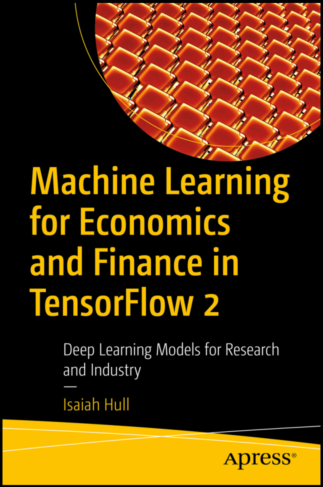
Lecture 1: Introduction to Machine Learning for Finance

Isaiah Hull[1,2]

[1]BI Norwegian Business School
[2]CogniFrame

October 24, 2023

Introduction



Lecture 1: Overview

1. Introduction to Machine Learning.
2. Machine Learning in Finance (and Economics).
3. Introduction to Python.
4. Introduction to TensorFlow.

Introduction

Course Materials

1. Introduction to Machine Learning

Introduction to Machine Learning

"Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome."

- Andrew Ng, Founder of Google Brain Team

Introduction to Machine Learning

- **What does it mean to be "explicitly programmed"?**
  - Expert systems: Define rules and apply symbolic logic.
  - Machine learning: Infer rules using examples and model.
- **Example: music production and classification.**
  - Explicit programming: Use expert-defined rules.

Introduction to Machine Learning

Timeline of AI Landmarks

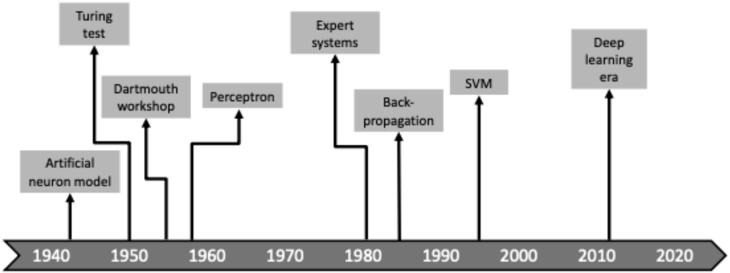


Figure 1: *A brief timeline of AI with some of the landmark advances.*

Source: Colliot (2023)

Summers and Winters of AI

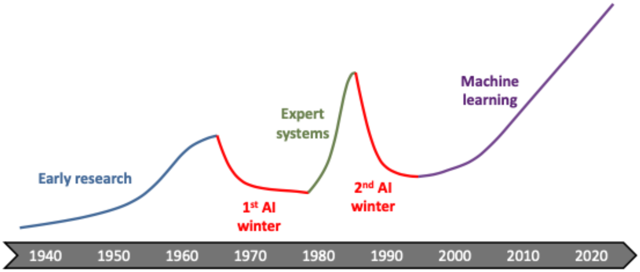
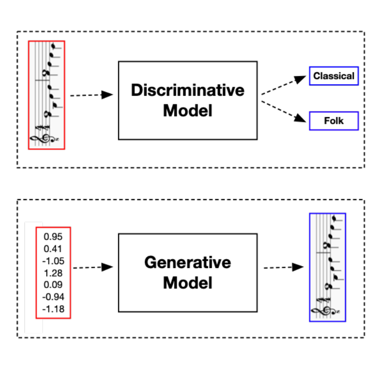


Figure 3: *Summers and winters of AI.*

Source: Colliot (2023)

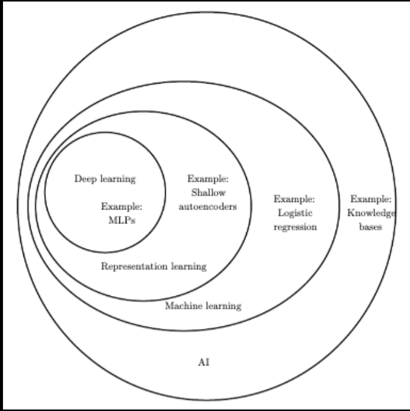
Machine Learning: Music Models



Source: “Using TensorFlow 2.0 to Compose Music” (DataCamp Tutorial)

Introduction to Machine Learning

Machine Learning and AI



Source: *Deep Learning* (Goodfellow et al., 2016).

Introduction to Machine Learning

Machine Learning: Definition

Definition: *A computer program is said to learn from experience E with respect to some class of tasks, T, and performance measure, P, if its performance at tasks in T, as measured by P, improves with experience E.*

Source: Mitchell (1997)

Machine Learning: Examples (1/3)

- **A checkers learning problem:**
  - Task T: playing checkers.
  - Performance measure P: percent of games won against opponents.
  - Training experience E: playing practice games against itself.

Source: Mitchell (1997)

Machine Learning: Examples (2/3)

- **A handwriting recognition learning problem:**
  - Task T: recognizing and classifying handwritten words within images.
  - Performance measure P: percent of words correctly classified.
  - Training experience E: a database of handwritten words with given classifications.

Source: Mitchell (1997)

Machine Learning: Examples (3/3)

- **A robot driving learning problem:**
  - Task T: driving on public four-lane highways using vision sensors.
  - Performance measure P: average distance traveled before an error (as judged by human overseer).
  - Training experience E: a sequence of images and steering commands recorded while observing a human driver.

Source: Mitchell (1997)

Introduction to Machine Learning

Machine Learning Use Cases

1. Unclear how to program a function that the human brain conducts.
    - Recognize 3D object from different viewpoints and lighting conditions in cluttered scene.
2. No simple and reliable rules for task.
    - Identifying fraudulent transactions may require application of large number of weak rules.

Source: Hinton et al. (2012)

Machine Learning Use Cases

- MNIST dataset.
  - Database of handwritten digits.
  - Publicly available.
  - Compare machine learning methods.
  - Strong performance in moderate-sized neural networks.

Source: Hinton et al. (2012)

Machine Learning Use Cases

MNIST

Introduced by LeCun et al. in Gradient-based learning applied to document recognition

The **MNIST** database (**Modified National Institute of Standards and Technology** database) is a large collection of handwritten digits. It has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger NIST Special Database 3 (digits written by employees of the United States Census Bureau) and Special Database 1 (digits written by high school students) which contain monochrome images of handwritten digits. The digits have been size-normalized and centered in a fixed-size image. The original black and white (bilevel) images from NIST were size normalized to fit in a 20x20 pixel box while preserving their aspect ratio. The resulting images contain grey levels as a result of the anti-aliasing technique used by the normalization algorithm. the images were centered in a 28x28 image by computing the center of mass of the pixels, and translating the image so as to position this point at the center of the 28x28 field.

