AI and Big Data Masterclass for International Management

AOM 2025 PDW 15383

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

universität innsbruck

Introduction to the PDW

Our Panelists

- Jakob Müllner
  - Department of Global Trade, WU Vienna
- Sheen Levine
  - Naveen Jindal School of Management, UT Dallas
  - Management, Technical University of Denmark
- Sima Yue Ling
  - Global Competitiveness Institute, University of Cork
  - Naveen Jindal School of Management, UT Dallas
- Harald Puhr
  - Department of Management and Marketing, University of Innsbruck
  - Department of Global Trade, WU Vienna
- Laurenz Tinhof
  - Department of Global Trade, WU Vienna

Agenda

- Part 1: What do Big Data, Machine Learning, and AI Mean for International Management?
    - 09:30 – 11:10
- Break
    - 11:10 – 11:20
- Part 2: How Can I Use Big Data, Machine Learning, and AI in My Research?
    - 11:20 – 13:30

Agenda – Part 1 (09:30-11:20)

- Big Data and Artificial Intelligence
  - Harald Puhr
- Big Data in Management Research
  - Jakob Müllner
- Strategic Decoupling in International Disputes: Sentiment and Topic Identification with LLMs
  - Sima Yue Ling
- Strategic Decoupling in International Disputes: Text-Embeddings to Discover Cognitive Patterns
  - Sheen Levine
- Valuing Public Goods in a Populist World: Identifying Network Ties
  - Harald Puhr

Agenda – Part 2 (11:30 – 13:30)

- Conceptual Foundations of AI and Machine Learning
  - Harald Puhr
- A Basic Machine Learning Workflow in R
  - Laurenz Tinhof
- Case Study: Predicting Foreign Subsidiary Profits
  - Laurenz Tinhof
- Coffee House Style Discussion

Big Data and Artificial Intelligence

Introduction to Big Data

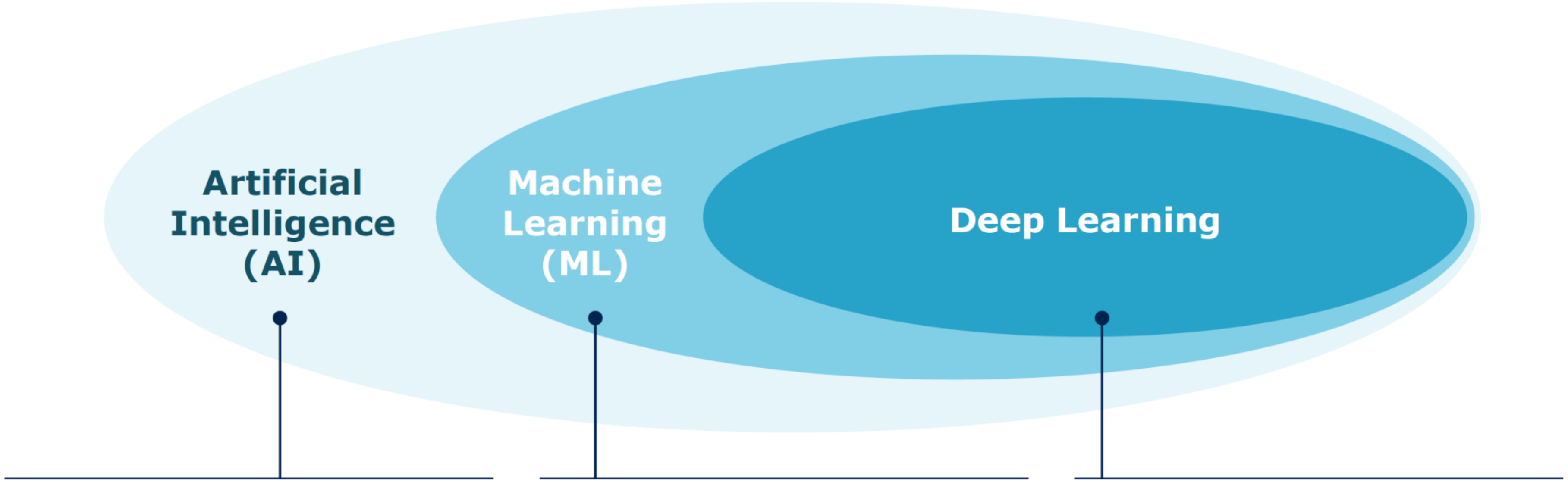
Big Data refers to extremely large, fast-growing, and diverse datasets, structured, semi-structured, and unstructured, that are too complex for traditional data management tools

- This creates challenges for researchers:
  - Data engineering: Big Data requires complex collection and feature engineering
  - Data storage: Data sets with multiple GB of data that slow down processing
  - Data analytics: Too many variables or observations for traditional methods

Examples in social sciences:

- Geo-spatial data
- Network data
- Text mining
- Video and image data
- Game data
- …

Introduction Artificial Intelligence

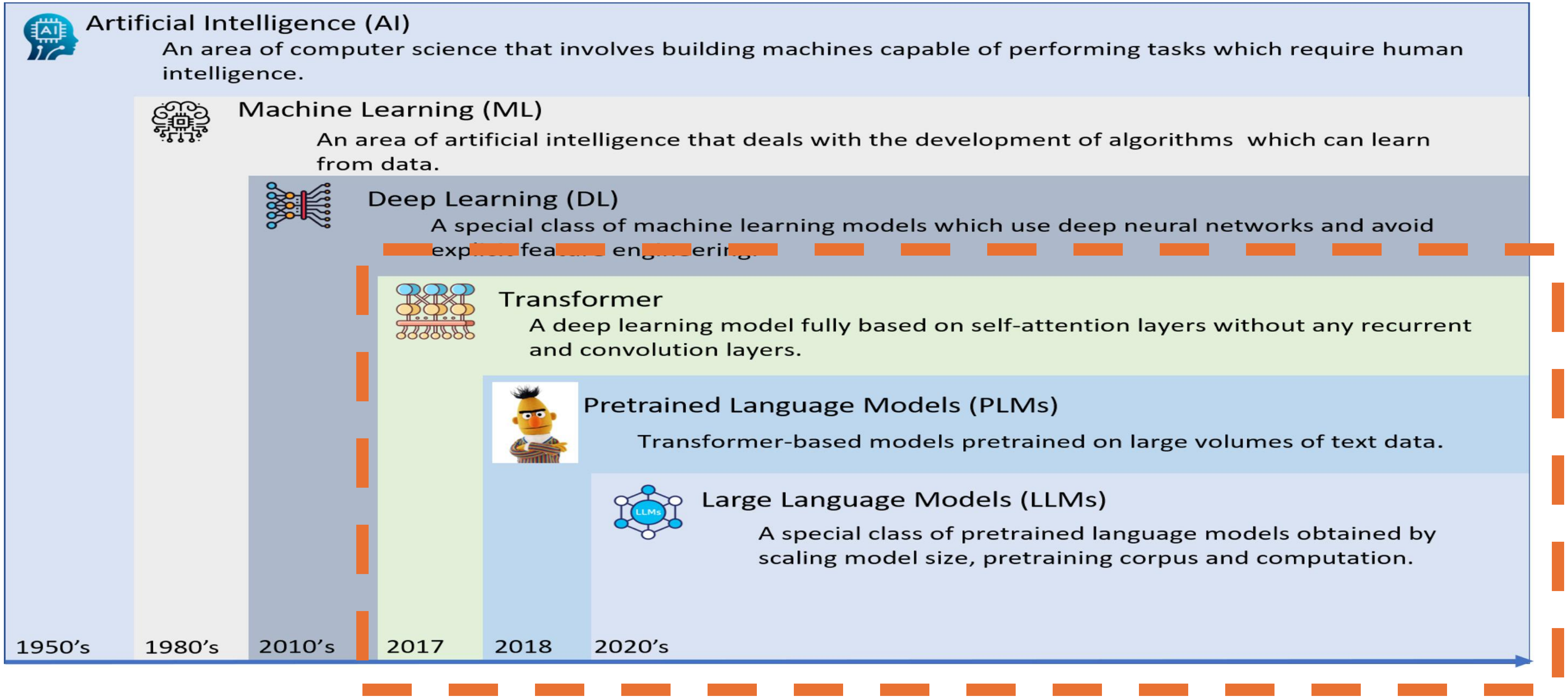


More generally describes **decision-support** that **automates certain thought processes** based on learning and intelligence *(McCarthy, 2007)*

Relates to a **specific set of methods** that **improve automatically through experience** *(Jordan & Mitchell, 2015)*

Refers to **specific methods** that compose of **multiple processing layers** to generate levels of abstraction of data *(LeCun, Bengio & Hinton, 2015)*

Introduction Artificial Intelligence



Artificial Intelligence (AI)
An area of computer science that involves building machines capable of performing tasks which require human intelligence.

Machine Learning (ML)
An area of artificial intelligence that deals with the development of algorithms which can learn from data.

Deep Learning (DL)
A special class of machine learning models which use deep neural networks and avoid explicit feature engineering.

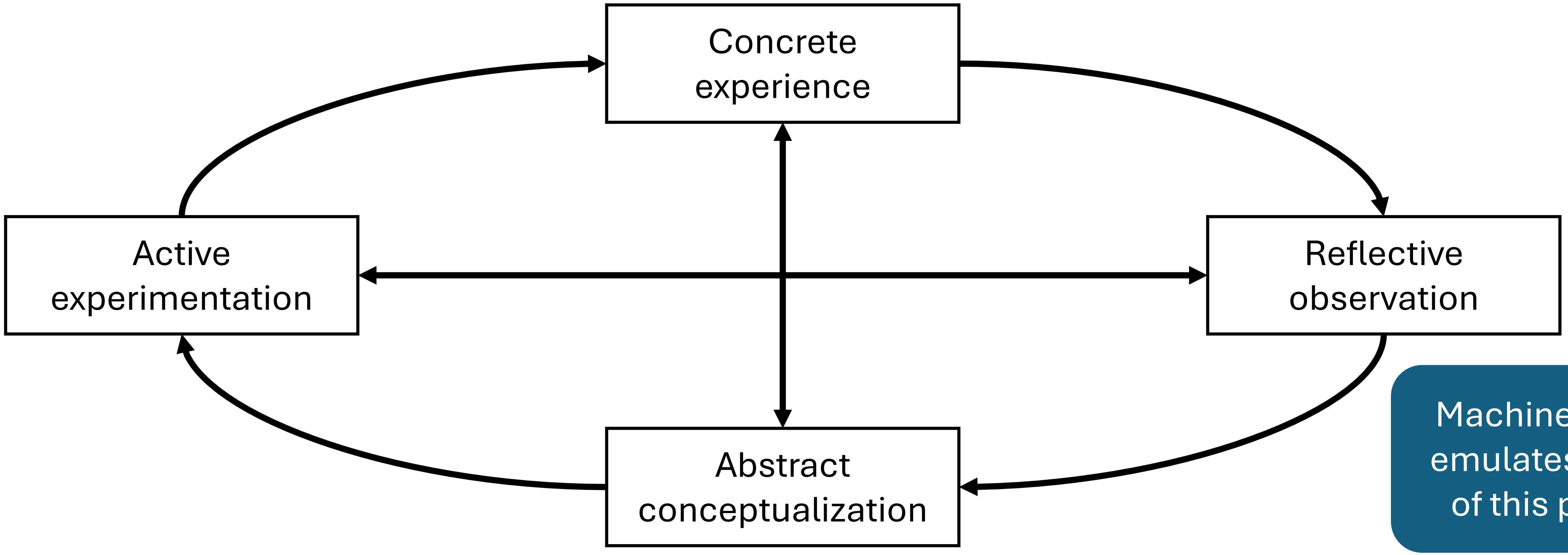
Transformer
A deep learning model fully based on self-attention layers without any recurrent and convolution layers.

Pretrained Language Models (PLMs)
Transformer-based models pretrained on large volumes of text data.

