CS-349: Machine Learning

Fall 2024

Introduction

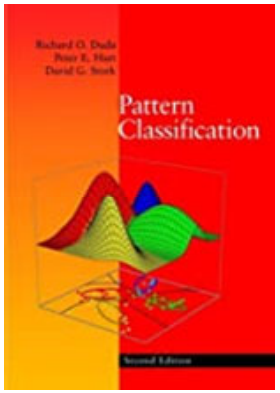
Course Overview: Introduction

Course Instructor:	David Demeter (Assistant Professor of Instruction)
Teaching Assistant:	Jacob John and Karan Garkel
Meeting Times:	Tuesdays/Thursdays 12:30PM to 1:50PM
Location:	Frances Searle, Room # 2107
Class Title:	Machine Learning
Format:	Bi-weekly lectures Four group homework assignments Final group project In-person attendance strongly encouraged Interactive classes preferred
Prerequisites:	Intermediate Python Programming (recursion, data structures, etc.) Basic familiarity with linear algebra, statistics and derivatives

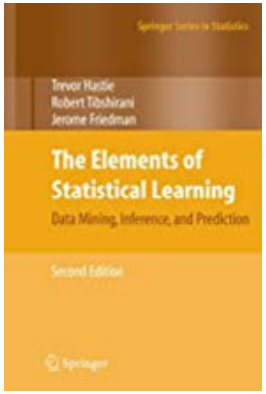
Course Overview: Resources

Text Books: None required, some recommended

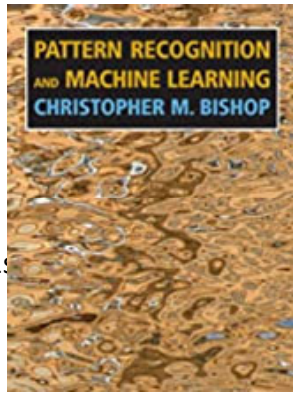
Richard O. Duda, Peter E. Hart, David G. Stock:
Pattern Classification, 2nd edition,
John Wiley & Sons, New York, 2001



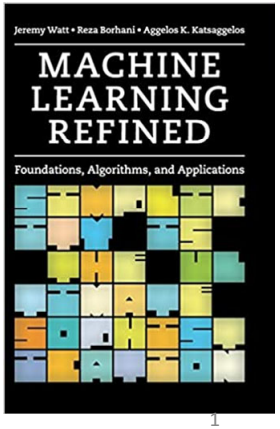
Trevor Hastie, Robert Tobshirani, Jerome Friedman:
The Elements of Statistical Learning –
Data Mining, Inference, and Prediction,
2nd edition, Springer, New York, 2009
http://www-stat.stanford.edu/~tibs/ElemStatLearn/



Christopher M. Bishop:
Pattern Recognition and Machine Learning,
Springer, New York, 2006



Jeremy Watt , Reza Borhani, Aggelos K. Katsaggelos
Machine Learning Refined: Foundations, Algorithms, and Applications 1



Resources: Machines in Wilkinson Lab have GPUs and students have access
(Note: Students may be also responsible for arranging their own GPU resources through AWS, GCP, **Colab**, etc.)

Course Overview: Grading

Deliverables:	HW #1: Decision Trees	10 pts
	HW #2: K-Nearest Neighbors and K-Means Clustering	10 pts
	HW #3: Neural Networks	10 pts
	HW#4: Fairness and Ethics	10 pts
	Final Exam	20 pts
	Final Project: Proposal	5 pts
	Final Project: Preliminary Results	5 pts
	Final Project: Presentation	20 pts
	Miscellaneous	10 pts