Source: http://yann.lecun.com/exdb/mnist/

Homepage

Source: https://www.paperswithcode.com

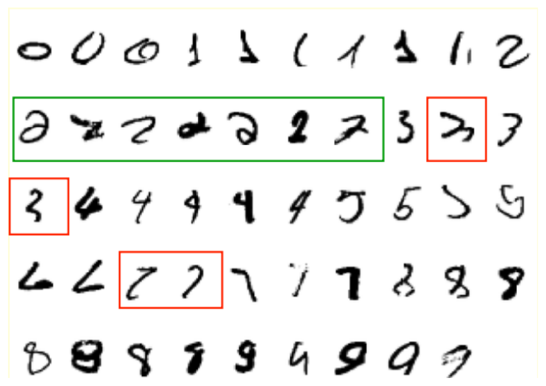
Introduction to Machine Learning

Machine Learning Use Cases

Rank	Model	Percentage error	Accuracy	Trainable Parameters	Error rate	Extra Training Data	Paper	Code	Result	Year	Tags
1	**Branching/Merging CNN + Homogeneous Vector Capsules**	0.13	99.87	1,514,187		×	No Routing Needed Between Capsules			2020	
2	**EnsNet** (Ensemble learning in CNN augmented with fully connected subnetworks)	0.16	99.84			×	Ensemble learning in CNN augmented with fully connected subnetworks			2020	
3	**Efficient-CapsNet**	0.16	99.84	161,824		×	Efficient-CapsNet: Capsule Network with Self-Attention Routing			2021	
4	**SOPCNN** (Only a single Model)	0.17	99.83	1,400,000		×	Stochastic Optimization of Plain Convolutional Neural Networks with Simple methods			2020	
5	**RMDL** (30 RDLs)	0.18	99.82			×	RMDL: Random Multimodel Deep Learning for Classification			2018	
6	**DropConnect**	0.21	99.77			×	Regularization of Neural Networks using DropConnect			2013	
7	**MCDNN**	0.23				×	Multi-column Deep Neural Networks for Image Classification			2012	

Source: https://www.paperswithcode.com

What Makes a 2?



Source: Hinton et al. (2012)

Machine Learning Tasks

1. **Supervised: Model trained using "targets" or "labels."**
   - E.g. deep learning, tree-based models, linear regression.
2. **Unsupervised: No target or self target.**
   - E.g. k-means clustering, autoencoders.
3. **Semi-Supervised: Combination of labeled and unlabeled data.**
   - E.g. Large language models (LLMs).
4. **Reinforcement Learning: Learning behavior to maximize reward.**
   - E.g. video games, high frequency trading.

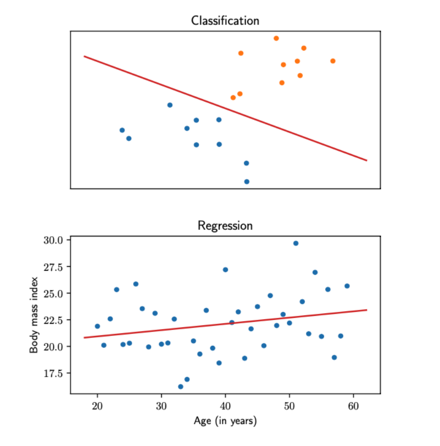
Supervised Learning: Definition

Definition: *We have* N *labelled training examples* $\mathcal{D}=\{(x_n, t_n)\}_{n=1}^{N}$, *where* x_n *represents a covariate, or explanatory variable, while* t_n *is the corresponding label, or response. ... The goal of supervised learning is to predict the value of the label* t *for an input* x *that is not in the training set. In other words, supervised learning aims at generalizing the observations in the data set* $\mathcal{D}$ *to new inputs.*

Source: Simeone (2018)

Introduction to Machine Learning

Supervised Learning: Classification vs. Regression



Source: Colliot (2023)

Unsupervised Learning: Definition

Definition: *Suppose now that we have an unlabelled set of training examples* $\mathcal{D}=\{x_n\}_{n=1}^N$. *Less well defined than supervised learning, unsupervised learning generally refers to the task of learning properties of the mechanism that generates this data set.*

Source: Simeone (2018)

Introduction to Machine Learning

Unsupervised Learning

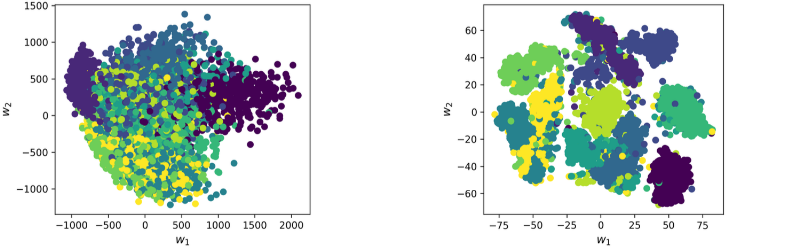
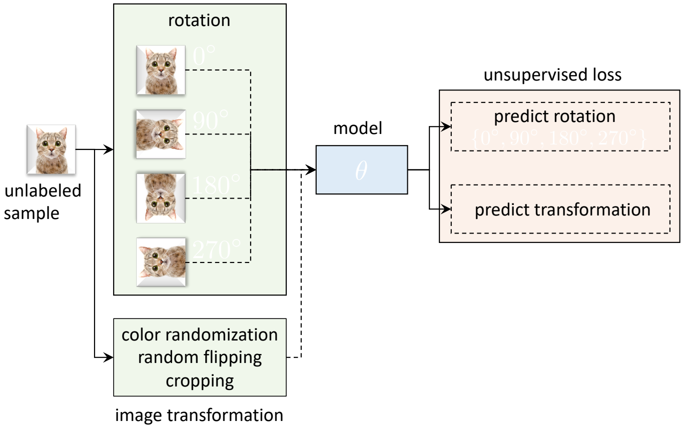


Figure 6: **PCA vs. t-SNE** Application of both methods on 5000 samples from the MNIST handwritten digit dataset. We see that perfect clustering cannot be achieved with either method, but t-SNE delivers the much better result.

Source: Chen et al. (2022).

Unsupervised Learning



Source: Chen et al. (2022).

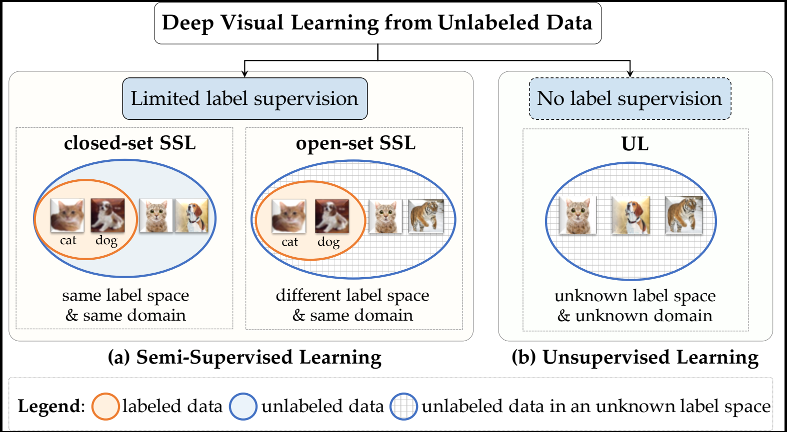
Semi-Supervised Learning: Definition

Definition: *Scenarios in which not all examples are labelled, with the unlabelled examples providing information about the distribution of the covariates* $\mathbf{x}$.

Source: Simeone (2018)

Introduction to Machine Learning

Semi-Supervised Learning



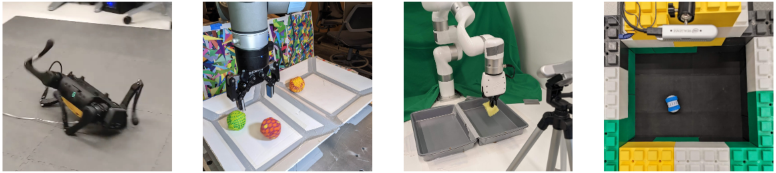
Source: Chen et al. (2022).

