3.1 Artificial Intelligence and Machine Learning

What is Biomedical and Health Informatics? - http://informatics.health/

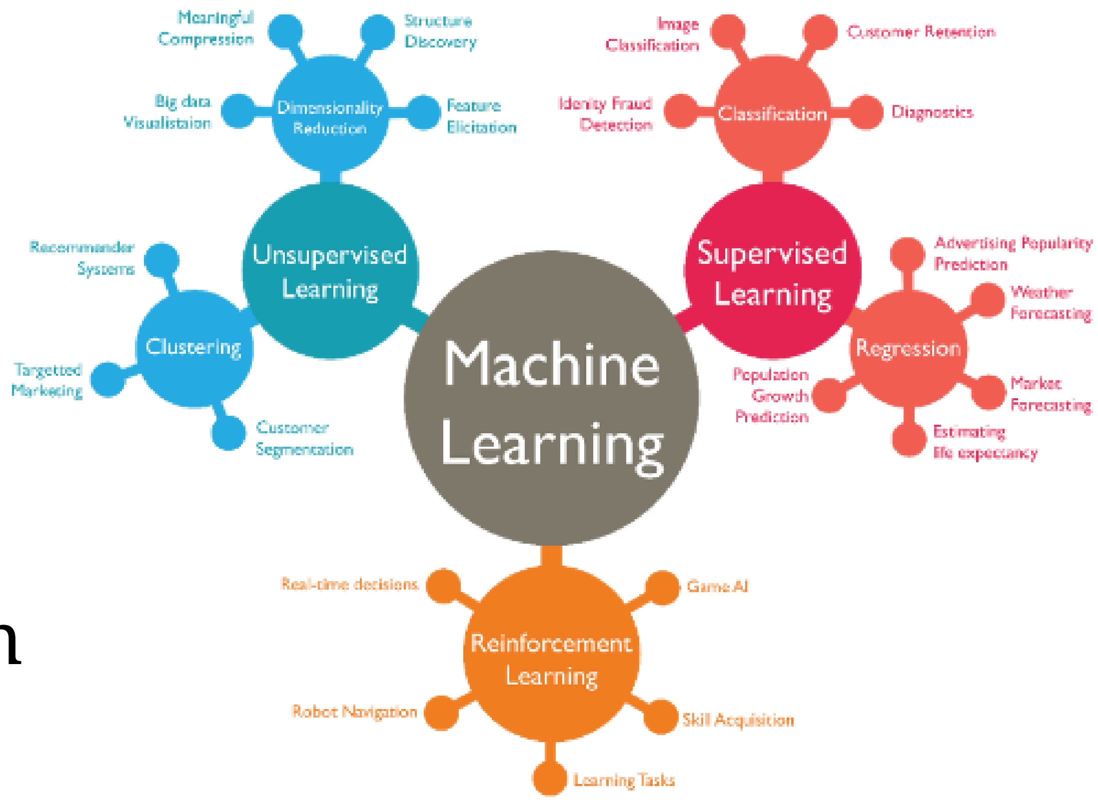
William Hersh, MD, FACMI, FAMIA, FIAHSI

Machine learning (ML)

- Methods
- Resources
- Tools

Methods of ML (Shah, 2022)

- Supervised – learn to predict a known output
  - Learns from labeled training data
  - Aim to assign labels to test data
    - To avoid overfitting
- Unsupervised – find naturally occurring patterns or groupings within data
- Reinforcement learning learns from new data and results, e.g., from ongoing use in a clinical setting (Gottesman, 2019; Murphy, 2024)



(Chugh, 2018)

Other types of learning in ML

- Semi-supervised learning – combination of supervised learning with (relatively) small amount of labelled data and unsupervised learning from (typically) larger amounts of other data
- Self-supervised learning – identify labels from patterns in data (unsupervised) and apply for supervised learning
  - Major application is in large language models (LLMs) (Harak, 2024)
- Transfer learning – applying learning trained for one task to another (Yang, 2020)

Tasks of supervised learning

- Classification – predict class from one or more features of data, e.g., diagnosis or patient outcome
  - k-Nearest Neighbors (kNN) – aim to find category having “closest” number of attributes
  - Naïve Bayes – derive conditional probabilities that classify into categories
  - Support vector machines (SVMs) – for binary classification, draw “line” that separates one category from other
  - Decision trees – develop set of rules that classify into categories
- Regression – predict numerical value from data, e.g., risk of disease or poor outcome or benefit from treatment
  - Linear – fit a line to data
  - Multivariate (polynomial) – fit many variables to model
  - Logistic regression – binary output