Standard scale: 93-100 = A, 90-93 = A-,

87-90 = B+, 83-87 = B, 80-83 = B-,

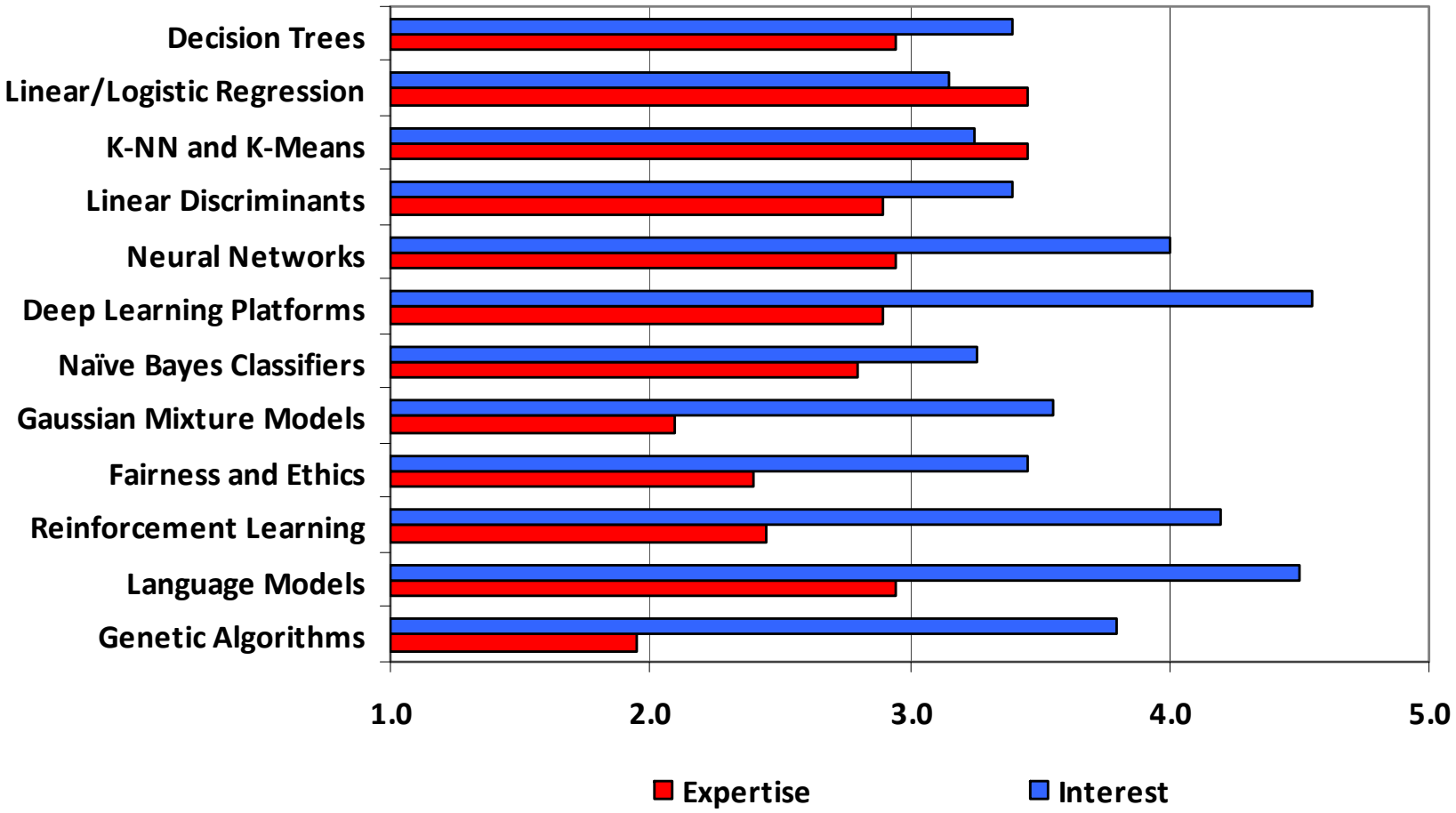
77-80 = C+, 73-77 = C, 70-73 = C-,

60-70 = D, and less than 60 = F

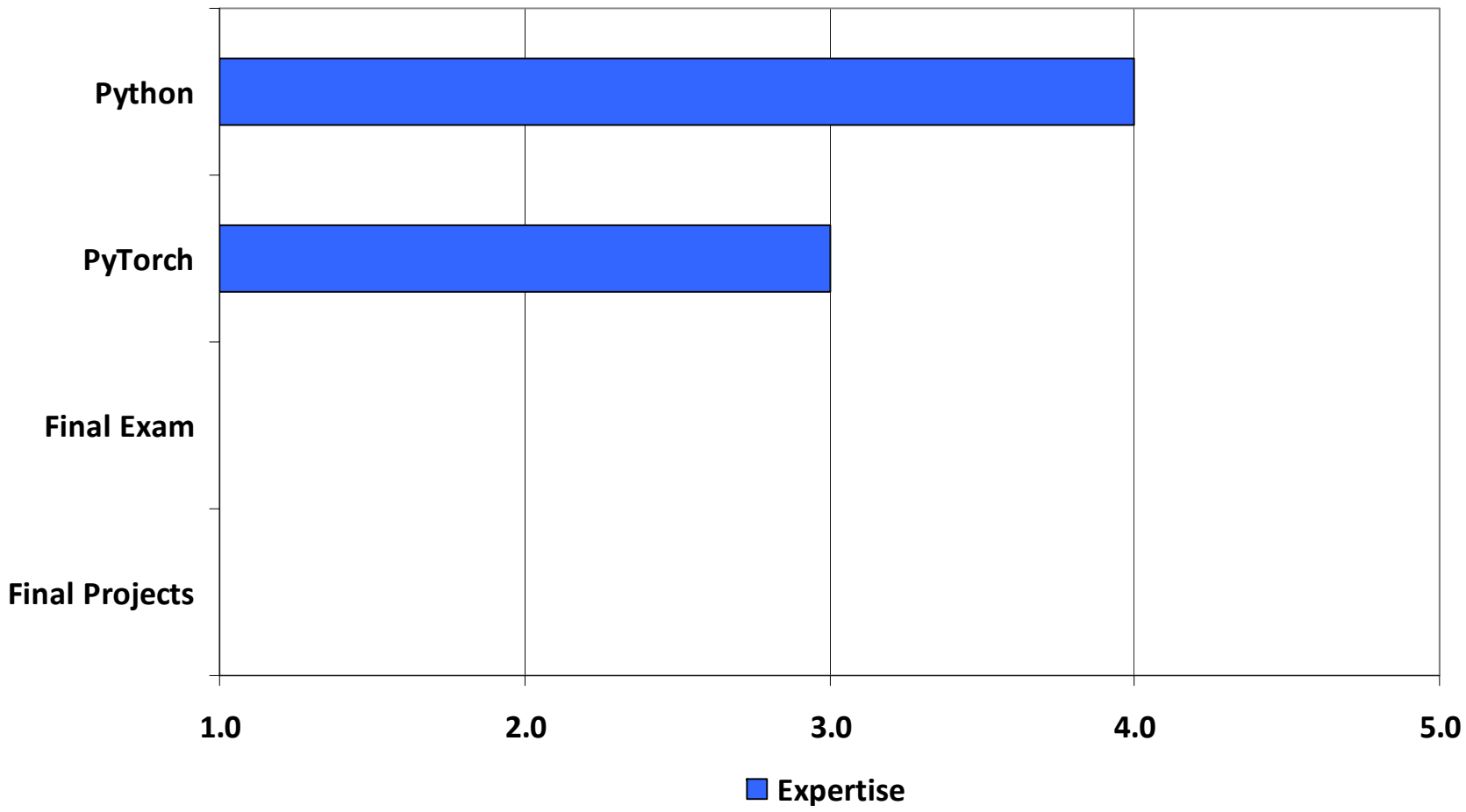
Extra Credit:

- To be determined

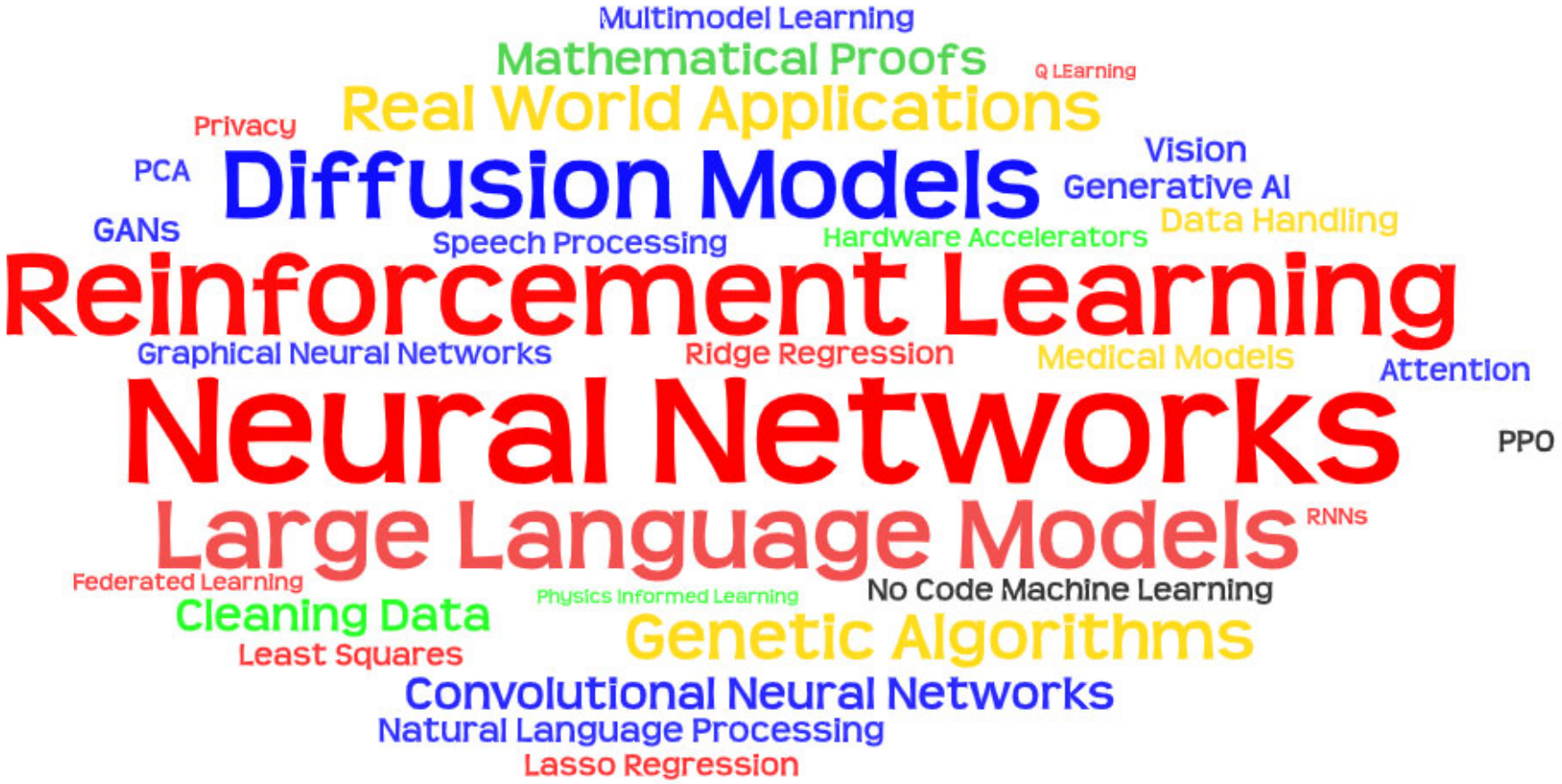
Survey Results - Interests and Expertise



Survey Results - Class Structure



Survey Results - Objectives



Survey Results - Favorite Quotes

Course Overview: Schedule

Week #1: Introduction

- Introductory remarks and class policies
- Machine learning in a nutshell

Week #2: Decision Trees

- Attribute-based representations and decision trees
- Learning decision trees with entropy and information gain
- ID3 algorithm, how to avoid over-fitting and pruning
- Random forests and regression

Week #3: Machine Learning Overview, Nearest Neighbors and K-Means Clustering

- Final project objectives and recommendations
- Overview of machine learning techniques
- Distance metrics
- K-nearest neighbors algorithm
- K-means clustering and collaborative filtering

HW #1: Decision Trees Due

Week #4: Regression, Linear Discriminants and Perceptrons

- Linear and polynomial regression
- Model complexity
- Linear discriminants and decision boundaries
- Perceptron algorithm and logistic regression

Final Project: Proposals Due

Course Overview: Schedule (*Cont.*)

Week #5: Neural Networks I

- Gradient descent, objective functions and regularization
- Feed forward neural network architecture
- Linear transformations, activation functions and output layers

HW #2: KNN and K-Means Due

Week #6: Neural Networks II

- Back propagation and differentiation rules
- Detailed walk-thru of a feed forward neural network
- Calculation graphs and parallel computation
- Deep learning platforms

Week #7: Hypothesis Testing and Gaussian Mixture Models

- Error estimates, normal distribution, t-tests, cross validation, evaluation metrics
- Parametric distribution, EM algorithm, multi-variate distributions
- Gaussian mixture models

Week #8: Language Modeling

- Introduction of language models and natural language processing tasks
- Embedding spaces and next token prediction
- Feed forward neural network language model (eg: Benio Model)
- Decoding and text generation

HW #3: Neural Networks Due

Course Overview: Schedule (*Cont.*)