Reinforcement Learning: Definition

Definition: *... the problem of inferring optimal sequential decisions based on rewards or punishments received as a result of previous actions. ... Upon taking an action t in a state x, the learner is provided with feedback on the immediate reward accrued via this decision, and the environment moves on to a different state. As an example, an agent can be trained to navigate a given environment in the presence of obstacles by penalizing decisions that result in collisions.*

Source: Simeone (2018)

Reinforcement Learning: Dreamer



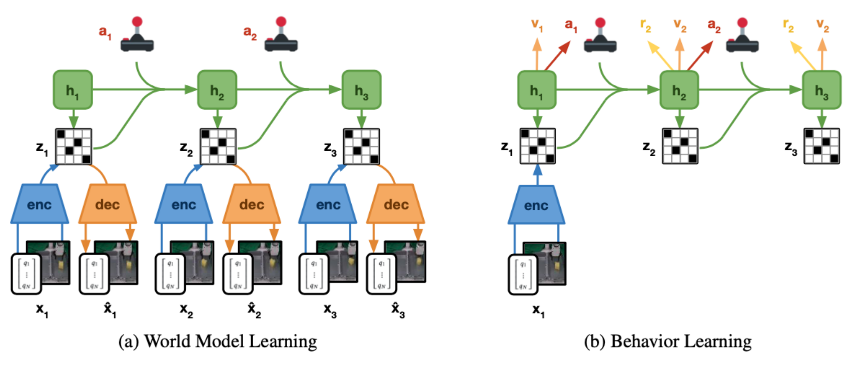
(a) A1 Quadruped Walking (b) UR5 Visual Pick Place (c) XArm Visual Pick Place (d) Sphero Navigation

Figure 1: To study the applicability of Dreamer for sample-efficient robot learning, we apply the algorithm to learn robot locomotion, manipulation, and navigation tasks from scratch in the real world on 4 robots, without simulators. The tasks evaluate a diverse range of challenges, including continuous and discrete actions, dense and sparse rewards, proprioceptive and camera inputs, as well as sensor fusion of multiple input modalities. Learning successfully using the same hyperparameters across all experiments, Dreamer establishes a strong baseline for real world robot learning.

Source: Chen et al. (2022).

Introduction to Machine Learning

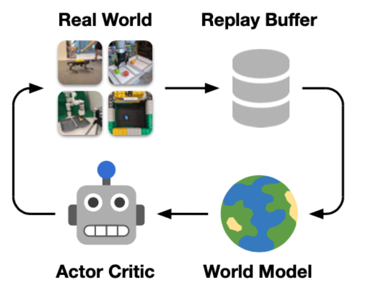
Reinforcement Learning: Dreamer



Source: Chen et al. (2022).

Introduction to Machine Learning

Reinforcement Learning: Dreamer



Source: Chen et al. (2022).

Introduction to Machine Learning

Reinforcement Learning: Dreamer

Figure 4: **A1 Quadruped Walking** Starting from lying on its back with the feet in the air, Dreamer learns to roll over, stand up, and walk in 1 hour of real world training time, without simulators or resets. In contrast, SAC only learns to roll over but neither to stand up nor to walk. For SAC, we also had to help the robot out of a dead-locked leg configuration during training. On the right we show training curves for both SAC and Dreamer. The maximum reward is 14. The filled circles indicate times where the robot fell on its back, requiring the learning of a robust strategy for getting back up. After 1 hour of training, we start pushing the robot and find that it adapts its behavior within 10 minutes to withstand light pushes and quickly roll back on its feet for hard pushes. The graph shows a single training run with the shaded area indicating one standard deviation within each time bin.

Source: Chen et al. (2022).

Reinforcement Learning: Dreamer

Figure 5: **UR5 Multi Object Visual Pick and Place** This task requires learning to locate three ball objects from third-person camera images, grasp them, and move them into the other bin. The arm is free to move within and above the bins and sparse rewards are given for grasping a ball and for dropping it in the opposite bin. The environment requires the world model to learn multi-object dynamics in the real world and the sparse reward structure poses a challenge for policy optimization. Dreamer overcomes the challenges of visual localization and sparse rewards on this task, learning a successful strategy within a few hours of autonomous operation.

Source: Chen et al. (2022).

Introduction to Machine Learning

Criteria for Usefulness of ML (Brynjolfsson and Mitchell, 2017)

1. Learning a function that maps well-defined inputs to well-defined outputs.
2. Large (digital) data sets exist or can be created containing input-output pairs.
3. The task provides clear feedback with clearly definable goals and metrics.
4. No long chains of logic or reasoning that depend on diverse background knowledge or common sense.

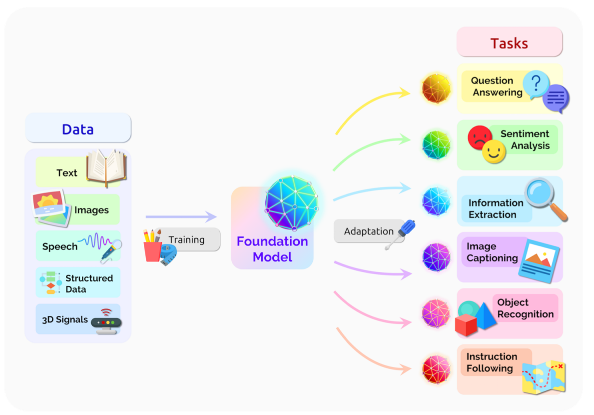
Introduction to Machine Learning

Criteria for Usefulness of ML (Brynjolfsson and Mitchell, 2017)

5. No need for detailed explanation of how the decision was made.
6. A tolerance for error and no need for provably correct or optimal solutions.
7. The phenomenon or function being learned should not change rapidly over time.
8. No specialized dexterity, physical skills, or mobility required.

Introduction to Machine Learning

Foundation Models



Source: Bommasani et al. (2022).

Masked Language Modeling

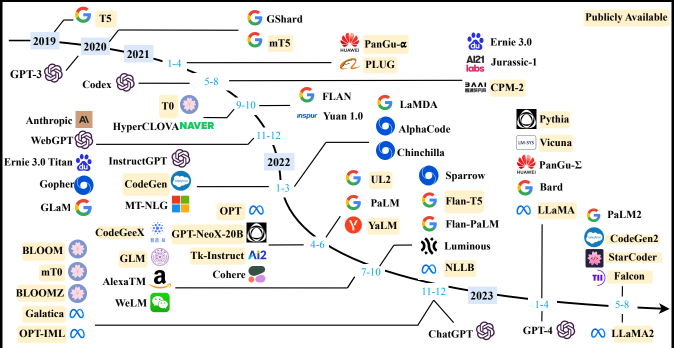
Sequence: *"Commercial [MASK]$_1$ have long been thought of, and indeed have functioned as, the backup source of [MASK]$_2$ for many other financial institutions and markets."*

Labels: *[MASK]$_1$ = banks, [MASK]$_2$ = liquidity.*

Source: Bertsch et al. (2022)

Introduction to Machine Learning

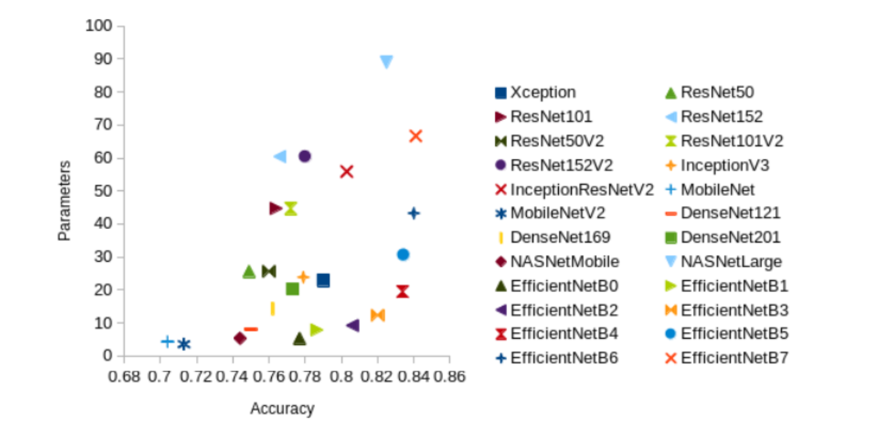
Large Language Models



Source: Zhao et al. (2023).

