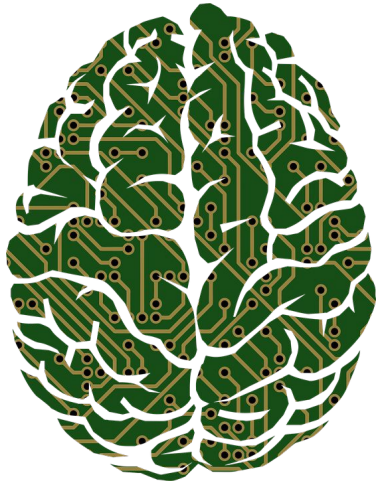
AI for Everyone



By: Ashish Salunkhe
At: AICVS, Cummins, Pune
13th March, 2020

About Me

Currently SWE at Persistent Systems, Pune

Organizations worked with: Taiyo AI Inc., Crysagi Systems, CoreView Systems and Infralytiks LLC.

Completed Undergraduate Degree in Computer Engineering from PCCOER, Pune

Founder of PCCOER ACM Student Chapter.

Working on ML, ML-Ops and Full Stack Web Applications.

Current Scenario

Culture

Img Src: Blog by Mashable

The AI Meme Generator Is Better At Making Memes Than Humans

Why?

How is AI helping the fight against COVID-19?

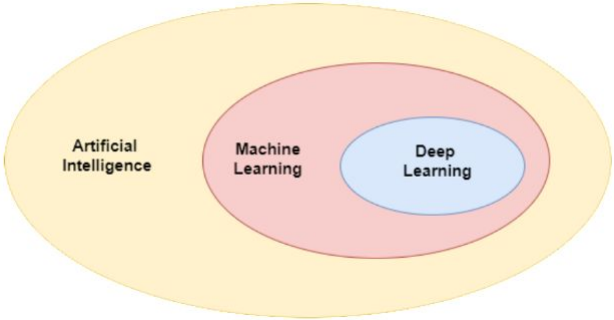
Spreading of Misinformation

Identifying Mask Violators

Early outbreak warning system

Peer-to-peer AI-tracing of COVID-19 (Contact tracing and Covid strain tracing)

What’s Machine Learning and AI?



Machine Learning:

Machine learning is about teaching computers how to learn from data and experiences to make decisions or predictions based on performance metrics and identify patterns without being explicitly programmed.

Artificial Intelligence:

Developing intelligent systems.