Tasks of other types of learning

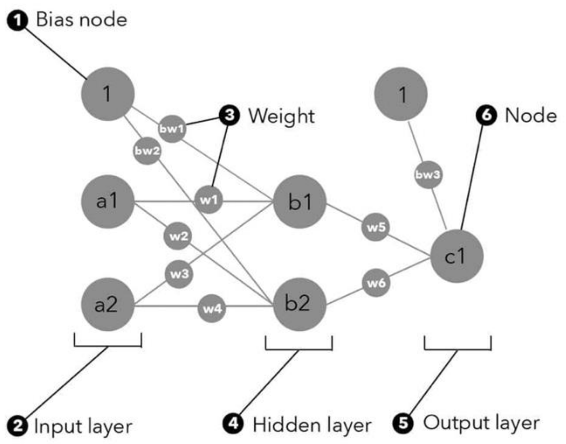
- Unsupervised learning
  - Clustering – group items together
  - Density estimation – find statistical values
  - Dimensionality reduction – reduce many to few features
- Growing use of self-supervised and transfer learning in LLMs
- With large models, sometimes describe zero/one/few-shot learning (Kadam, 2020; Sarojag, 2023)
  - Zero-shot – use existing model as is
  - One-shot – add some material or context just one time
  - Few-shot – add material up to a few times

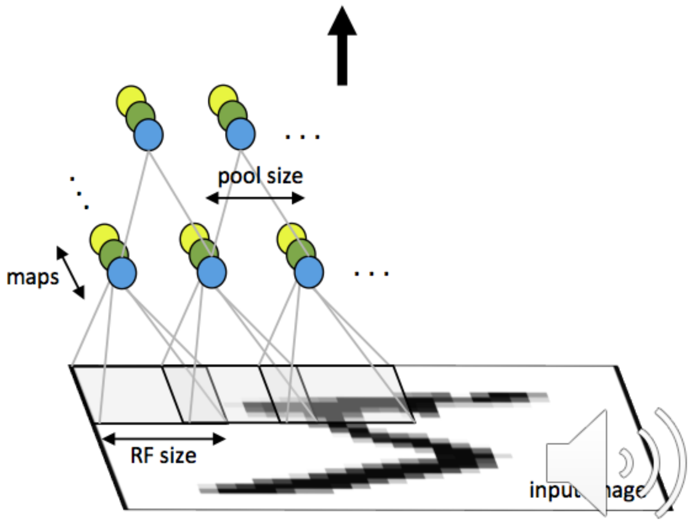
Artificial neural networks (ANNs)

- Have come to fore as main approach for ML with large amounts of data and increased modern computing power (Shah, 2022)
  - Particular success has been achieved with deep learning, with much internal complexity to networks
  - ANNs had been around for many decades (McCulloch, 1943), but deep learning successes often attributed to work of Hinton (2006)
- Mathematics complex, but can understand what they do in context of ML tasks

Anatomy and physiology of neural networks (Taylor, 2017)

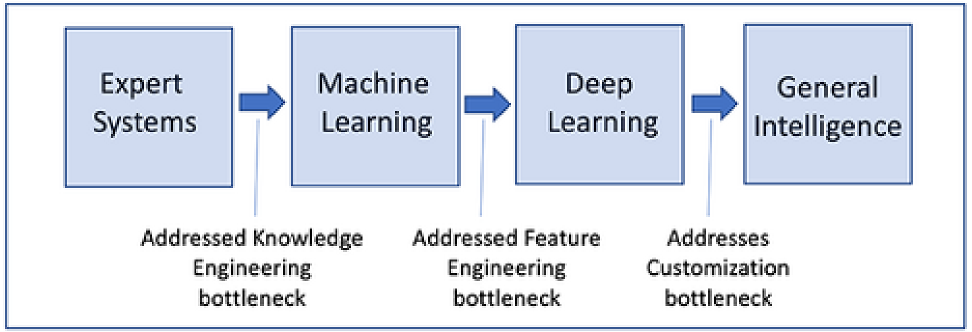
- Anatomy
  - Layers
  - Nodes and weights – connected like neurons
- Physiology
  - Feedforward – processing from input to output
    - Convolutional neural networks (CNNs) particularly effective for image analysis
  - Feedback – processing loops backwards
    - Sometimes called recurrent neural networks (RNNs), most useful for sequential data, such as text





