (2024)

- Forecasting market bubbles and crashes focusing the US and EU stock indices and the EURUSD exchange rate during the COVID-19 pandemic-induced crash in March 2020

- Employ the time series econometric Diebold-Mariano (DM) test in conjunction with GANs

Potential Uses of GANs in Economics

GAN architecture

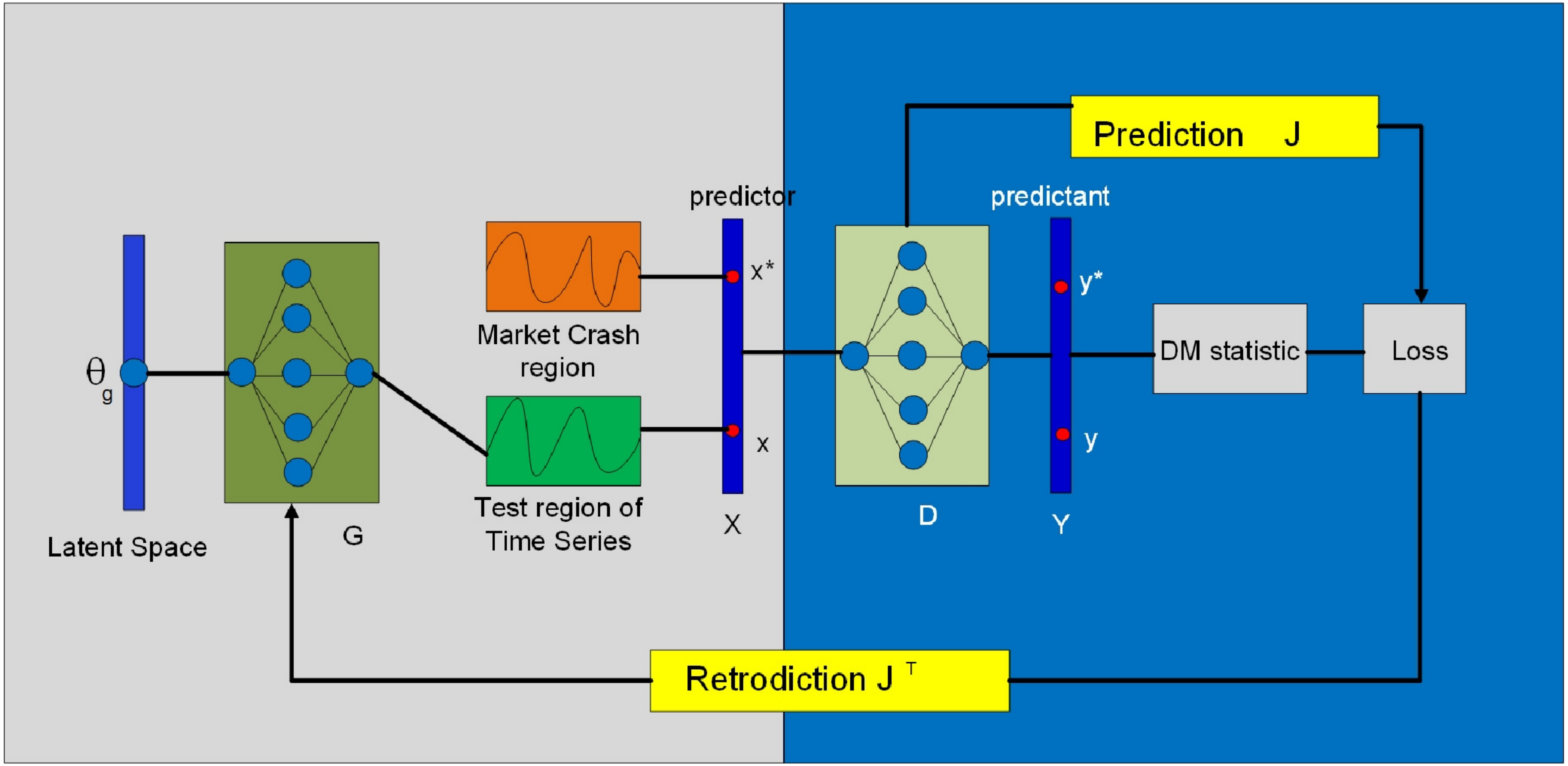
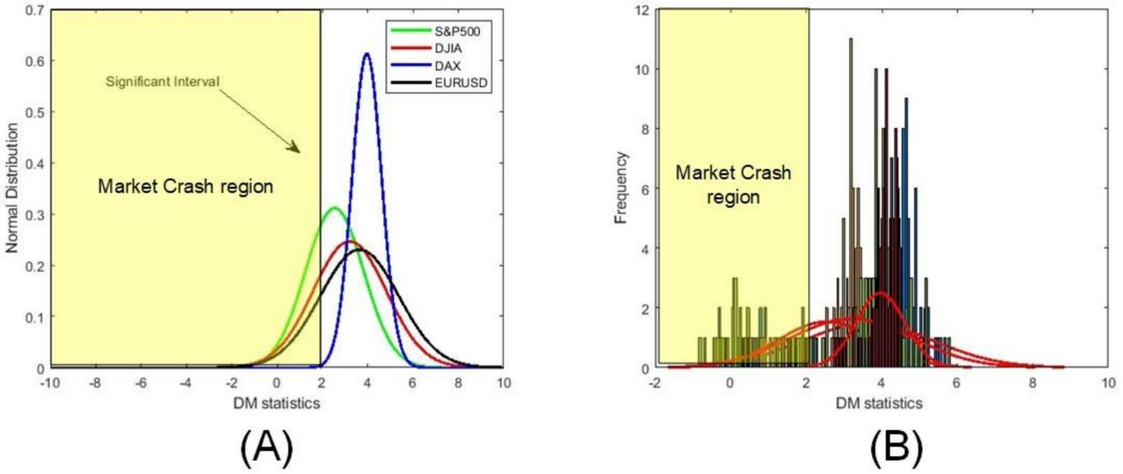
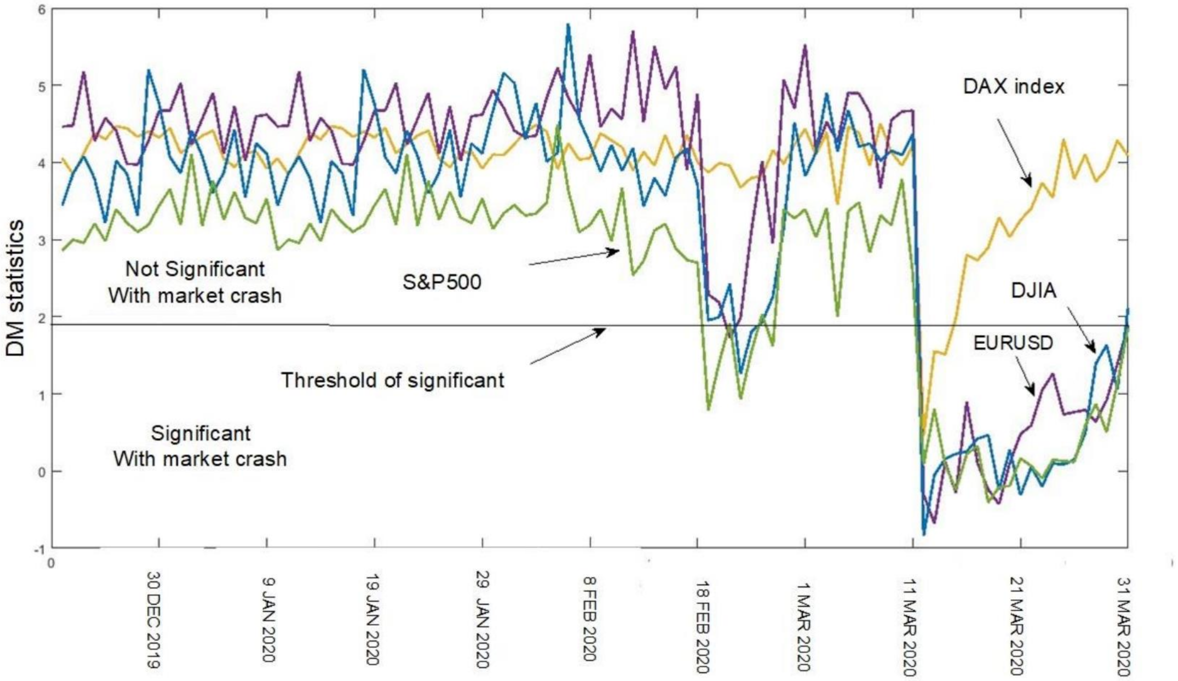


