Welcome to the Intro to AI and Mini-Hack

https://forms.gle/pgsqnzBntKeNw2mC9

Ohio Supercomputer Center

California State University San Bernardino

ASU Knowledge Enterprise Arizona State University

CI REN Cross-Institutional Research Engagement Network

ACCESS Support

CAMPUS CHAMPIONS

ACCOUNT SIGN-UP

Sign-up link https://my.osc.edu/acprod/odb_osc/r/osc/portal/signup?clear=202

Project code:
PZS1142

Access code:
624239

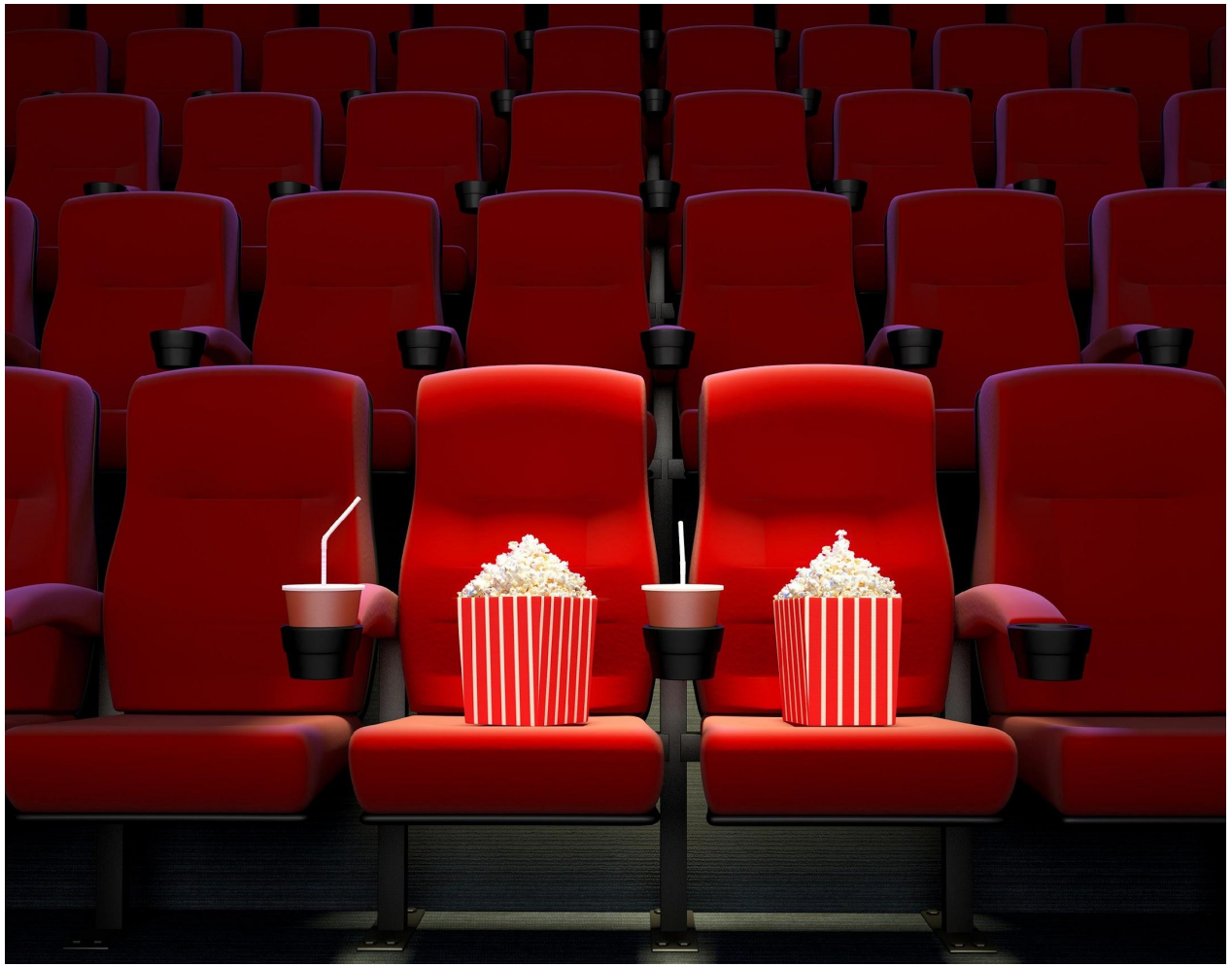
Expires:
31 July 2024

Account sign-up instructions
https://www.osc.edu/supercomputing/portals/client_portal/self_signup_for_accounts

Ohio Supercomputer Center

PEARC24 ACCESS AI MINI-HACKATHON

Machine Learning Movie Night



Evan Jaffe, PhD

Machine Learning Engineer

Ohio Supercomputer Center

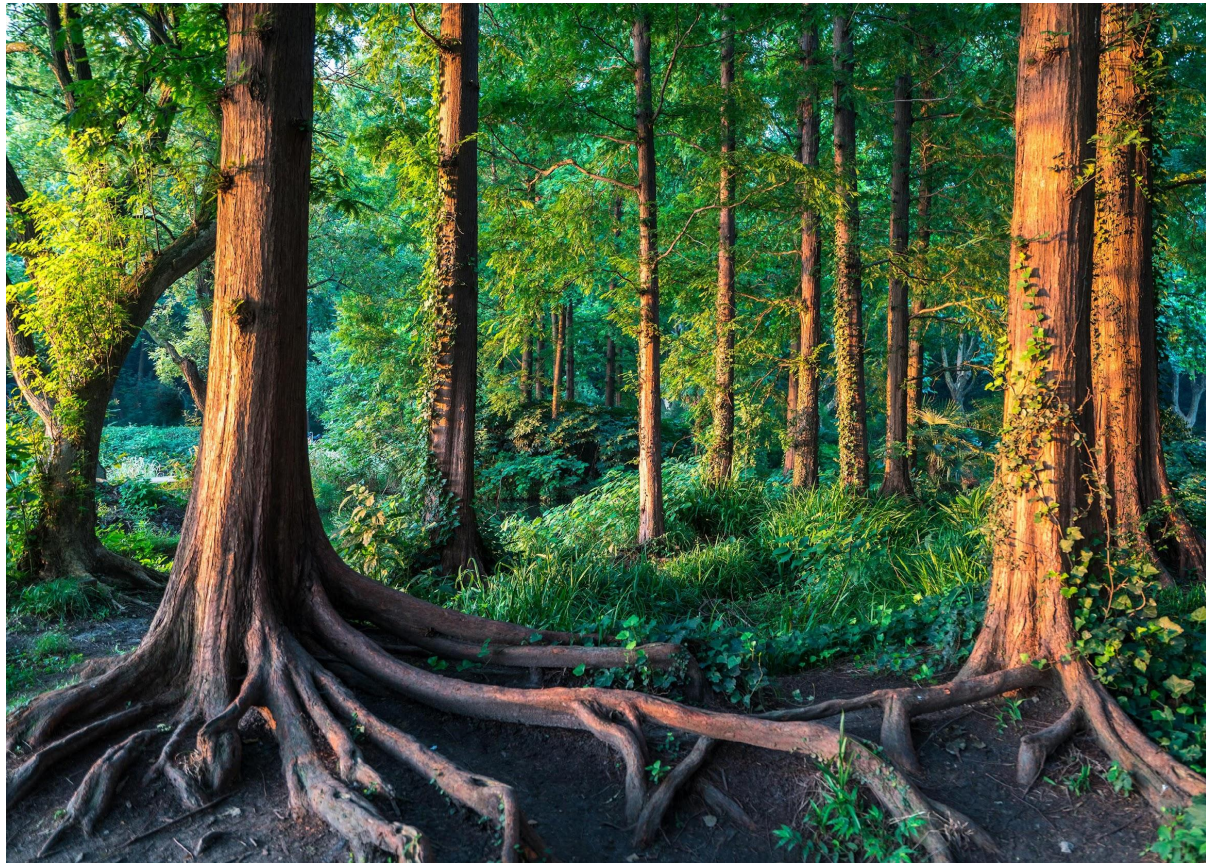
SENTIMENT ANALYSIS

- Given a review, predict user sentiment
- Useful in e-commerce, client management, customer service, social media analysis, political polling, etc.