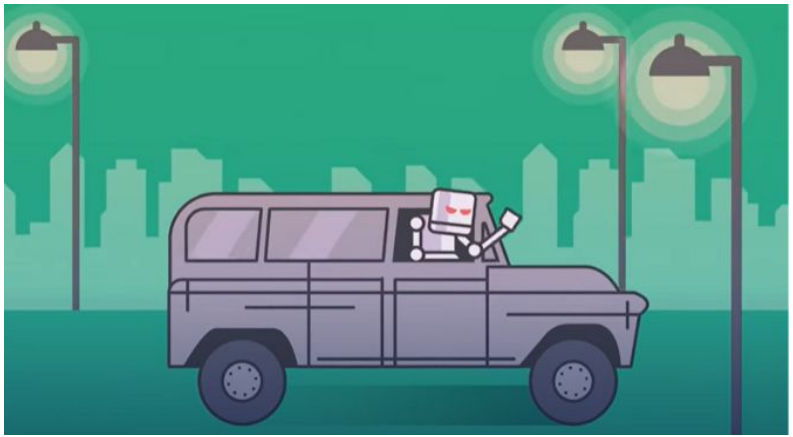
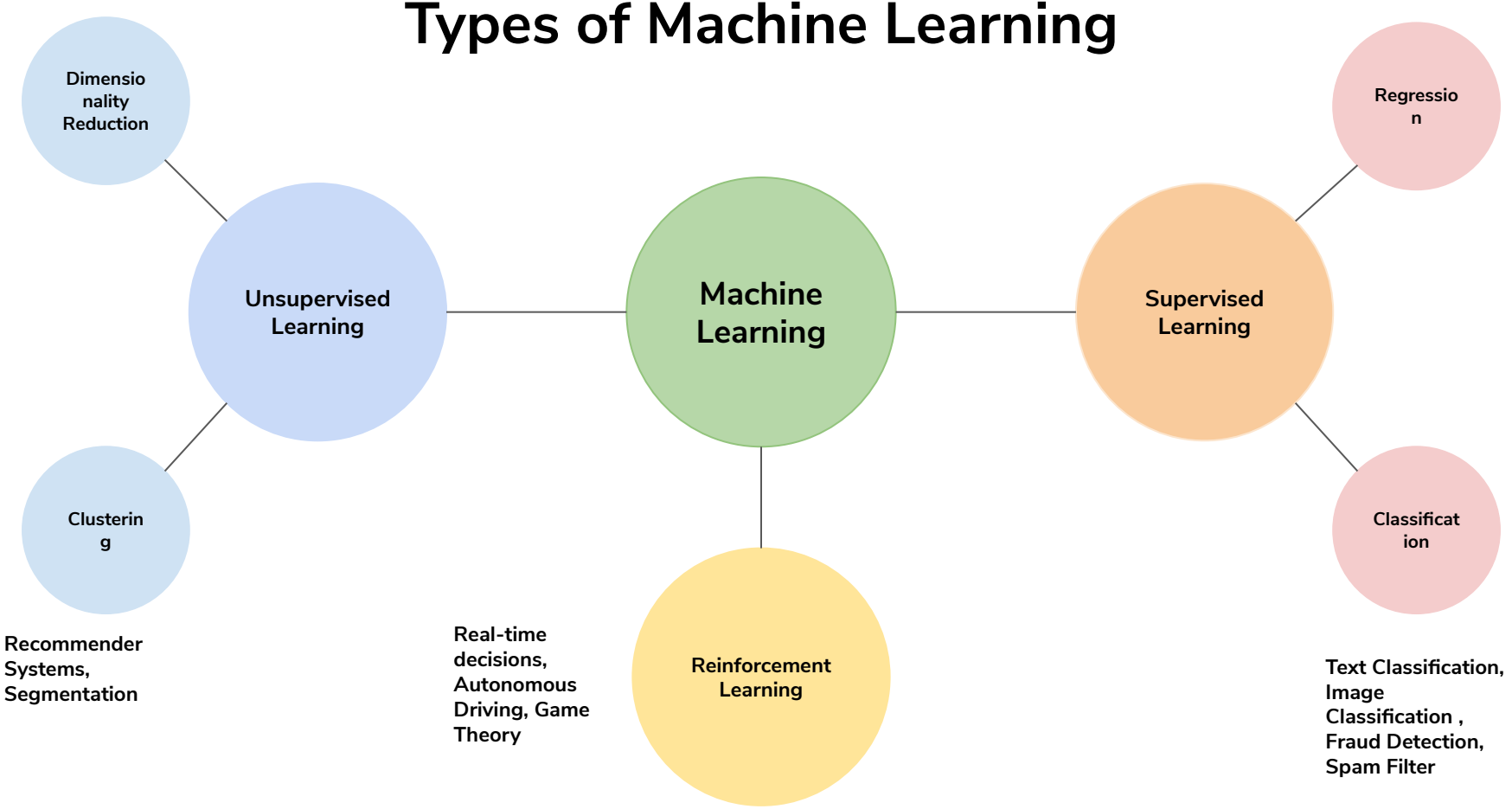


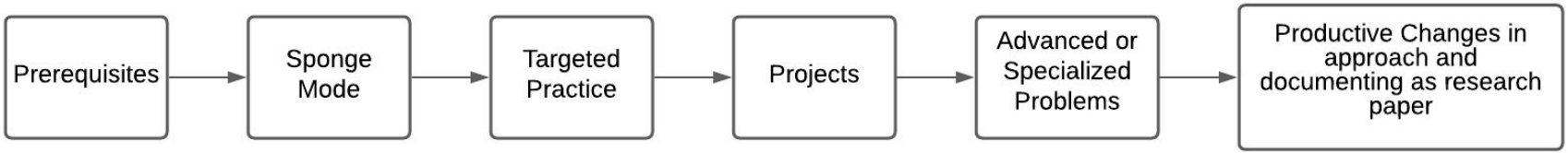
Image from Artificial Intelligence HubSpot

Types of Machine Learning



How to Start?