Large Language Models (LLMs)
A special class of pretrained language models obtained by scaling model size, pretraining corpus and computation.

https://doi.org/10.1016/j.nlp.2023.100048

Experiential Learning by Humans



Concrete experience

Reflective observation

Abstract conceptualization

Active experimentation

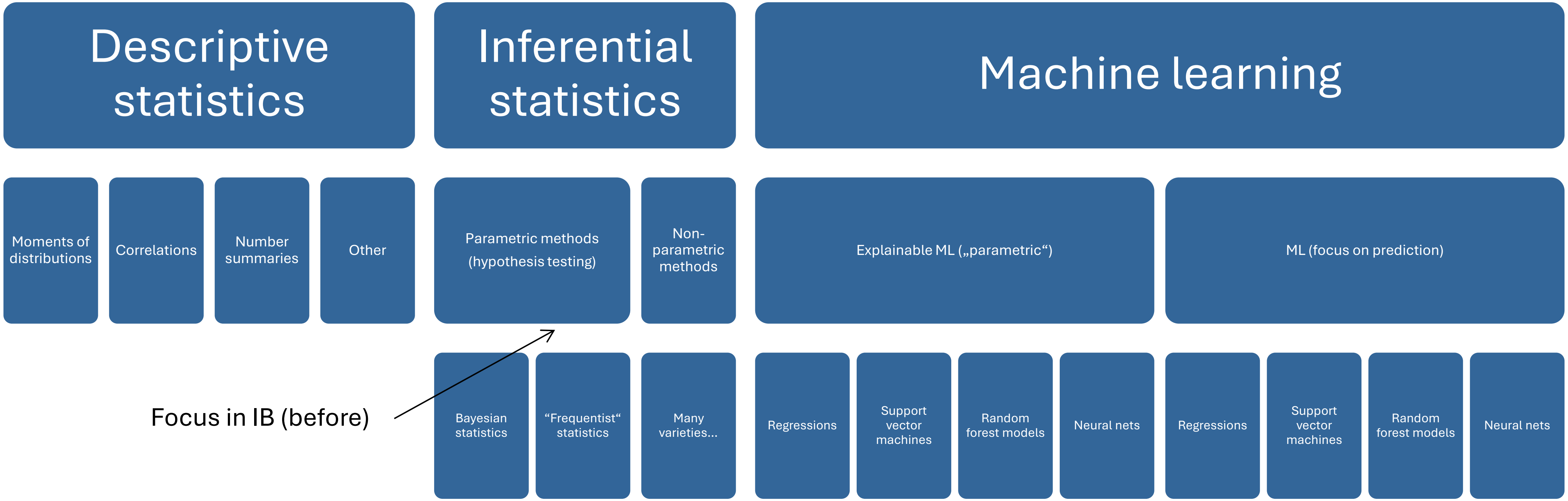
Machine learning emulates aspects of this process.

https://doi.org/10.1177/1350507605058130

Human Learning vs. Machine Learning

- Concrete experience
  - Human: Real-life experiences and encounters made by humans.
  - Machine: Collect the raw material (observations or measurements) from which the model learns.
- Reflective observation
  - Human: Learners reflect on their experience, observing and noting outcomes, patterns, or anomalies.
  - Machine: Analyze the data and understand patterns from the data (model fitting).
- Abstract conceptualization
  - Human: Learners form theories or conceptual frameworks to explain what was observed.
  - Machine: Generalize by defining a functional form that describes the identified relations.
- Active experimentation
  - Human: Learner test their new knowledge by applying it in different situations
  - Machine: Model testing and updating. In machine learning, this is part of the algorithm.

ML as a Method for Quantitative Data Analysis (before)



- Descriptive statistics
  - Moments of distributions
  - Correlations
  - Number summaries
  - Other
- Inferential statistics
  - Parametric methods (hypothesis testing)
    - Bayesian statistics
    - “Frequentist“ statistics
  - Non-parametric methods
    - Many varieties...
- Machine learning
  - Explainable ML („parametric“)
    - Regressions
    - Support vector machines
    - Random forest models
    - Neural nets
  - ML (focus on prediction)
    - Regressions
    - Support vector machines
    - Random forest models
    - Neural nets

Focus in IB (before)

This Perspective Is Changing Rapidly!

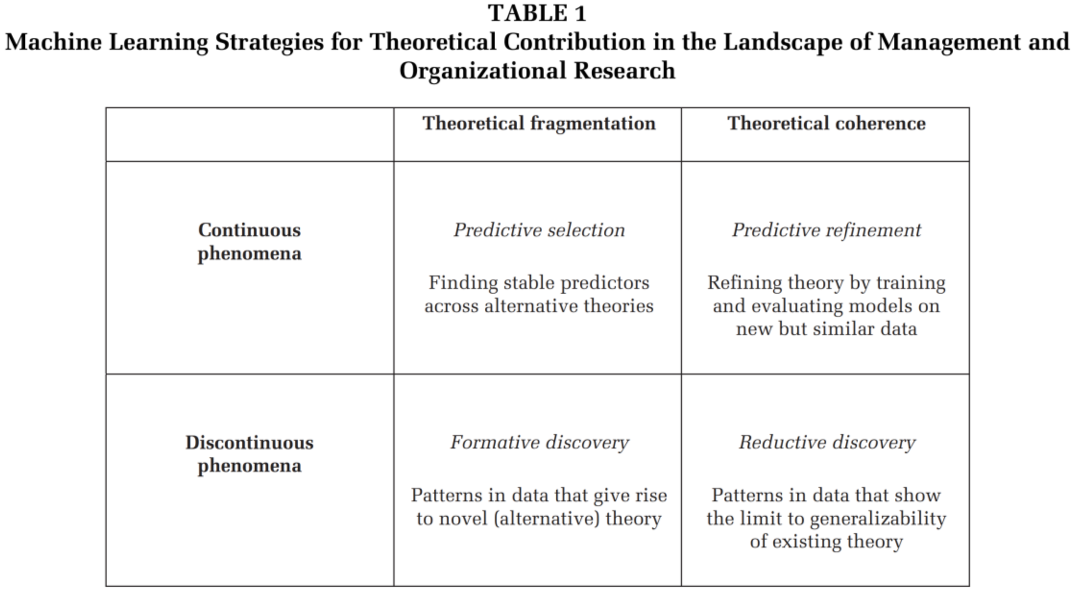
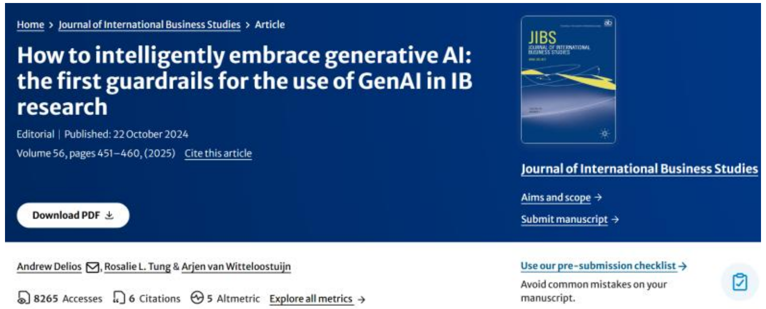


TABLE 1
Machine Learning Strategies for Theoretical Contribution in the Landscape of Management and Organizational Research

	Theoretical fragmentation	Theoretical coherence
Continuous phenomena	*Predictive selection* Finding stable predictors across alternative theories	*Predictive refinement* Refining theory by training and evaluating models on new but similar data
Discontinuous phenomena	*Formative discovery* Patterns in data that give rise to novel (alternative) theory	*Reductive discovery* Patterns in data that show the limit to generalizability of existing theory

Von Krogh, G., Roberson, Q., & Gruber, M. (2023). Recognizing and Utilizing Novel Research Opportunities with Artificial Intelligence. Academy of Management Journal, 66(2), 367-373.



Home > Journal of International Business Studies > Article

How to intelligently embrace generative AI: the first guardrails for the use of GenAI in IB research

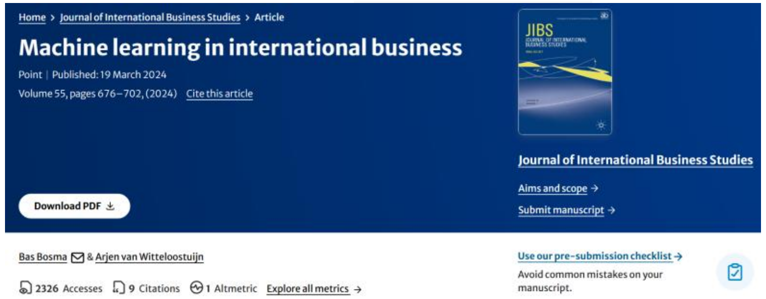
Editorial | Published: 22 October 2024

Volume 56, pages 451–460, (2025) Cite this article

Andrew Delios, Rosalie L. Tung & Arjen van Witteloostuijn

8265 Accesses 6 Citations 5 Altmetric Explore all metrics →

Download PDF
Journal of International Business Studies
Aims and scope →
Submit manuscript →
Use our pre-submission checklist →
Avoid common mistakes on your manuscript.



Home > Journal of International Business Studies > Article

Machine learning in international business

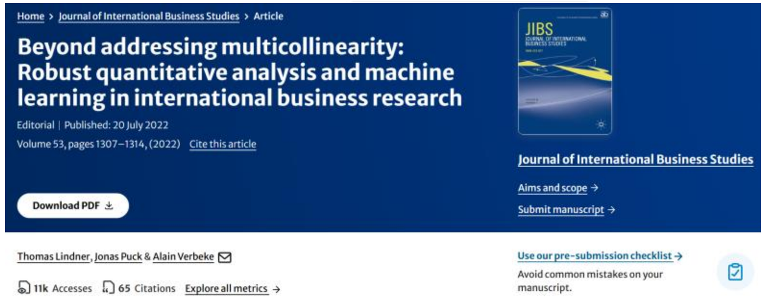
Point | Published: 19 March 2024

Volume 55, pages 676–702, (2024) Cite this article

Bas Bosma & Arjen van Witteloostuijn

2326 Accesses 9 Citations 1 Altmetric Explore all metrics →

Download PDF
Journal of International Business Studies
Aims and scope →
Submit manuscript →
Use our pre-submission checklist →
Avoid common mistakes on your manuscript.



Home > Journal of International Business Studies > Article

Beyond addressing multicollinearity: Robust quantitative analysis and machine learning in international business research

Editorial | Published: 20 July 2022

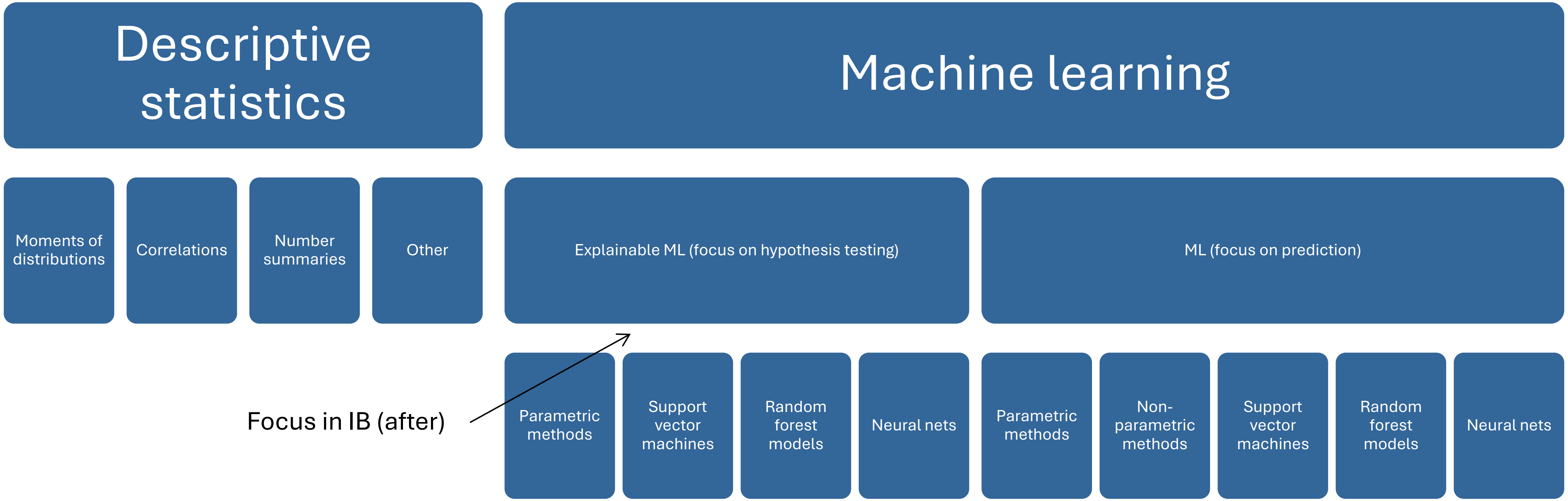
Volume 53, pages 1307–1314, (2022) Cite this article

Thomas Lindner, Jonas Puck & Alain Verbeke

11k Accesses 65 Citations Explore all metrics →

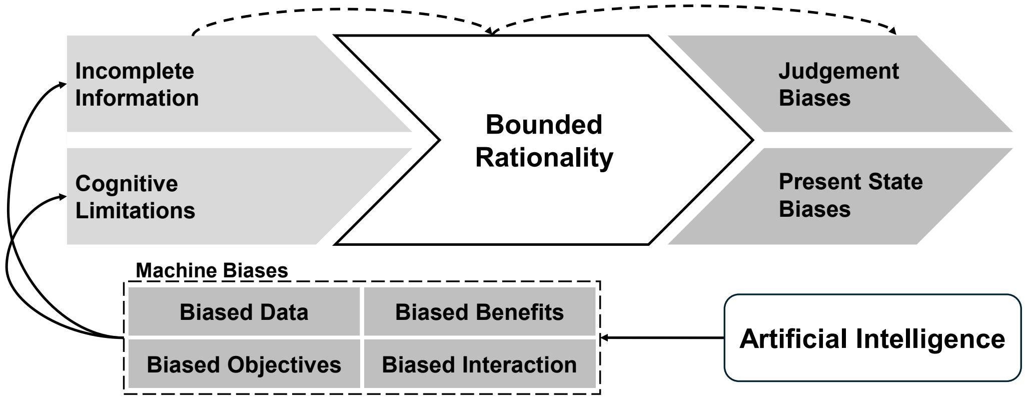
Download PDF
Journal of International Business Studies
Aims and scope →
Submit manuscript →
Use our pre-submission checklist →
Avoid common mistakes on your manuscript.

