Cracking the Code: Machine Learning and Medical Diagnosis

Darin Tsui*

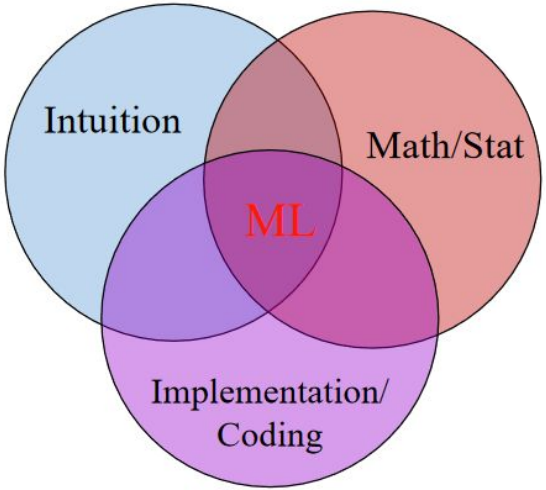
February 27th, 2023

University of California, San Diego

** Department of Bioengineering*

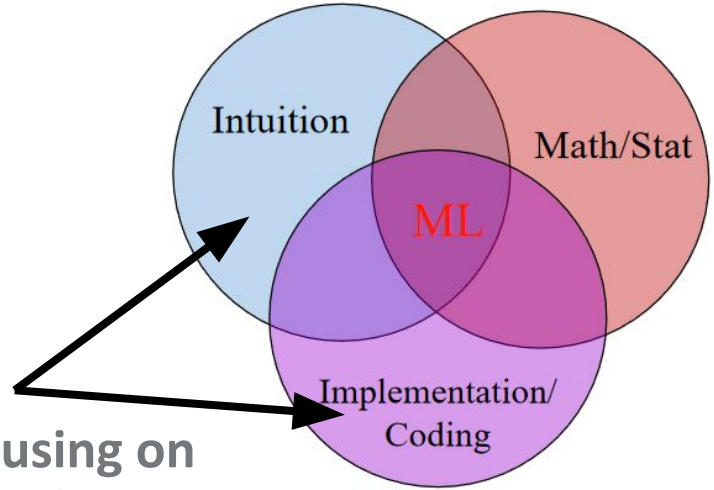
Machine Learning

The capability of a machine to imitate intelligent human behavior



Machine Learning

The capability of a machine to imitate intelligent human behavior

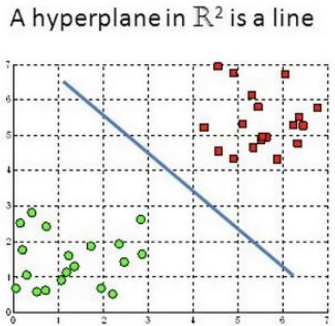


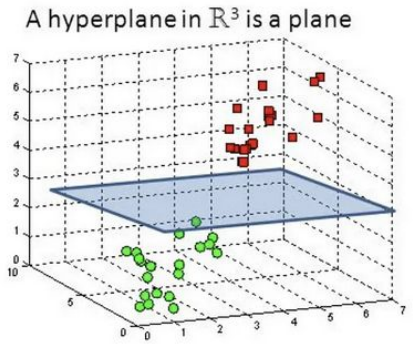
We will be focusing on these two today!

Machine Learning Classification

Intuition

Dividing data by a “decision boundary”



