Machine Learning with Python

A Session to Get You Started

Himanshu Mittal

himanshu.mittal224@gmail.com
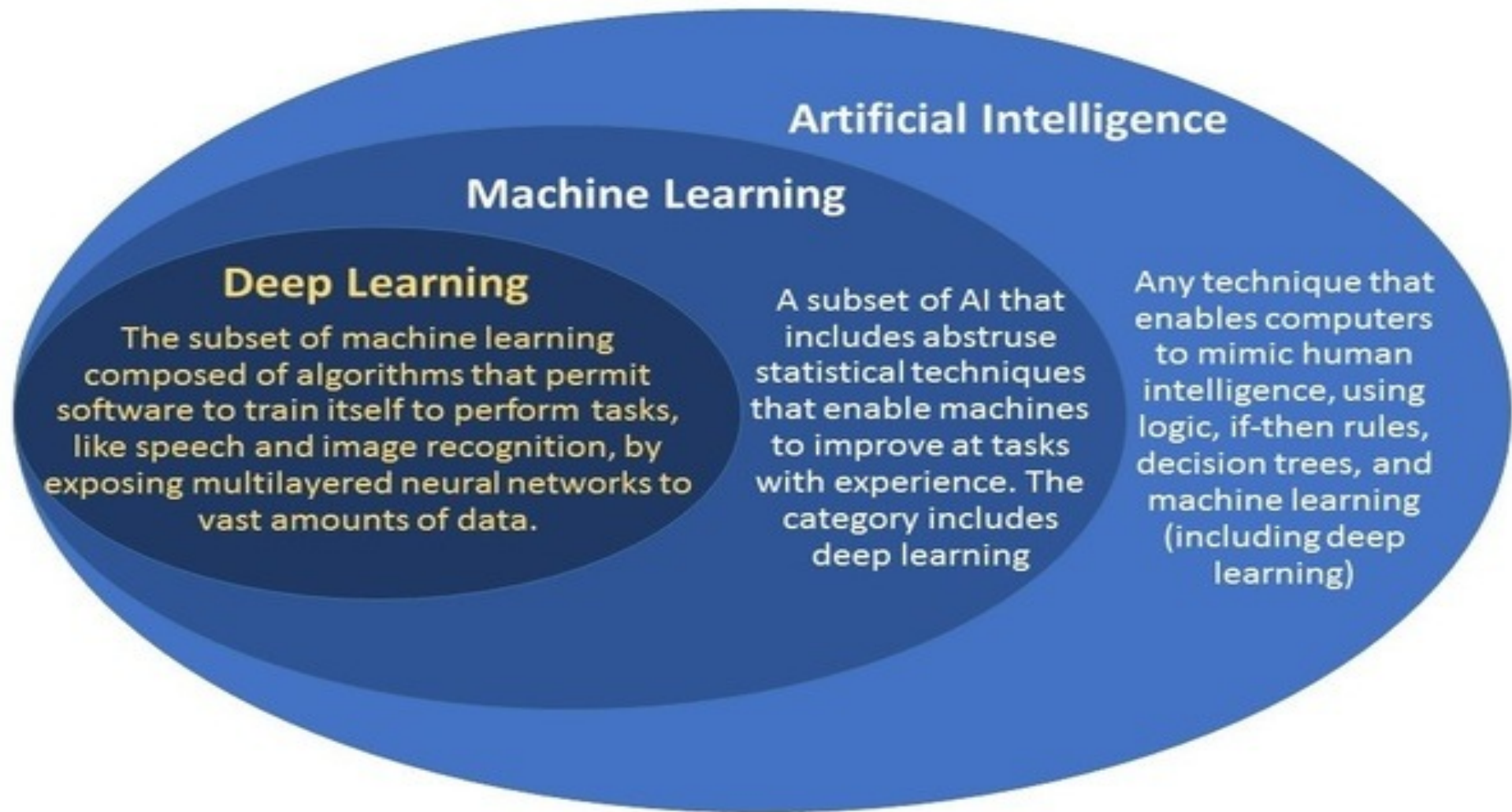
Artificial Intelligence

Machine Learning

Deep Learning

The subset of machine learning composed of algorithms that permit software to train itself to perform tasks, like speech and image recognition, by exposing multilayered neural networks to vast amounts of data.