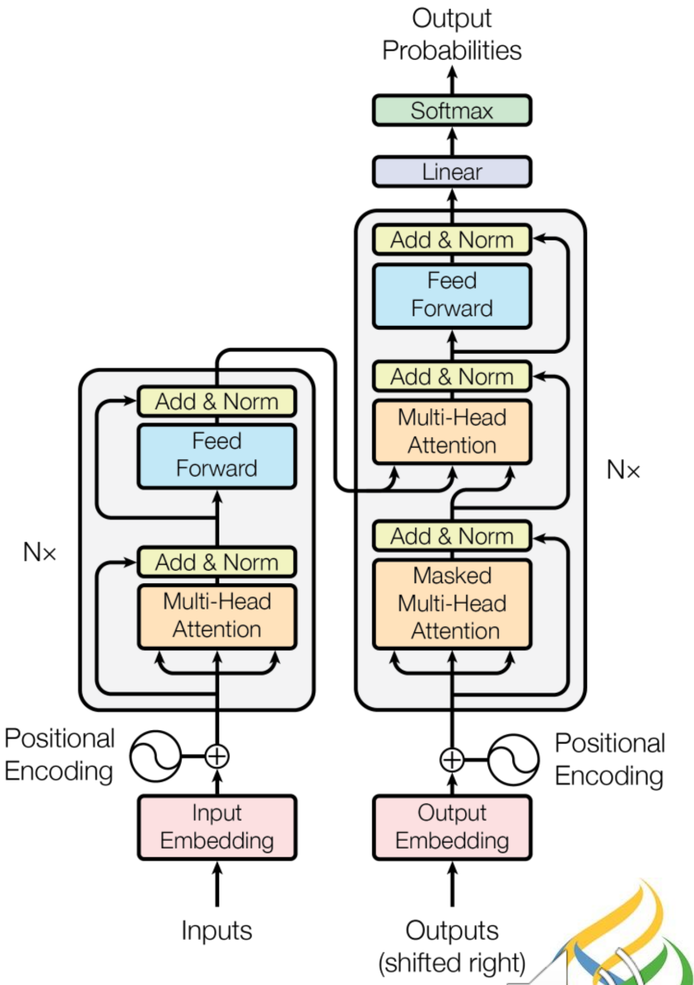
LLMs based on ANNs

- Next evolution of paradigms of AI (Dhar, 2024)
- Recent overviews
  - LLMs generally (Shao, 2024; Bahree, 2024)
  - LLMs in biomedicine
    - Clinical (Omiye, 2024)
    - Technical (Shanmugam, 2024; Xiao, 2024)