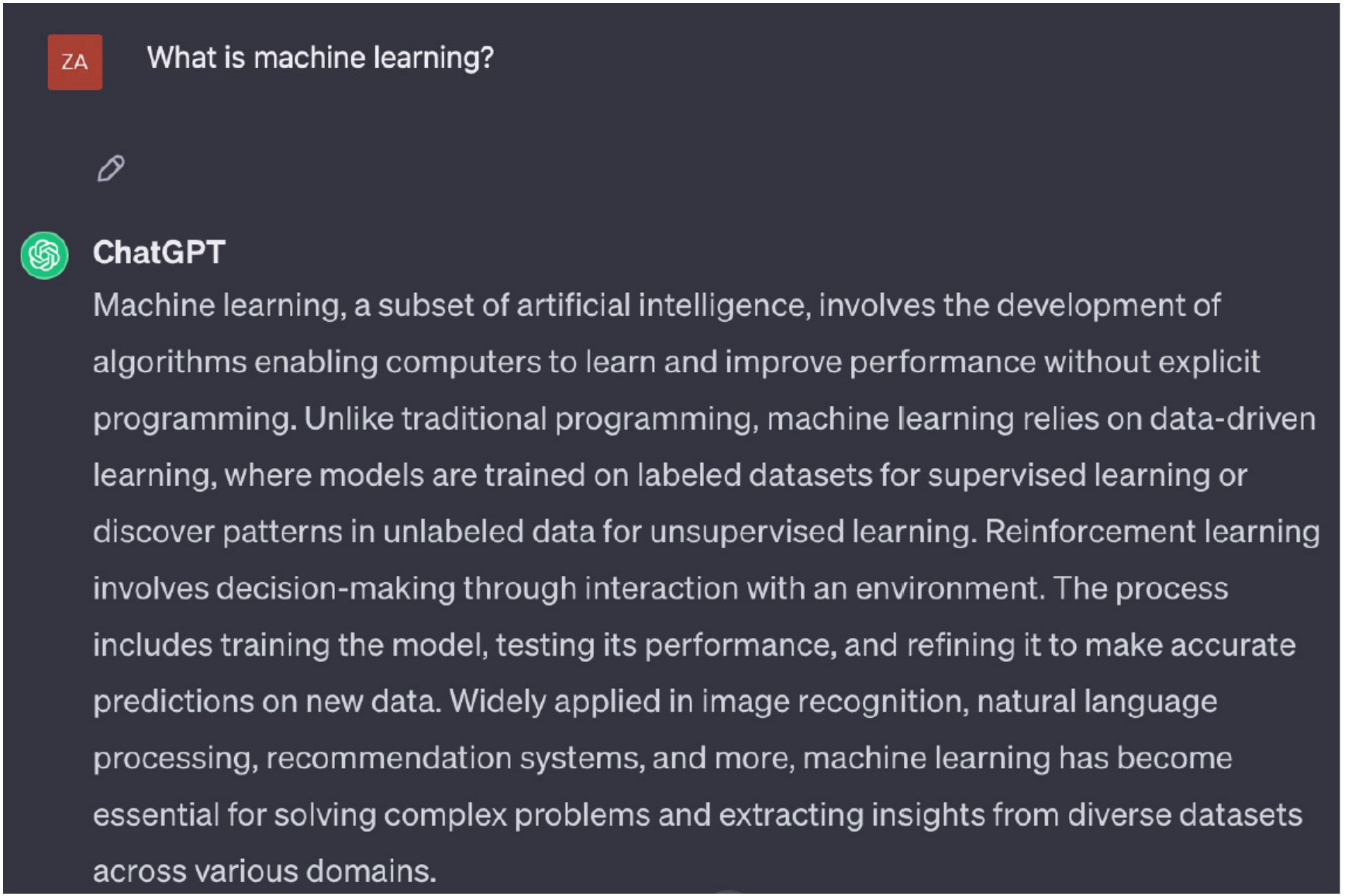
Week #9: Fairness in Machine Learning and Miscellaneous Topics
• Sources and types of bias in datasets, impact on model training, metrics to measure bias • Techniques to identify and address bias • Recurrent architectures, GANs, convolution neural networks • Support vector machines, genetic algorithms, active learning boosting **Final Project: Preliminary Results Due**
Week #10: Catch-up and Thanksgiving Break
• Course review • In-person final exam
Week #11: Review and Final Exam
• To be determined
Finals Week:
HW #4: Fairness and Ethics Due **Final Project: Presentations Due**

Course Overview: Policies

Video Recordings:	Video lectures will be posted to Canvas (for now) May not be shared outside of the class May not be used in lieu of in-person attendance (per University policy)
Communication:	Please use NorthwesternCS349@gmail.com
Groups:	Statics for the quarter. More details to follow.
Academic Integrity:	Your code and answers must be your own work. We will use a combination of automated and manual methods for comparing your code and free-response answers. Violations will be reported.
AccessibleNU:	Please make me aware of any accommodations in a timely manner
COVID-19:	Please be aware of and follow University policies
Positive Environment:	Please make me aware of anything that creates an uncomfortable environment.
Other:	Please see the course syllabus for a longer list of University policies.

Questions?

What is Machine Learning?



What is Machine Learning?

- "A set of methods that can automatically detect patterns in data, then use the uncovered patterns to predict future data."
- "Using experience to gain expertise."
- "Fitting a function to data."

Source: Murphy, *Machine Learning: A Probabilistic Perspective*
Shalev-Shwartz and Ben-David, *Understanding Machine Learning*

What is Machine Learning?