How should your path look like?



Typical Machine Learning Workflow

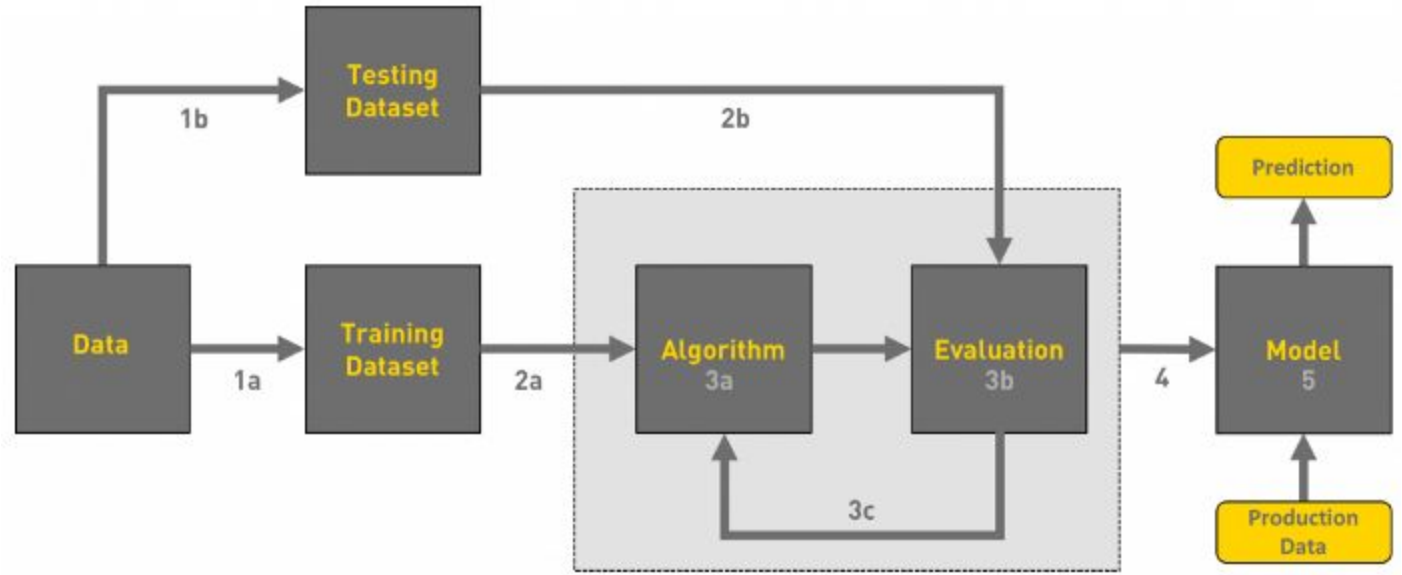


Image from [Workflow Blog by Ayush Pant]

Prerequisites

Python

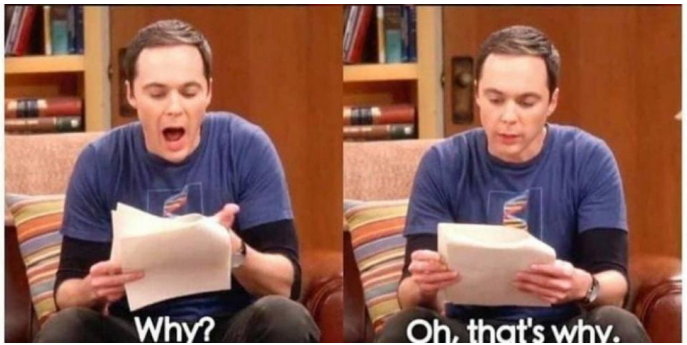
Understanding statistics, especially Bayesian probability

Linear algebra and multivariable calculus.

