AI & ML Club IITGN

Why did we start this club?

The discipline of AI/ML has had tremendous growth recently, with several researchers and engineers entering both academia and industry. With improved and more affordable computational power and a colossal amount of data generated every day, AI and ML are finding broader and deeper applications in every area. However, at IIT Gandhinagar, we lacked a dedicated club to cater to the large community of students interested in this field. This club strives to address this void by being the one-stop destination for IITGN students passionate about AI/ML.

The club recognizes the diverse aspirations of students in academia, industry, and entrepreneurship, and thus it focuses on catering to the various career directions in AI/ML. The vision of the club is to provide content, guidance, and practical experience to students of IITGN in AI/ML, building a community of like-minded individuals.

Before getting into the Core ML:

Let's get the Basics Done!

1. Learn the basics in a minute - What's AI - YouTube Playlist
2. For the people who prefer ***Reading***:
    - 5 Beginner Friendly Steps to Learn Machine Learning and Data Science with Python - Daniel Bourke: [LINK]
    - What is Machine Learning? - Roberto Iriondo: [LINK]
    - Machine Learning for Beginners: An Introduction to Neural Networks - Victor Zhou: [LINK]

Now you know a general idea in the field, and it's the time to start:
*(**Don't neglect Math or Statistics and Probability!** It's important to know the general math to become an expert in Machine Learning)*

3. Machine Learning Math Resources:

- Linear Algebra - Khan Academy: [LINK]
- Statistics and probability - Khan Academy: [LINK]
- Multivariable Calculus - Khan Academy: [LINK]

Resources

Python

Python is a general-purpose language that is it is designed to be used in a variety of applications, including machine learning, data science, software and web development, automation, and generally getting stuff done.

Python is the Lingua Franca of AI (at the moment). To get started, the Google guide or the Sentdex tutorial series can be helpful. The official documentation of Python and the books Think Python and Learn Python the Hard way are also good references.

Python, unlike other languages, supports data manipulation through many libraries to make our job easier. It is robust in working and easy to understand and apply, making it the primary language for data science and machine learning.
To check how good you are at Python, you can try **this quiz** before you move on to the next section :)

Software and Frameworks

Jupyter Notebook

The Jupyter Notebook is an open-source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. This YouTube video and this tutorial contain useful information on how, to begin with the Jupyter Notebook. Install the Jupyter Notebook on your local PC using these guidelines.

Google Colab

Google Colab is similar to Jupyter Notebook, except that it can be used when multiple people are working on a project to collaborate efficiently in the code. Just like we collaborate on Google Docs, you can co-code with various people using the Google Colab notebook. Since colab

enables multiple people working on the same code, it requires an internet connection. This video will help you get started with Google Colab. .

Git

An integral part of most software projects, Git is a tool for version control and managing changes to code. This is a great overview guide for a better thorough introduction, check out this course for a more detailed intro. Make sure to register on GitHub for the Student Developer Pack.

Conda

When working on several projects at once, managing various packages might be difficult (especially in Python). Conda is a very functional package manager that makes life considerably simpler (along with numerous more utilities of a similar nature). Learn how to use it here or here, and check out the documentation.

Linux Terminal

The Linux terminal is a text-based interface used to operate Linux computers. It's just one of the many tools offered to Linux users for accomplishing any given task, but it is often regarded as the most effective approach.
Although using any other terminal won't pose many problems, using a Linux terminal is advised for efficient workflow. The following will help you to get started (1, 2).

Data Manipulation and Analysis

NumPy

NumPy is a Python library used for working with arrays. This is essentially MATLAB for Python. It is advisable to learn NumPy after completing the course by Andrew NG. To start with, the official tutorial of NumPy or learn it with a hands-on deep learning experience via the first course of deeplearning.ai: Neural Networks and Deep Learning. Resources: Part-1, Part-2

Pandas

Pandas is a software library written for the Python programming language for data manipulation and analysis. This is one of the most crucial and potent libraries in data science.You eventually learn how to work with Data Frame rows, columns, and contents as well as how to read and

write CSV and JSON files. Again Sentdex and the official documentation of Pandas will be of great help here.

Data Visualization

Matplotlib

Matplotlib is a library useful in data visualization and plotting graphs. The pyplot submodule ought to be enough for the beginning. As always, Sentdex and the documentation will assist you here.

Seaborn

Seaborn is based on Matplotlib and is closely integrated with NumPy and Pandas. It has various dataset-oriented plotting functions that work with data frames and arrays that contain entire datasets. The documentation and this tutorial will significantly help the initial journey.

Resources for ML and AI

We suggest Bishop's Pattern Recognition and Machine Learning Textbook for Core Machine Learning and Russel and Norvig's Artificial Intelligence: A Modern Approach for Core AI. Both of them are very comprehensive textbooks that provide clear explanations of the concepts in both fields. You can use them more as reference materials and learn about particular subjects one at a time rather than having to read them from start to end in one sitting.

Recommended Courses on Machine Learning

- Most AI enthusiasts start with this course, which also introduces you to the mathematics underlying machine learning. The assignment calls for the use of ML methods in Octave/MATLAB. It is recommended to try the Python versions of these exercises instead, which are available here. (Andrew Ng).

- Introduction to machine learning: YouTube Playlist (Stanford)
- Intro to Machine Learning: Kaggle
- Want to get in-depth knowledge about Machine Learning for Engineering and Science Applications, then do refer to this NPTEL playlist.

Deep Learning & Computer Vision:

For those new to Deep Learning, we recommend you start with the Stanford CS231 (**Computer Vision**), available both on YouTube and on the official website course, which covers everything from loss functions and basics of Neural Networks to Computer Vision. It is advised to go through the video lectures and attempt the respective assignments to understand various fundamental concepts thoroughly. The Deep Learning Book is also an excellent resource for detailed explanations of various subjects.

For those interested in getting into specific areas -

Natural Language Processing:

Stanford CS224 [YouTube Link] [Course Link]

Reinforcement Learning:

Berkeley CS285: [YouTube Link] [Course Link]

NPTEL [YouTube Link]

Sutton and Barto's Classic Introductory Text [Book Link]

Important Frameworks:

Tensorflow

Tensorflow is an open-source library for machine learning and artificial intelligence. It majorly focuses on the training and inference of deep neural networks. It basically simplifies the process of numerical computing. The Introduction to Tensorflow for Artificial Intelligence, Machine Learning, and Deep Learning is a helpful course on Coursera to learn Tensorflow. Also, this YoutTube Tutorial can be of great help.

PyTorch

PyTorch is a Machine Learning framework created to provide tools for writing Deep Learning research and application code. The official tutorials can give an excellent hands-on experience with PyTorch.

Points to follow:

- It is highly recommended to discuss with your peers and mentors different approaches and concepts to solve a particular problem.
- If you don't understand the concepts, revisit them from the resources mentioned or just google them online. This is a vast and growing field with lots of constant development. Have a look at a few examples.
- There will be a lot of shortcuts that can make the task easier. Here lies the art of googling. You'll always need to google things to get a different perspective on how the world solves a particular problem. Don't think of it as plagiarism. Stackoverflow is a good resource for that.
- For more insightful research articles and blog posts on Machine Learning and Data Science, read **towardsdatascience.com** and **medium.com** articles.