- Binary classification
- 25k each, positive and negative movie reviews (Maas et al., 2011)

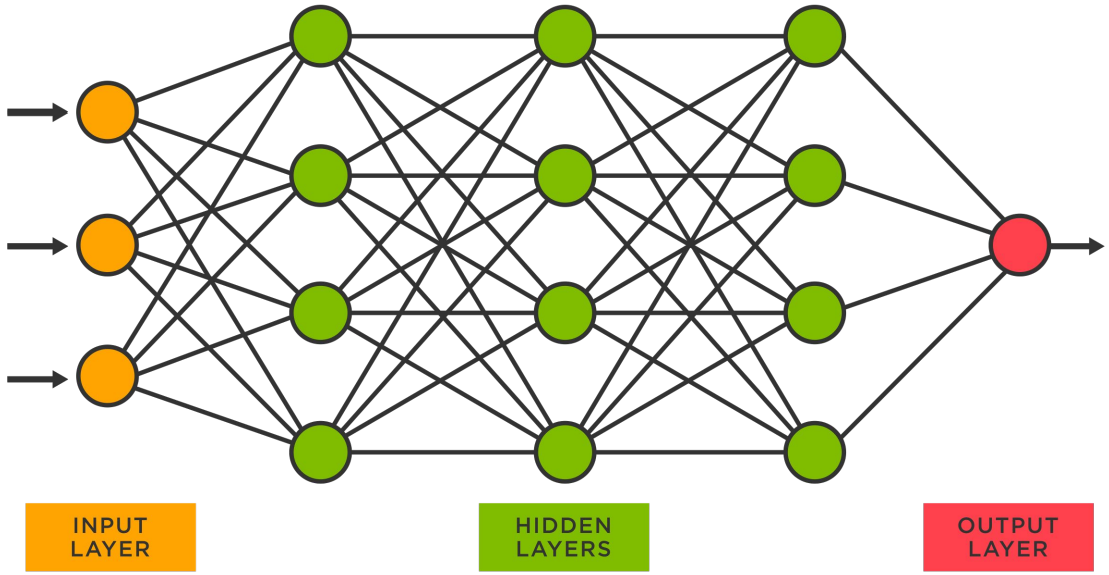
SUPERVISED CLASSIFICATION – 2 EXAMPLES

- Random Forest
  - Classic method
  - Splits data according to features
  - Interpretable



SUPERVISED CLASSIFICATION – 2 EXAMPLES

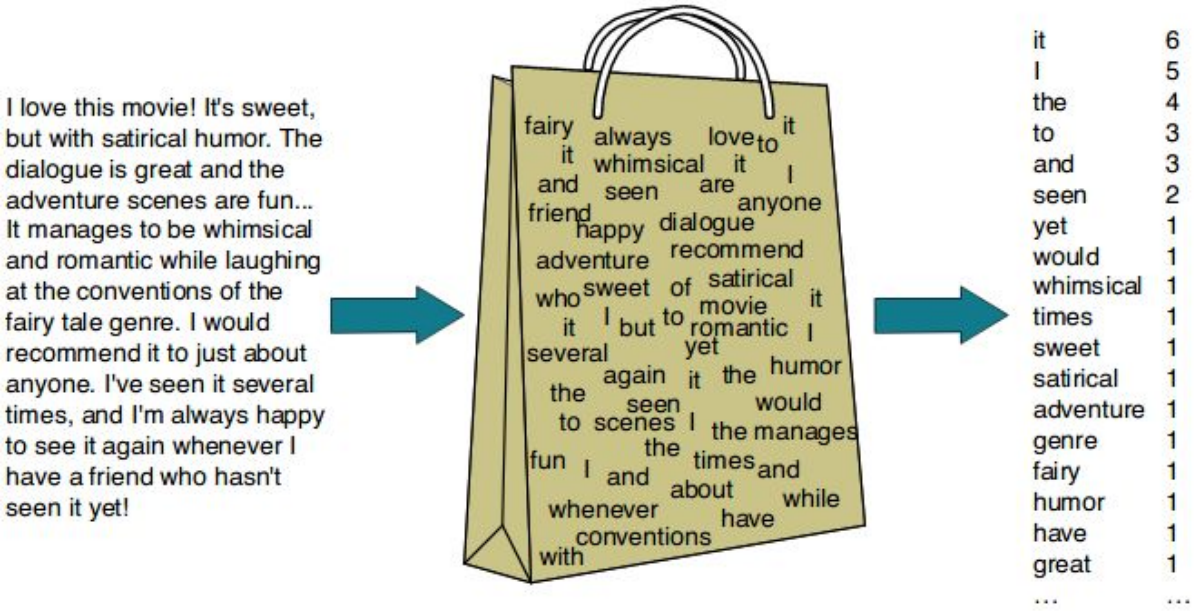
- Multilayer Perceptron
  - Simple artificial neural network
  - Basis for complex ANNs, Deep Learning



https://www.tibco.com/reference-center/what-is-a-neural-network

DATA PREPROCESSING

- Vectorization
  - Converts data into vector, e.g. [6,5,4,3,3,2,1,1,1,1,…]
- Bag of Words
  - Count words in document
- Other concerns
  - Vocabulary size, e.g. 5000 most frequent
  - Tokenization
  - Case normalization

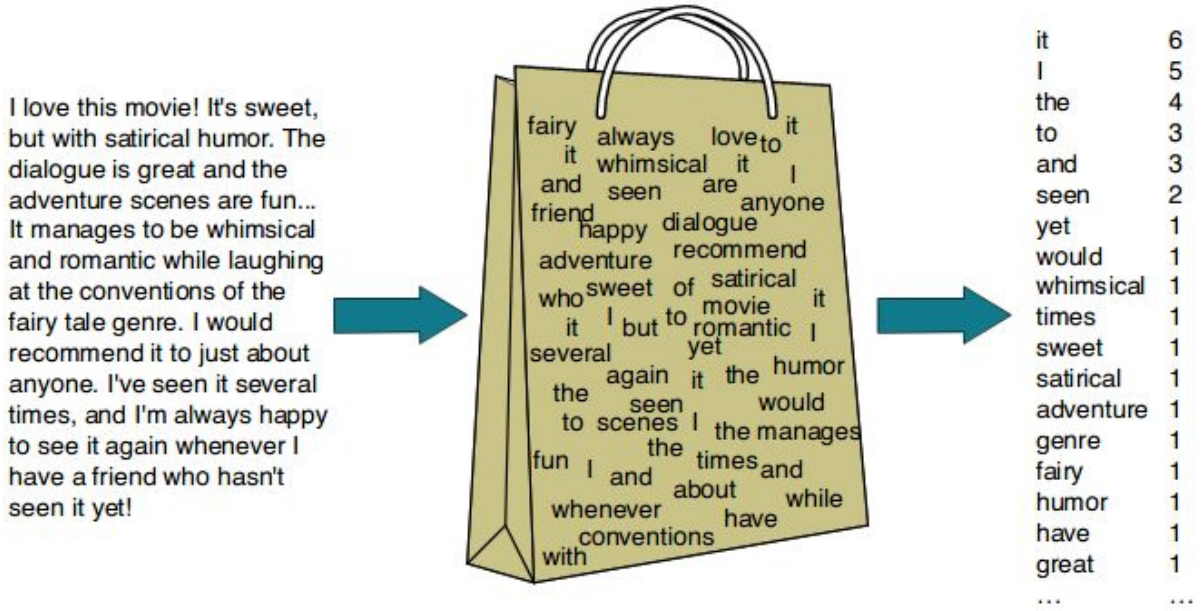


https://dudeperf3ct.github.io/lstm/gru/nlp/2019/01/28/Force-of-LSTM-and-GRU/

DATA PREPROCESSING

review id	Whimsical	Satirical	great	…	sentiment
1	1	1	1	…	1
2	0	1	0	…	0
3	0	0	0	…	0
4	0	0	2	…	1

Some features correspond more strongly with one category or another!

