Data Science Online Resources

Mathematics :

- Mathematics for Machine Learning Specialization (Coursera)
- Linear Algebra - MIT
- Calculus - MIT
- Data structure & Algorithm - Coursera
- Introduction to Probability - Edx
- Algorithm Design and Analysis- Edx
- Hasti/Tibshirani's Statistical Learning
- Daphne Koller's Probabilistic Graphical Models
- Strang's Linear Algebra Lectures
- Kolter/Do's Linear Algebra Review and Reference Notes
- Calculus 1

Machine Learning :

- Machine Learning by Andrew Ng (Coursera)
- Machine Learning Engineering Nanodegree (Udacity)
- Machine Learning Specialization series (Coursera)
- Machine Learning by MathematicalMonk (Youtube)
- Neural Network classes by Hugo Larochelle (Youtube)
- Applied Data Science with Python Specialization (Coursera)
- Intro to Machine Learning (Udacity)
- Probabilistic graphical model Specialization (Stanford)
- Machine Learning - Fast.ai
- Recommender Systems Specialization (Stanford)
- Explanation of Random Forest
- Explanation/Demonstration of Gradient Boosting
- Example of kNN

Data Science :

- Data Analyst Nanodegree (Udacity)
- Data Science Courses (Datacamp)
- Business Analyst Nanodegree - Udacity
- Yaser Abu-Mostafa's Learning From Data

Deep Learning :

- Deep Learning Nanodegree by Siraj Raval (Youtube)
- The Math of Intelligence by Siraj Raval (Youtube)
- Deep Learning Part-1 and Part-2 (fast.ai)
- Advance Machine Learning Specialization (Coursera)
- Deep Learning Specialization (Coursera)
- 6.S191: Introduction to Deep Learning (MIT)
- Neural Networks for Machine Learning by Geoffrey Hinton (Coursera)
- CS-20Si Tensorflow for Deep Learning Research (Stanford)
- CMU Deep Learning (Youtube)
- Deep Learning by Google (Udacity)
- Deep Learning Book - Youtube
- PyTorch ZeroToAll (Youtube)
- Karpathy's Stanford CS231n: Convolutional Neural Networks for Visual Recognition (Lecture Notes)
- Video Lecture's Stanford CS231n: Convolutional Neural Networks for Visual Recognition
- Colah's Informational Blog
- Bruna's UC Berkeley Stat212b: Topics Course on Deep Learning Overview of Neural Network Architectures
- GANs
- Geoff Hinton's Neural Nets for Machine Learning
- Hugo Larochelle's Neural Net lectures

Computer Vision :

- Introduction to Computer Vision (Udacity)
- 6.S094: Deep Learning for Self-Driving Cars (MIT)
- Fundamentals of Digital Image and Video Processing (Coursera)

NLP :

- NLP with Deep Learning CS224d (Stanford)
- Deep Learning for Natural Language Processing by Oxford (Youtube)

Reinforcement Learning :

- CS 294: Deep Reinforcement Learning (UC Berkeley)
- RL Course by David Silver (Youtube)

Robotics :

- Artificial Intelligence for Robotics (Udacity)

Projects Ideas :

- Analytics Vidhya for Data Science Projects
- Machine Learning Mastery by Jason Brownie
- Deep Learning Project Ideas

Books:

- An Introduction to Statistical Learning by G.L. Devore
- CLRC
- Grokking Deep Learning
- Deep Learning Book

- Machine Learning Book v2 - Packt
- Tensorflow machine Learning cookbook - Packt
- Deep Learning Tensorflow cookbook - Packt
- Analyticsvidhya resource
- Hasti/Tibshirani/Friedman's Elements of Statistical Learning *FREE*
- Barber's Bayesian Reasoning and Machine Learning *FREE*
- Murphy's Machine Learning: a Probabilistic Perspective
- MacKay's Information Theory, Inference and Learning Algorithms *FREE*
- Goodfellow/Bengio/Courville's Deep Learning *FREE*
- Nielsen's Neural Networks and Deep Learning *FREE*
- Graves' Supervised Sequence Labelling with Recurrent Neural Networks *FREE*
- Sutton/Barto's Reinforcement Learning: An Introduction; 2nd Edition *FREE*
- Bishop's Pattern Recognition and Machine Learning
- An Introduction to Statistical Learning *FREE*
- Probabilistic Programming and Bayesian Methods for Hackers *FREE*

Datasets and competitions and Resources:

Twitter:

- **@drfeifei : Fei-Fei Li - Director of Stanford AI Lab, creator of ImageNet**
- **@ylecun: (Yann LeCun) Leading Facebook AI/ML Research**
- **@karpathy : Andrej Karpathy - Teaches Deep Learning at Stanford, Research Scientist at OpenAI**
- **@AndrewYNg (Andrew Ng) : Led ML research in google, now leads AI research in Baidu**
- **@Kdnuggets (Gregory Piatetsky) : Leading KDnuggets - tweets/retweets lots of relevant stuff.**
- **@OpenAI**
- **@googleresearch**
- **@BaiduResearch**
- **@AndrewYNg**
- **@demishassabis (DeepMind)**
- **@elonmusk**
- **@distillpub (Clear and crisp explanations)**
- **@ch402 (Chris olah)**
- **@goodfellow_ian (GAN's)**
- **@jeremyphoward (fastai)**
- **@OpenAI (Reinforcement Learning)**

- **Reddit: r/machinelearning**
- **Kaggle has a lot of challenges to sink your teeth into. Some even offer prize money!**
- **The UCI Machine Learning Repository is a collection of a lot of good datasets**
- **/r/datasets has a nice place to ask for data**
- **http://blog.mortardata.com/post/67652898761/6-dataset-lists-curated-by-data-scientists lists some more datasets**
- **Here is a very extensive list of large-scale datasets of all kinds.**
- **Another dataset list**

Research Oriented Datasets:

- **MNIST A short handwriting dataset that is often used as a sanity check in modern research**
- **SVHN Similar to MNIST, but with color numbers. A sanity check in most cases.**
- **CIFAR-10/0 CIFAR 10 and 100 are two natural color images that are often used with convolutional neural networks for image classification.**

Communities

- **http://www.datatau.com/ is a data-science centric hackernews**
- **http://metaoptimize.com/qa/ and http://stats.stackexchange.com/ are Stackoverflow-like discussion forums**

Additional Tools

- **Vowpal Wabbit repository**
- **XGBoost repository**
- **LightGBM repository**
- **Interactive demo of simple feed-forward Neural Net**
- **Frameworks for Neural Nets: Keras,PyTorch,TensorFlow,MXNet, Lasagne**
- **Example from sklearn with different decision surfaces**
- **Arbitrary order factorization machines**
- **Basic SciPy stack (ipython, numpy, pandas, matplotlib)**
- **Jupyter Notebook**
- **Stand-alone python tSNE package**

- **Libraries to work with sparse CTR-like data: LibFM, LibFFM**
- **Another tree-based method: RGF (implemetation, paper)**
- **Python distribution with all-included packages: Anaconda**
- **Blog "datas-frame" (contains posts about effective Pandas usage)**

StandCloud Computing:

- **AWS, Google Cloud, Microsoft Azure**
- **AWS spot option:**
  - **Overview of Spot mechanism**
  - **Spot Setup Guid**
  - **GCP cheat sheet**

Good journals for ML papers are the Journal of Machine Learning Research, the Journal of Machine Learning and arxiv.

Public access to Research Papers: Sci-Hub