A subset of AI that includes abstruse statistical techniques that enable machines to improve at tasks with experience. The category includes deep learning

Any technique that enables computers to mimic human intelligence, using logic, if-then rules, decision trees, and machine learning (including deep learning)

Machine Learning with Python

1. Why learn Machine Learning?

2. Machine Learning – The Past, The Present and Future

3. Machine Learning Process

4. Supervised and Unsupervised Learning

5. Deep Learning

6. How and Where to Start?

7. Popular Repositories, Tools and Frameworks

8. Q & A and Hands On

Why Machine Learning?

1. **Employability** – Data science and machine learning jobs are growing in the market. Nearly all organizations operating at a global scale have introduced ML into their process and every passing quarter new organizations are getting added to this list. So, ML will not take away jobs but create many more.

2. **Social and Business Impact** – Machine learning can help eliminate a lot of hassle from the lives of the general public. Particularly in the domains of healthcare, governance, cyber-security, etc.

3. **Fast processing** – We live in the age of social media. Social media is known to generate huge amounts of data which carries valuable information impossible to process and use without the use of techniques like machine learning.

Why Machine Learning?

4. **Big Data** – While social media is generating a lot of data, companies like Amazon and Google are already housing datastores of the size of petabytes.

5. **Hobby** – Learn Machine Learning because you want to learn something new.

6. **Research** – Take it up as a challenge and delve into the finer details because machine learning is complex and has a lot of scope for research in terms of optimization and performance.
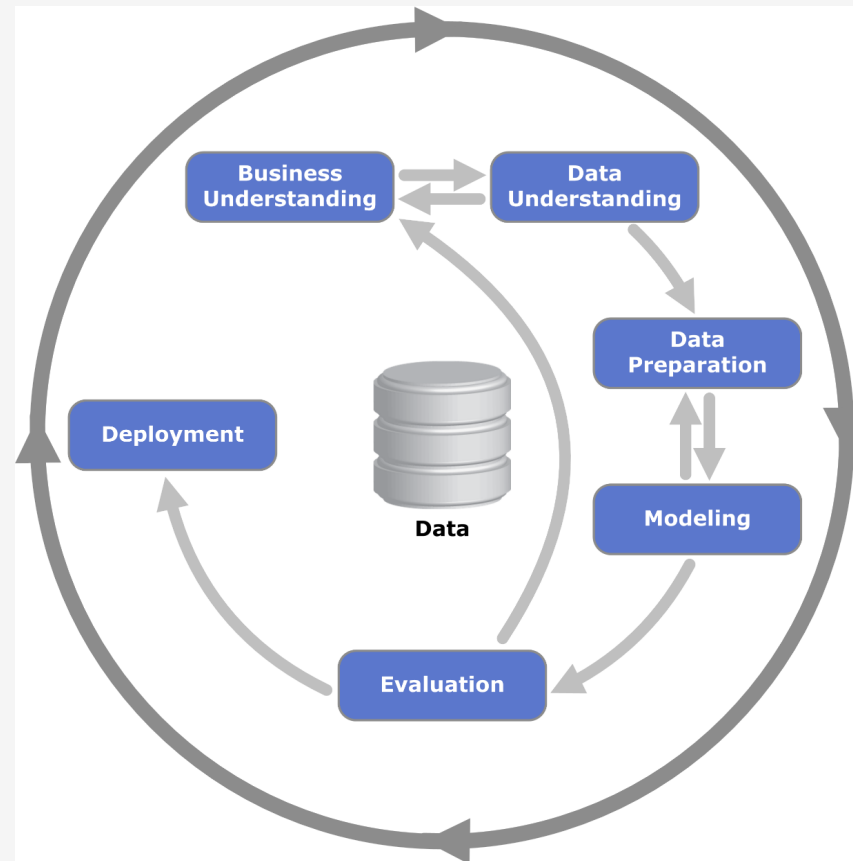
Machine Learning with Python

History of Machine Learning

1. 1952 – First learning program written.

2. 1957 – The perceptron designed.