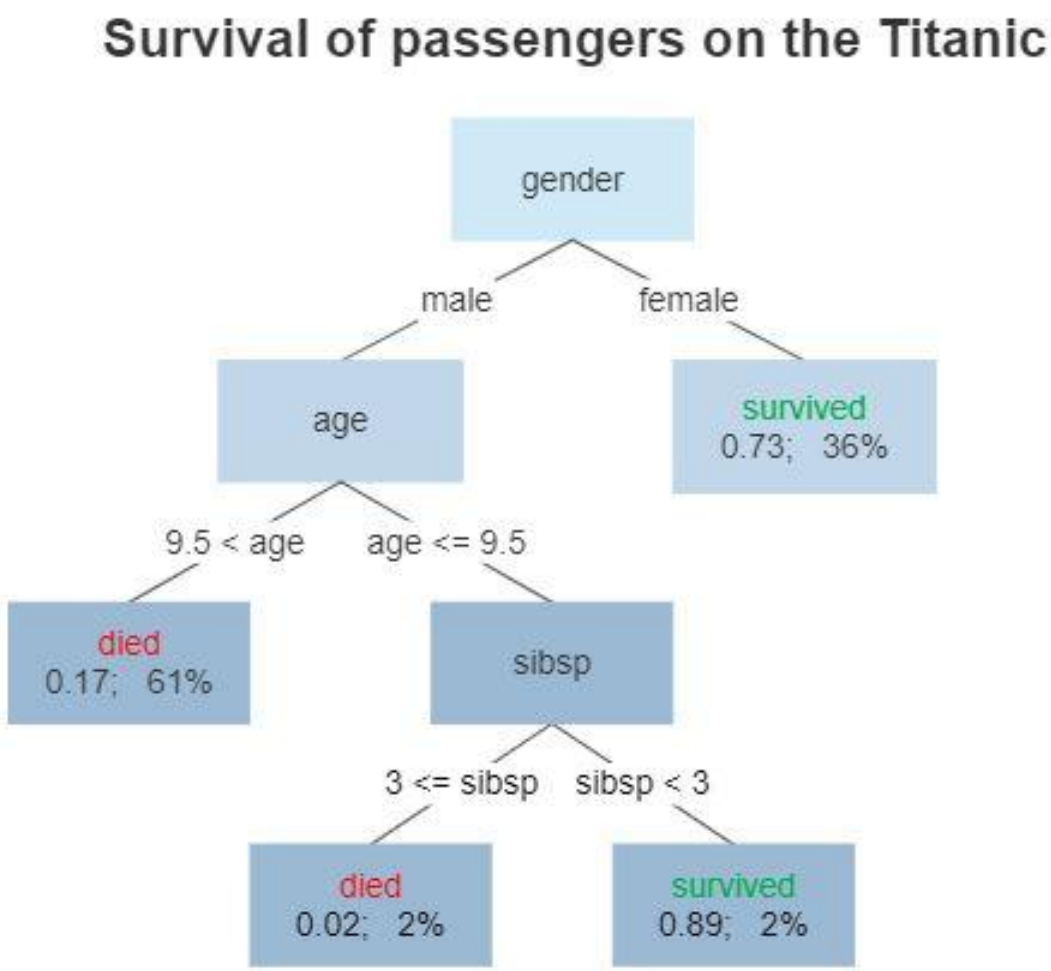


https://dudeperf3ct.github.io/lstm/gru/nlp/2019/01/28/Force-of-LSTM-and-GRU/

1. RANDOM FOREST

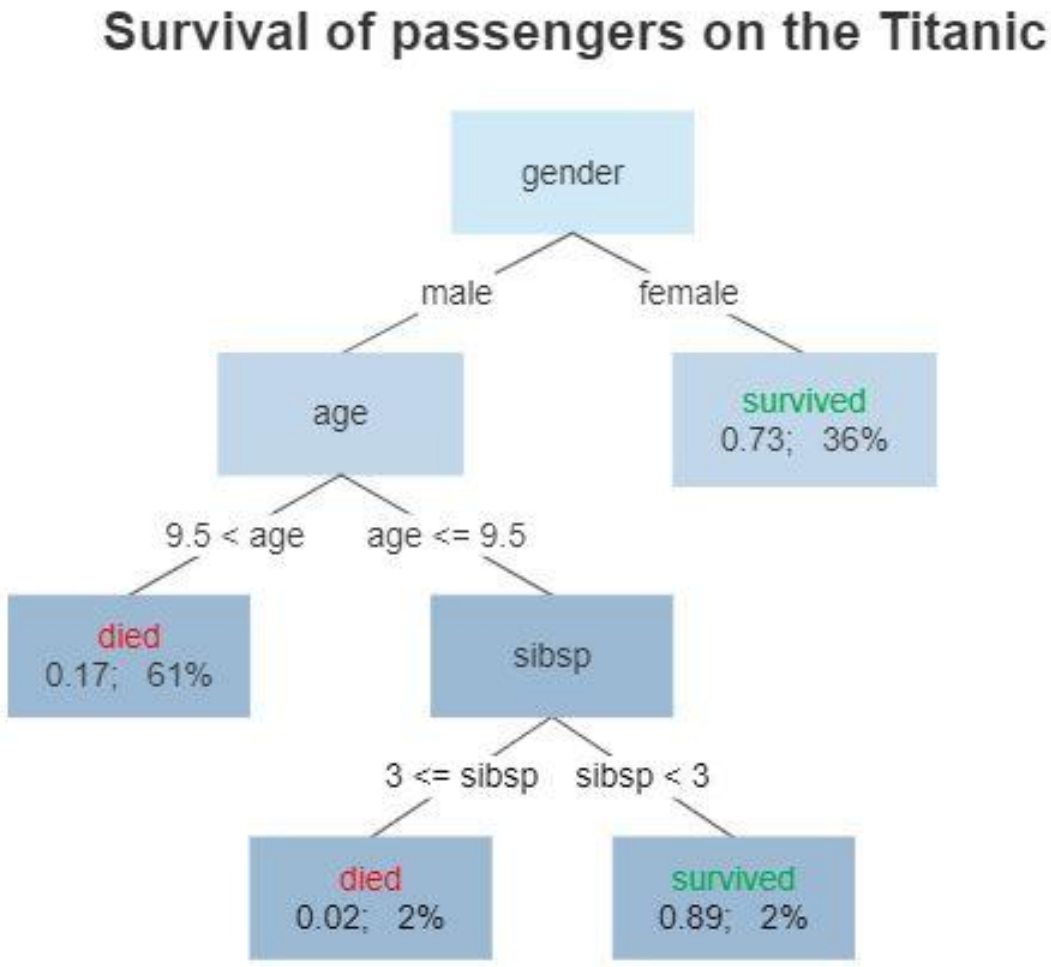
DECISION TREE

- Find split conditions that best separate data into categories/values
  - CART - Breiman et al. 1984
- Once constructed, data filters through tree, is assigned label at leaf node