Introduction to Machine Learning

Image Classification Foundation Models



Source: Plested and Gedeon (2022).

2. Machine Learning in Finance (and Economics)

Overview

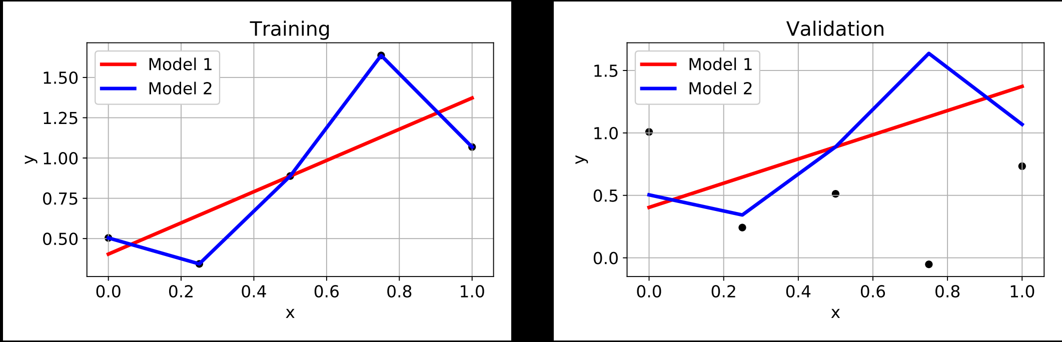
- **Objectives: Prediction and classification, rather than hypothesis testing and causality.**
  - Strong agreement on objectives in forecasting exercises.
- **Models: Concentration on non-linearities, model architecture, transfer learning, and feature extraction.**
  - Deep learning versus shallow learning.
- **Training: Typically stochastic and in batches, rather than deterministic.**

"Big Data: New Tricks for Econometrics"
Varian (2014)

- **Early examination of potential ML uses in econometrics.**
  - ML takes different approach to model uncertainty and validation than econometrics.
- **Machine learning approach to validation.**
  - Model evaluation primarily performed using out-of-sample prediction (cross validation).
  - Alternative to goodness-of-fit measures in econometrics.

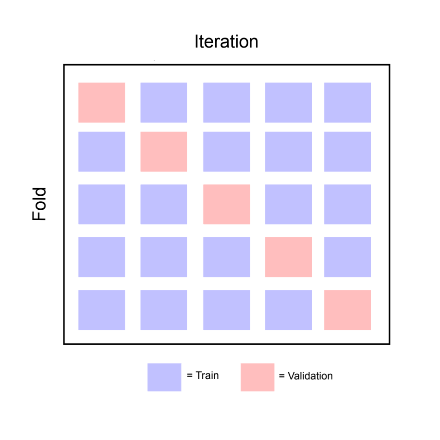
Machine Learning in Economics and Finance

Model Selection



Source: Hull (2021)

Cross Validation



Source: Hull (2021)

Machine Learning in Economics and Finance

"Big Data: New Tricks for Econometrics"

Varian (2014)

- **ML methods that might be applied in econometrics.**
  - Classification and regression trees.
  - Variable selection techniques: LASSO, spike-and-slab regression.
- **Combining models into ensembles.**
  - Economists typically try to find "true" model.
  - ML combines weak models with bagging, boosting, and bootstrapping.

Machine Learning in Economics and Finance

"Prediction Policy Problems"

Kleinberg et al. (2015)

- **Prediction policy problem:** Accurate predictions more important than demonstration of causality.
- **Relevance:** Impact of policy is known, but unknown whether policy is needed.
- **Application:** Differentiating scenarios where prediction is more important than causal inference.

Machine Learning in Economics and Finance

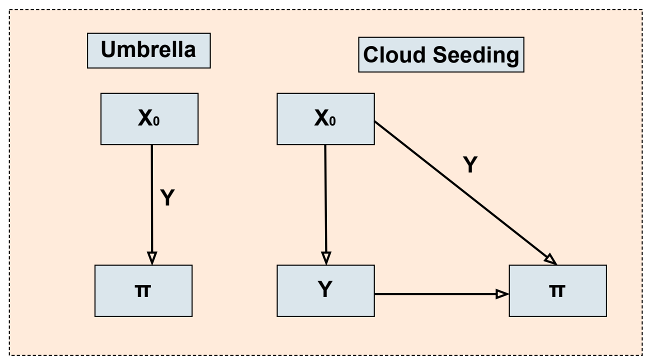
"Prediction Policy Problems"

Kleinberg et al. (2015)

- **Drought Policy:** Cloud seeding to increase probability of rain.
  - Causal impact unknown.
- **Umbrella Choice:** Anticipating the likelihood of rain.
  - Weather unknown.
- **Outcome:** The impact of rainfall intensity in both policy decisions.
  - Causality not always most important issue.

"Prediction Policy Problems"

Kleinberg et al. (2015)



Source: Hull (2021)

Machine Learning in Economics and Finance

“Prediction Policy Problems”

Kleinberg et al. (2015)

- **Implications:** A new subfield in economics
  - Utilizing prediction over causal inference in policy problems
  - Facilitating policy decisions using machine learning techniques

Machine Learning in Economics and Finance

“Machine Learning: An Applied Econometric Approach”

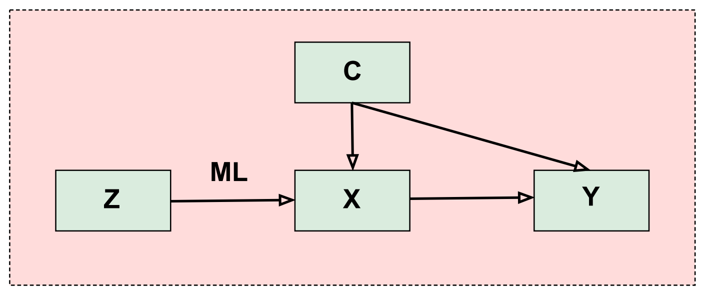
Mullainathan and Spiess (2017)

- Similar to Kleinberg et al. (2015), argue that ML should be used in economics for predictive tasks.
  - **Measuring economic activity.** Using image or text datasets to measure economic activity in developing countries or at a higher frequency.
  - **Inference tasks with prediction step.** Certain inference tasks, such as instrumental variables regression, involve intermediate step where fitted values are generated.
  - **Policy applications.** Recommendations for policy often depend on prediction.

Machine Learning in Economics and Finance

<u>"Machine Learning: An Applied Econometric Approach"</u>

<u>Mullainathan and Spiess (2017)</u>



Source: Hull (2021)

"The Impact of Machine Learning on Economics"
Athey (2018)

- **Background:**
  - Discussion on the influence of ML in economics.
  - Comparison between ML and traditional econometric methods.
- **Focus Areas:**
  - Evaluation of ML routines for economics.
  - Policy prediction problems discussion.

Machine Learning in Economics and Finance

Functional Forms and Data Handling (Athey, 2018)

- **Econometric Methods:**
  - Concentration on causal inference in linear regression models.
  - Limited in capturing nonlinear relationships.
- **ML Advantages:**
  - Better suited for handling big data.
  - Allows for nonlinearities between features.