3. 1967 – Nearest neighbor algorithm was written.

4. 1979 – Stanford Cart could detect obstacles.

5. 1990 – Machine learning moved from knowledge-driven to data-driven.

6. 1997 – IBM Deep Blue beats the world champion in chess.

7. 2012 – AlexNet competed in ImageNet challenge.

Machine Learning with Python

The Machine Learning Process

The Machine Learning Process

1. **Business Understanding** – What is the objective of the project? What goals are to be met with the final product?

2. **Data Understanding** – Procure the right data if possible. Regardless of whether the data available is self procured or pre-made. Understanding the information available within the data is important. Evaluate each parameter for correlation with other parameters and the outcome.

3. **Data Preparation** – Redundancy in any computing scenario is a performance bottleneck. Also, redundancy can sometimes impact the results of a computing process. So is the case with machine learning.

The Machine Learning Process

4.   **Modeling** – Finding the right algorithm for your problem and training the algorithm using the optimized data obtained in step 3. Incorrect choices for learning algorithms can have severe impact on the outcome.

5. **Evaluation** – Testing the model fitment using a dataset previously unseen by the model.

6. **Deployment** – Once the accuracy of the model has been evaluated and found to be acceptable for the given problem and business case, the model is deployed to production use.

Choosing the Right Algorithm
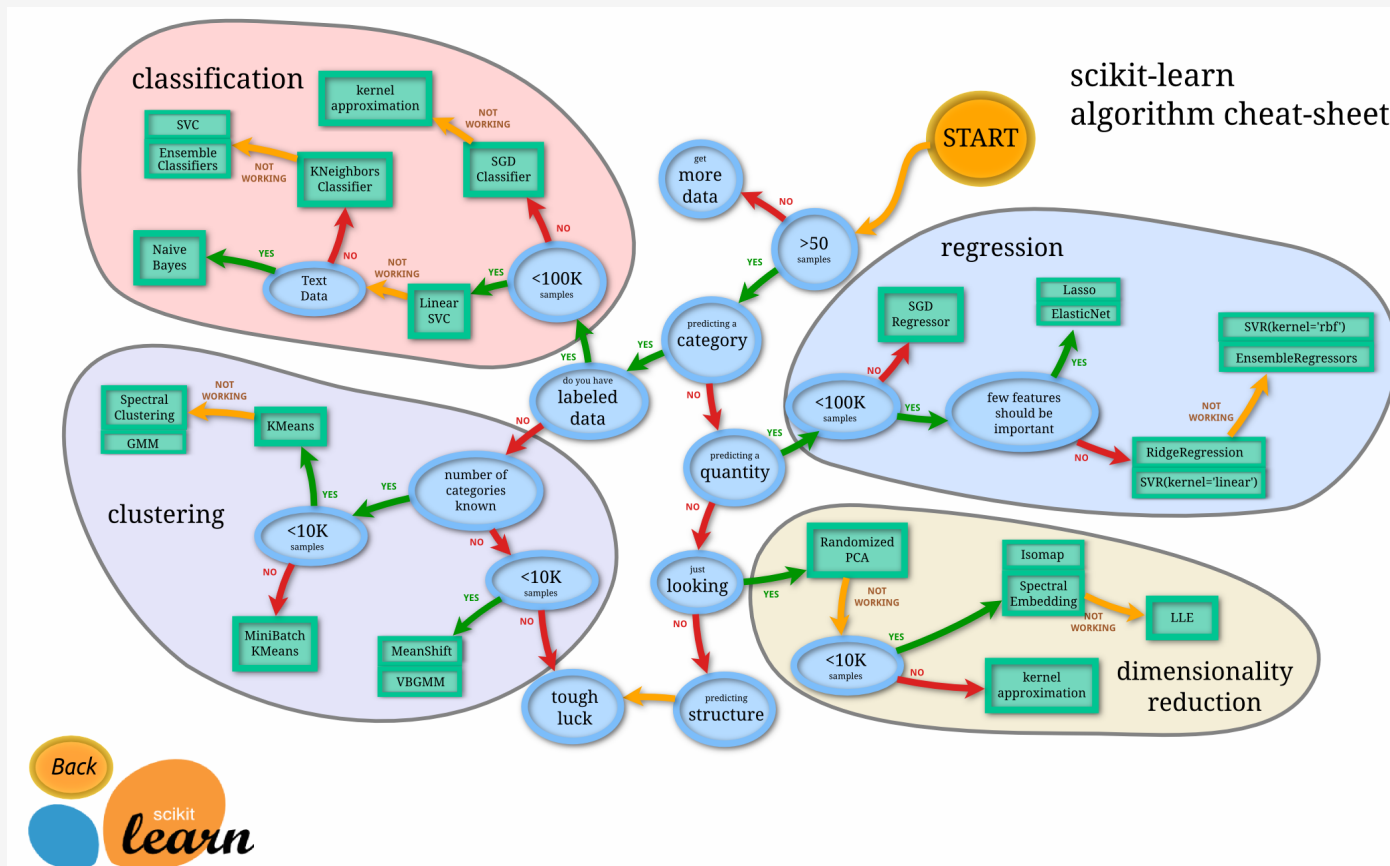
Choosing the Right Algorithm