DECISION TREE

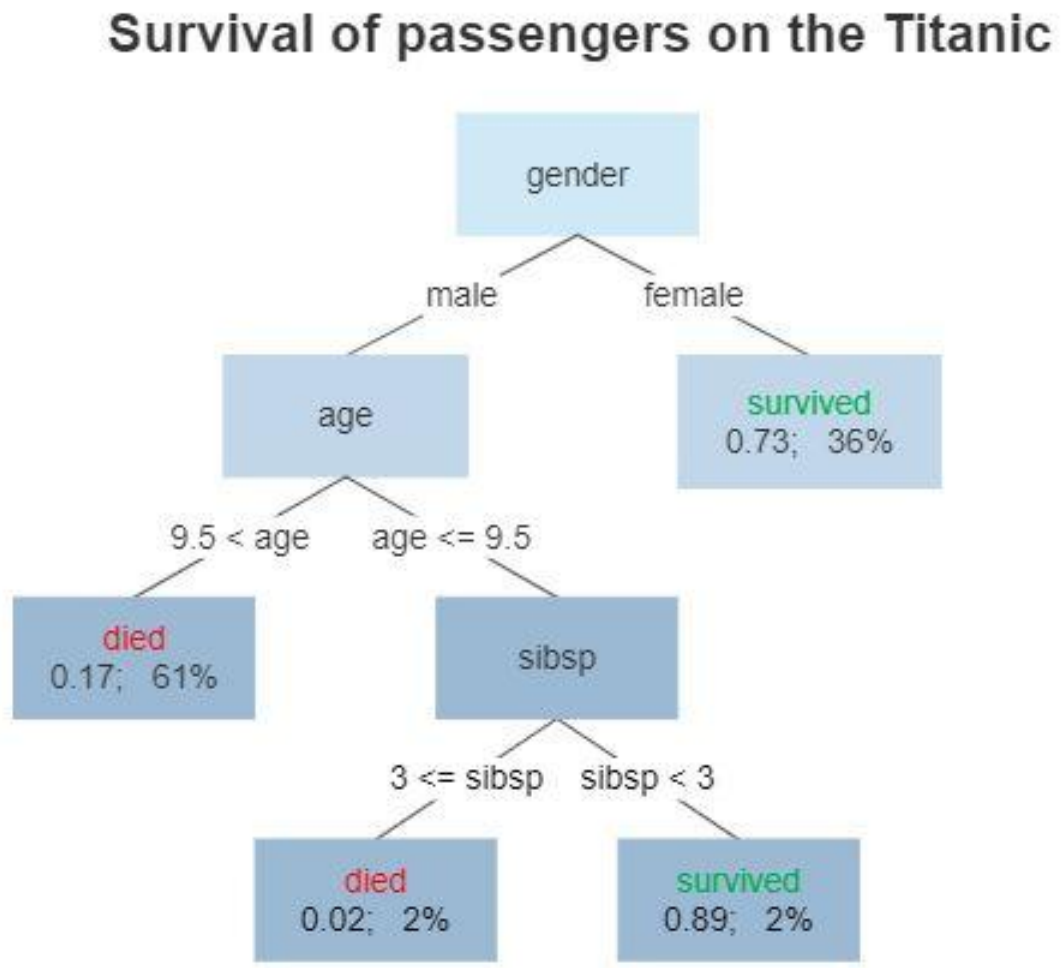
id	sex	age	siblings	survived
1	f	26	3	yes
2	f	52	0	yes
3	m	9	4	no



DECISION TREE

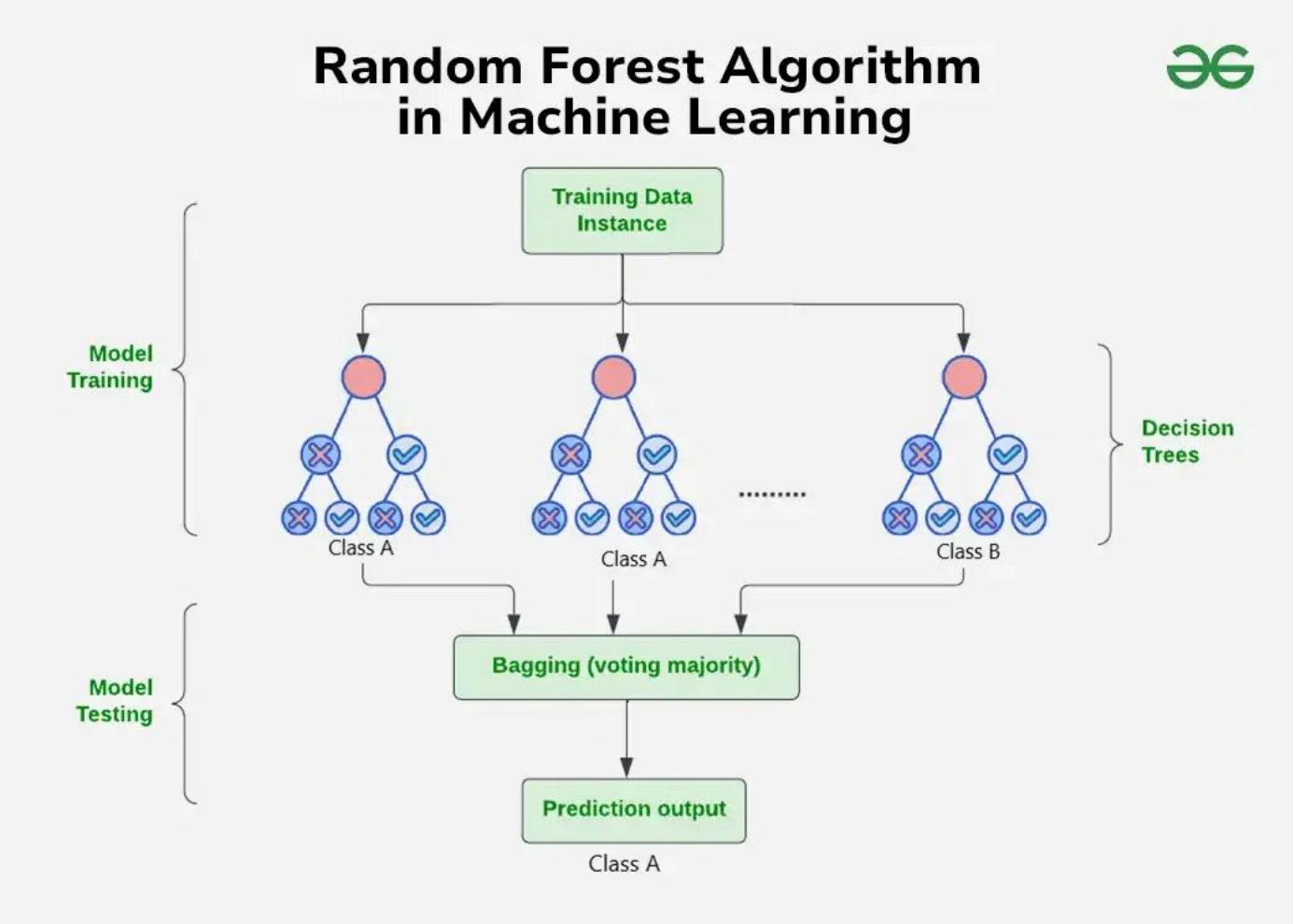
- Nodes with high purity are desirable
- Information gain or Gini coefficient often used to calculate best split
- Greedy, recursive search

- Probability of survival; percentage of total datapoints
- Decision process is interpretable