Machine Learning in Economics and Finance

Empirical Analysis (Athey, 2018)

- **Traditional Approach:**
  - Model selection based on principles and theory.
  - Single estimation process.
- **ML Approach:**
  - Iterative approach with tuning and cross-validation.
  - Focus on improving performance through empirical analysis.

Challenges and Opportunities (Athey, 2018)

- **Evaluation Metrics:**
  - Simple and measurable performance evaluation in ML.
  - Causality remains immeasurable.
- **Concerns:**
  - Difficulty in improving causality dimension through ML.
  - The challenge of optimizing for “causality” in ML.

Machine Learning Challenges (Athey, 2018)

- **Confidence Intervals:**
  - The complication of deriving valid confidence intervals in ML.
  - Dependence on advanced methods to overcome restrictions.
- **Hypothesis Testing:**
  - Centered around statistical significance of parameters in economics.
  - Not a focal point in ML due to high parameter count.

Machine Learning in Economics and Finance

Unsupervised ML Methods (Athey, 2018)

- **Benefits:**
  - Avoidance of spurious relationships.
  - Generation of a dependent variable through methods like clustering.
- **Supervised ML Methods:**
  - Evaluation and classification according to adoption to econometrics.
  - Neural networks gaining increased acceptance.

ML Models and Trade-offs (Athey, 2018)

- **ML Models:**
  - Inclusion of regularized regression, SVM, and matrix averaging.
  - Avoiding overfitting while allowing for increased flexibility with feature extraction.
- **Applicability:**
  - Suitable for high number of covariates.
  - The necessity of non-standard routines for confidence intervals.

Machine Learning in Economics and Finance

Policy Prediction and Economics (Athey, 2018)

- **Focus:**
  - Understanding the role of prediction in economic policy analysis.
  - The relevance of non-causal associations in economic forecasting.
- **Real-world applications:**
  - Example of a financial institution's decision-making process.
  - Insight into policy prediction tools and mechanisms.

ML for Policy Prediction (Athey, 2018)

- **Prospects:**
  - Potential of ML in risk assessment and prediction.
  - Understanding how prediction affects policy analysis.
- **Limitations:**
  - The restrictions of using ML in policy analysis.
  - Balancing prediction accuracy and causality.

Machine Learning in Economics and Finance

"Machine Learning Methods Economists Should Know About"
Athey and Imbens (2019)

1. **Local linear forests.**

2. **Neural networks.**

3. **Boosting.**

4. **Classification trees and forests.**

"Machine Learning Methods Economists Should Know About"
Athey and Imbens (2019)

5. **Unsupervised learning with k-means clustering and GANs.**
6. **Average treatment effects under the confoundedness assumption.**
7. **Orthogonalization and cross-fitting.**
8. **Heterogeneous treatment effects.**

"Machine Learning Methods Economists Should Know About"
Athey and Imbens (2019)

9. **Experimental design and reinforcement learning.**
10. **Matrix completion and recommender systems.**
11. **Synthetic control methods.**
12. **Text analysis.**

Natural Language Processing

Sentiment Analysis

September 18, 2019

Federal Reserve issues FOMC statement

For release at 2:00 p.m. EDT

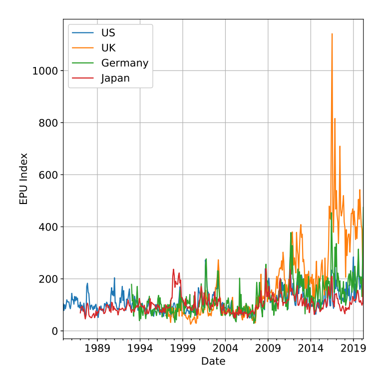
Share

Information received since the Federal Open Market Committee met in July indicates that the labor market remains strong and that economic activity has been rising at a moderate rate. Job gains have been solid, on average, in recent months, and the unemployment rate has remained low. Although household spending has been rising at a strong pace, business fixed investment and exports have weakened. On a 12-month basis, overall inflation and inflation for items other than food and energy are running below 2 percent. Market-based measures of inflation compensation remain low; survey-based measures of longer-term inflation expectations are little changed.

Source: Hull (2021)

Natural Language Processing

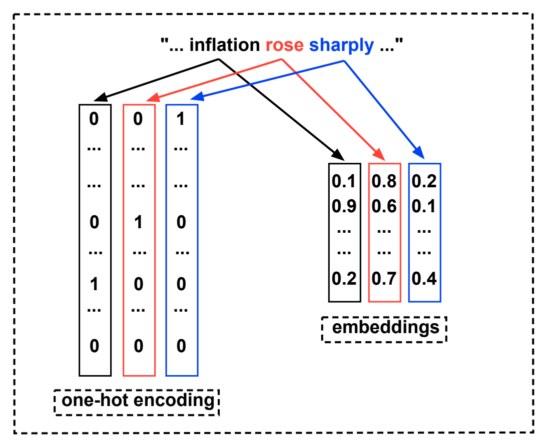
Economic Policy Uncertainty



Source: Hull (2021)

Natural Language Processing

Representing Text as Data



Source: Hull (2021)

Natural Language Processing

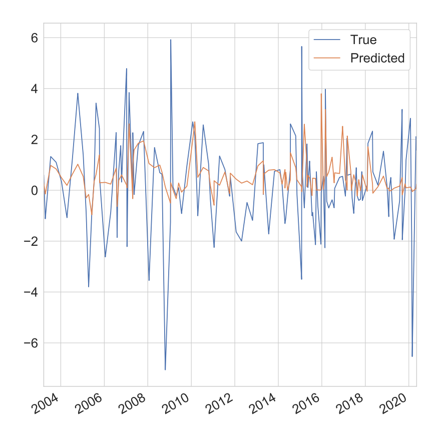
Word Embeddings

Dimension	king	queen	prince	man	woman	child
Royalty	0.99	0.99	0.95	0.01	0.02	0.01
Masculinity	0.94	0.06	0.02	0.99	0.02	0.49
Age	0.73	0.81	0.15	0.61	0.68	0.09
...						

Source: Gentzkow et al. (2019)

Natural Language Processing

Stock Price Prediction with Text Features



Source: "Machine Learning for Economics and Finance in TensorFlow 2"

Tree-Based Models

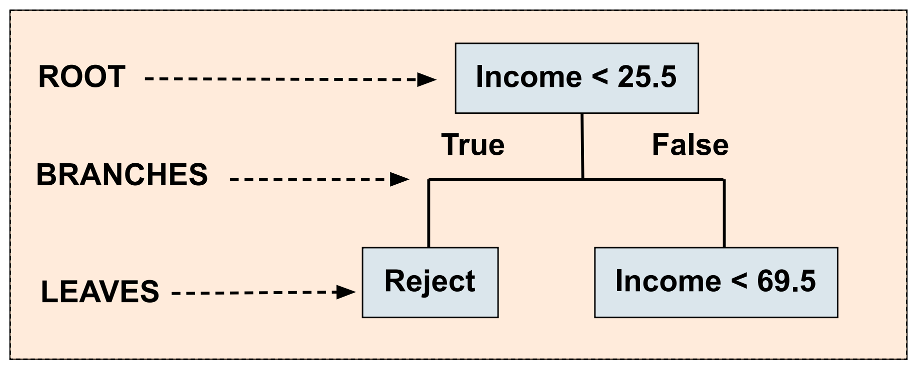
Tree-Based Models

<u>Overview</u>

1. **Perform sequential partition of data.**
   - Move from “root” to “leaves.”
2. **Achieve state-of-the-art forecasting performance.**
   - Gradient boosted trees.
   - Random forests.
3. **More interpretable than neural networks.**
   - Legally required for some problems.