ML as a Method for Quantitative Data Analysis (after)



AI/ML also Affect International Business Theory

- We can make predictions about firm decisions, frequently because people have biases
- Large-scale AI applications may change these biases
- We can probably change our theoretical assumptions to incorporate these new biases, but we need to know which biases matter for which decision



Lindner, T., Puck, J., & Puhr, H. (2025). Artificial Intelligence in International Business: IB Theory under Augmented Decision-Making. Working Paper.

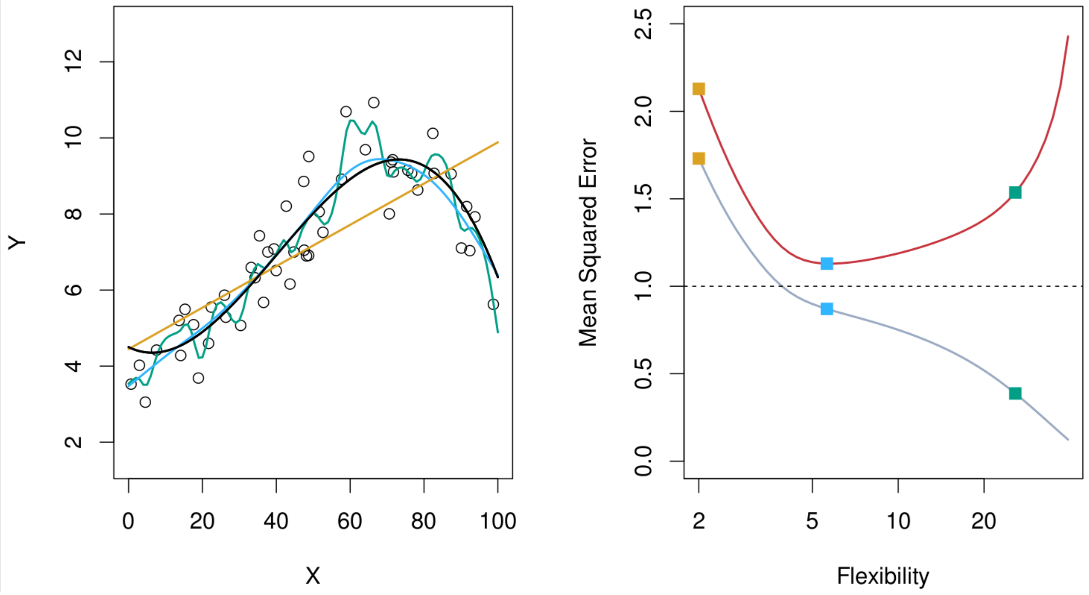
Research Applications of Big Data and AI in IB

Break

Conceptual Foundations of AI and Machine Learning

Train/Test Im Machine Learning models

- In ML, datasets are split into training and testing data (something like 60/40 or 70/30)
  - Models are built on the training data, then fit is assessed on the testing data (see figure below)
- In cross-validation, we do this repeatedly
  - K-fold cross-validation takes k splits of the data into train-test.



https://doi.org/10.1007/978-3-031-38747-0

Cross-Validation

- k-fold cross-validation
  - Split the training data into k groups („folds“)
  - Train the model on k-1 folds and test on the remaining fold
  - Repeat the process for each of the k folds
  - Average the model performance (e.g., MSE) over the k folds

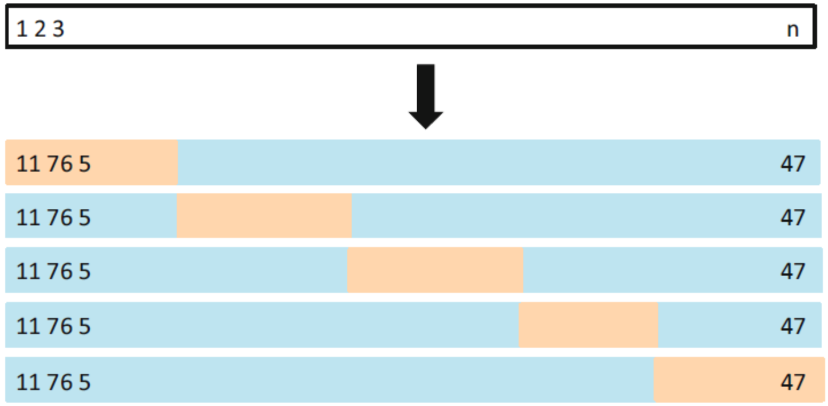
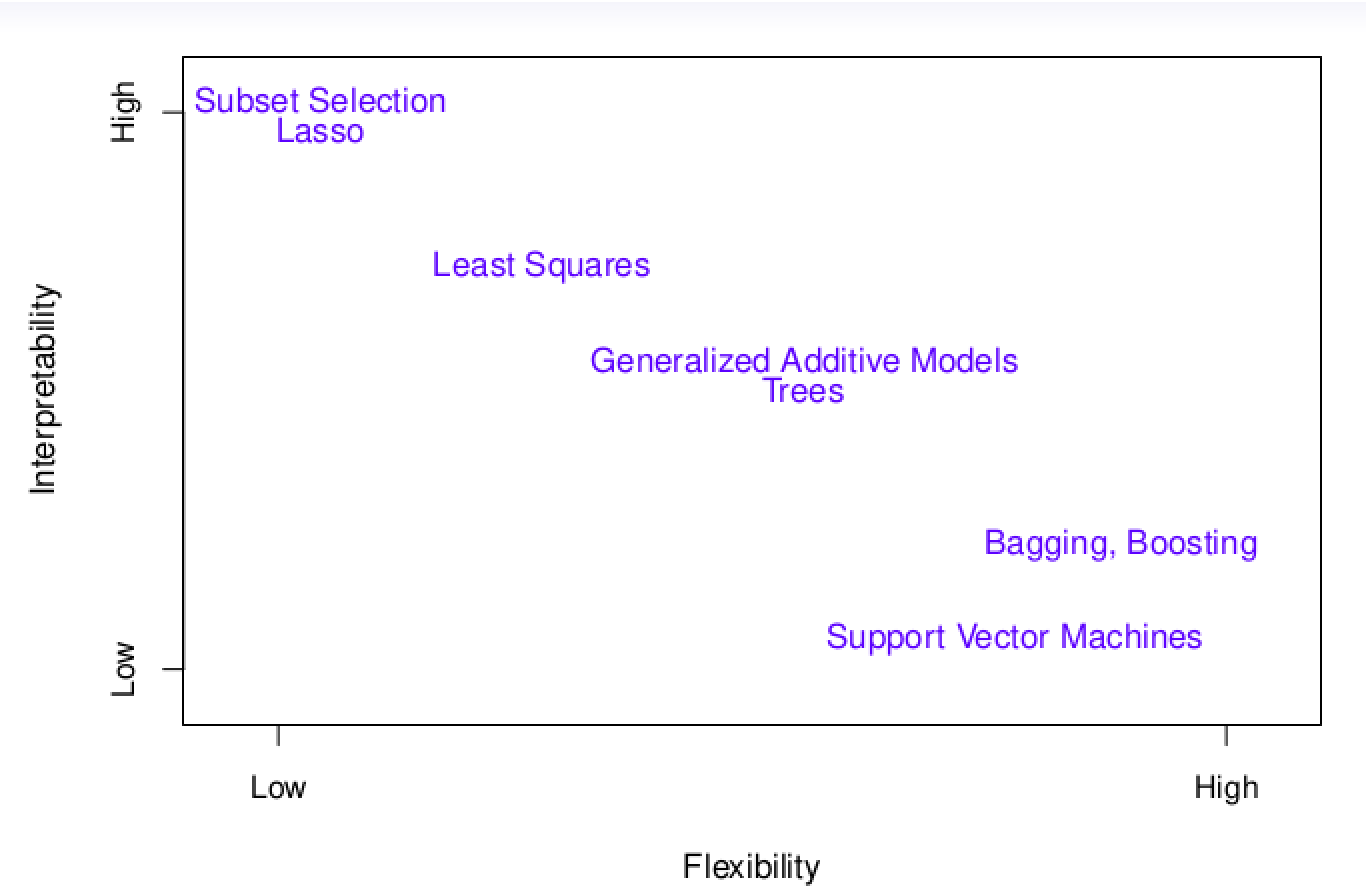


FIGURE 5.5. *A schematic display of 5-fold CV. A set of n observations is randomly split into five non-overlapping groups. Each of these fifths acts as a validation set (shown in beige), and the remainder as a training set (shown in blue). The test error is estimated by averaging the five resulting MSE estimates.*

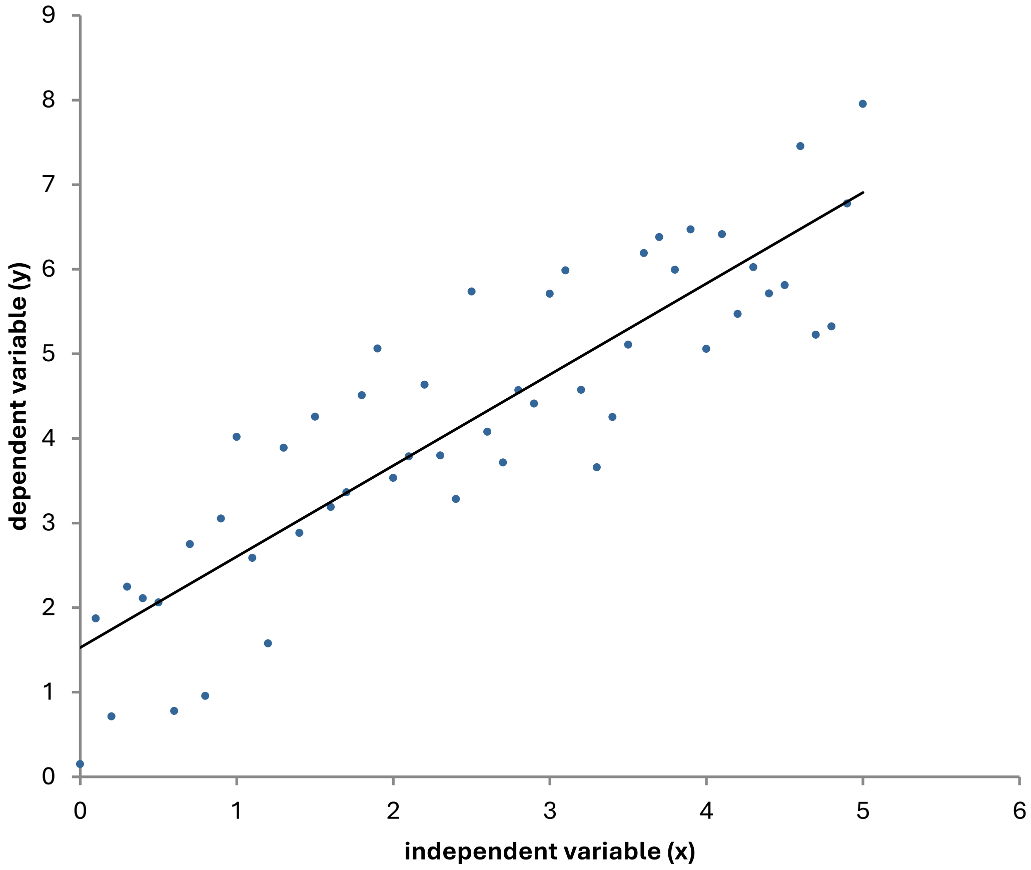
https://doi.org/10.1007/978-3-031-38747-0