RANDOM FOREST

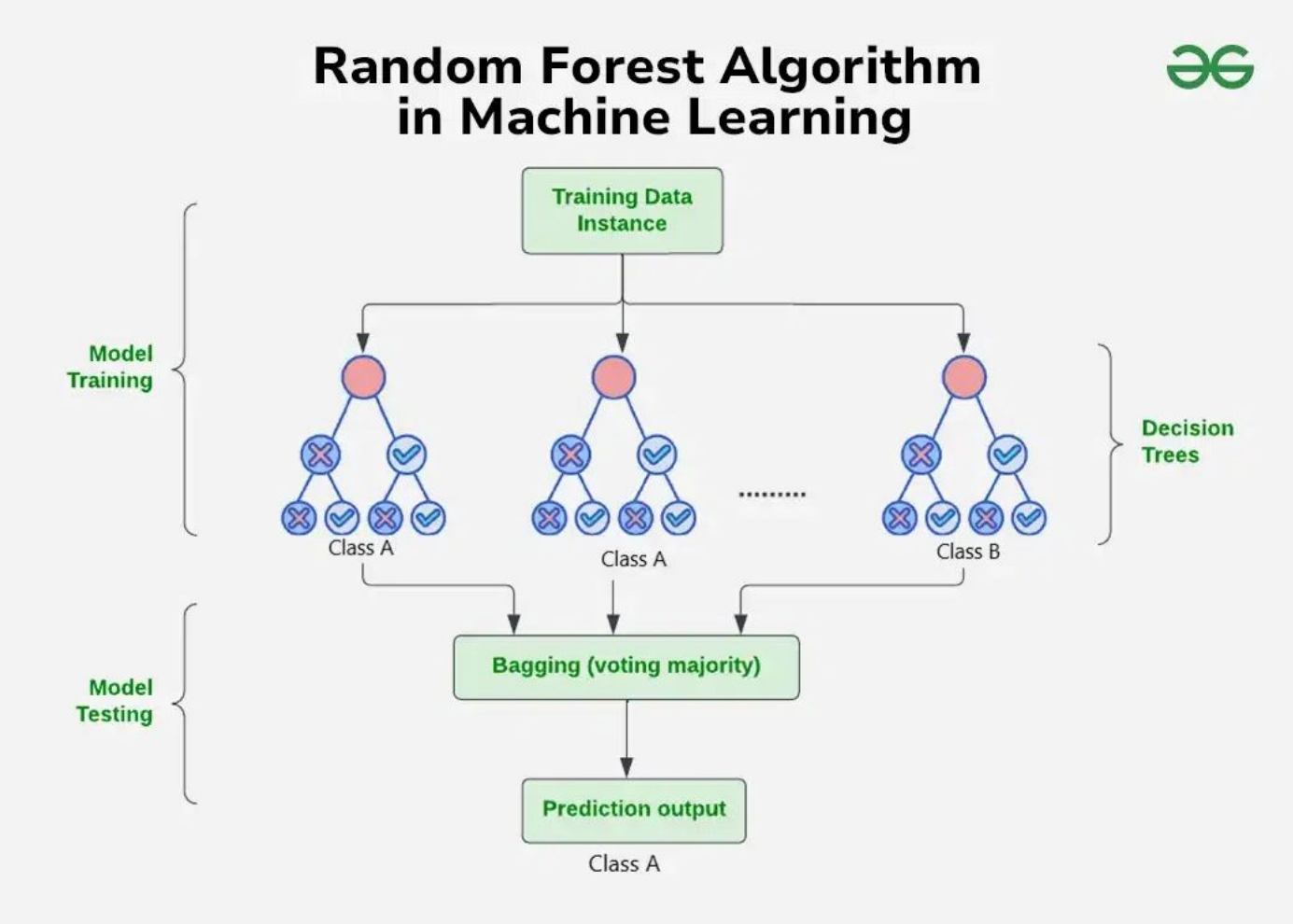
- Random forests are ensembles of decision trees
  - Salzberg and Heath, 1993
  - Ho, 1995
  - Breiman and Cutler, 2006
- Bagging
  - Different samples of data produce different trees
  - Majority voting
  - Reduces overfitting



https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/

RANDOM FOREST

- Can set number of trees, depth of trees
- Parallelizable
- Small loss of interpretability compared to individual decision tree
- Reduces overfitting (memorizing), which improves generalization

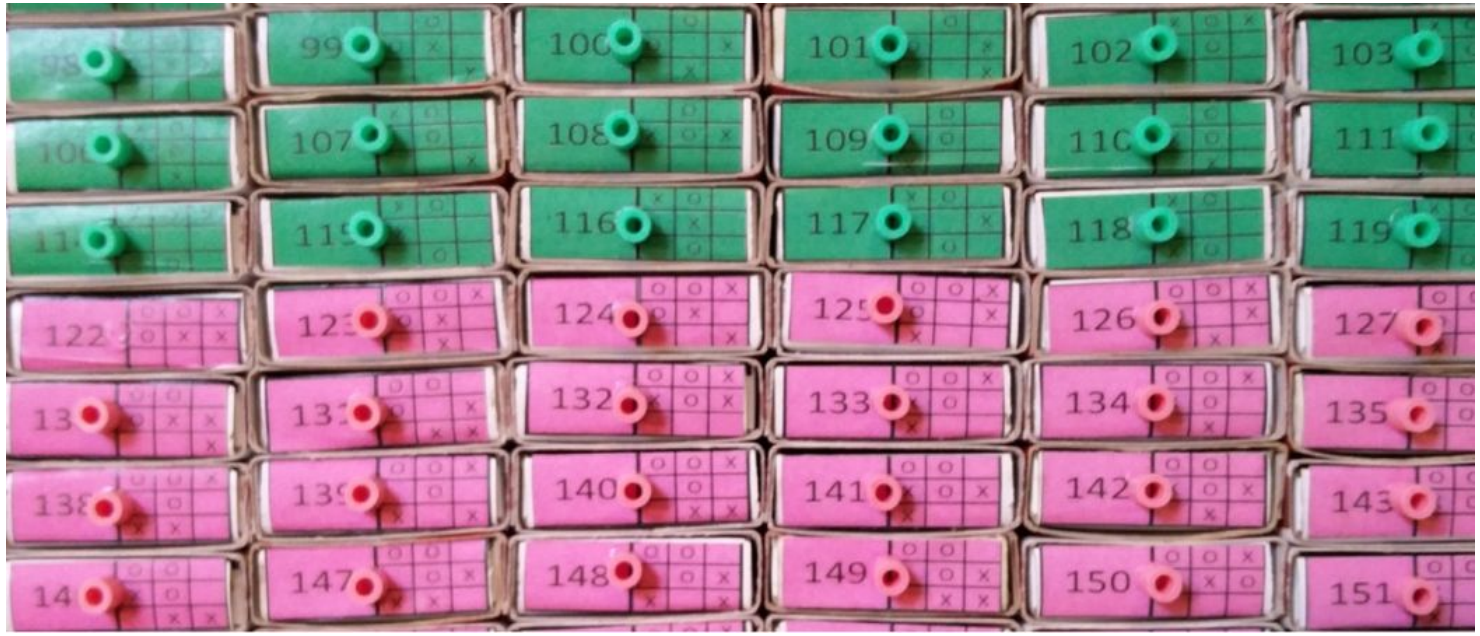


https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/

TIC-TAC-TOE INTERLUDE

INTUITION

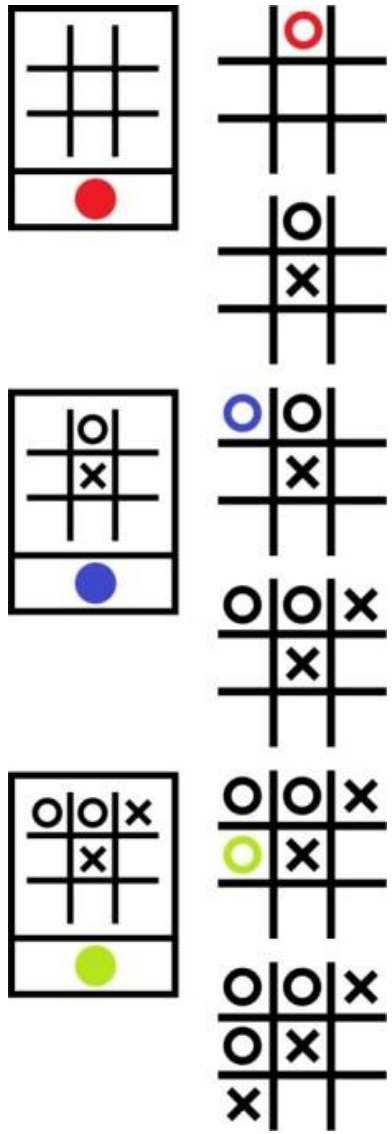
- MENACE - Matchbox Educable Noughts and Crosses Engine (Michie 1960)
- 304 matchboxes, each representation a game state
- 9 colored beads in each matchbox, each representing a different next move



https://chalkdustmagazine.com/features/menace-machine-educable-noughts-crosses-engine/

INTUITION

- Play games, sampling randomly to decide next move
- If win, add beads corresponding to moves sampled
- If lose, remove beads
- Reinforcement learning – reward desired behavior, punish undesired behavior, adjust system to reflect changes