The cheat sheet is available at,

http://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Machine Learning with Python

1. Why learn Machine Learning?

2. Machine Learning – The Past, The Present and Future

3. Machine Learning Process

4. **Supervised and Unsupervised Learning**

5. Deep Learning

6. How and Where to Start?

7. Popular Repositories, Tools and Frameworks

8. Q & A and Hands On

Supervised Learning

1. Models are trained before deployment.

2. Supervised learning needs well constructed training data to be effective.

3. The goal of supervised learning is to assign an unseen data element to one of the labels the model is trained for.

4. Hence the supervised learning problem is also called classification.

5. Supervised learning is used to make decisions. Like detecting a spam email.

6. Some popular supervised learning techniques are Decision Tree, Support Vector Machines.

Unsupervised Learning

1. Models are not trained before deployment.

2. Unsupervised learning does not require training data.

3. The goal of unsupervised learning is to create clusters or segments of similar objects within any data elements input to the model.

4. Hence the unsupervised learning problem is also called clustering.

5. Unsupervised learning is used for creating patterns. Like finding outliers in a data set.

6. A very popular algorithm in this space are k-means clustering.

Deep Learning

1. Establishes a deeper connect between data and outcome by use of multiple hidden layers.

2. Deep learning is nothing but Neural Networks with multiple hidden layers.

3. Each hidden layer is dependent on the layer below for input thereby removing the dependency on actual data.

4. Hidden layers establish an abstract relationship between data and outcome.

5. This learning approach is mostly inspired by the human learning process.

Machine Learning with Python

1. Why learn Machine Learning?

2. Machine Learning – The Past, The Present and Future

3. Machine Learning Process

4. Supervised and Unsupervised Learning

5. Deep Learning

6. **How and Where to Start?**

7. Popular Repositories, Tools and Frameworks

8. Q & A and Hands On

How and Where to Start?

1. **Start with Jupyter** - A python data science platform that enables code writing and execution. It supports inline plotting and output. A very good place to start getting your hands dirty.

2. **Read the book** – Introduction to Statistical Learning (http://www-bcf.usc.edu/~gareth/ISL/).

3. **Learn Python** – Follow the official documentation of Python (https://docs.python.org/3/).

Popular Python Packages

1. **numpy** – High level mathematical computation with Python

2. **scipy** – Python library for scientific and engineering computation

3. **pandas** – Python library for data analysis

4. **scikit-learn** – Python ML package for ML models (algorithms)

5. **TensorFlow** – Google project for Deep Learning (https://www.tensorflow.org/)

6. **Anaconda** – fully functional Python ML framework the installs most of the popular packages and dependencies.

Popular Repositories and Resources

1. UCI Machine Learning Repository - https://archive.ics.uci.edu/ml/datasets.html

2. Scikit-learn datasets – available with scikit-learn

3. Kaggle - https://www.kaggle.com/

INDY / TECH

FACEBOOK'S ARTIFICIAL INTELLIGENCE ROBOTS SHUT DOWN AFTER THEY START TALKING TO EACH OTHER IN THEIR OWN LANGUAGE

Questions?

Thank You !