Figure 3: The diagram illustrates the architecture of the generator G and discriminator D for retrodiction in a Wilson loop generative adversarial network. It incorporates the loss function derived from the prediction of DM statistics for null hypothesis testing.

Jithitikulchai and Kanjamapornkul (2024)

Potential Uses of GANs in Economics

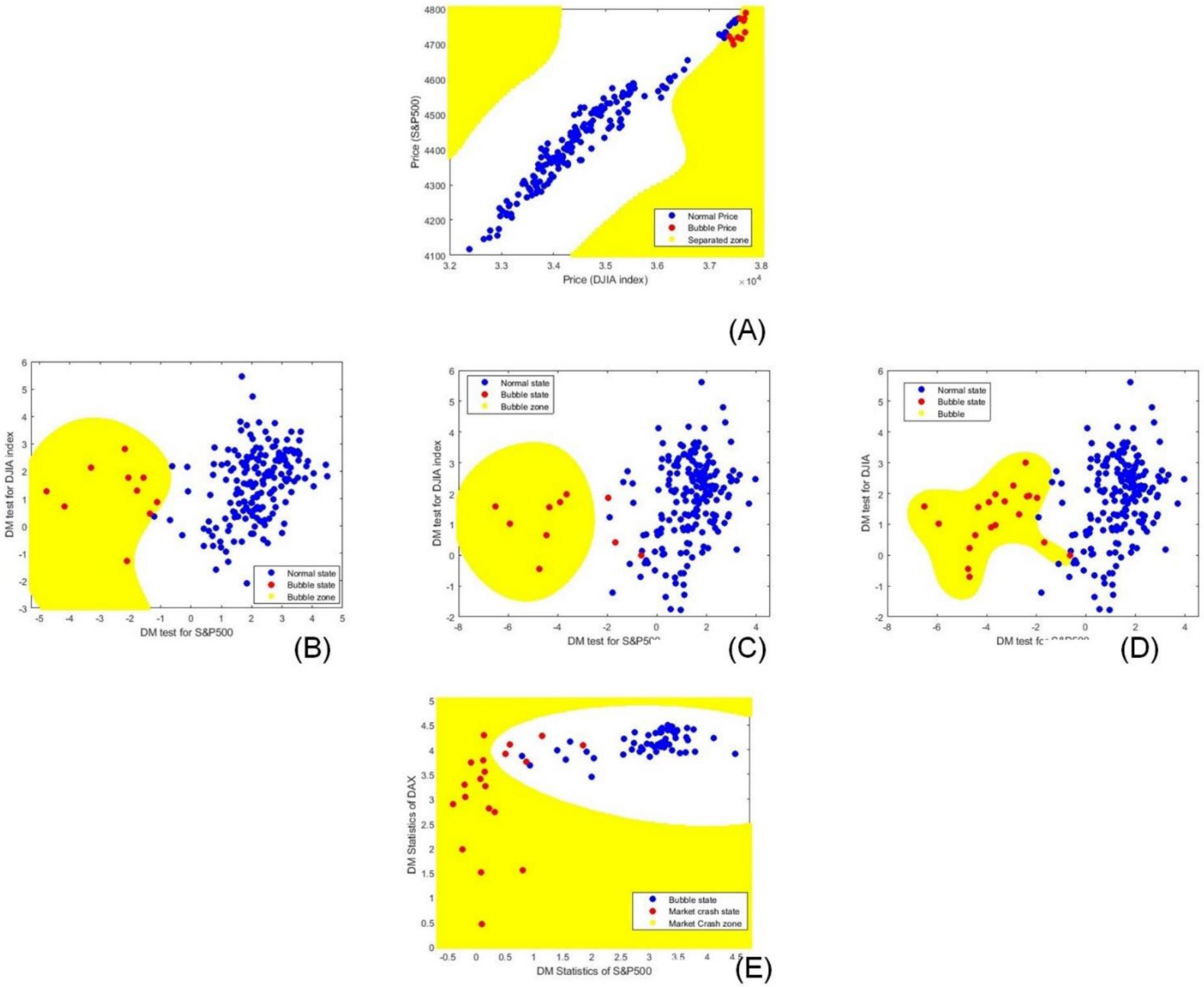




Jithitikulchai and Kanjamapornkul (2024)

- During market crashes, the DM statistics of the null hypothesis tests for EURUSD, S&P 500, Dow Jones Average Index, and DAX INDEX tend to decrease together with a notable level of significance.

Potential Uses of GANs in Economics



- Using Support Vector Machine (SVM) to classify market bubble and crash regions

Jithitikulchai and Kanjamapornkul (2024)

Potential Uses of GANs in Economics

- Data Augmentation
  - Increasing the amount of training data can improve model accuracy, especially when real data is limited.
  - For example, it can be used to study consumer behavior and forecasting, or financial time series data.

Potential Uses of GANs in Economics

- Synthetic Control Group
  - Synthetic control groups can be created when there is no actual control group.
  - This can help create balanced datasets to analyze the impact of policies.
  - For example, Professor Susan Athey's research (Athey, S., Imbens, G.W., Metzger, J. and Munro, E., 2021) applied GANs to various imbalanced data sets, both in terms of sample size (N) and observable characteristics.

Potential Uses of GANs in Economics

- Analyzing Complex Economic Relationships
  - GANs can be used to analyze complex economic relationships, such as nonlinear relationships between economic variables.
  - They can also be used to analyze the impact of social networks and behavior on economic decisions.
  - Research on developing GANs for nonlinear causal relationships is growing.

Potential Uses of GANs in Economics

- Synthetic Data for Privacy Protection