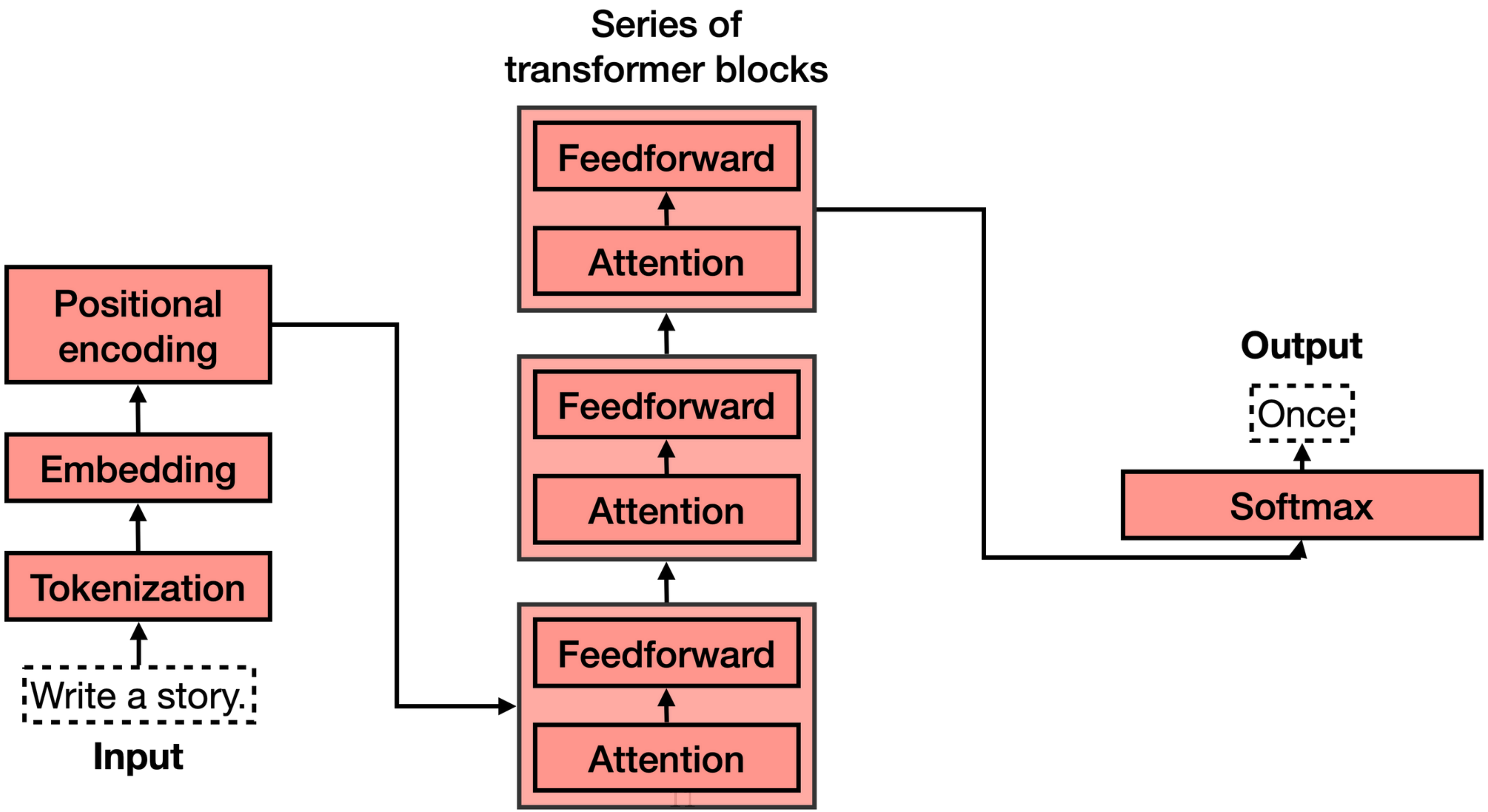


LLMs

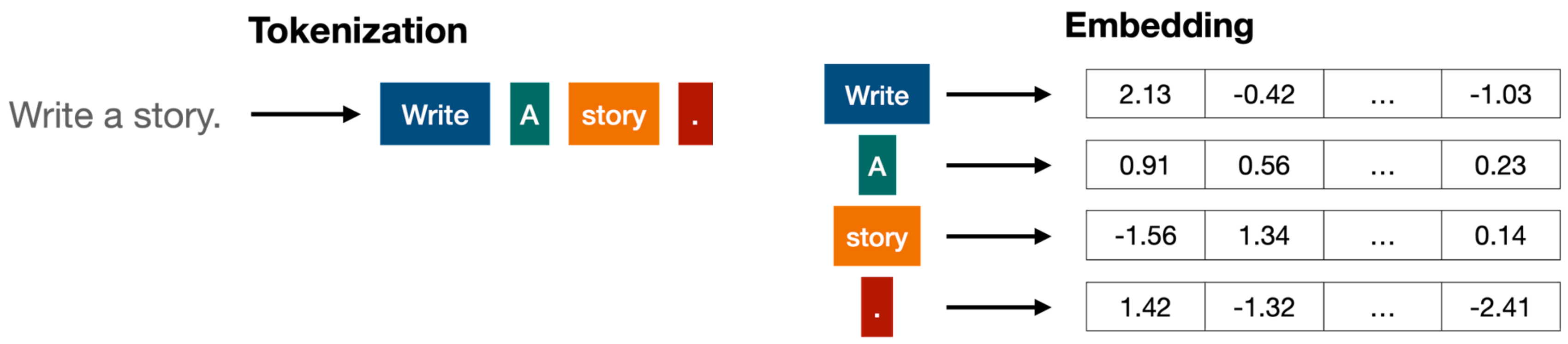
- LLMs based on transformer architectures – key concepts include (Vaswani, 2017; Amatrain, 2023)
  - Predicting next word in sequence
  - Measuring “attention” – “calculating” context of words
- Transformer models may include encoders and/or decoders
  - Encoder models transform text into LLMs for tasks such as document classification and sentiment analysis
    - Earliest and one of best-known encoder models is Bidirectional Encoder Representations from Transformers (BERT) (Devlin, 2019; Muller, 2022)
  - Decoder models generate text from LLMs for tasks such as answering questions
    - One of best-known decoder models is Generative Pre-trained Transformer (GPT) (Brown, 2020; Piper, 2020; OpenAI, 2024)