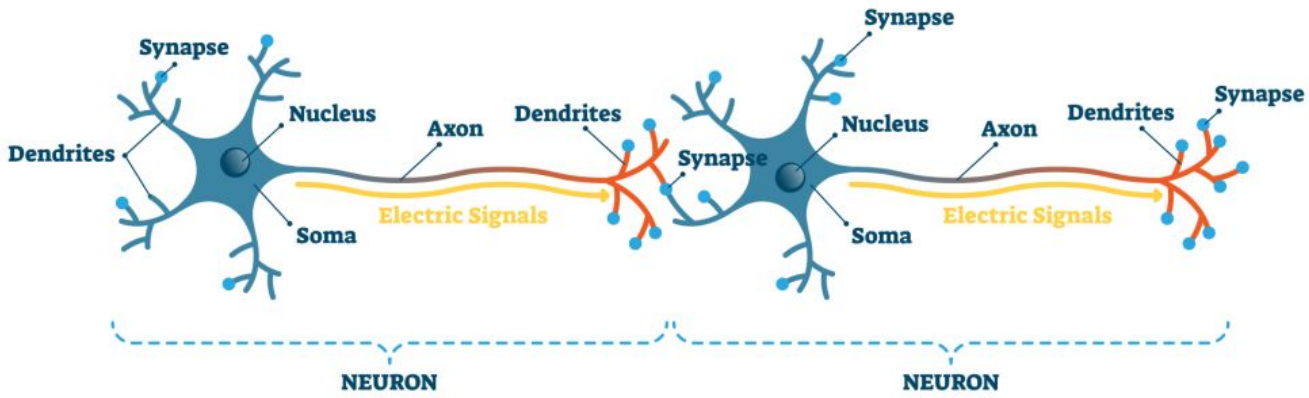


https://chalkdustmagazine.com/features/menace-machine-educable-noughts-crosses-engine/

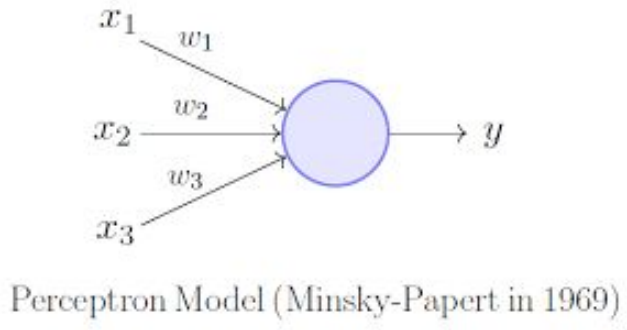
2. MULTILAYER PERCEPTRON

ARTIFICIAL NEURAL NETWORK (ANN)

- Inspired loosely by biological neuron cell
  - McCullock and Pitts, 1943
- Perceptron
  - Rosenblatt, 1957; Minsky-Papert, 1969
  - Expresses a linear equation
    - $x_1w_1+x_2w_2+x_3w_3 = y$
    - **Wx** = **y**
  - Could solve handwritten digit recognition!
  - Limited utility for more complex, non-linear data



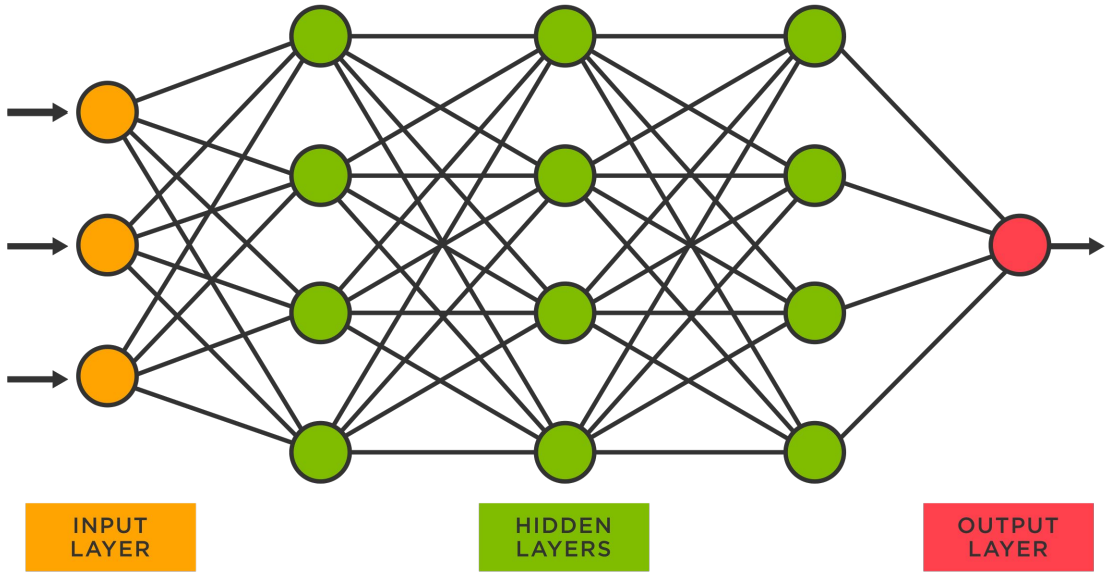
https://www.simplypsychology.org/neuron.html



Perceptron Model (Minsky-Papert in 1969)

MULTILAYER PERCEPTRON (MLP)

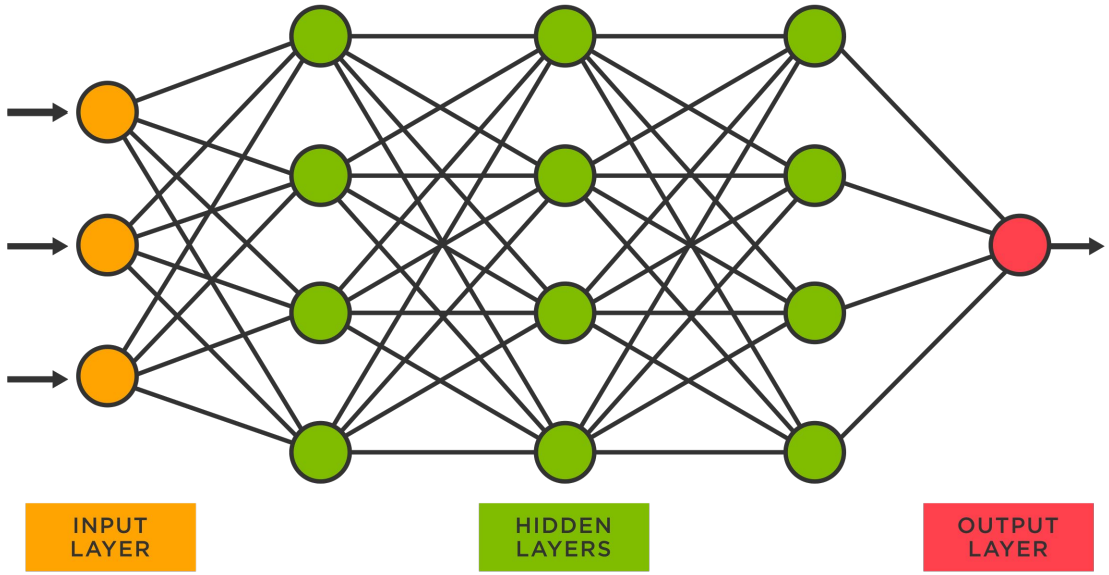
- Series of functions, transforming input to output – f(g(h(x)))
- Feedforward – unidirectional
- Addition of non-linear functions allows depth, greater expressivity
- Parameterized learning



https://www.tibco.com/reference-center/what-is-a-neural-network

MULTILAYER PERCEPTRON (MLP)

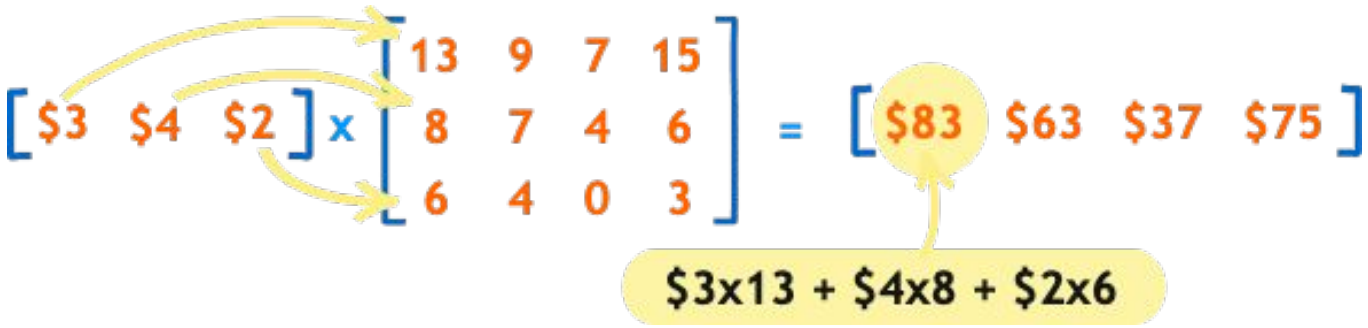
review id	Whimsical	Satirical	great	…	sentiment
1	1	1	1	…	1
2	0	1	0	…	0
3	0	0	0	…	0
4	0	0	2	…	1



https://www.tibco.com/reference-center/what-is-a-neural-network

MATRIX MULTIPLICATION

- Linear transformation
- **Wx = y**
- Can be interpreted algebraically or geometrically