A variety of models with different flexibility and interpretability



https://doi.org/10.1007/978-3-031-38747-0

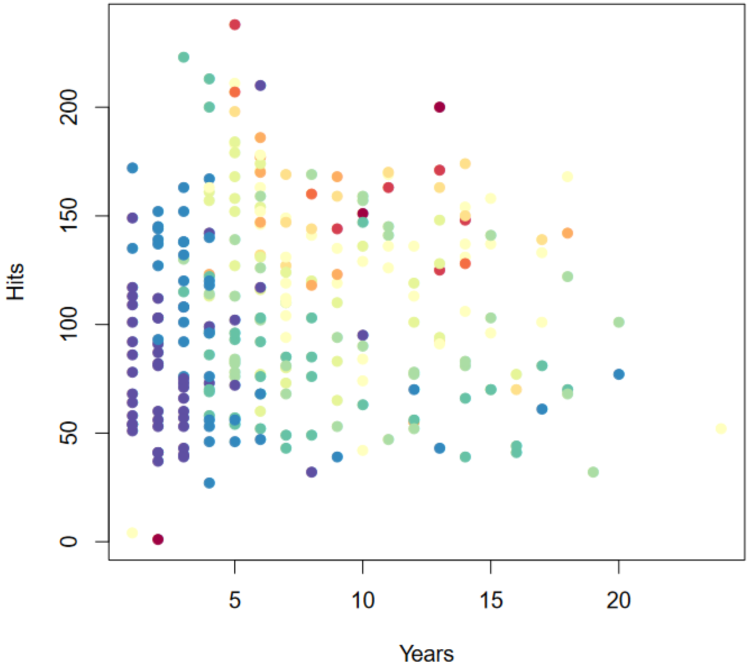
Least Squares: The Least Intelligent AI

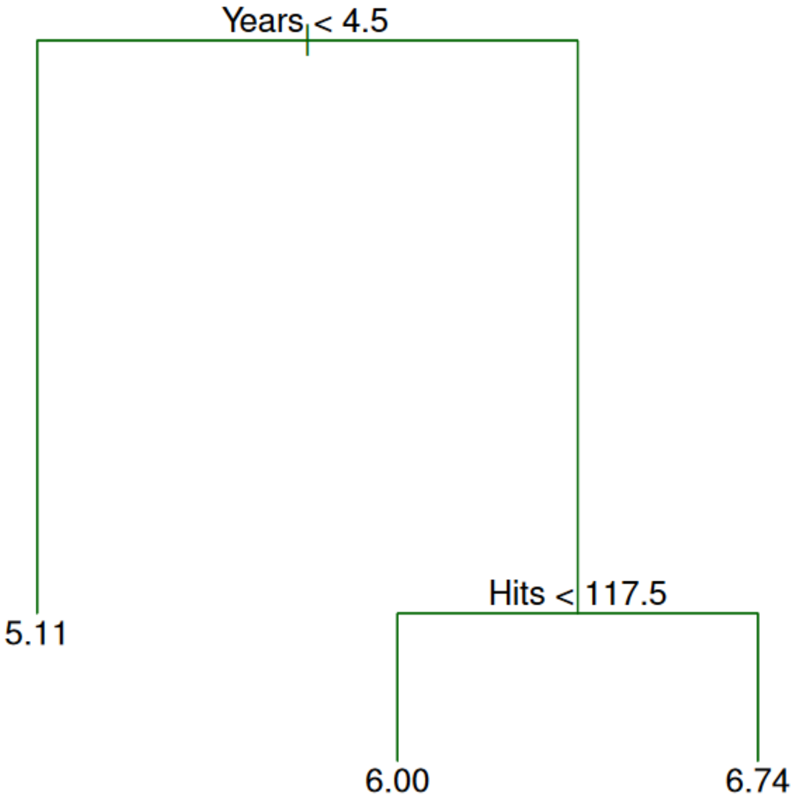


- In the simplest form, we want to understand how a variable y changes with a variable x
- Assumed functional form:
  - $y_i = b_0 + b_1 \cdot x_i + e_i$
- Learning problem:
  - Find b_0 and b_1 such that the squared sum of e_i is minimal.

Decision trees

- As the name suggests, decision trees are related to sequential decision-making problems





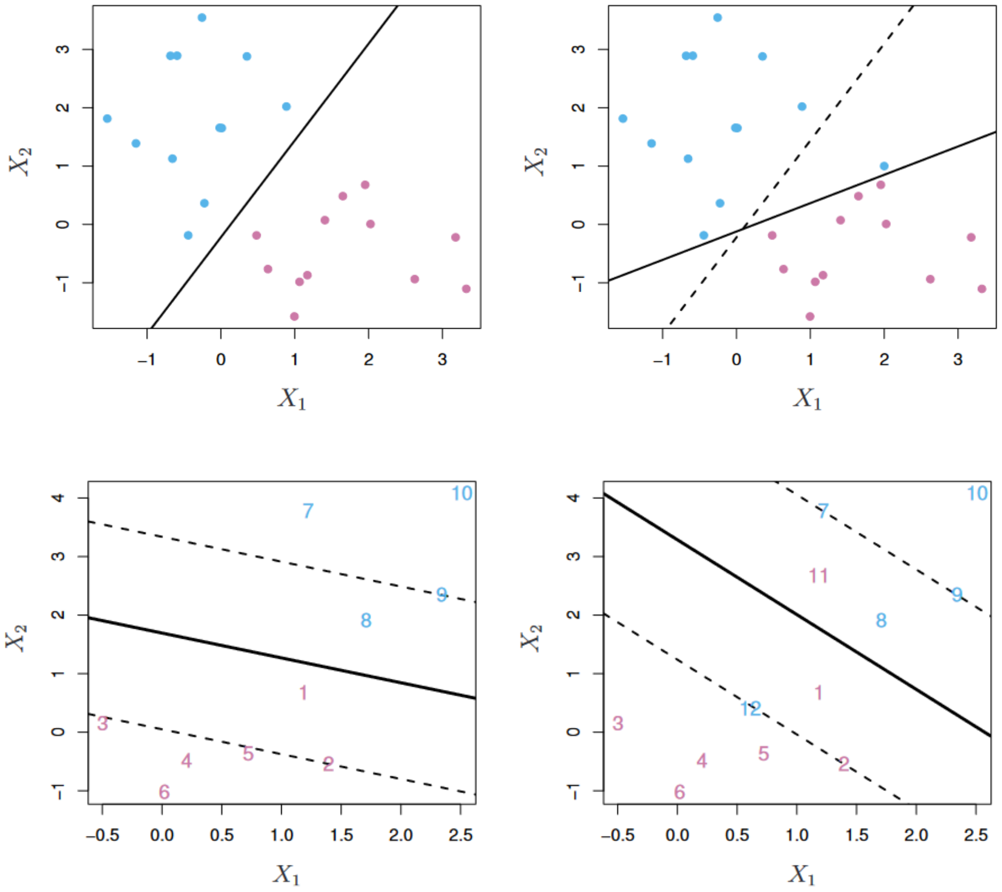
https://doi.org/10.1007/978-3-031-38747-0

Random forests aggregate repeated runs of decision trees

- For aggregation, we draw random samples from the training data for repeated estimation of the decision tree
- In this approach (bagging), we reduce the variance in our prediction by re-running the training model on different subset
- To maximize predictive power, we de-correlate predictions, using a method analogous to the Mahalanobis correction

Support Vector Machines

- SVMs let us relax some rigid assumptions
- Support vectors let us introduce a soft boundary in classification problems
  - Noise in the data can give very different results
- If we allow for some error, we can get more consistent results
- In extensions, we can also use non-linear (and higher dimensional) separators