Tree-Based Models

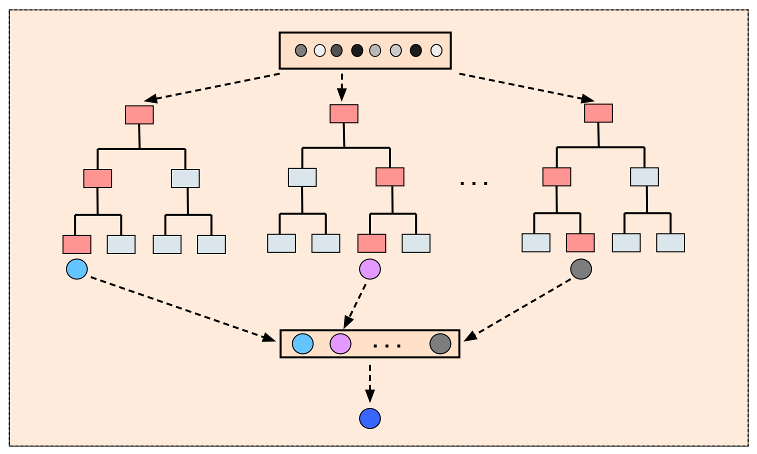
Tree-Based Model: Loan Originations



Source: "Machine Learning for Economics and Finance in TensorFlow 2"

Tree-Based Models

Random Forest Model



Source: "Machine Learning for Economics and Finance in TensorFlow 2"

Time Series Forecasting

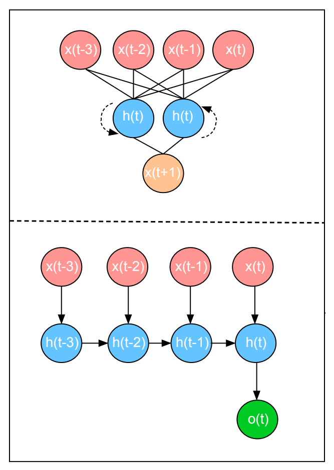
Time Series Forecasting

Overview

1. **Short time series not ideal application for ML models.**
   - Non-stationary series, time-varying relationships.
2. **Deep learning and tree-based models most promising.**
   - Transformer models, RNNs, LSTMs, RFs, gradient boosting.
3. **New features may be more useful than new models.**
   - Text and images.

Time Series Forecasting

Sequential Deep Learning Models



Source: "Machine Learning for Economics and Finance in TensorFlow 2"

Dimensionality Reduction

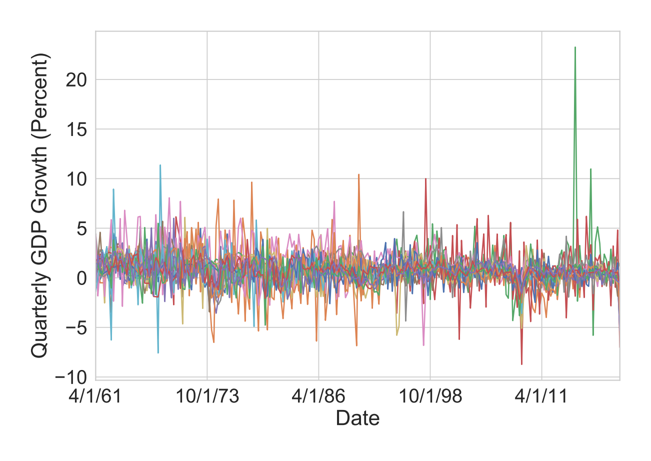
Dimensionality Reduction

Overview

1. **Need to reduce dimensionality of data.**
   - Text is naturally high dimensional.
2. **Identify common component in multiple series.**
   - Sectoral component of stock price movements.
   - Cross-country variation in GDP growth.
3. **Typically done with PCA or PLS.**
   - Also possible with autoencoder.

Dimensionality Reduction

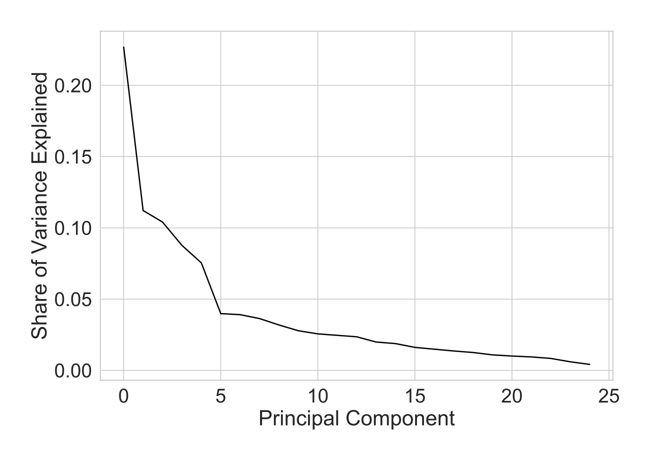
GDP Growth for 25 Countries



Source: “Machine Learning for Economics and Finance in TensorFlow 2”

Dimensionality Reduction

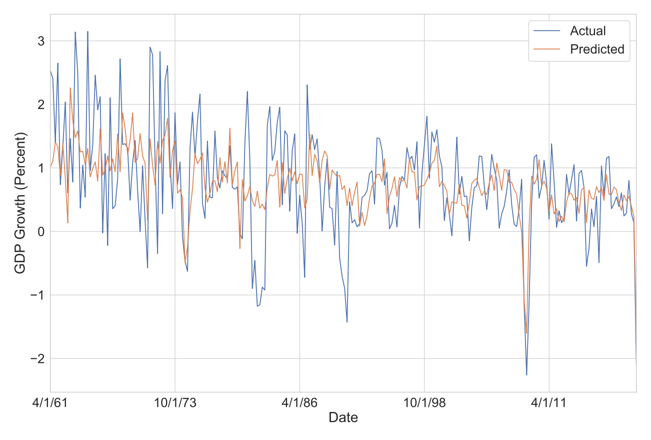
Principal Component Analysis



Source: "Machine Learning for Economics and Finance in TensorFlow 2"

Dimensionality Reduction

PCR-Predicted Growth for Canada

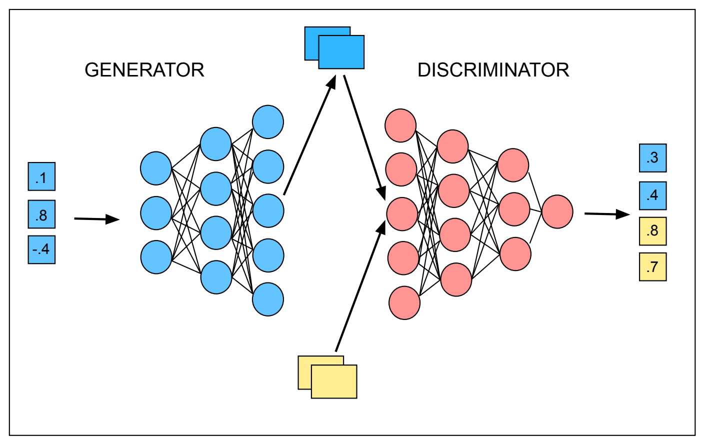


Source: “Machine Learning for Economics and Finance in TensorFlow 2”