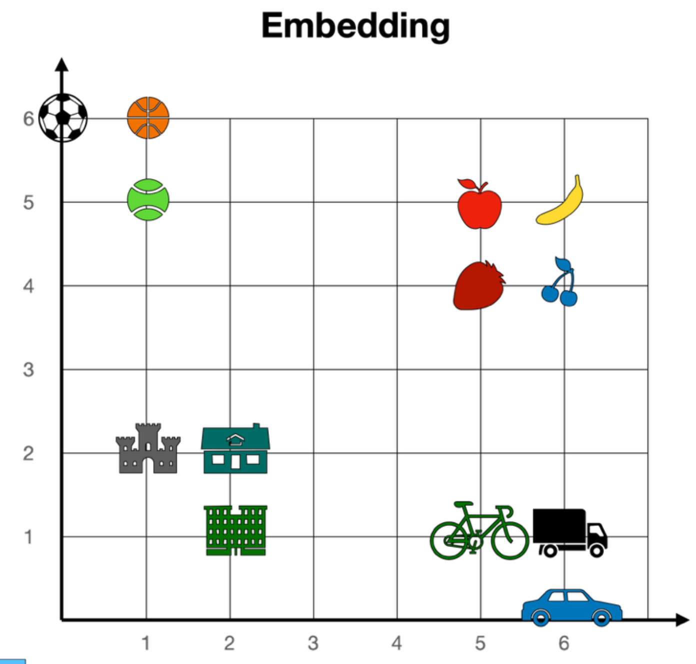


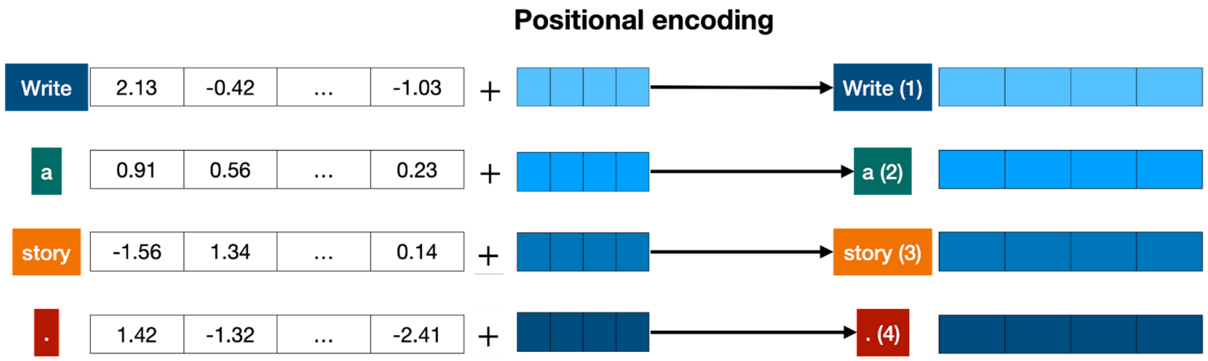
Basic overview of how transformer decoders work (Serrano, 2023)



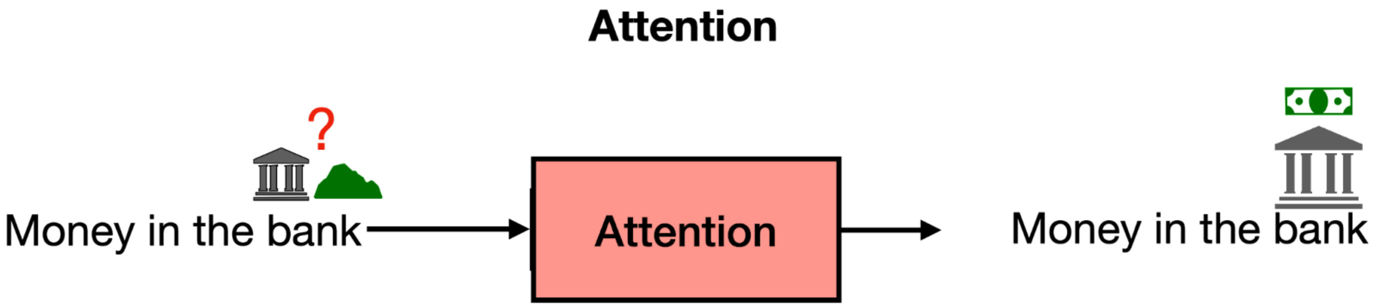
Tokenization, embedding, and positional encoding