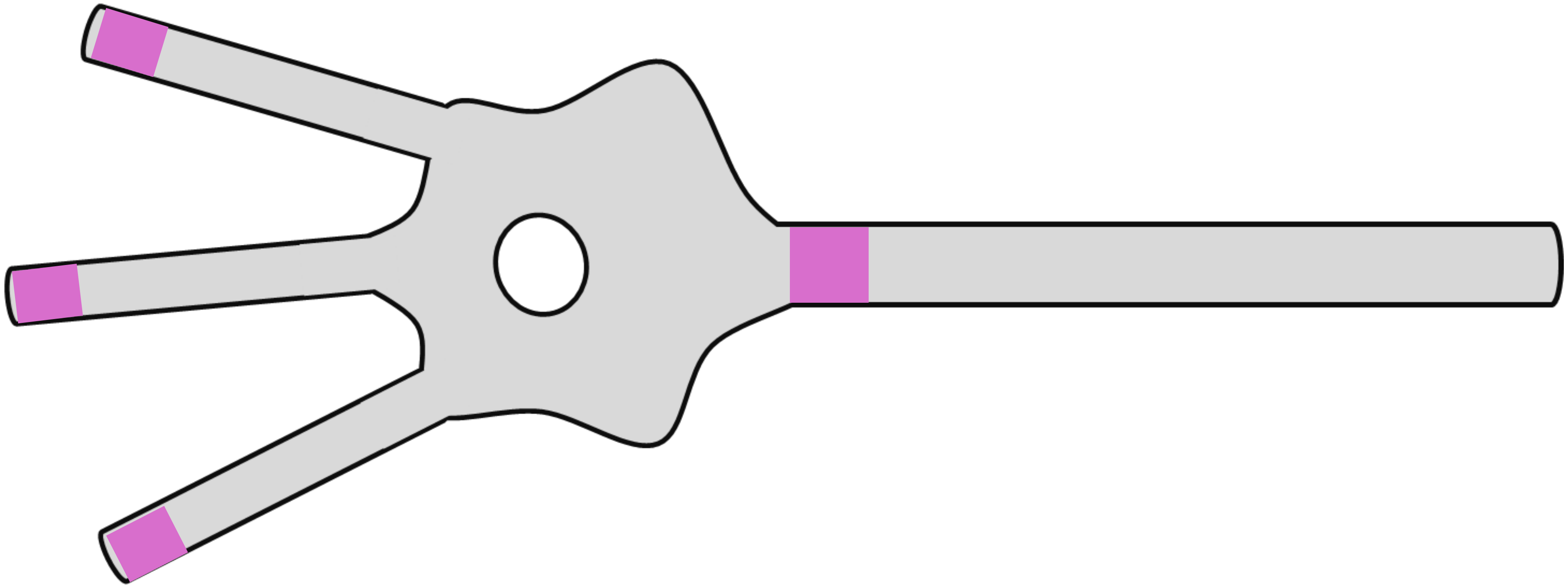


https://doi.org/10.1007/978-3-031-38747-0

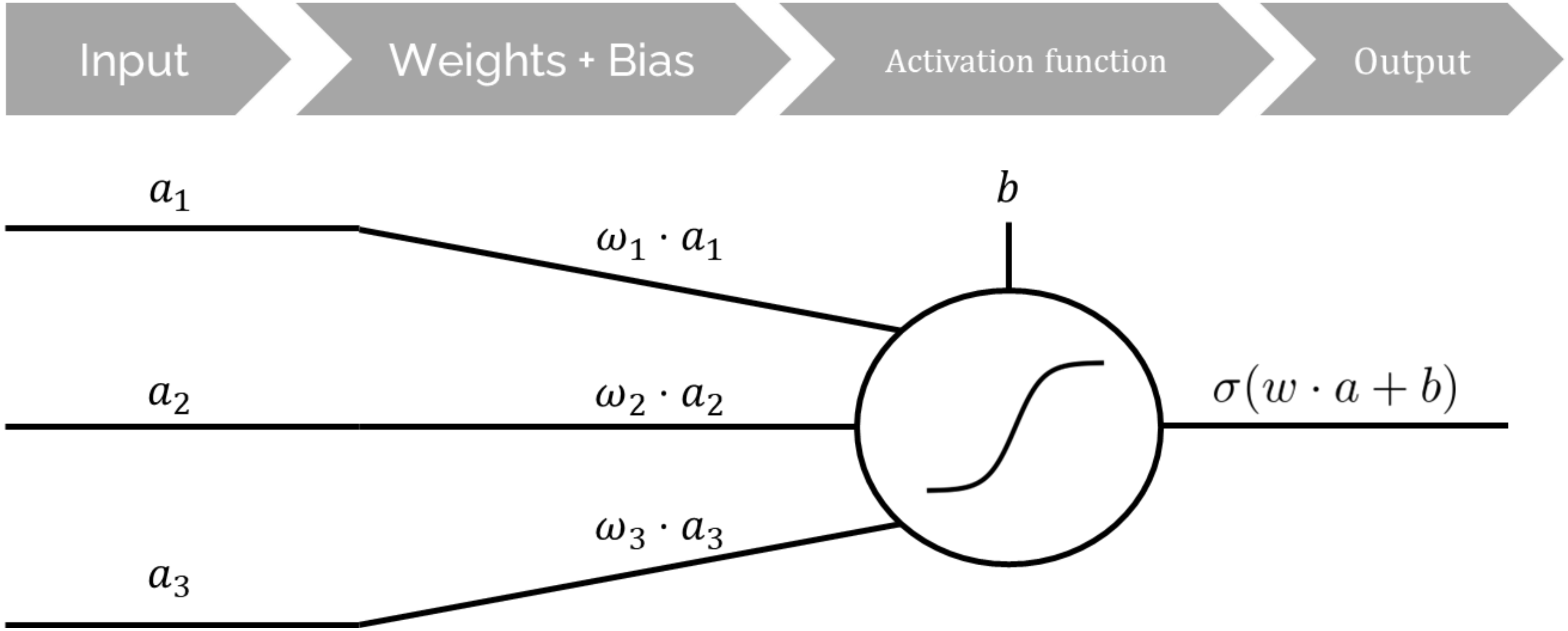
Deep Learning and Neural Nets

1. **Neurons and neural nets**
2. Learning
3. Progress

Biological Neuron



Artificial Neuron



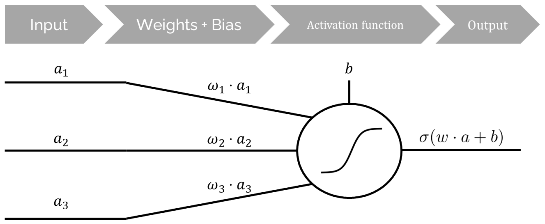
Input

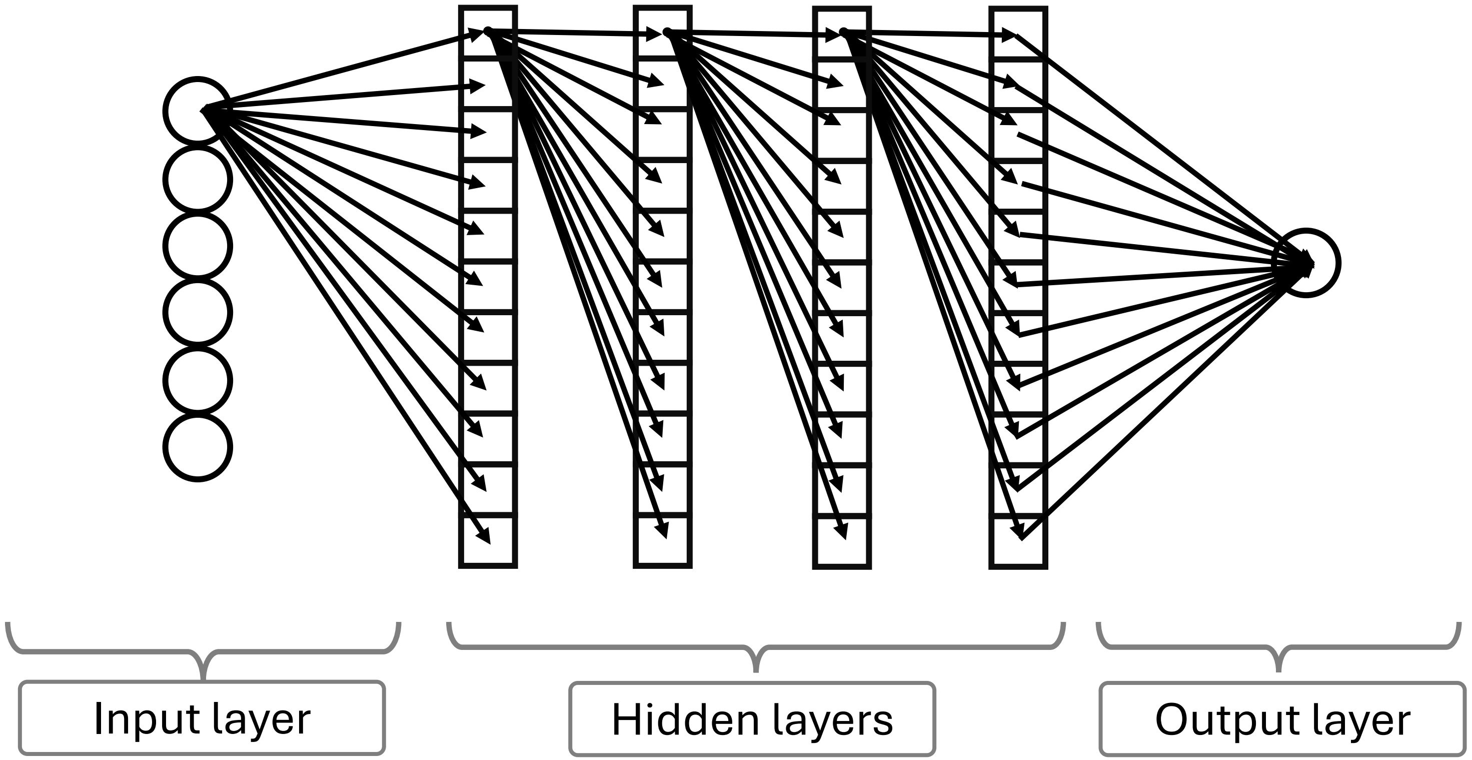
Weights + Bias

Activation function

Output

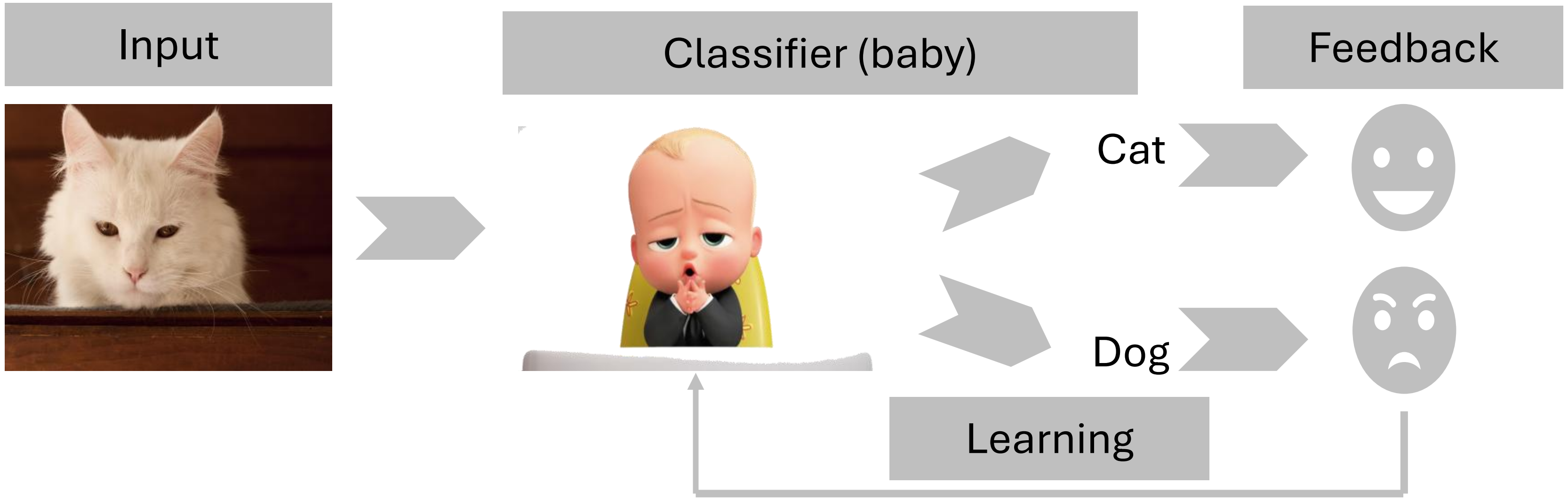
Artificial Neural Net





Deep Learning and Neural Nets

1. Neurons and neural nets
2. **Learning**
3. Progress



Input

Classifier (baby)

Feedback

Cat

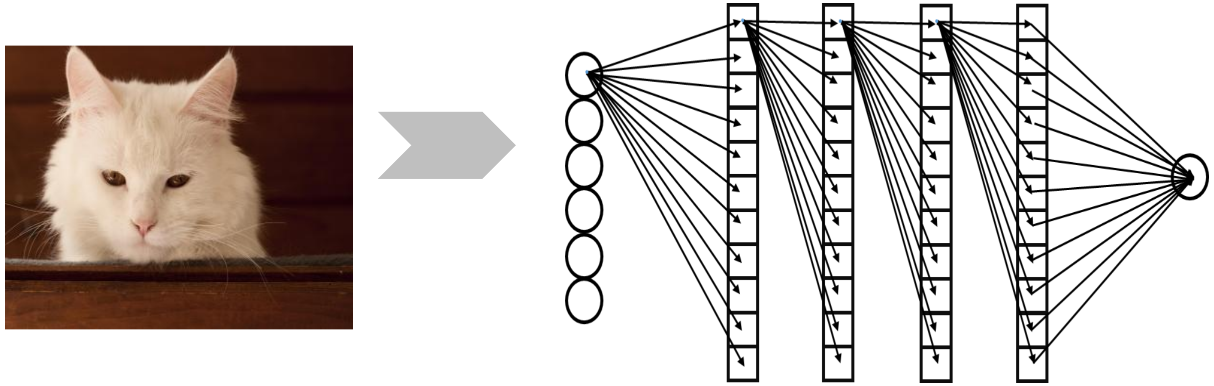
Dog

Learning

Input

Classifier (Neural Network)

Loss function



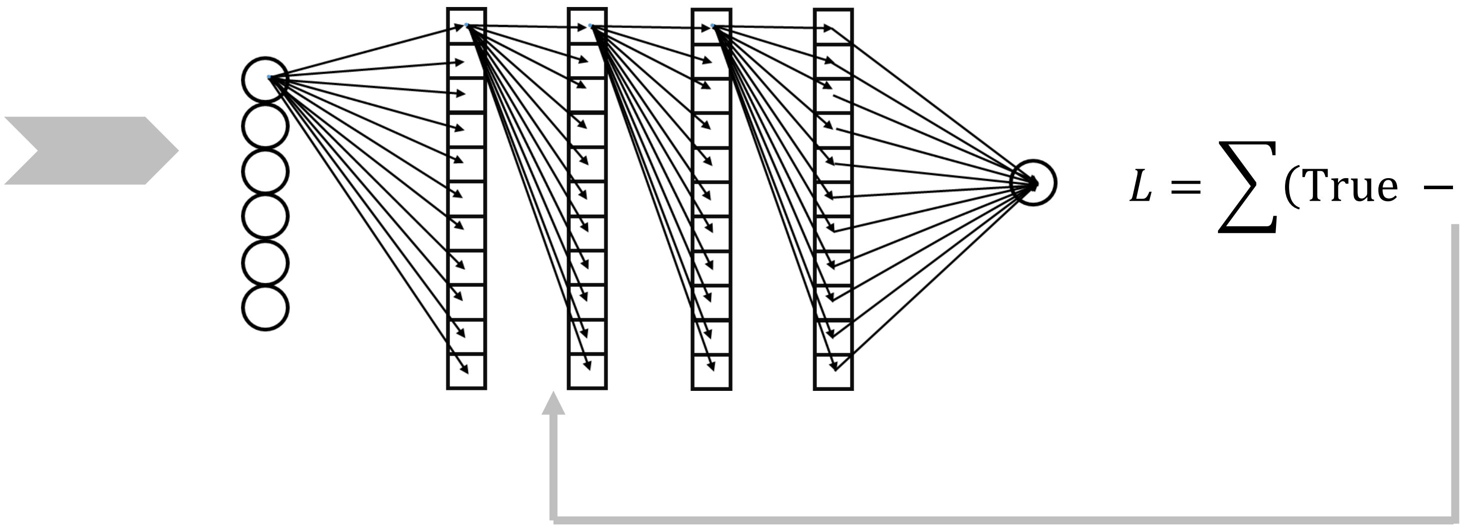
$$L = \sum (\text{True} - \text{Predicted})^2$$

Training a neural network

Input

Classifier (Neural Network)

Loss function



$$L = \sum (\text{True} - \text{Predicted})^2$$

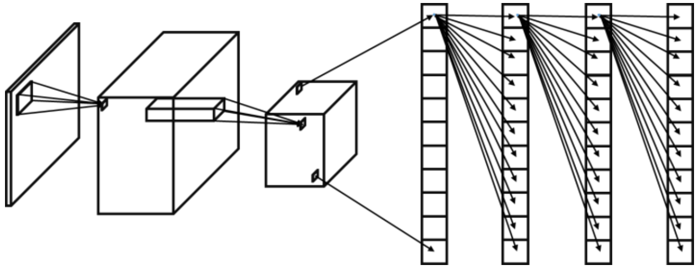
Adjust weights and biases

Deep Learning and Neural Nets

1. Neurons and neural nets
2. Learning
3. **Progress**

1997: Deep blue vs. Kasparov

2012: AlexNet (A. Krizhevsky et al.)