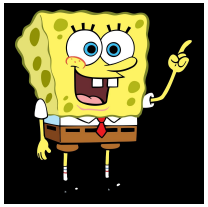
You don't need to be a professional mathematician or veteran programmer to learn machine learning, but you do need to have the core skills in those domains

"If I don't plan to perform original research, why would I need to learn the theory when I can just use existing ML packages?"

- Data Collection
- Data Preprocessing
- Interpretation of Model Results
- Optimization and Parameter Tuning
- Understanding business problems and driving the business value



Sponge Mode



~~Spongebob~~
Sponge mode

1. Doing Appropriate Coursework
2. Studying Reference Books (only if you want to learn more)
3. Data Mining : Concepts and Techniques, by Jiawei Han et al
4. Andrew Ng's course on Machine Learning
5. Elements of Statistical Learning, by Freidman et al
6. Deep Learning, by Ian Goodfellow

Stanford's Machine Learning Course by Andrew Ng

Machine Learning with Python

Later, also you "can" enroll into Deep Learning Specialization — Coursera

Apply for GitHub Student Pro, you get access to DataCamp - Has wonderful hands-on exercises

“I’ve completed the courses, got myself certified. What next? Do certifications matter?”

Targeted Practice

Exercises to hone your skills!

1. Implementing PoCs
2. Practice the entire machine learning workflow
3. Practice on real datasets - Google Datasets, Kaggle
4. Deep dive into specific topics

What About Tools?

1. Pandas
2. Scikit-Learn
3. Numpy
4. Visualization Packages like Matplotlib, seaborn

Topics that you must understand and have a hands-on practice

Data Preprocessing	Sampling and Splitting	Supervised Learning	Unsupervised Learning
Model Evaluation	Ensemble Learning	Bias-Variance Tradeoff	Hyperparameter Tuning and Optimizations

The Business Application

“You’ve covered the prerequisites, completed the courses, implemented PoCs, let’s get into the fun part!”

Projects

1. *Complete the projects below. (In the given order - order is based on difficulty)*
    - *Classification Problem*: Titanic Challenge
    - *Regression Problem*: House Price Prediction
    - *Computer Vision*: Hand Digit Recognition
    - *Image Processing*: Facial Key Points
    - *Natural Language Processing*: NLP Tutorial
2. *Implement Algorithms from Scratch*
3. *Participate in Kaggle Competitions and online hackathons*
4. *Based on your own expertise as you progress read and implement notable arXiv papers through "PapersWithCode"*

Refer Projects implemented under Stanford CS229 Course

Research

- Make your own productive incremental changes in existing work and document the work in the form of research paper.
- Read Papers and try to replicate results
  - arXiv
  - NeurIPS
  - ICML / ICLR

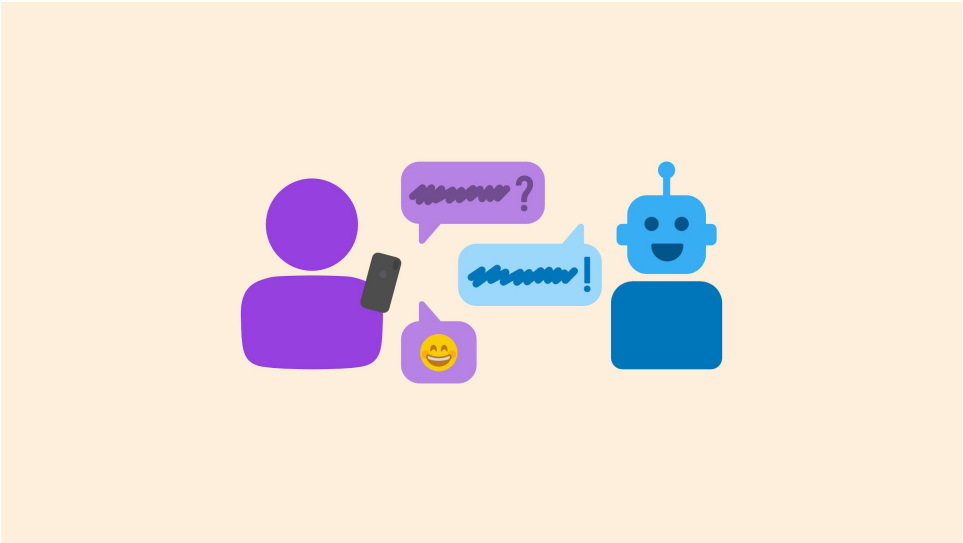
My Experiences with ML

- Visual Assistant for Visually Impaired
- Detection of Fake Reviews on Online Review Platforms
- Dickens Speaks
- Customer Support Intelligent Chatbot

Customer Support Chatbot

Motivation:

- Overhead of support tickets not being addressed.
- A solution to assist customers 24*7
- Resolve problems faster
- Automate answering repetitive questions to let the support team focus on more challenging cases.