Generative Machine Learning

Generative Machine Learning

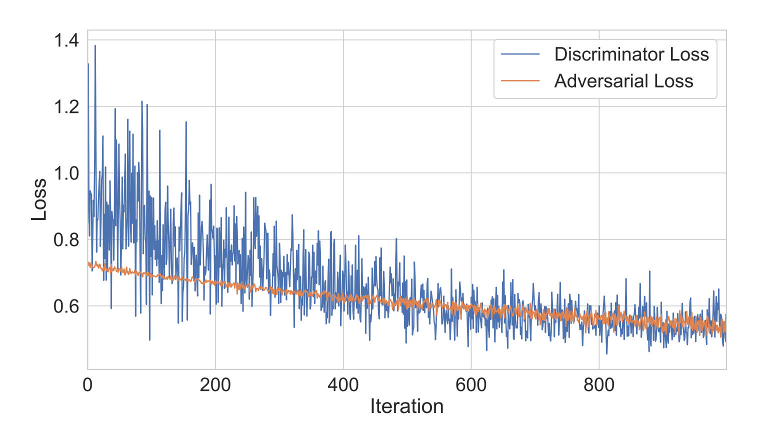
Generative Adversarial Networks (GANs)



Source: "Machine Learning for Economics and Finance in TensorFlow 2"

Generative Machine Learning

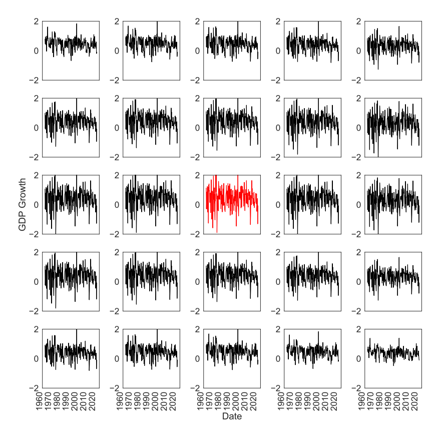
Stable Evolutionary Equilibrium



Source: "Machine Learning for Economics and Finance in TensorFlow 2"

Generative Machine Learning

Simulated GDP Growth Series



Source: “Machine Learning for Economics and Finance in TensorFlow 2”

3. Introduction to Python

Introduction to Python

Colab Tutorial

- Introduction to Python

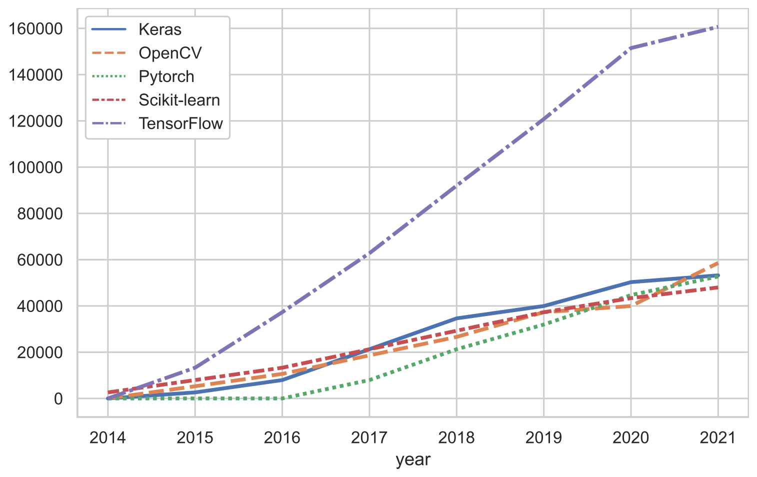
4. Introduction to TensorFlow

Introduction to TensorFlow

TensorFlow Overview

- **Open source framework for ML introduced by Google's Brain Team.**
  - Large community, tools for production settings, capacity for distributed training.
- **Built for neural networks, but can be used for any graph-based model.**
  - Tree-based models, theoretical models in economics and finance, reinforcement learning.

Introduction to TensorFlow



GitHub stars by ML framework (Perrault et al., 2022).

Introduction to TensorFlow

TensorFlow for Economics and Finance

1. Causal Inference
2. Feature Extraction
3. Non-linear Modeling
4. Simulation
5. Dimensionality Reduction
6. Reinforcement Learning
7. Model Uncertainty

Introduction to TensorFlow

Advantages of TensorFlow

1. **Flexibility.** Suitable for non-standard and causal inference tasks.
2. **Distributed Training.** Automated GPU and TPU detection.
3. **Production Quality.** Transition from experimental to stable production codes.
4. **High-Quality Documentation.** Detailed and user-friendly with Google Colab integration.
5. **Extensions.** Diverse tools tailored for specific needs.

Flexibility and Distributed Training

- Flexibility is crucial for non-standard tasks in economics.
- TensorFlow provides a mix of high and low-level APIs.
- Example: Nesting a deep neural network within an econometric estimation routine.
- Distributed Training essential for tasks like satellite image-based trade flow predictions.

Introduction to TensorFlow

Production Quality and Documentation

- Transition from experimental to stable, bug-free code.
- TensorFlow offers Estimators API for high-quality model development.
- TensorFlow Serving: Deploy models for end-user applications.
- Improved documentation with Google Colab notebooks for hands-on experience.

Introduction to TensorFlow

TensorFlow Extensions

1. **TensorFlow Hub.** Pre-trained models library.
2. **TensorFlow Probability.** Supports MCMC, BFGS, etc.
3. **TensorFlow Federated.** Training models with decentralized data.
4. **TensorFlow Lite.** Model deployment in resource-constrained environments.

Introduction to TensorFlow

tf.keras

1. High-level submodule for neural networks.
2. Sequential model, functional model, custom (subclassing).
3. Provides TF-related integration not included in standalone Keras.

tf.estimator

1. Tree-based models, linear models, neural networks.
2. Restricted framework with small number of choices.
3. Eliminates common errors and is ideal for production settings.

Introduction to TensorFlow

TensorFlow 1

```
>>> import tensorflow as tf
>>> c = tf.constant(1.0)
>>> print(c)

Tensor("Const_2:0",
shape=(), dtype=float32)

>>> with tf.Session() as sess:
        print(c.eval())

1.0
```

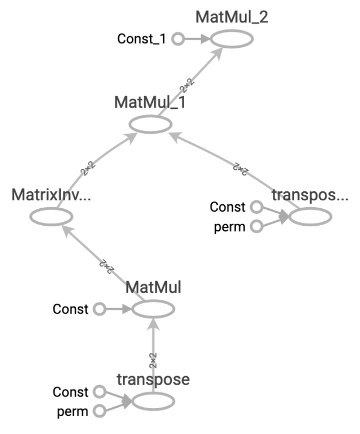
TensorFlow 2

```
>>> import tensorflow as tf
>>> c = tf.constant(1.0)
>>> print(c)

<tf.Tensor: shape=(),
dtype=float32, numpy=1.0>
```

Introduction to TensorFlow

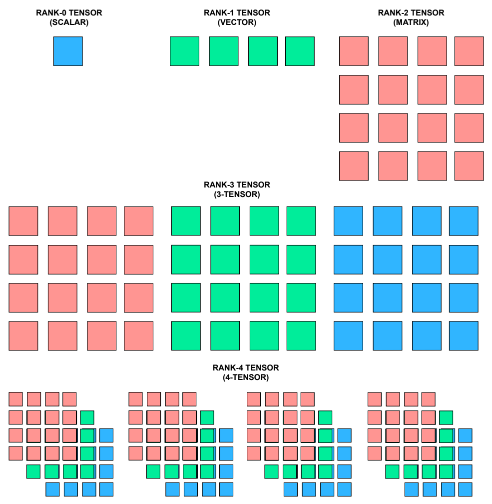
Static Graph



TensorFlow 2

```
>>> @tf.function
>>> def ols_predict(X, beta):
        yhat = tf.matmul(X, beta)
        return yhat
```

