OVERVIEW

• Background
  • Why Machine Learning?
  • Why Python?
  • My Current Machine Learning Projects

• Part 1: Introduction to Machine Learning in Python: `scikit-learn`
• Part 2: Introduction to Deep Learning in Python: `TensorFlow & Keras`
• Part 3: Introduction to Deep Learning in Python: `PyTorch`
WHY MACHINE LEARNING?

- Easily build exceptionally flexible and powerful models to:
  - Predict values
  - Optimize systems
  - Classify, categorize, and cluster data
  - Parse and generate natural language
  - Generate or denoise sounds, images, video, etc.

- Use data rather than explicit programming to create models
  - ML can perform tasks not explicitly programmed
  - Generalize / robust

- Applications
  - Science of all kinds
  - Agriculture
  - Medicine
  - Robotics / Manufacturing
  - Banking / Business / Economics
  - Entertainment

AI/ML is in its infancy and will come to define an era
WHY PYTHON?

- Easy to learn
- Consistent, concise, readable code
- Collaborative
- Extensive libraries
- Platform independent
- High performance (cython)
- Massive head start (TensorFlow, PyTorch)
Sheneman L, Vasdekis AE, et al.

*Deep learning classification of lipid droplets in quantitative phase images.*

PLoS ONE 2021 16(4)

github.com/sheneman/deep_lipid

Tristin Sanchez, Luke Sheneman, Andreas E. Vasdekis et al.

*Photon-Sparse, Poisson Light-Sheet Microscopy*

ACS Photonics 2021 8 (10)
NanoDetector Workflow

MegaDetector AI Model

NanoDetector AI Model

encoded inferences

On timer

University of Idaho
Developing the NanoDetector Edge AI Device

Moultrie Camera

9W Solar Panel

Arduino Microcontroller

Jetson Nano

RockBlock Satellite Modem

19,200mAh, 75 Watt Hour Always-On Li-Ion Battery

External Antenna

Clock and Temperature
Integrative imaging of Plant Roots during symbiosis with Mycorrhizal Fungi

DOE Award #: DE-SC0022282

Sparse Photon Microscopy

“How Low Can You Go?”
NEPALESE TIGER

ACCELEROMETER DATA

Tri-Axial Accelerometer
4 months of 16 Hertz Data


Associate Professor
Ecosystem Science and Management
Geospatial Data Sciences
Conservation + Restoration

University of Michigan
Research Computing and Data Services

Welcome

Welcome to the RCDS website. RCDS was established by merging the former IBEST Computational Resources Core (CRC) with the IBEST Northwest Knowledge Network (NKN) to provide a single campus-wide resource for multi-disciplinary research computing support.

We have updated our help documentation, and plan to continue to add new information and tutorials. Please contact us with any specific documentation requests or other inquiries.

Overview

Research Computing and Data Services (RCDS) is the central provider of research computing infrastructure and services for the University of Idaho. We provide investigators with state-of-the-art high performance computing (HPC) for use in modeling, analysis, and management of research data. We support an advanced mix of parallel clusters, customized virtual machines, and powerful stand-alone servers. RCDS can manage complex data storage.

hpc.uidaho.edu
SCIKIT-LEARN

https://scikit-learn.org/

- Datasets
- Data Splitting
- Linear/Logistical Regression
- Decision Trees
- Bagging / Boosting
- Random Forest
- Gradient Boost
- Confusion Matrices / Scoring
- K-Means Clustering
- K-Fold Cross Validation
- Principal Component Analysis
- Feature Extraction
- Neural Networks
TRAIN A HANDWRITING RECOGNITION MODEL

Supervised Machine Learning
Use scikit-learn Datasets
Perform Data Splitting
Perform Feature Extraction
Use Multiple Machine Learning Methods
   1. Random Forest
   2. Gradient Boost
   3. Neural Network (MLP)
Confusion Matrices and Scoring
Compare Methods
MNIST

THE "HELLO WORLD" OF IMAGE CLASSIFICATION DATASETS

70,000 28x28 grayscale images (with labels)