Design

- A utility implemented to build models from the database content.
- The chatbot framework loads a prebuilt predictive model and connects to database to retrieve documents which contain possible responses and context information.
- Use database to store documents containing
  - a defined classification or name of user input. this is the intent of the input/response interaction
  - a list of possible responses to send back to the user
  - a context value of the intent used to guide or filter which response lists makes sense to return
  - a set of patterns of potential user input.
  - the patterns are used to build the model that will predict the probabilities of intent classifications used to determine responses.

Fun with ML/AI

Here are a few sites you can look at:

1. https://www.machinelearningisfun.com/
2. https://medium.com/creative-ai
3. https://experiments.withgoogle.com/collection/ai

Need of AI + X: Don’t Switch Careers, Add AI

- Companies look for AI+X.
- Solving real-world problems with AI requires subject-matter expertise.
  - Kian Katanforoosh and Andrew Ng reflect that 2/3rd of students taking their class are from majors outside the computer science department — everything from chemical engineering to astrophysics to law.
  - Why are these students delivering high quality AI projects:
    - Subject-matter expertise
    - Access to Data
    - Time and Passion
- Takeaways -
  - Spot when a problem is suited for an AI solution.
  - Solve real-world problems with AI.
  - Get an edge in the job market.
  - Keep doing what you love while learning new technologies.

Topic Ref.: Blog by Kian Katanforoosh

Profile Building and Networking in ML

While a resume matters, having a portfolio of public evidence of your data science skills can do wonders for your job prospects

How do you get experience if you need experience to get your first job? - Projects!

--

Make a Portfolio Website

Master Resume and a Job Resume

Participate in Hackathons

Publish Papers in Good journals and top tier conferences

Blogs and Talks

Networking?

- **Twitter**
- **LinkedIn**
- **Conferences**
- Contact PhD Students to get in touch with Professors

Good Blogs for Portfolios and Profile Building

- Advice on Building Data Portfolio Projects
- How to Build a Data Science Portfolio

Jobs in DS/ML/AI

- Data Scientist
  - Explore data, validate hypotheses, statistical modeling, PoC
  - Skills: Python, Scikit-Learn, Pandas, Numpy, Statistics
- Data Analyst
  - Business Decisions, Visualization
  - Skills: SQL, Tableau, Python
- Data Engineer
  - Storage, Microservices, Optimization
  - Skills: Database Theory, Distributed Systems, Scala, SQL
- Machine Learning Engineer
  - Productionizing ML systems at scale
  - Skills: Python, TensorFlow, PyTorch, Algorithms, DS

How could you land up with these jobs?

1. Higher Studies Abroad
2. Work at Startups in India
3. Foreign Transfer - Work with Tech Giant in India and based on project requirements sponsor visa abroad

Slide Reference: Rajvardhan Oak, Talk on MS Session.

Recommended Plan during undergraduate degree

- FE: Explore domains and clubs - learn Excel, Latex and Python
- SE: Stick to one tech and non-tech club, pick up an application development skill
- TE: Depending on your goal - placement or MS, start doing competitive coding or depth specialization respectively and participate in hackathons.

  Add projects to your portfolio, develop GitHub profile and go for internships.

- BE: Continue internships online / offline (if college permits), independent projects or freelance, present papers at conference and your final BE Project
- If aiming for MS immediately after bachelor's take GRE in TE Sem 6 and TOEFL in BE Sem 7 and apply to universities before December start (deadlines Dec 15th - Jan 15th every year for Fall Intake)

Questions? Comments?