2016: AlphaGO vs. Lee Sedol (Google)

2022: Dall E 2 (OpenAI)

2022: Copilot (Github)

2022: ChatGPT (OpenAI)

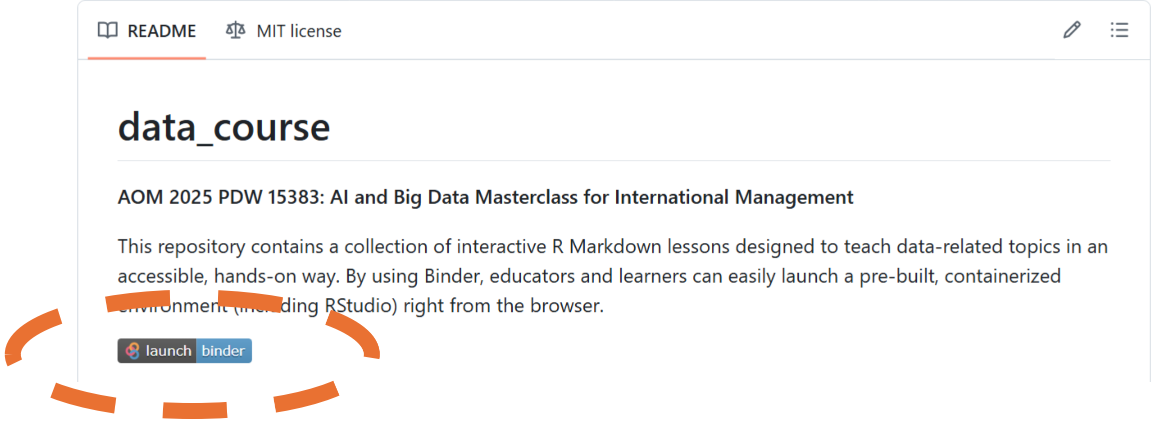
'Welcome to a world, where data is the new oil, and neural networks are the refineries that turn it into insights and predictions.'

by ChatGPT

A Basic Machine Learning Workflow in R

Hands-On Exercise

- Go to: https://tinyurl.com/3ub5famm
- Click on the „launch binder“ button
- This starts an RStudio session in your browser (takes 1-2 minutes)
- Let us know if there are issues



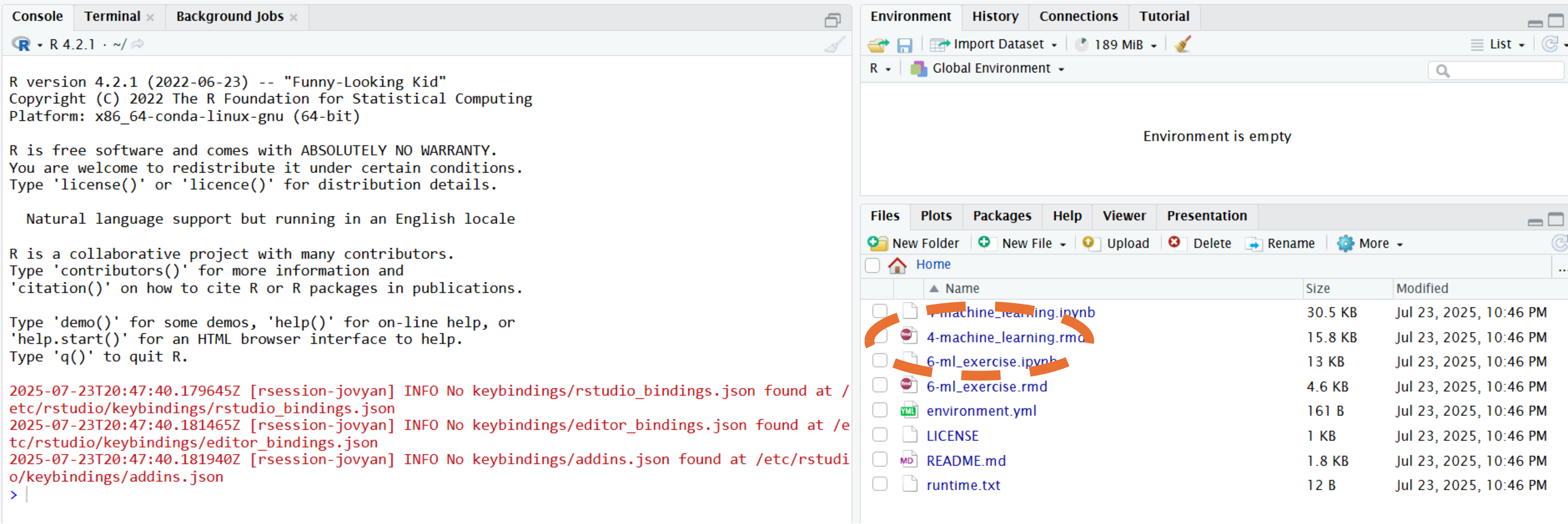
data_course

AOM 2025 PDW 15383: AI and Big Data Masterclass for International Management

This repository contains a collection of interactive R Markdown lessons designed to teach data-related topics in an accessible, hands-on way. By using Binder, educators and learners can easily launch a pre-built, containerized environment (including RStudio) right from the browser.

Hands-On Exercise

- Once the RStudio has opened, click on „4-machine_learning.rmd“ on the bottom right

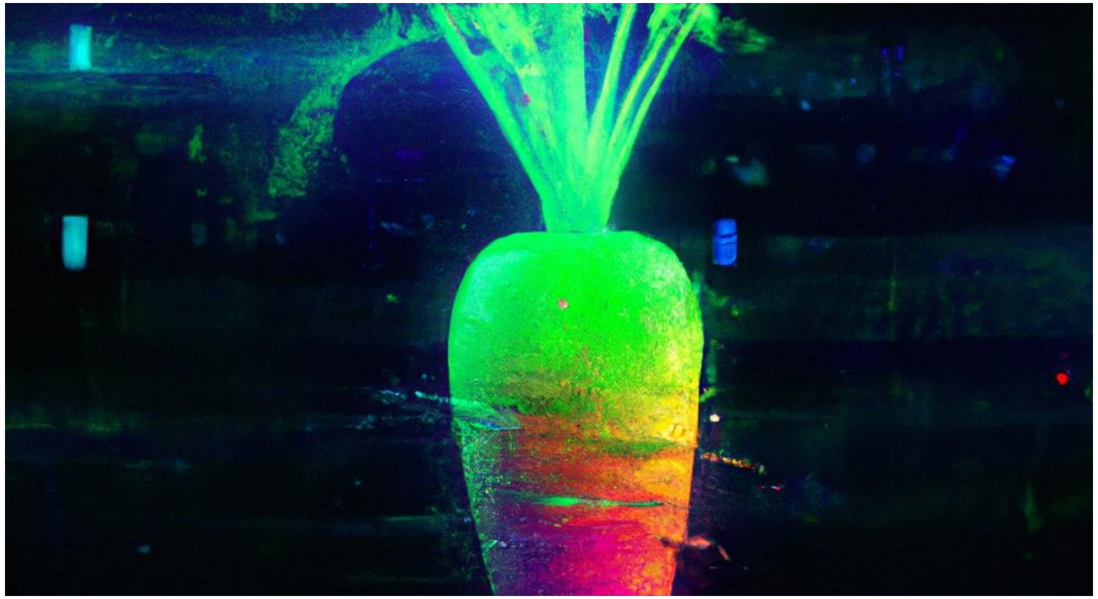


```
R version 4.2.1 (2022-06-23) -- "Funny-Looking Kid"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-conda-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

2025-07-23T20:47:40.179645Z [rsession-jovyan] INFO No keybindings/rstudio_bindings.json found at /
etc/rstudio/keybindings/rstudio_bindings.json
2025-07-23T20:47:40.181465Z [rsession-jovyan] INFO No keybindings/editor_bindings.json found at /e
tc/rstudio/keybindings/editor_bindings.json
2025-07-23T20:47:40.181940Z [rsession-jovyan] INFO No keybindings/addins.json found at /etc/rstudi
o/keybindings/addins.json
> 
```

Name	Size	Modified
4-machine_learning.ipynb	30.5 KB	Jul 23, 2025, 10:46 PM
4-machine_learning.rmd	15.8 KB	Jul 23, 2025, 10:46 PM
6-ml_exercise.ipynb	13 KB	Jul 23, 2025, 10:46 PM
6-ml_exercise.rmd	4.6 KB	Jul 23, 2025, 10:46 PM
environment.yml	161 B	Jul 23, 2025, 10:46 PM
LICENSE	1 KB	Jul 23, 2025, 10:46 PM
README.md	1.8 KB	Jul 23, 2025, 10:46 PM
runtime.txt	12 B	Jul 23, 2025, 10:46 PM

Caret

- Caret is a library the builds a common interface for many machine learning algorithms from different libraries in r
- As the name suggests it is specialized for supervised learning (the outcome variable is known) so for **C**lassification **A**nd **Re**gression **T**asks.
- Many different types of algorithms
  - Trees & forests
  - Regressions
  - Support vector machines
  - Neural nets
  - Gradient boosting
  - …



The ML Workflow

Step I:
Preparing your data

- Formatting
- Test-Train split
- Preprocessing

Step II:
Training a learning algorithm

- Algorithm selection
- Validation
- Tuning
- Testing

Step III:
Building a learning architecture

- Model lists
- Ensemble models
- Auto-ML

Workflow for management research (very similar):

Choudhury P, Allen RT, Endres MG. 2021. Machine learning for pattern discovery in management research. *Strategic Management Journal* **42**(1): 30–57.

Step I: Preparing your data

- To prepare our data for machine learning we need to:
  - Make sure it is the right format
  - Split it into training and test data
  - Conduct necessary pre-processing steps

Preparing your data I

- The right format
  - Data comes in many formats
  - Not all are equally easy to process
  - Continuous numeric data is the easiest
  - Rank data is also okay (*)
  - Categorical data needs to be transformed into dummy variables
  - Text and images (videos) have their own approaches and will not be covered here

Preparing your data II

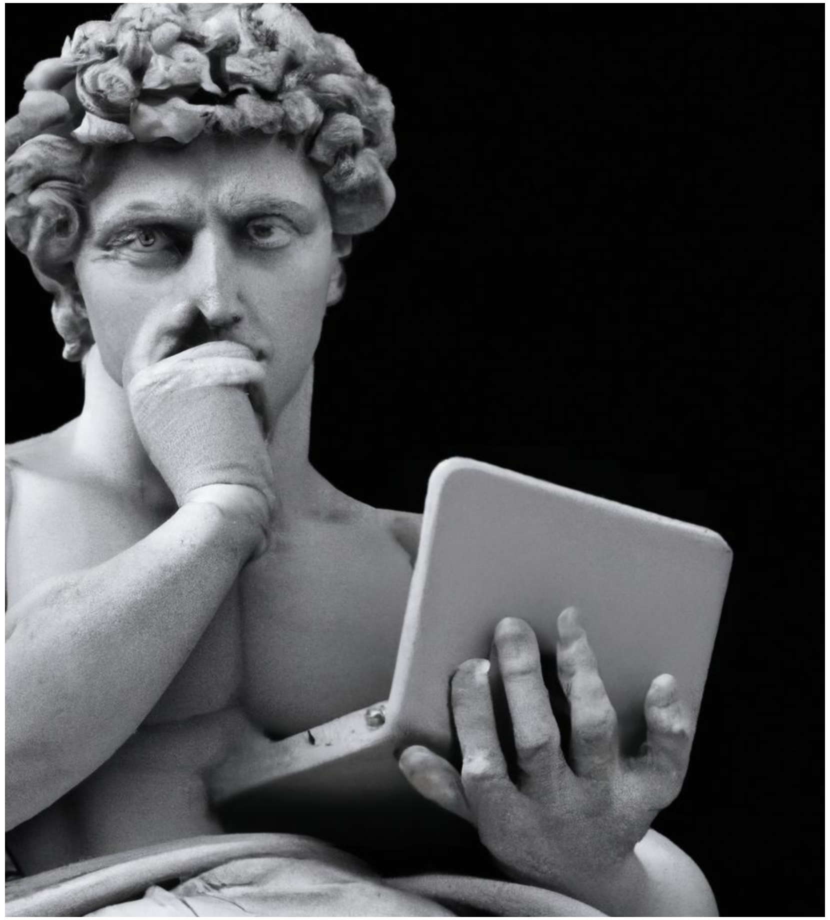
- To assess the fit of our models we need to split our data into a Test and a Training set
    - This is done because ML algorithms are prone to overfitting to the data
    - The performance on a data set the algorithm has not "seen" is more indicative of real world performance
    - In the real world the train-test split is often around 80:20-95:5 depending on how much data you have
    - When creating a train-test split we need to make sure that those datasets are sufficiently similar
    - We accomplish this by holding the distribution of the outcome variables as close to the initial distribution as possible

Preparing your data III

- Information content
  - Sometimes variables contain little or no variance (e.g., most/all customers come from the same country)
  - These variables contain little or no information but cost compute and could lead to overfitting
- Correlation
  - Some data points might be highly correlated
  - These variables contain redundant information and therefore, cost compute and might “confuse” models
- Imputation
  - Some variables might not be complete but still contain valuable information
  - We don’t want to drop observations or these variables -> imputation
- Centering and scaling
  - For some algorithms it can be useful if the variables are centered and scaled (e.g. clustering)

Preparing your data IV

- Things that can go wrong during data prep
  - Technical stuff
  - Wrong order: Sometimes it is tempting to preprocess all the data and then splitting it.
  - This is ok as long as the preprocessing does not involve looking at the whole dataset (as for example with imputation and correlation tests)
  - As a general rule:
    - Changing the format is ok (e.g., creating dummies)
    - Computing something is not (imputation, scaling, ...)