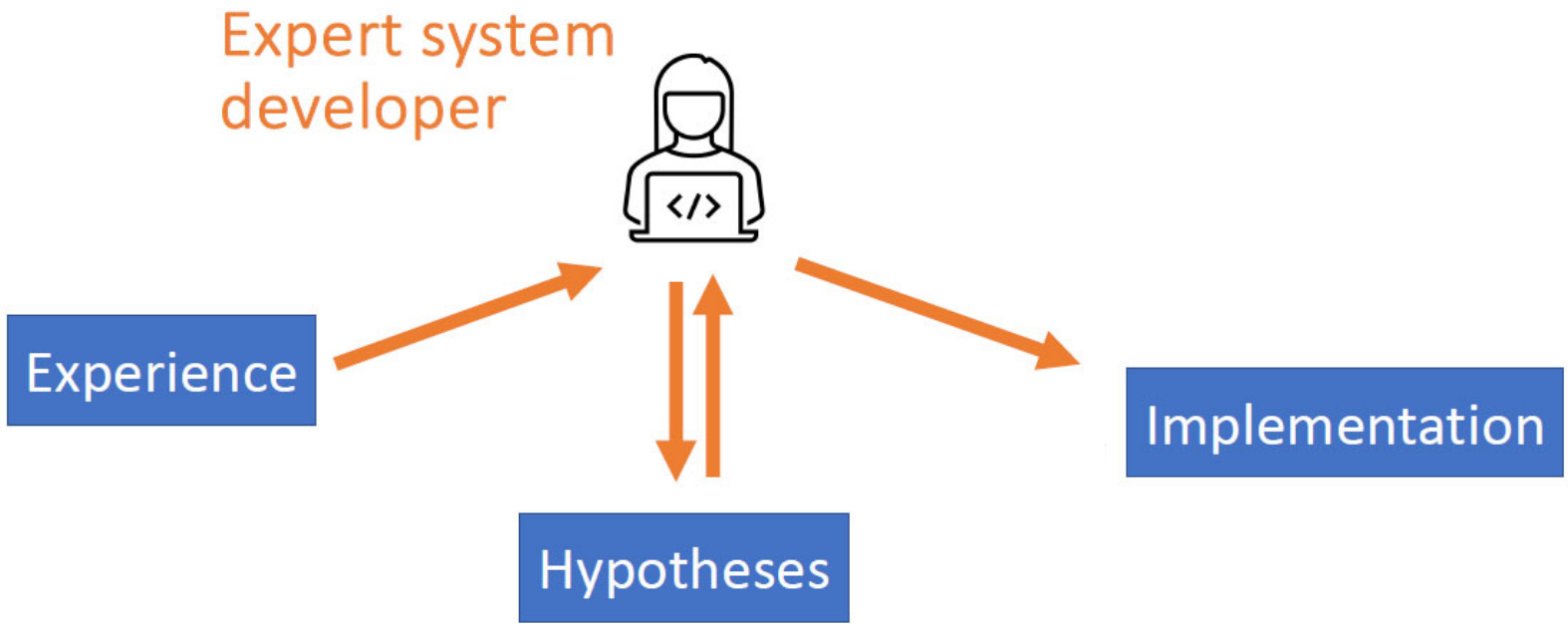
Is Machine Learning different from an Expert System?

- "Expert system" for predicting your grade:

```
if points >= 93.0:
  return "A"
elif points >= 90.0:
  return "A-"
elif points >= 87.0:
  return "B+"
elif points >= 83.0:
  return "B"
...
```

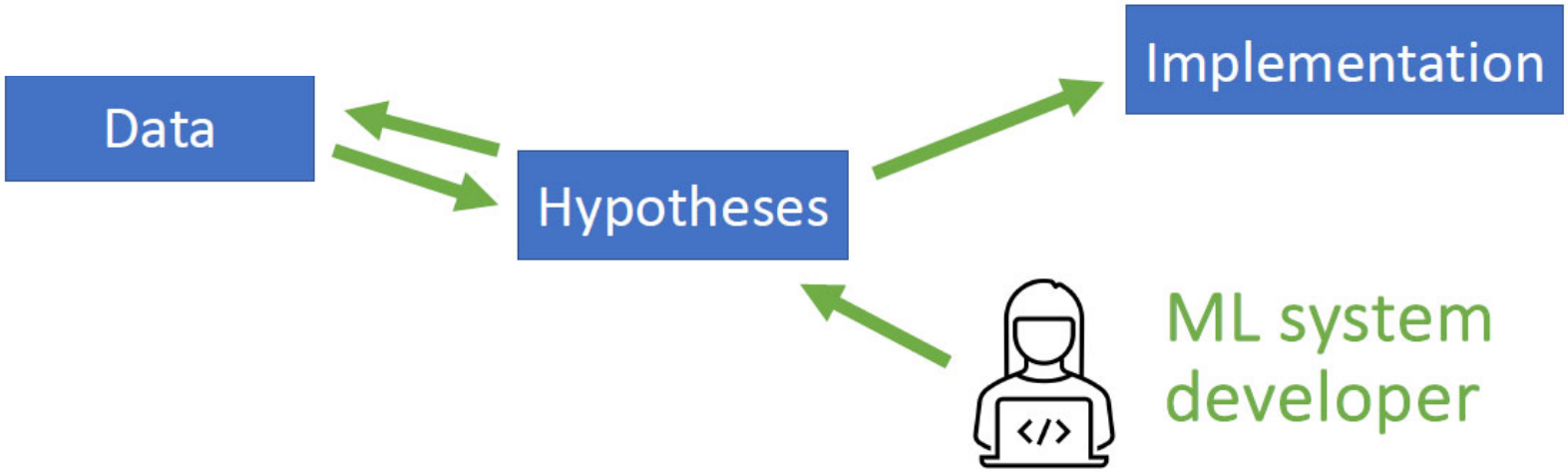
What is Machine Learning?

Is Machine Learning different from an Expert System?