Introduction to TensorFlow



Source: “Machine Learning for Economics and Finance in TensorFlow 2”

Introduction to TensorFlow

Tensor Definitions

```
>>> scalar =
tf.constant(1., tf.float32)

>>> matrix = tf.Variable(
[[1., 2.], [3., 4.]], tf.float32)

>>> tensor =
tf.random.normal((2, 4,
6, 3))
```

Operation Definitions

Operation	Example
tf.add()	tf.add(scalar, tensor)
tf.multiply()	tf.multiply(scalar, matrix)
tf.matmul()	tf.matmul(matrix, matrix)

Introduction to Tensor Multiplication

Types of Tensor Multiplication

1. Elementwise multiplication.
2. Dot product.
3. Matrix multiplication.

Introduction to Tensor Multiplication

Elementwise Multiplication

- Only defined for tensors with identical dimensions.

$$C_{ijr} = A_{ijr} * B_{ijr} \tag{1}$$

$$\begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \bigodot \begin{bmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{bmatrix} = \begin{bmatrix} a_{00} \times b_{00} & a_{01} \times b_{01} \\ a_{10} \times b_{10} & a_{11} \times b_{11} \end{bmatrix} \tag{2}$$

Dot Product and its Implementation

Dot Product

- Can be performed between two vectors with the same number of elements.
- Transforms two vectors into a scalar.

$$c = \sum_{i=0}^{n} a_i \times b_i \tag{3}$$

Matrix Multiplication and its Implementation

Matrix Multiplication

- Tensors don't need the same shape but need to be matrices.
- Number of columns in A = Number of rows in B.
- Result: Number of rows in A x Number of columns in B.

$$C_{ij} = A_{i:} \cdot B_{:j} \tag{4}$$

$$C = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \cdot \begin{bmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{bmatrix} \tag{5}$$

Introduction to TensorFlow

Automatic Differentiation

- TensorFlow employs “automatic differentiation” Abadi and Plotkin (2019).
- Method is neither purely symbolic nor purely numerical.
- Especially efficient for deep learning model training.
- Many applications don’t require explicit computation of derivatives.

Symbolic Differentiation

- Directly yields an exact, algebraic expression.
- Manually: Time-consuming and error-prone.
- Computationally: Complex and may lack a closed-form expression.

Introduction to TensorFlow

Numerical Differentiation

$$f'(x) \approx \frac{(f(x+h) - f(x))}{h} \tag{6}$$

- Based on limit definition; uses a small step size.
- Doesn't aim for exact expression; evaluates function at different points.
- Quality depends on the size of step.

Automatic Differentiation

- Achieves increased accuracy over numerical differentiation.
- Stable for deep learning problems with many parameters.
- No need for a single derivative expression, unlike symbolic differentiation.
- Breaks down the symbolic computation into elementary parts.
- Evaluates the derivative at a single point, using a forward or backward chain of partial derivatives.

Introduction to TensorFlow

Automatic Differentiation Example

- Compute $\partial g(f(x))/\partial x$.
  1. $g(y) = 3y$
  2. $f(x) = x^2$
  3. $x = 2$

Introduction to TensorFlow

Symbolic

1. $g(f(x)) = 3x^2$
2. $\frac{\partial g(f(x))}{\partial x} = 6x$
3. $\frac{\partial g(f(x))}{\partial x}_{|x=2} = 12$

Numerical

1. $g(f(x)) = 3x^2$
2. $\frac{\partial g(f(x))}{\partial x} \approx \frac{3(x+h)^2 - 3x^2}{h}$
3. $\frac{\partial g(f(x))}{\partial x} \approx 6x + h$
4. $\frac{\partial g(f(x))}{\partial x}_{|x=2} \approx 12 + h$

Auto

1. $\frac{\partial g(f(x))}{\partial x} = \frac{\partial g(y)}{\partial y} \frac{\partial f(x)}{\partial x}$
2. $\frac{\partial g(y)}{\partial y} = 3$
3. $\frac{\partial f(x)}{\partial x} = 2x$
4. $\frac{\partial f(x)}{\partial x}_{|x=2} = 4$
5. $\frac{\partial g(y)}{\partial y}_{|x=2} = 3$
6. $\frac{\partial g(f(x))}{\partial x}_{|x=2} = 12$

Introduction to TensorFlow

Colab Tutorial

- Introduction to TensorFlow - Part 1
- Introduction to TensorFlow - Part 2

References I

Abadi, Martín and Gordon D. Plotkin (2019) "A Simple Differentiable Programming Language," *Proc. ACM Program. Lang.*, 4 (POPL), 10.1145/3371106.

Athey, Susan (2018) *The Impact of Machine Learning on Economics*, 507–547: University of Chicago Press, 10.7208/chicago/9780226613475.001.0001.

Athey, Susan and Guido W. Imbens (2019) "Machine Learning Methods That Economists Should Know About," *Annual Review of Economics*, 11 (1), 685–725, 10.1146/annurev-economics-080217-053433.

Bertsch, Christoph, Isaiah Hull, Robin L. Lumsdaine, and Xin Zhang (2022) "Central Bank Mandates and Monetary Policy Stances: through the Lens of Federal Reserve Speeches," *SSRN Electronic Journal*, 10.2139/ssrn.4255978.

References II

Bommasani, Rishi, Drew A. Hudson, Ehsan Adeli et al. (2022) "On the Opportunities and Risks of Foundation Models."

Brynjolfsson, Erik and Tom Mitchell (2017) "What can machine learning do? Workforce implications," *Science*, 358 (6370), 1530–1534, 10.1126/science.aap8062.

Chen, Yanbei, Massimiliano Mancini, Xiatian Zhu, and Zeynep Akata (2022) "Semi-Supervised and Unsupervised Deep Visual Learning: A Survey."

Colliot, Olivier (2023) *A Non-technical Introduction to Machine Learning*, 3–23, New York, NY: Springer US, 10.1007/978-1-0716-3195-9_1.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy (2019) "Text as Data," *Journal of Economic Literature*, 57 (3), 535–74, 10.1257/jel.20181020.

References III

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016) *Deep Learning*: MIT Press, http://www.deeplearningbook.org.

Hinton, Geoffrey, Nitish Srivastava, and Kevin Swersky (2012) "Neural Networks for Machine Learning."

Hull, Isaiah (2021) *Machine Learning for Economics and Finance in TensorFlow 2*: Apress, 10.1007/978-1-4842-6373-0.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer (2015) "Prediction Policy Problems," *American Economic Review*, 105 (5), 491–95, 10.1257/aer.p20151023.

Mitchell, Tom M (1997) *Machine learning*, 1: McGraw-hill New York.

References IV

Mullainathan, Sendhil and Jann Spiess (2017) "Machine Learning: An Applied Econometric Approach," *The Journal of Economic Perspectives*, 31 (2), 87–106, http://www.jstor.org/stable/44235000.

Plested, Jo and Tom Gedeon (2022) "Deep transfer learning for image classification: a survey."

Simeone, Osvaldo (2018) "A Brief Introduction to Machine Learning for Engineers."

Varian, Hal R. (2014) "Big Data: New Tricks for Econometrics," *Journal of Economic Perspectives*, 28 (2), 3–28, 10.1257/jep.28.2.3.

Zhao, Wayne Xin, Kun Zhou, Junyi Li et al. (2023) "A Survey of Large Language Models."