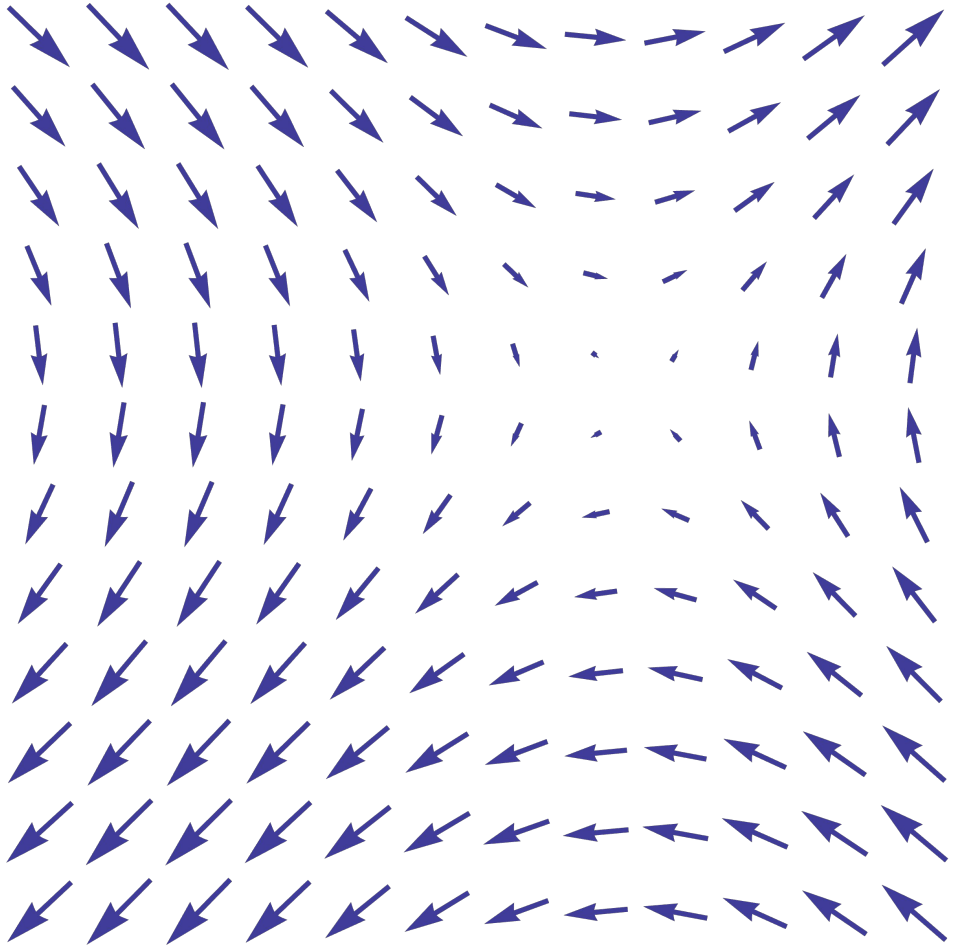


https://www.mathsisfun.com/algebra/matrix-multiplying.html

- Input is transformed into different space

MULTILAYER PERCEPTRON (MLP)

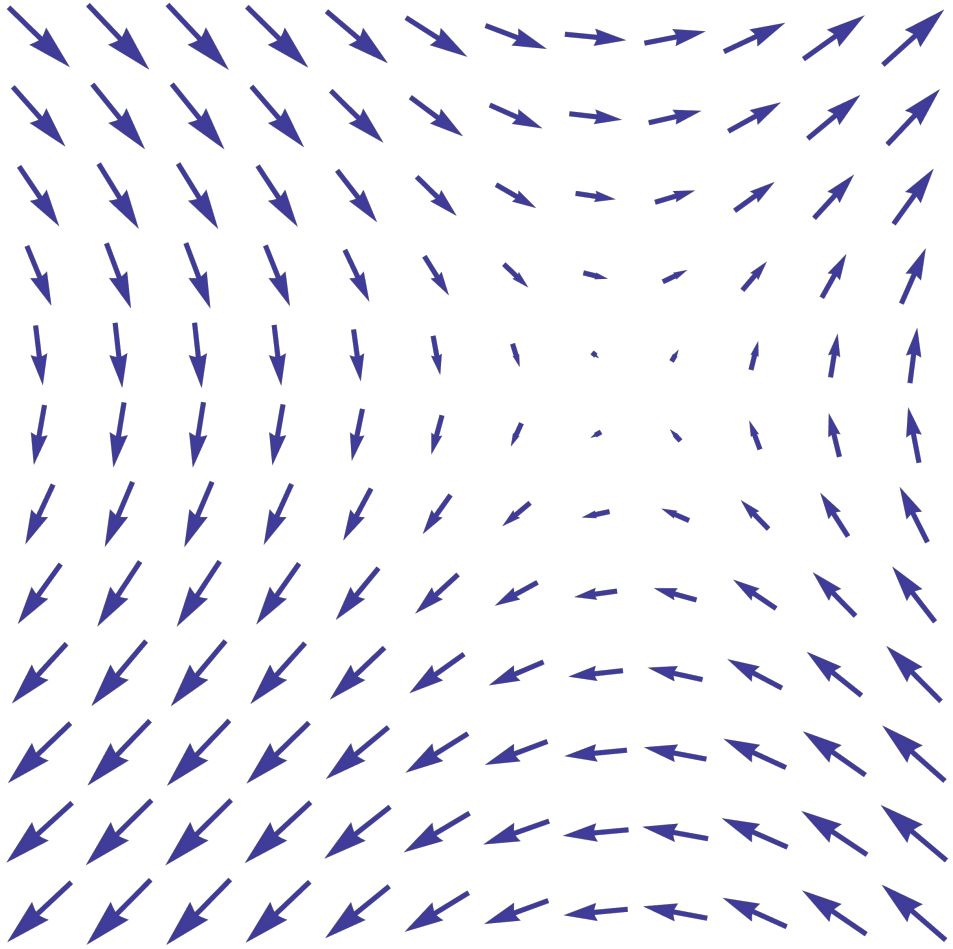
- Random weight initialization
- Error is back-propagated through network
  - Rumelhart, Hinton, Williams; 1986
  - Chain rule allows per-layer gradients
- Gradient Descent
  - Cauchy, 1847; Hadamard, 1907; Curry, 1944
  - Find parameters that minimize error
  - Define differentiable loss function (e.g. MSE)
  - Gradient is vector of partial derivatives



Jim Belk https://upload.wikimedia.org/wikipedia/commons/b/b9/VectorField.svg

MULTILAYER PERCEPTRON (MLP)

- Gradient Descent (cont.)
  - Each partial derivative represents the influence from a given parameter on the error
  - Update parameters to push them in opposite direction of gradient
  - More specific blame than RL
  - Not guaranteed to find global optimum



Jim Belk https://upload.wikimedia.org/wikipedia/commons/b/b9/VectorField.svg

REFLECTIONS

- Neural networks **more difficult to interpret**
  - Decision trees can be read off
  - MLPs are learning latent representations
- Neural networks **can learn complex patterns**
  - E.g., tic-tac-toe boards vs. variable length text
- Neural networks **require more resources** (training data and processing power), but can potentially perform better