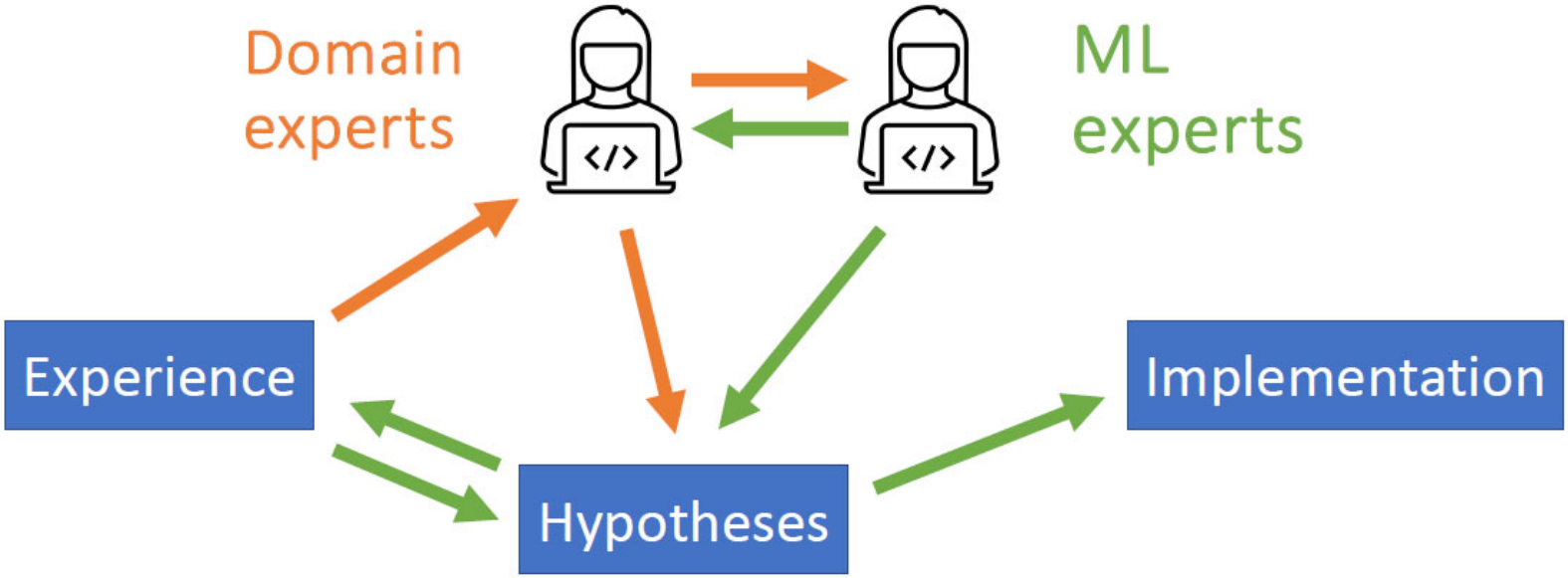
What is Machine Learning?

Is Machine Learning different from an Expert System?

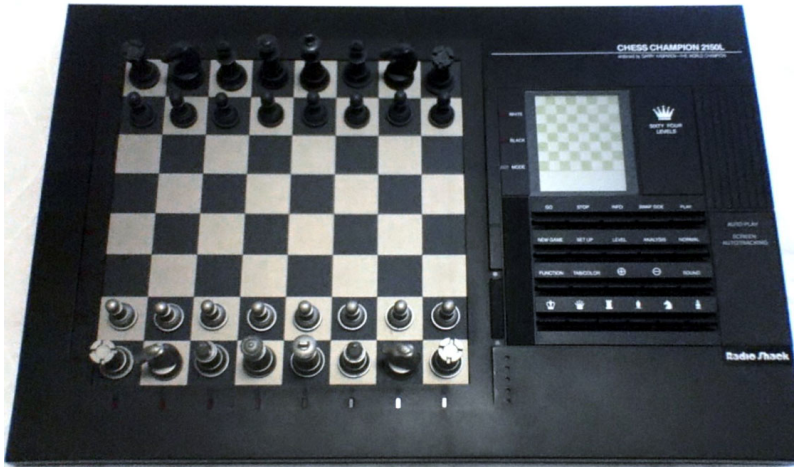


What is Machine Learning?

Is Machine Learning different from an Expert System?

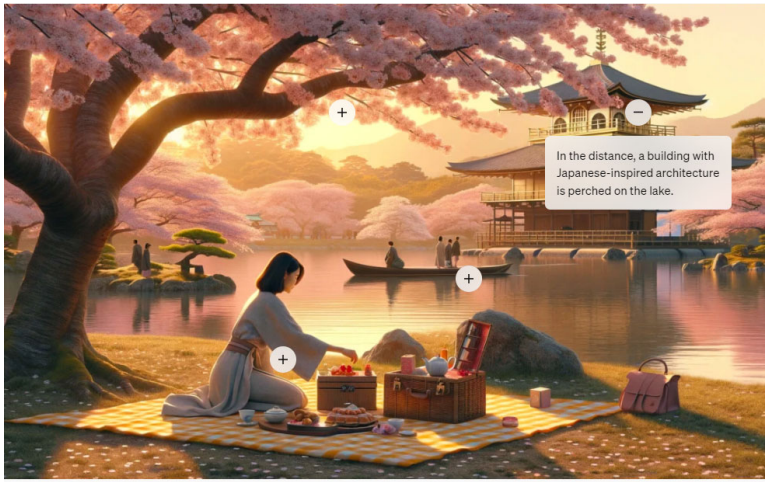


What Can Machine Learning Do?



Computer Chess

Computer Go



DALL·E 3

Self-Driving Car

What Can Machine Learning Do?

- Recognizing patterns:
  - Facial identities or facial expressions
  - Handwritten or spoken words, sentiment
  - Medical images
- Recognizing anomalies:
  - Unusual credit card transactions
  - Unusual patterns of sensor readings in a nuclear power plant
- Prediction:
  - Next word in a sequence of natural language text
  - Future stock prices or currency exchange rates
- Generating patterns
  - Generating text, images or audio
  - Question answering (factual recall, reasoning, planning)?