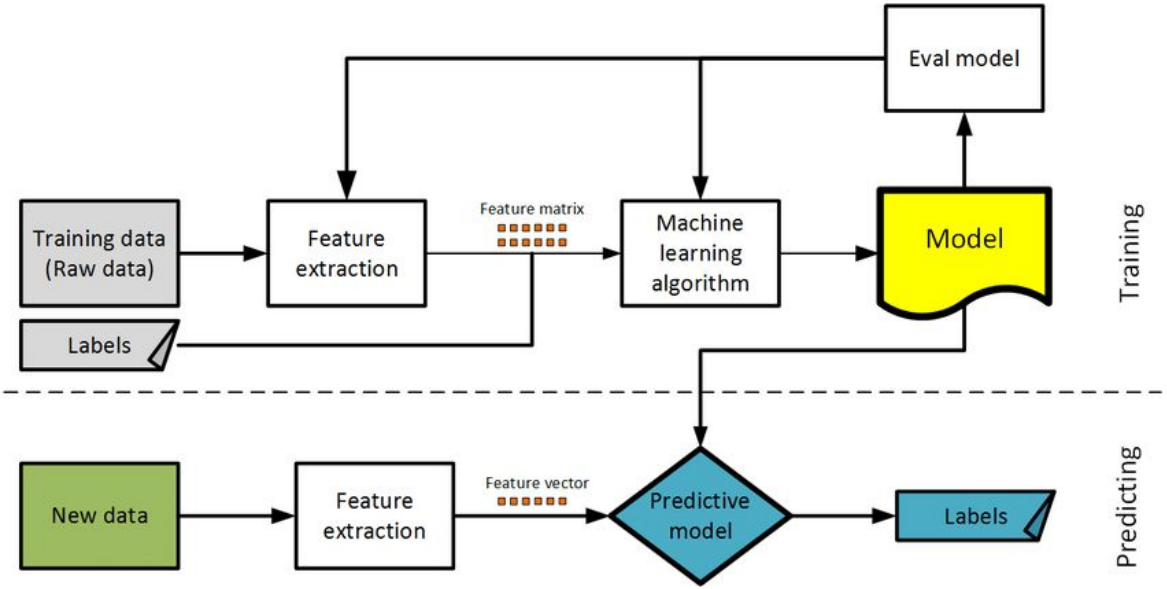


Implementation/Coding

Python packages

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import ruptures as rpt
import scipy.optimize as scp
from sklearn.preprocessing import StandardScaler
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis as LDA
from sklearn.model_selection import train_test_split
from sklearn.utils import shuffle
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import RepeatedStratifiedKFold
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
from sklearn import svm
```

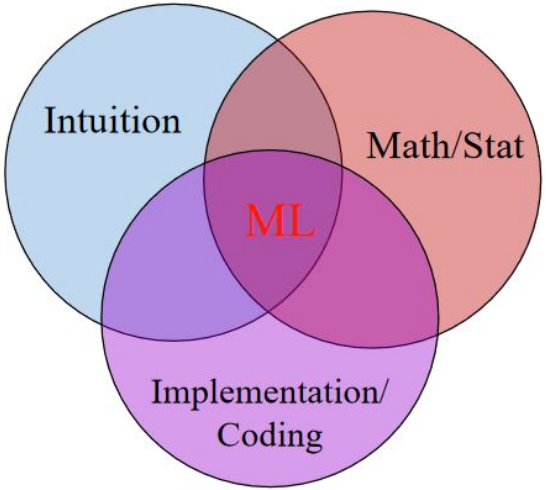
Machine Learning in a Nutshell



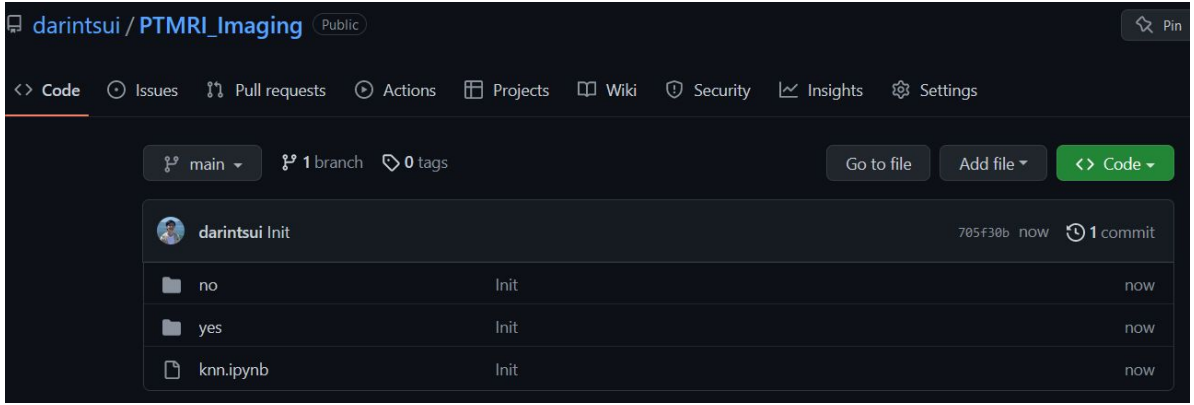
1. Raw data
2. Feature Extraction
3. Training
4. Testing

Today: Machine Learning in MRI imaging

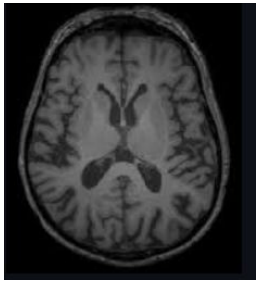
https://github.com/darintsui/PTMRI_Imaging



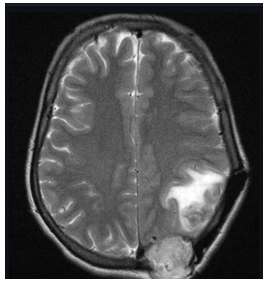
1. Raw Data


darintsui / PTMRI_Imaging
Public
Pin
Code
Issues
Pull requests
Actions
Projects
Wiki
Security
Insights
Settings
main
1 branch
0 tags
Go to file
Add file
Code
darintsui Init
705f30b now
1 commit
no Init now
yes Init now
knn.ipynb Init now


PTMRI_Imaging / no / 10 no.jpg




PTMRI_Imaging / yes / Y102.jpg



2. Feature Extraction

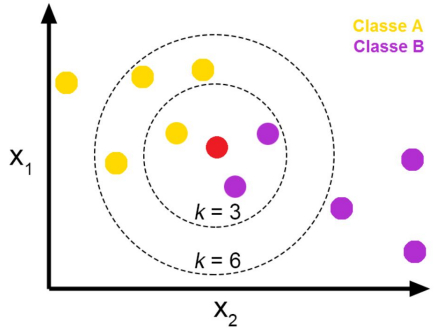
For each training image:

- Assign a label value (0 for no tumor, 1 for tumor)
- Get the RGB values of each pixel
- Resize images to be 128x128 pixels (ensures feature space is equal)
- 49152 features for each image
    - (128*128 pixels)* 3 RGB values = 49152 features

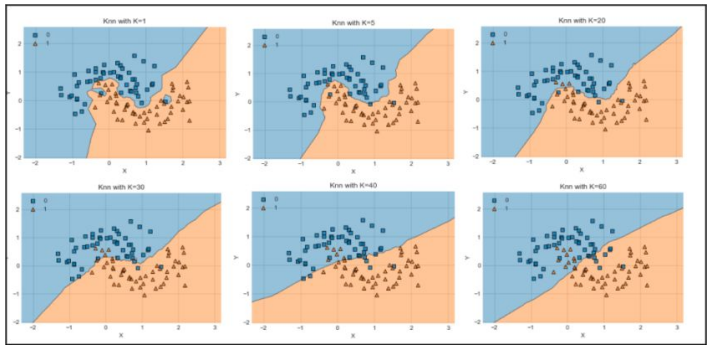
3. Train - K-Nearest-Neighbors

Method: Assigns to the majority vote of the k nearest neighbors

- Supervised machine learning algorithm
- One of the simplest machine learning algorithms



1. Calculate the Euclidean distance between all points
2. Take the nearest N neighbors
3. Pick a class by majority vote



4. Test - K-Nearest-Neighbors

Method:

1. Perform KNN on training data
2. Test model accuracy on testing data
3. Test a bunch of “neighbors” to find the optimal number
    a. In this instance: all odd numbers from 1 to 101
    b. Odd numbers are used to ensure no ties happen

So… what's next?

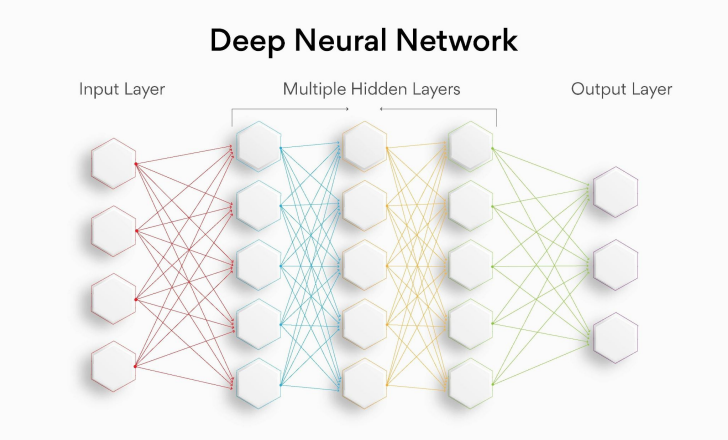
KNN was discovered in 1951, how does machine learning look now?

- In theory: grounded in mathematics, linear algebra, and probability
- In practice: much more difficult to understand and interpret

Two discoveries that drive modern-day machine learning

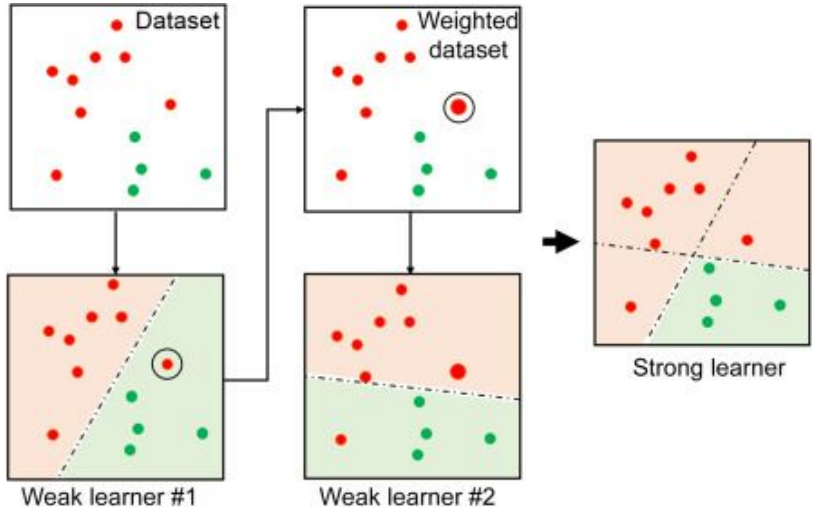
1941: Start of Neural Networks

- Biology-propelled ML