Step II: Training a Learning Algorithm

- To train a machine learning algorithm we need to:
  - Select (an) appropriate algorithm(s)
  - Select a way to validate our algorithm to avoid overfitting
  - Tune the algorithm’s hyper-parameters to find the best learner
  - Test the out-of-sample (OOS) performance of our trained algorithm

Training a Learning Algorithm I

- Choosing the right algorithms for a problem is as much art as it is science. Some important classes of algorithms include:
    - https://topepo.github.io/caret/available-models.htmlt
    - Decision tree based
    - Linear models
    - Neural networks
- Experience shows that many weak predictors can be combined to create strong predictors. Those ideas are key to:
    - Boosting
    - Bagging
    - (“StatQuest with Josh Starmer” on YouTube for intuitive explanations)

Training a Learning Algorithm II

- ML algorithms are very prone to overfitting.
- This is why we use validation during the training process to find the parameter tuning which optimize OOS accuracy
- Common methods are cross validation and bootstrapping:
  - CV: We split our data into k-Folds and train the model on k-1 of them. Then we estimate OOS performance on the remaining fold. We do this on all k combinations.
  - Bootstrapping: here we draw observations form the existing data set with repetition and test performance on the original dataset
  - Time series sometimes need other methods

Training a Learning Algorithm III

- ML algorithms try to fit the model parameters to the data.
- But the models themselves have (hyper-) parameters which determine how well they can "learn" from the data.
- There are different approaches to hyper parameter tuning:
  - Grid search
  - Random search
  - Adaptive resampling
  - Others (not yet implemented in caret; e.g. evolutionary algorithms in the "mlr" package)

Training a Learning Algorithm IV

- AFTER building our models we can test and compare their performance on the test data.
- Theoretically, test data is “burnt” once we check a models OOS performance on it. If we engage in further optimization, we start to (implicitly) fit to the test data.
- This is not very practical. Rule of thumb: Don’t do statistics to the test data performance. Comparing 10 models is probably ok, 1000 is not.

Step III: Building a Learning Architecture

- Model lists allow us to train multiple models at the same time with just a single line of code
- Auto-ML (e.g., in H2O)
- Ensemble models
  - Many weak learners can together create a strong learner
  - Simple approaches use averages or votes to aggregate the results of multiple instantiations of the same algorithm
  - But we can also use different algorithms
  - And use ML-algorithms to find the best way to combine them
  - “It’s turtles all the way down” – Hawking S. 1988. *A brief history of time*.

Other Resources

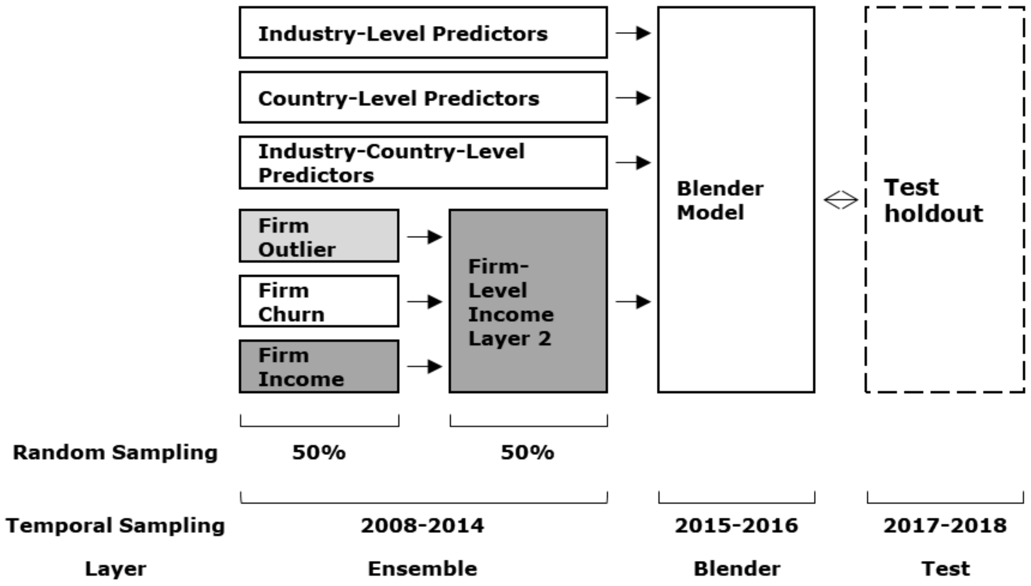
- DataCamp
- YouTube
- Caret documentation
- Other ML libraries for R
  - mlr
  - H2O
- ChatGPT et al.

Case Study: Predicting Foreign Subsidiary Profits

Forecasting Austria's International GDP with ML Techniques

- **Goal:** Explore the applicability of machine learning techniques to GDP forecasting utilizing firm-level micro data
- Approach
  - Ensemble model of 60 ML-algorithms
  - Utilizing information embedded on different scales of the data (firm-level, country-level, industry-level,…)
  - Special attention to outlier prediction
- Data
  - OENB Active Direct Investments
  - OECD Composite Leading Indicator

Simplified model and data sampling framework



Industry-Level Predictors

Country-Level Predictors

Industry-Country-Level Predictors

Firm Outlier

Firm Churn

Firm Income

Firm-Level Income Layer 2

Blender Model

Test holdout

Random Sampling 50% 50%

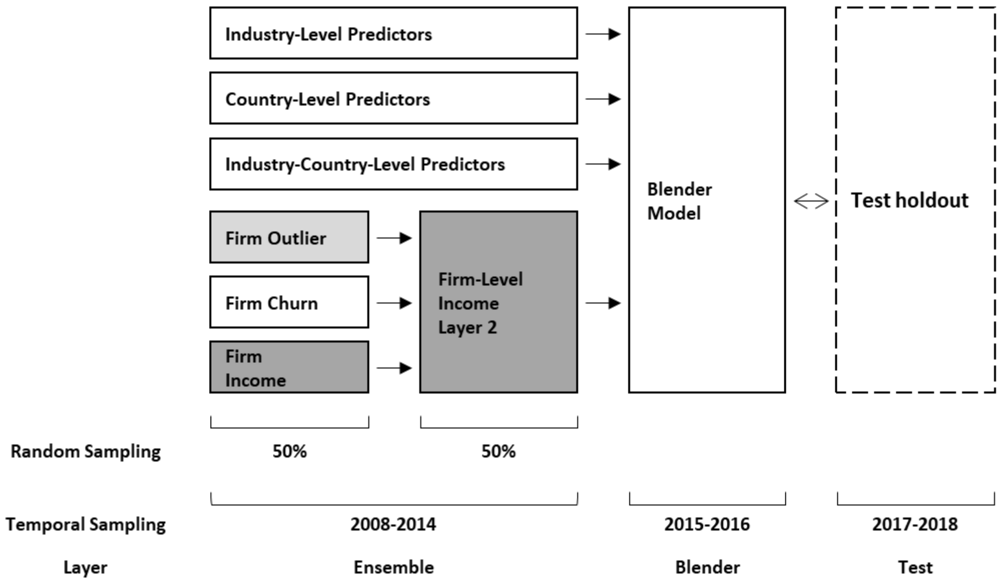
Temporal Sampling 2008-2014 2015-2016 2017-2018

Layer Ensemble Blender Test

Characteristics of the Data Shape the Training Architecture

- Two important characteristics
  - Several information rich scales/ levels (firm, country-industry, country, industry)
  - Short time series
- To account for both we implemented
  - two **layered firm level sub ensemble** allowing the estimation of firm level characteristics, based on traditional observation bases random sample splits
  - A single layered architecture for the **aggregate models** accounting for the time series nature of the data

Simplified model and data sampling framework



Industry-Level Predictors

Country-Level Predictors

Industry-Country-Level Predictors

Firm Outlier

Firm Churn

Firm Income

Firm-Level Income Layer 2

Blender Model

Test holdout

Random Sampling	50%	50%		
Temporal Sampling	2008-2014		2015-2016	2017-2018
Layer	Ensemble		Blender	Test

New Features Reflect the Information Found at Different Scales

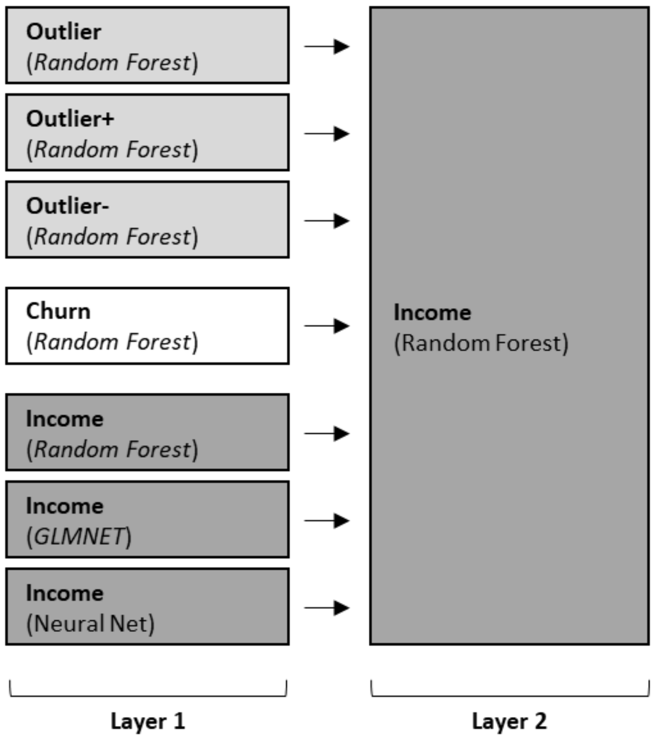
- Firm Level Features
  - Year of entry
  - Year of exit (churn)
  - Outlier scores:
    - Isolation score (tree based)
    - Local outlier factor (clustering/density based)
    - Top/bottom Quantile
- Temporal Features
  - All data points were lagged by up to 6 years to cover temporal effects

- Aggregation Level Features
  - Aggregation Levels:
    - Industry – Country
    - Industry
    - Country
  - Features (absolute & relative):
    - Income
    - Churn
    - Entry
    - Number of firms

Firm Level Ensemble Architecture

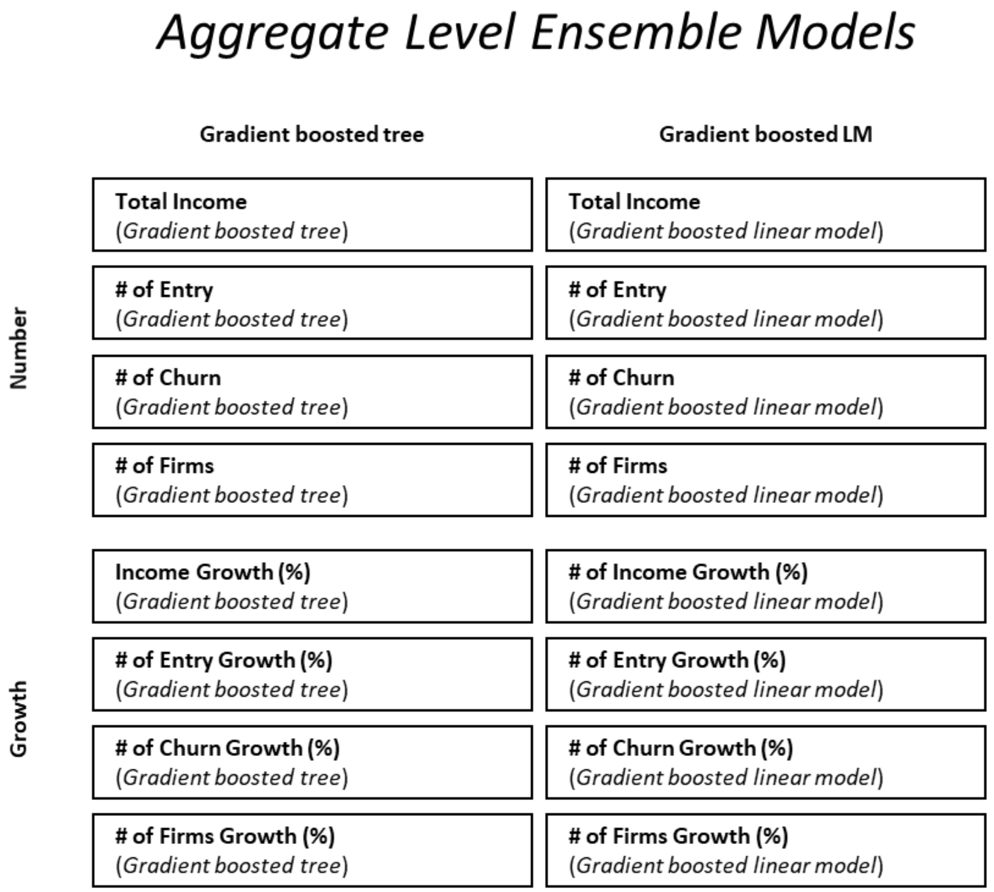
- Outlier prediction
  - Random forests predict (some of the) outlier features
- Churn
  - A random forest predicts the exit (/churn) feature
- Income
  - Income is used as the label in training of a random forest, a GLMNET and a neural net model.
- Layer 2
  - Layer 2 predicts income with a random forest, but also takes the predictions of the previous models as input