Types of Learning

Supervised (inductive) learning

- training data + desired outputs (labels)

Unsupervised learning

- training data (without desired outputs)

Semi-supervised learning

- training data + a few desired outputs

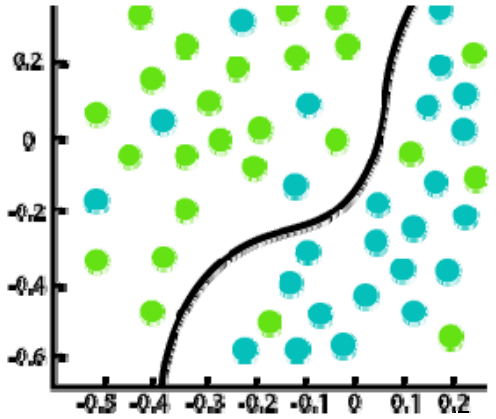
Reinforcement learning

- rewards from sequence of actions

Types of Learning

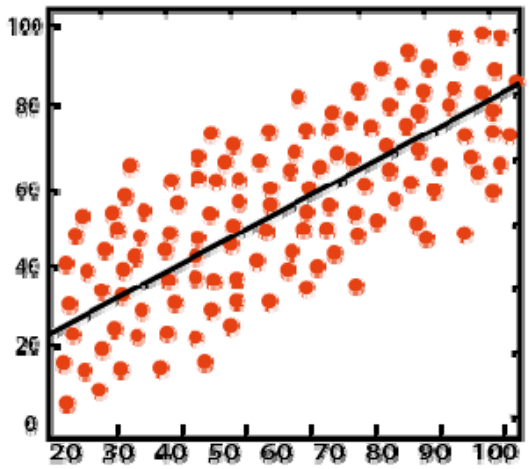
- Classification:

  Learning a function to map from a n-tuple to a ***discrete*** value from a finite set



- Regression:

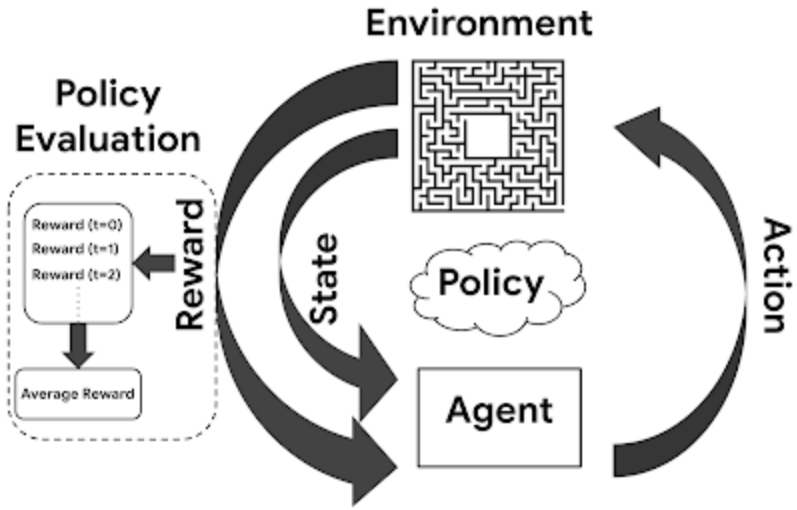
  Learning a function to map from a n-tuple to a ***continuous*** value



Types of Learning

- Reinforcement Learning:

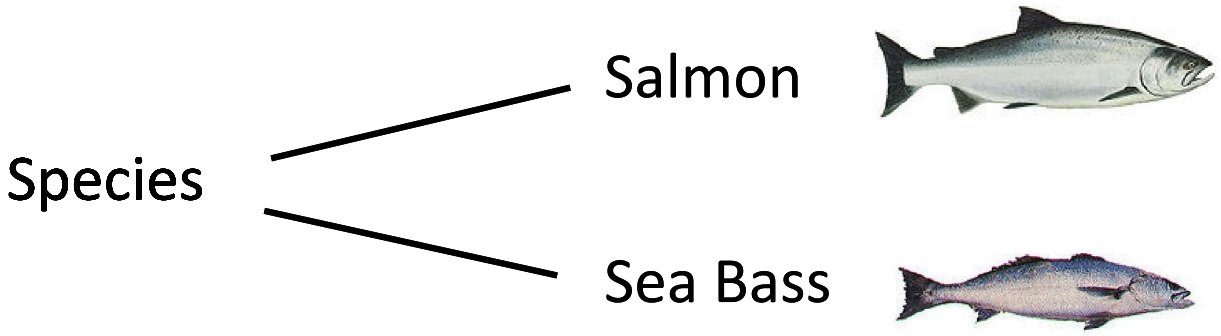
  Learning a ***policy*** to maximize a reward from an agent interacting with its environment through actions and state transitions



Source: Google Research

Fish Example: Task

"Sorting incoming fish on a conveyor according to species using optical sensing"



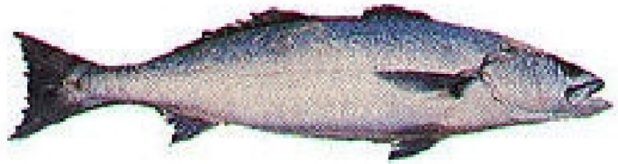
Risk: Finding sea bass in a package of salmon annoys customers!

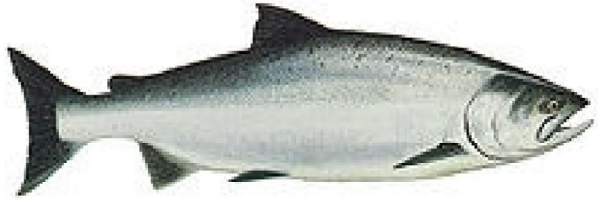
Fish Example: Problem Analysis

Set up a camera and take some sample images

Extract characteristics that make distinction between species possible

- Length
- Lightness
- Width
- Number and shape of fins
- Position of the mouth, etc.