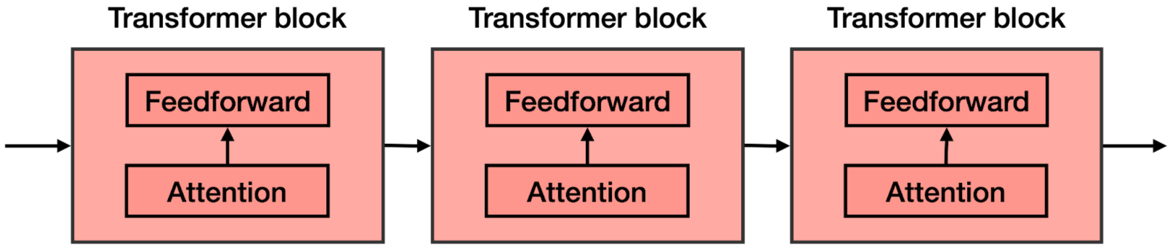


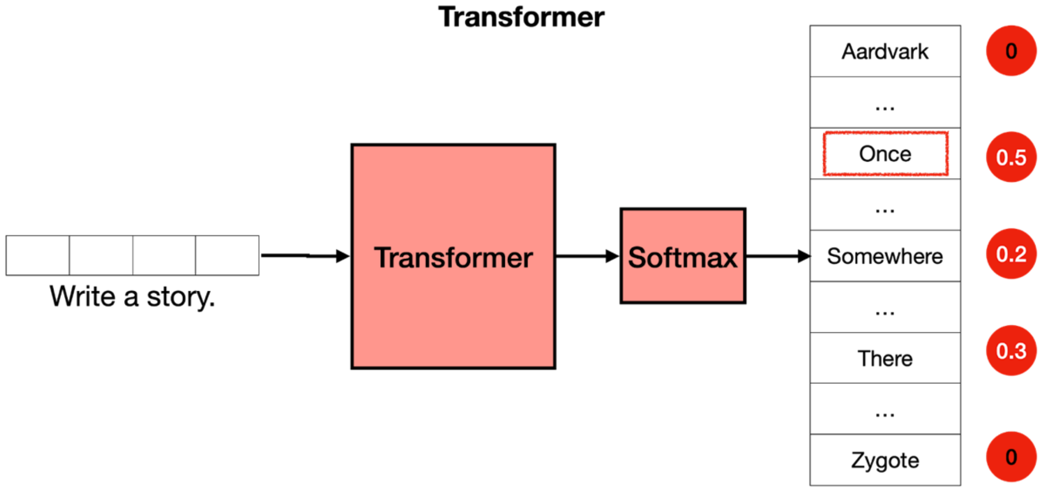


Attention, transformers, and softmax



Sentence 1: The **bank** of the river.
Sentence 2: Money in the **bank.**





Post-training – from models to applications

- Reinforcement learning from human feedback (RLHF) – additional training of model to reward desired behavior, e.g., (Ziegler, 2020; Lambert, 2022)
  – Carrying out desired tasks, such as answering questions or writing computer code
  – Disincentivizing inappropriate behavior, e.g., sexism, racism, etc.
- Context windows – allow additional content to be added to augment model after initial training, e.g.,
  – Retrieval-augmented generation (RAG) can add new content to existing model (Gao, 2024; Ng, 2025)

Resources for ML

- Biomedical Data Science (Hoyt, 2019)
- No-code data science (Patrishkoff, 2023)
- ML algorithms in depth (Smolyakov, 2024)
- Hands-on introduction (Shah, 2022)
- Biomedical ML and AI (Simon, 2024)
- Math important but not necessary for understanding big picture
  - Statistical learning (James, 2017)
  - Math for ML (Deisenroth, 2020)
  - Causal inference (Hernán, 2023)
  - Math behind AI (Ananthaswamy, 2024)

How much do we need to know about what happens under the hood in order to drive a car?

Tools for ML

- Programming
- Models
- Datasets

ML programming

- Many programming languages but 2 most widely used (both open-source)
  - Python – easy to use and read language has gained popularity for data science and ML (Downey, 2024)
  - R – focused on statistical computing and graphics, especially with “tidy” data (Wickham, 2023)
- Computational notebooks – locally run Web pages that contain live code, equations, figures, interactive apps, and Markdown text (Vaughan, 2023)
  - Original and most commonly used are Jupyter notebooks, initially developed for Python but expanded to other languages, including R

ML programming (cont.)