  - Synthetic data can be created to protect privacy, especially when data is sensitive or confidential.
  - For example, this can be used in research that requires access to population-level health or financial data.

Potential Uses of GANs in Economics

Software

Pytorch, e.g. StudioGAN
https://pytorch.org/tutorials/beginner/dcgan_faces_tutorial.html
https://www.coursera.org/specializations/generative-adversarial-networks-gans * * * * *

R, e.g. RGAN
https://cran.r-project.org/web/packages/RGAN/RGAN.pdf

MATLAB, e.g. dlarray, dlnetwork, Matlab-GAN
https://www.mathworks.com/help/deeplearning/ug/train-generative-adversarial-network.html

https://www.mathworks.com/matlabcentral/fileexchange/74865-matlab-gan

Tensorflow
https://github.com/tensorflow/gan

Potential Uses of GANs in Economics

Hardware

GPU - Graphics Processing Unit
- Most important hardware for training GANs
- Designed for parallel processing - computationally intensive tasks involved in training deep learning models
- Parallel training on a single system and several GPUs

CPU - Central Processing Unit
- Data preprocessing and feeding data to the GPU

RAM: Random Access Memory
- (Virtually) store the data that is being processed by the CPU and GPU

Storage, e.g. HDD, SDD, cloud
- Solid-state drives (SSDs) are recommended for faster data access times

Potential Uses of GANs in Economics

- Concluding Remarks
  - GANs are a powerful tool for economic research, offering solutions for data scarcity and enhancing research methods.
  - They can generate synthetic data, create counterfactuals, and address selection bias.
  - Research on GANs in economics is a growing field with continuous advancements.
  - GANs are a promising tool for economic research with the potential to address various challenges and improve the accuracy and efficiency of economic analysis.

APPENDIX

Relevant article (in Thai)

https://setthasarn.econ.tu.ac.th/blog/detail/676