Hackathon Hands-On

1. Sign-up for OSC account - my.osc.edu
2. Request project access

ACCOUNT SIGN-UP

Sign-up link https://my.osc.edu/acprod/odb_osc/r/osc/portal/signup?clear=202

Project code:
PZS1142

Access number:
624239

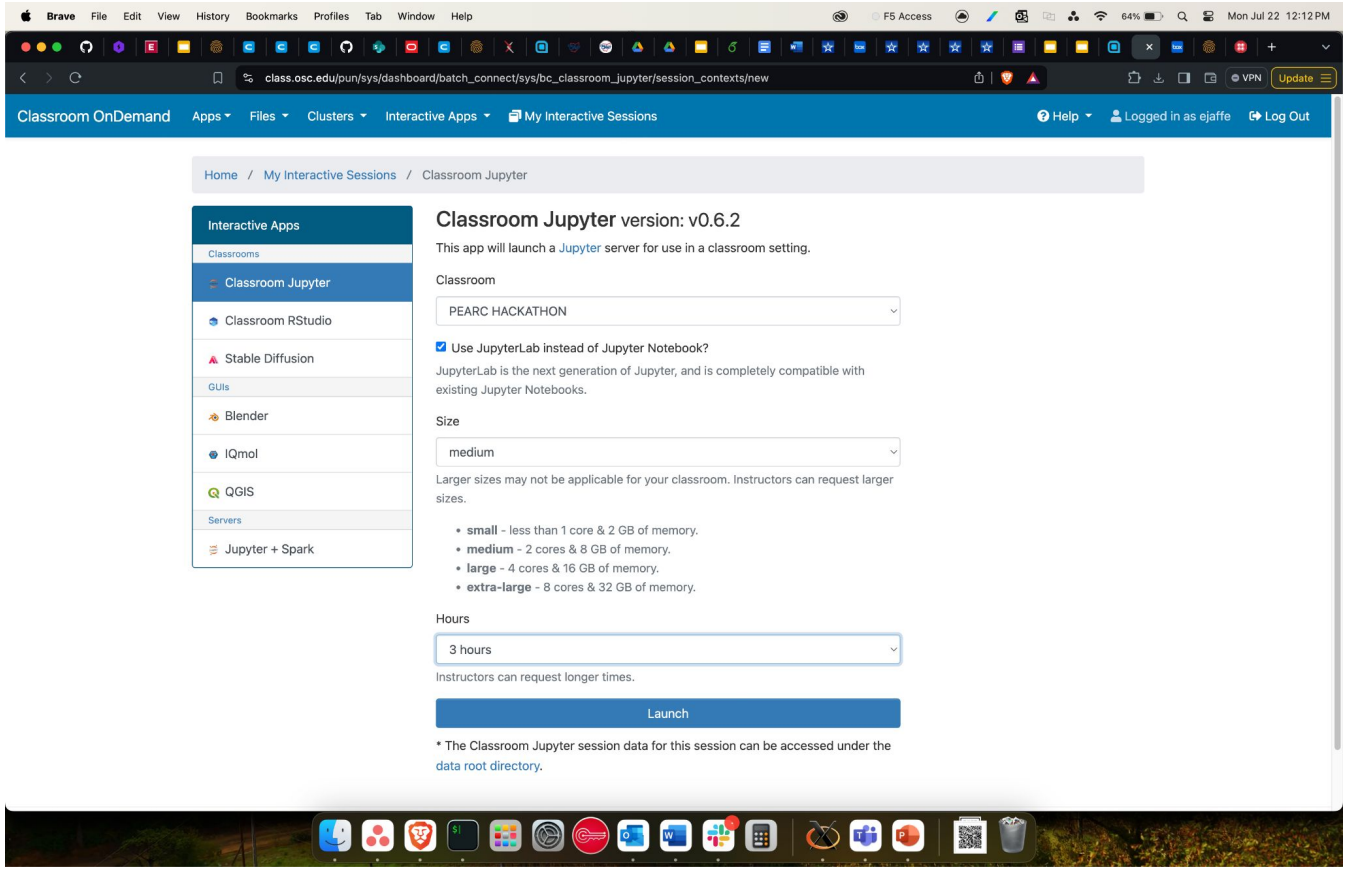
Expires:
31 July 2024

Account sign-up instructions
https://www.osc.edu/supercomputing/portals/client_portal/self_signup_for_accounts

Ohio Supercomputer Center

Hackathon Hands-on - Run the lab

1. Navigate to **class.osc.edu**
2. Click Classroom Jupyter
3. Select PEARC_HACKATHON
4. Click Launch
5. Run pearc_hackathon.ipynb



Hackathon Instructions

Extend the Jupyter notebook to perform the sentiment analysis task using your own features, model, etc. Get creative!

For consistency across teams, please use a 20% random split for test data - honor system! (no cross-validation, test only once, etc.)

Remaining 80% can be used for training and validation/dev.

Final submission link:

Prizes!

3K Travel CCEP for Most Creative Solution (Advanced Category)

3K Travel CCEP for Highest Accuracy (Advanced Category)

3K Travel CCEP for Most Probable Solution (Experiential Category)

Everyone who submits a complete form to the Experiential Category will be entered in a random drawing for a Gift Card to Amazon.

Everyone who attends the full Intro to AI, Hackathon, AND takes the after-event survey (which will come in a week or so after PEARC) will be part of a random drawing for a Gift Card to Amazon. No final submission is required.

THANK YOU

OSC.EDU

CWRU

NSF Award #: 2320952

University of CINCINNATI

NSF Award #: 2320953

Ohio Department of Higher Education

Ohio Supercomputer Center

NSF Award #: 2320954

Ohio Supercomputer Center

REFERENCES

Breiman, L. et al. “Classification and Regression Trees.” *Biometrics* 40 (1984): 874.

Breiman, Leo, and Adelle Cutler. "Random Forests.” *Berkeley*. (2005).

Cauchy, Augustin. "Méthode générale pour la résolution des systemes d’équations simultanées." *Comp. Rend. Sci. Paris* 25.1847 (1847): 536-538.

Curry, H.B. “The method of steepest descent for nonlinear minimization problems.” *The Quarterly Journal of Mechanics and Applied Mathematics* 2 (1944) 258-261.

Hadamard, Jacques. "Mémoire sur le problème d'analyse relatif à l'équilibre des plaques élastiques encastrées". *Mémoires présentés par divers savants éstrangers à l'Académie des Sciences de l'Institut de France*. (1908): 33.

Ho, T.K. Random Decision Forest. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, 14-16 August (1995): 278-282.

Maas, Andrew L. and Daly, Raymond E. and Pham, Peter T. and Huang, Dan and Ng, Andrew Y. and Potts, Christopher. “Learning Word Vectors for Semantic Analysis.” *The 49th Annual Meeting of the Association for Computational Linguistics* (ACL 2011).

REFERENCES

McCulloch, Warren S., and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity." *The bulletin of mathematical biophysics* 5 (1943): 115-133.

Michie, Donald. “Experiments on the Mechanization of Game-Learning Part I. Characterization of the Model and its parameters.” *The Computer Journal*, Volume 6, Issue 3, November (1963): 232-236, https://doi.org/10.1093/comjnl/6.3.232

Minsky, Marvin, and Seymour A. Papert. "Perceptrons, reissue of the 1988 expanded edition with a new foreword by Léon Bottou: an introduction to computational geometry.” *MIT Press*, (2017).

Rosenblatt, Frank. "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological Review* 65.6 (1958): 386.

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." *Nature* 323.6088 (1986): 533-536.

Heath, David, Simon Kasif, and Steven Salzberg. "Committees of decision trees." *Advances in Psychology*. Vol. 113. North-Holland, (1996): 305-317.

BRAND COLORS

HEX #0E3F75

HEX #C12637

HEX #FFFFFF

Example uses:

- Bar charts
- Highlight text
- Subheadings
- Icons

  Feel free to use the Microsoft stock icon library

HEX #0098D3

HEX #729364

HEX #EBA70E

HEX #69C2C6

HEX #BBD36F

HEX #F3DF89

HEX #8A5E32

HEX #B0B3AF

HEX #000000