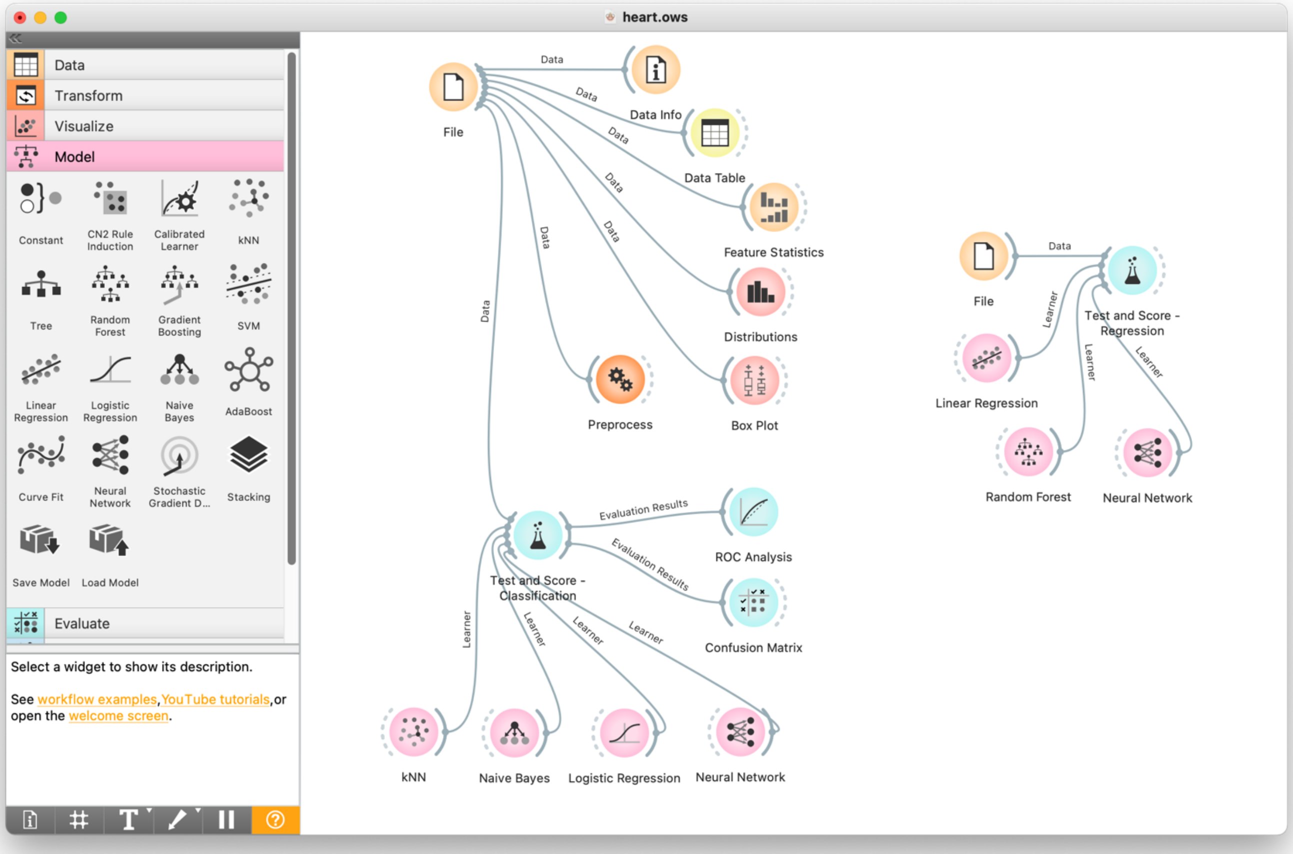
- Code libraries – several open-source
  - TensorFlow – Google
  - Scikit-learn – for Python
  - Tidyverse – libraries for analyzing (dplyr) and visualizing (ggplot) “tidy” data in R
  - Langchain – for LLMs (Lim, 2023)
- Integrated environment – Anaconda
  - Access to Python, R, computational notebooks, and other systems
  - Maintains updates for all systems
  - Downloadable versions for PC and Mac