1988: Boosting

- AI automatically extracts features



Machine learning in MedTech

REVIEW PAPER

Artificial Intelligence in Surgery: Promises and Perils

Hashimoto, Daniel A. MD, MS*; Rosman, Guy PhD†; Rus, Daniela PhD†; Meireles, Ozanan R. MD, FACS*

Author Information

Annals of Surgery 268(1):p 70-76, July 2018. | *DOI:* 10.1097/SLA.0000000000002693

Metrics

Review

Neuroscience-Inspired Artificial Intelligence

Demis Hassabis [1 2], Dharshan Kumaran [1 3], Christopher Summerfield [1 4], Matthew Botvinick [1 2]

Show more

Artificial intelligence biosensors: Challenges and prospects

Xiaofeng Jin [a], Conghui Liu [a], Tailin Xu [a b], Lei Su [a], Xueji Zhang [a b]

Show more

Review | Published: 04 September 2018

Machine learning: applications of artificial intelligence to imaging and diagnosis

James A. Nichols, Hsien W. Herbert Chan & Matthew A. B. Baker

Biophysical Reviews **11**, 111–118 (2019) | Cite this article

How to get involved?

I’m BENG: BENG (mechanical engineering track), but I do primarily machine learning research

- Research!!
  - Hands-on experience with minimal prereqs
- Classes
  - COGS 18 (Introduction to Python)
  - COGS 108 (Data Science in Practice)
  - COGS 118a and 118b (Introduction to Machine Learning)
- Projects
  - Kaggle

Thank you!

dtsui@ucsd.edu