Firm Level Ensemble Model



Aggregate Level Ensemble Architecture

- Aggregation levels
  - Industry-Country Years
  - Industry-Years
  - Country-Years
- Predicted Labels
  - Income
  - Number of firms, entries, & exits
- Models
  - A gradient boosted tree and linear model each
- Key issues
  - Low number of observation in country and industry aggregates

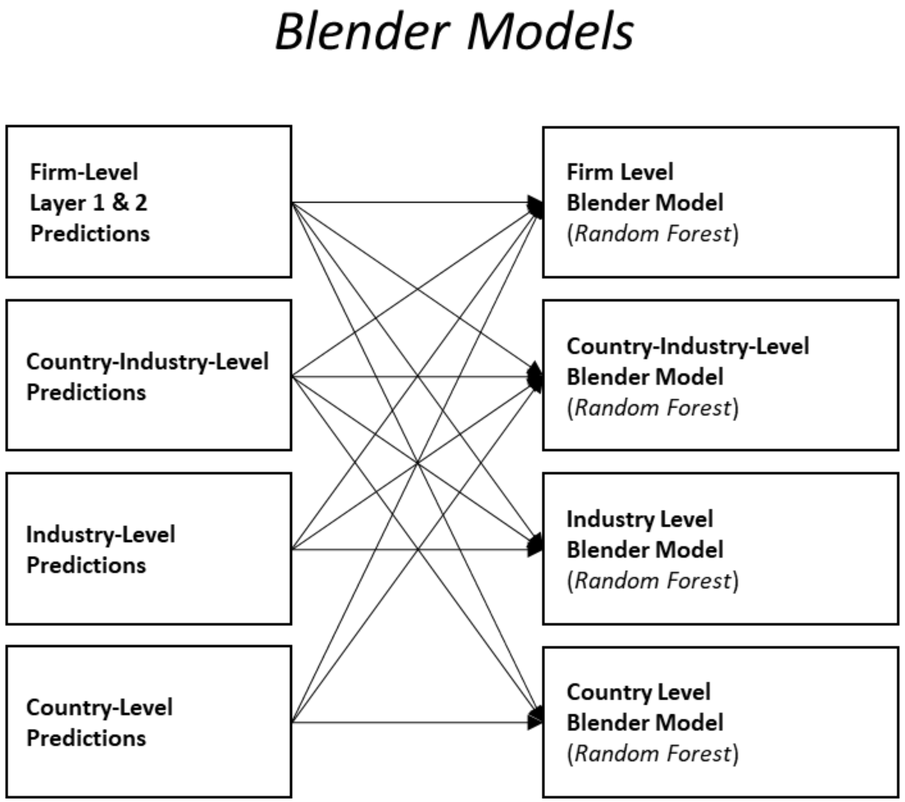


Aggregate Level Ensemble Models

	Gradient boosted tree	Gradient boosted LM
Number	**Total Income** (*Gradient boosted tree*)	**Total Income** (*Gradient boosted linear model*)
	# of Entry (*Gradient boosted tree*)	**# of Entry** (*Gradient boosted linear model*)
	# of Churn (*Gradient boosted tree*)	**# of Churn** (*Gradient boosted linear model*)
	# of Firms (*Gradient boosted tree*)	**# of Firms** (*Gradient boosted linear model*)
Growth	**Income Growth (%)** (*Gradient boosted tree*)	**# of Income Growth (%)** (*Gradient boosted linear model*)
	# of Entry Growth (%) (*Gradient boosted tree*)	**# of Entry Growth (%)** (*Gradient boosted linear model*)
	# of Churn Growth (%) (*Gradient boosted tree*)	**# of Churn Growth (%)** (*Gradient boosted linear model*)
	# of Firms Growth (%) (*Gradient boosted tree*)	**# of Firms Growth (%)** (*Gradient boosted linear model*)

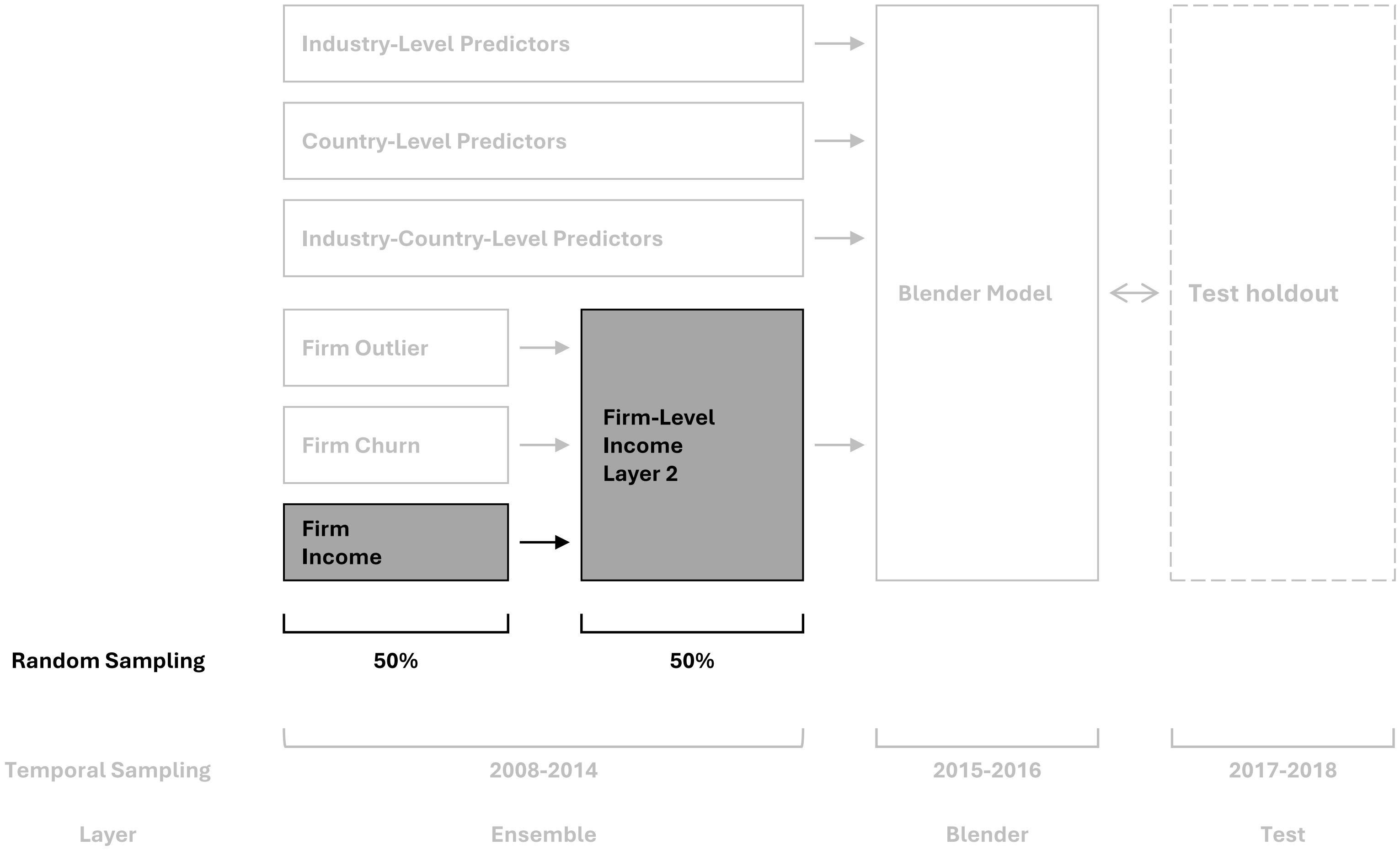
Blender Models

- Aggregation levels
    - Firm
    - Industry-Country Years
    - Industry-Years
    - Country-Years
- Predictors
    - Previous years income
    - Predictions of all the models
    - Country and industry level predictions can not be combined



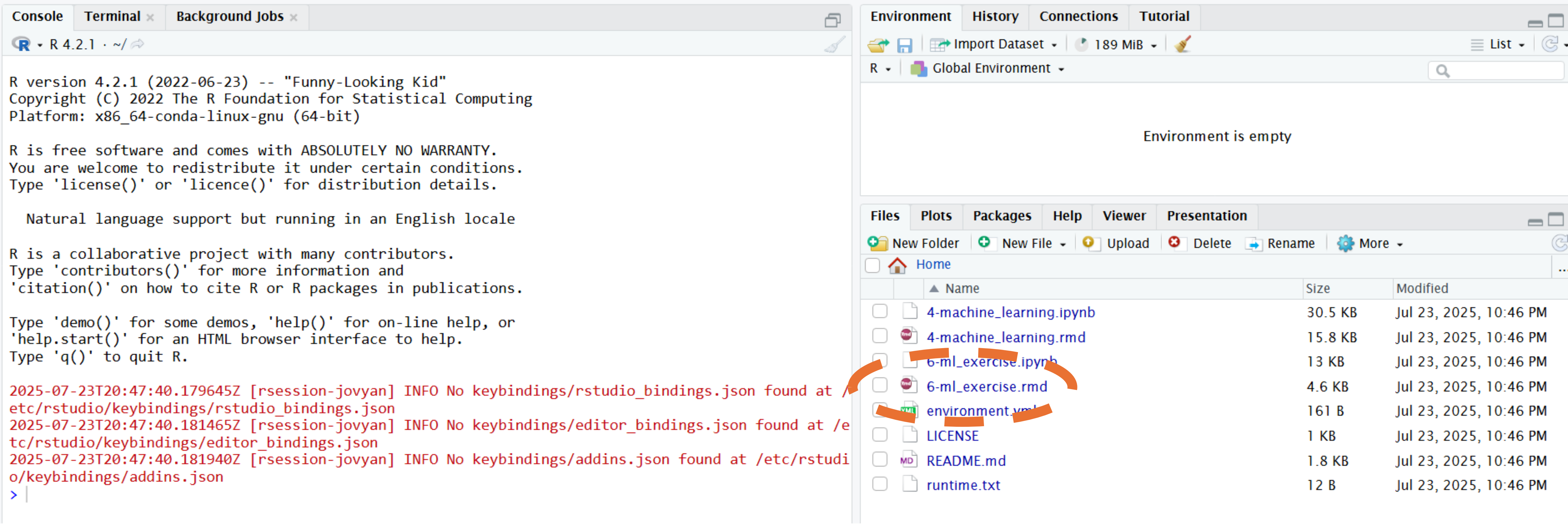
Blender Models
Firm-Level Layer 1 & 2 Predictions
Country-Industry-Level Predictions
Industry-Level Predictions
Country-Level Predictions
Firm Level Blender Model (*Random Forest*)
Country-Industry-Level Blender Model (*Random Forest*)
Industry Level Blender Model (*Random Forest*)
Country Level Blender Model (*Random Forest*)

Your Turn!


Industry-Level Predictors
Country-Level Predictors
Industry-Country-Level Predictors
Firm Outlier
Firm Churn
Firm Income
Firm-Level Income Layer 2
Blender Model
Test holdout
Random Sampling
50%
50%
Temporal Sampling
2008-2014
2015-2016
2017-2018
Layer
Ensemble
Blender
Test

Hands-On Exercise

- Once the RStudio session has opened, click on „6-ml_exercise.rmd“ on the bottom right



```
R version 4.2.1 (2022-06-23) -- "Funny-Looking Kid"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-conda-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

2025-07-23T20:47:40.179645Z [rsession-jovyan] INFO No keybindings/rstudio_bindings.json found at /
etc/rstudio/keybindings/rstudio_bindings.json
2025-07-23T20:47:40.181465Z [rsession-jovyan] INFO No keybindings/editor_bindings.json found at /e
tc/rstudio/keybindings/editor_bindings.json
2025-07-23T20:47:40.181940Z [rsession-jovyan] INFO No keybindings/addins.json found at /etc/rstudi
o/keybindings/addins.json
> |
```

Name	Size	Modified
4-machine_learning.ipynb	30.5 KB	Jul 23, 2025, 10:46 PM
4-machine_learning.rmd	15.8 KB	Jul 23, 2025, 10:46 PM
6-ml_exercise.ipynb	13 KB	Jul 23, 2025, 10:46 PM
6-ml_exercise.rmd	4.6 KB	Jul 23, 2025, 10:46 PM
environment.yml	161 B	Jul 23, 2025, 10:46 PM
LICENSE	1 KB	Jul 23, 2025, 10:46 PM
README.md	1.8 KB	Jul 23, 2025, 10:46 PM
runtime.txt	12 B	Jul 23, 2025, 10:46 PM

Coffee House Style Discussion