No-code/visual programming – Orange data mining

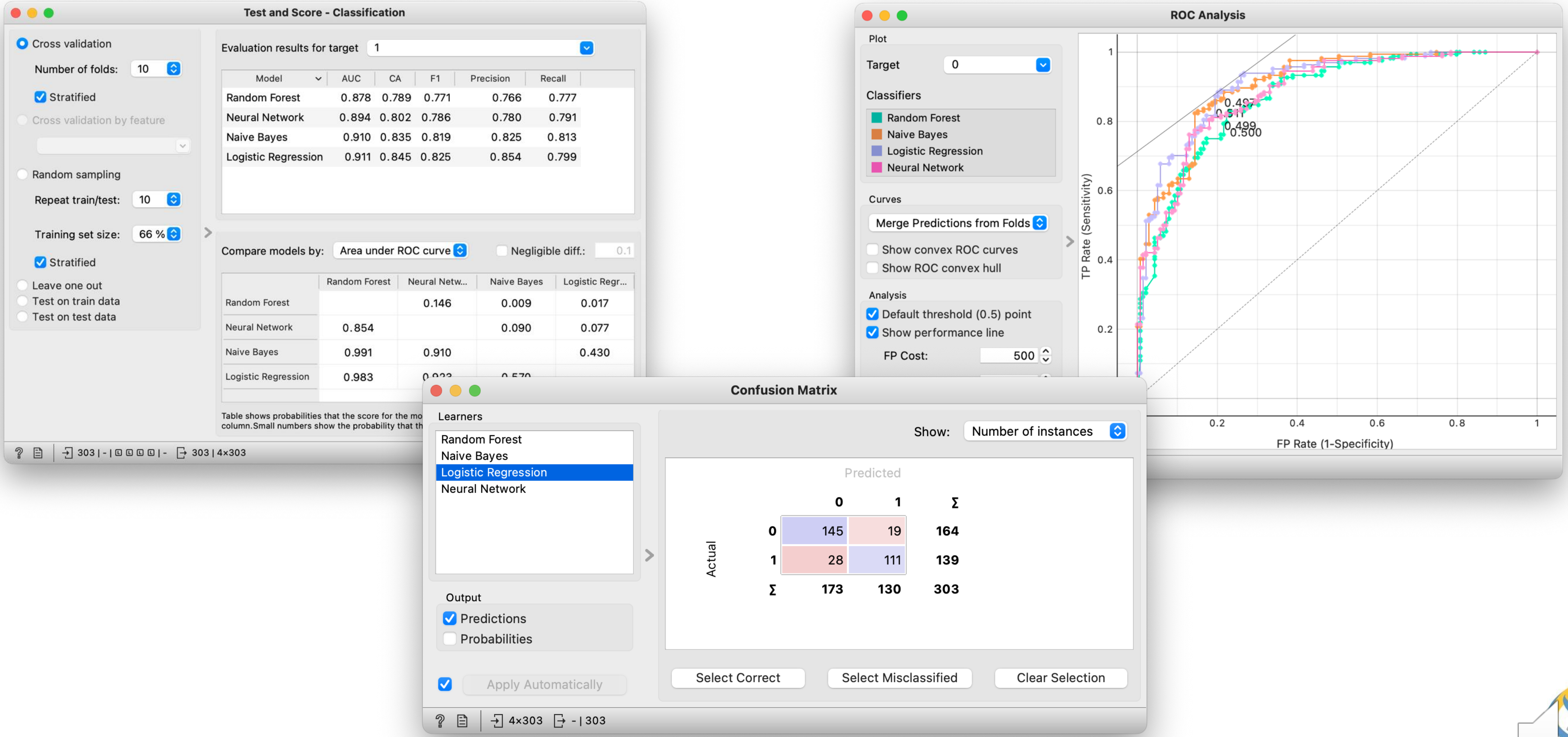
- "No-code" model development – visual programming package
- Open-source with large supporting community
  - Downloadable versions for PC and Mac
- Written in Python and extensible
  - Comes with "built-in" datasets and capability to import others
- Extensive series of introductory YouTube videos
- No-Code Data Science textbook and datasets (Patrishkoff, 2023)

Functionality in Orange

- Data loading, transformation, and visualization
- Predictive models
  - Classification
  - Regression
- Image classification
- Text mining



Orange – using and evaluating models



ML models

- Kaggle – original site for “competitions” but has expanded focus to repository for datasets, models, and notebooks with code
- Hugging Face – access to models, datasets, and compute
  - Educational toolkit
- Bringing models to data for federated evaluation, e.g., MedPerf (Karargyris, 2023)
  - Expanded into ML Commons

Model cards (Ozoani, 2023)

- Metadata about models, including
  - Description – developer(s), funding, type, etc.
  - Sources – repository, paper, demo
  - Uses – direct, downstream, out-of-scope
  - Bias, risks, and limitations
  - Training details – training data and procedures
  - Evaluation
  - Technical specifications – how to implement
- Used by model aggregation sites, such as Hugging Face and Kaggle, and others
- Proposed by Mitchell (2019), based on similar approaches used by natural language processing (Bender, 2018) and dataset (Gebru, 2020) communities

Additional ML datasets

- Overview and concerns (Altexsoft, 2022)
- UCI ML Repository
- Papers with code
- Physionet.org, including Medical Information Mart for Intensive Care (MIMIC) – (Johnson, 2023)
- Bridge2AI – NIH project to develop high-quality, ethically sourced data sets for ML research
  - Voice AI – voice as a biomarker (Bensoussan, 2024)
  - Artificial Intelligence Ready and Equitable Atlas for Diabetes Insights (AI-READI) – salutogenesis (AI-READI, 2024)
  - Patient-Focused Collaborative Hospital Repository Uniting Standards (CHoRUS) for Equitable AI – AI/ML for clinical care
  - Cell Maps for AI (CM4AI) – functional genomics

OHSU