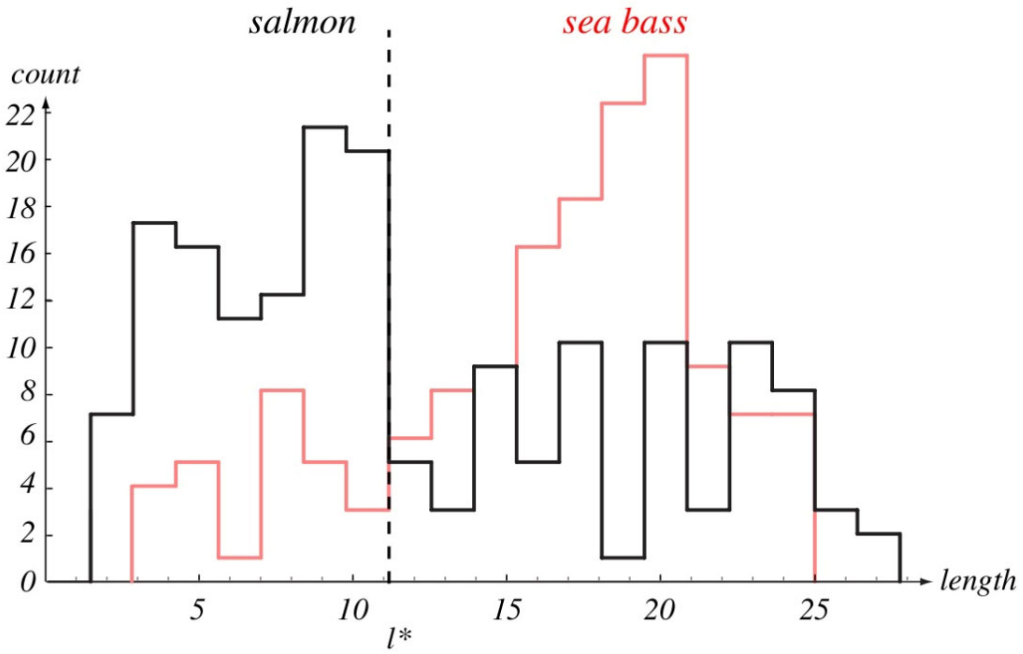




This is the set of all suggested features to explore for use in our classifier!

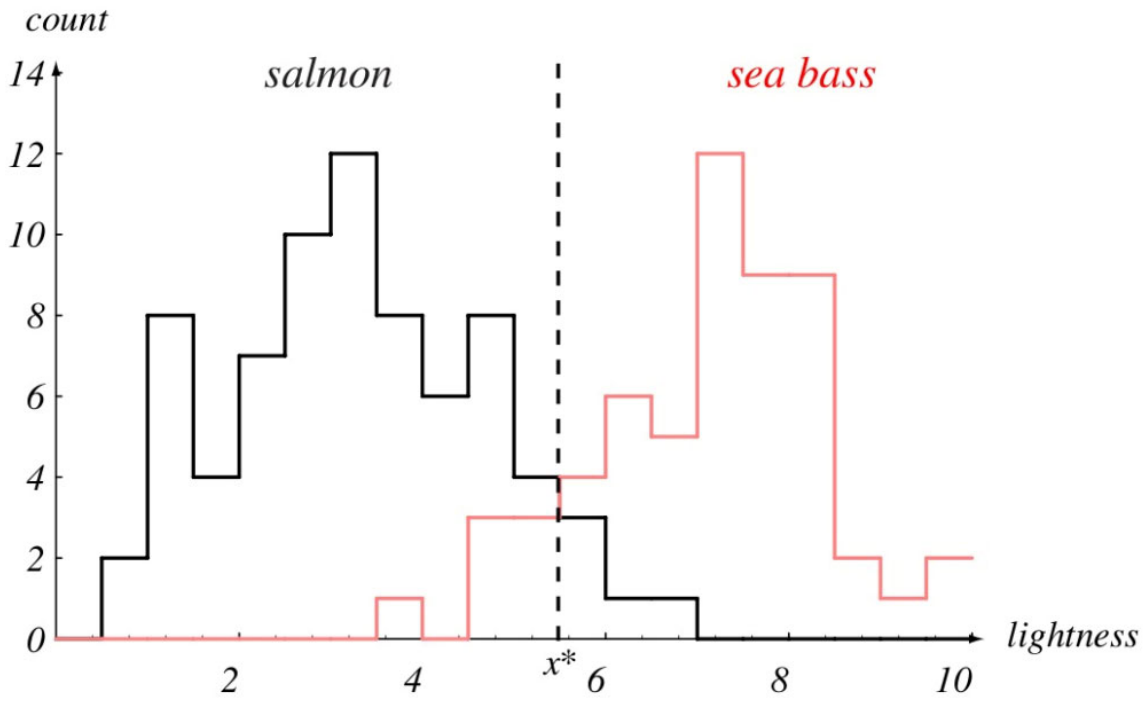
Fish Example: Feature Extraction

Possible feature for discrimination: length of a fish



Fish Example: Consider other Features

Another possible feature: lightness of a fish



Relationship between decision boundary and costs!

- Move decision boundary in response to lightness to reduce costs
- Reduces number of sea bass that are classified as salmon

Fish Example: Use Two Features

Adopt lightness (x_1) and width of the fish (x_2)

$$\text{Fish} \longrightarrow \boldsymbol{x} = [x_1, x_2]$$

$$\vec{\mathbf{x}} = \begin{bmatrix} 1.8 & 14.5 \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 1.8 & 14.5 \\ 2.5 & 15.6 \\ 3.6 & 15.4 \\ 4.5 & 15.5 \\ 5.0 & 14.8 \\ \vdots & \vdots \\ 6.7 & 21.9 \end{bmatrix}$$